

<b>Data Collection</b>	<GAIAESO/batch_1212>
<b>Release Number</b>	<3>
<b>Data Provider</b>	<G.Gilmore, S.Randich>
<b>Date</b>	<25.08.2016>

## Abstract

Gaia-ESO is a large public spectroscopic survey<sup>1</sup> carried out with FLAMES, targeting  $\geq 10^5$  stars, systematically covering all major components of the Milky Way, from halo to star-forming regions, providing the first homogeneous overview of the distributions of kinematics and elemental abundances. This alone will revolutionise knowledge of Galactic and stellar evolution: when combined with *Gaia* astrometry the survey will quantify the formation history and evolution of young, mature and ancient Galactic populations. With well-defined samples, we will survey the bulge, thick and thin discs and halo components, and open star clusters of all ages and masses. The UVES and Giraffe spectra will: quantify individual elemental abundances in each star; yield precise radial velocities for a 4-D kinematic phase-space; map kinematic gradients and abundance - phase-space structure throughout the Galaxy; follow the formation, evolution and, dissolution of open clusters as they populate the disc, and provide a legacy dataset that adds enormous value to the *Gaia* mission and on-going ESO imaging surveys.

## Overview of Observations

This release of the Gaia-ESO Survey (GES)<sup>2</sup> covers observations obtained by the Survey in the period 31.12.2011-19.07.2014. These include Milky Way field observations, Open Cluster observations, and calibration observations of different targets, such as radial velocity standard stars, benchmark stars, globular clusters, COROT red giants and more. Also included are complementary observations extracted from the ESO archive and processed with the GES pipelines. These encompass mostly cluster observations retrieved to benefit calibrations, as well as some bulge observations. These archival observations span a wider range of dates (31.01.2003-11.07.2012) than the GES observations and are denoted by use of the prefix 'AR' rather than 'GE' in the GES\_TYPE header keyword of the spectra. See Table 1 for a list of GES\_TYPES and the corresponding field types. Figure 1 shows the location of the fields on the sky.

The MW targets survey the Bulge, Halo, Thick Disc and Thin Disc populations of the Milky Way. Three primary instrumental setups were used for these observations: UVES 580 for brighter objects and Giraffe HR10 and HR21 for fainter ones.

For the Bulge survey observations of K giants were carried out for the brighter objects (GK stars) using UVES 580, otherwise Giraffe HR10 and/or HR21 were used. For the Halo/Thick disc survey, the primary targets are F+G stars, where bluer fainter F stars probe the halo, and brighter F/G stars probe the thick disc. The outer thick disc is probed using distant F/G stars, as well as K giants to sample the far outer disc. For the solar neighbourhood, G stars were observed using UVES 580 only.

Table 1: The list of GES\_TYPE header keywords used within the Survey to denote the observation and field types, and their definitions.

GES_TYPE prefix	Observation type
GE	Observed by GES
AR	ESO archive observation
GES_TYPE <sup>3</sup>	Field type

<sup>1</sup>. ESO programmes 188.B-3002, 193.B-0936 described in 2012Msngr.147...25G

<sup>2</sup>. ESO programme 188.B-3002(A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P,Q,R,S,T,U,V,W)

<sup>3</sup> In the following list, '\*' denotes either of the strings 'GE' or 'AR', which complete the GES\_TYPE keyword.

*_MW	Milky Way programme field
*_MW_BL	Milky Way programme field: bulge field
*_CL	Open Cluster programme field
*_SD_BM	Standard field <sup>4</sup> : FGKM benchmark stars
*_SD_CR	Standard field: CoRoT field
*_SD_GC	Standard field: globular cluster
*_SD_MC	Standard field: miscellaneous <sup>5</sup>
*_SD_OC	Standard field: calibrating open clusters
*_SD_PC	Standard field: peculiar star templates
*_SD_RV	Standard field: radial velocity standards
*_SD_TL	Standard field: telluric standards

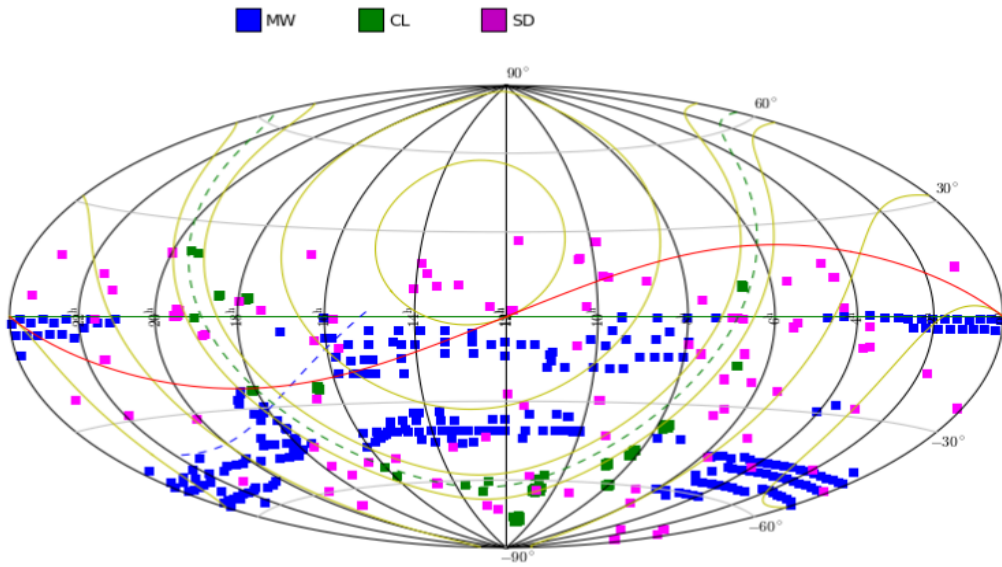


Figure 1: Location of target fields in the sky

The Standard fields included in this release are: calibration observations of stars in the globular clusters M15, M2, NGC104, NGC1851, NGC1904, NGC2808, NGC362, NGC4372, NGC4833, NGC5927, NGC6752 and M67 (archival data processed by the GES pipelines for the latter) which meet our selection threshold for inclusion (see **Data Quality** section), as well as *Gaia* benchmark stars, COROT giants, and radial velocity standards.

The open cluster survey aims to cover the age-metallicity-distance-mass parameter space. Depending on the stellar spectral type, open cluster stars are observed with different Giraffe gratings (HR03/5A/6/9B/14A/15N), and two UVES settings (UVES520 and UVES580). Recently, HR04 has been also introduced for the very hot stars. This data release includes spectra and advanced products for 23 science open clusters (Berkeley 25, Berkeley 44, Berkeley 81, Chamaeleon I, IC2391, IC2602, IC4665, NGC2243, NGC2264, NGC2451, NGC2516, NGC2547, NGC3293, NGC4815, NGC6005, NGC6530, NGC6633, NGC6705, NGC6802, Pismis 18, rho Ophiucus, Trumpler 20, Trumpler 23, gamma 2 Velorum). For each of them different products are delivered, based on our quality control criteria (see below).

Normally, the faint cluster members ([pre-]main sequence or turn-off stars) are observed using Giraffe, while for the brighter stars (typically evolved giants or bright [pre-]main sequence cluster candidates) UVES parallels are employed. Limiting magnitudes for cool stars (later than A-type) are  $V=16.5$  and  $V=19$  mag for UVES and Giraffe respectively. Different magnitude ranges are cov-

<sup>4</sup> Standard fields are observed/extracted from the ESO archive for calibration purposes

<sup>5</sup> Stars which no longer meet the definition criteria of their original classification (e.g. BM) and which may otherwise contaminate the selection functions of the observing programmes have been reclassified as miscellaneous.

ered in clusters where hot stars are observed with the blue gratings. An overlap in magnitude between the Giraffe and UVES samples is present normally and a number of stars were observed with both instruments for inter-calibration purposes.

Within each cluster, the target selection procedure was implemented slightly differently between Giraffe and UVES, but uniformly across clusters. Namely, for Giraffe, with which we aim to observe unbiased and inclusive samples, cluster candidates are selected on the basis of photometry. We used proper motions and other membership indicators (like e.g., X-ray emission) only to define the photometric sequences and the spatial extent of the clusters. In general, we did not use proper motions to select the targets, although in some cases they were employed to discard secure non-members. For UVES, with which we aim to target more secure cluster members, we instead employed membership information from the literature (e.g.,  $v_{\text{rad}}$ , Li,  $H\alpha$ ), when available. More details on the target selection within clusters can be found in Bragaglia et al. (2016, to be submitted). For both MW and open clusters the range of observations are restricted to  $+10^\circ \geq \text{Dec} \geq -60^\circ$  whenever possible to minimise airmass limits (in practice a few target clusters are outside of this range). Figure 2 shows the seeing distribution, for the combined MW and CL dataset. Figure 3 shows instead the range of observing conditions during which the observations were taken.

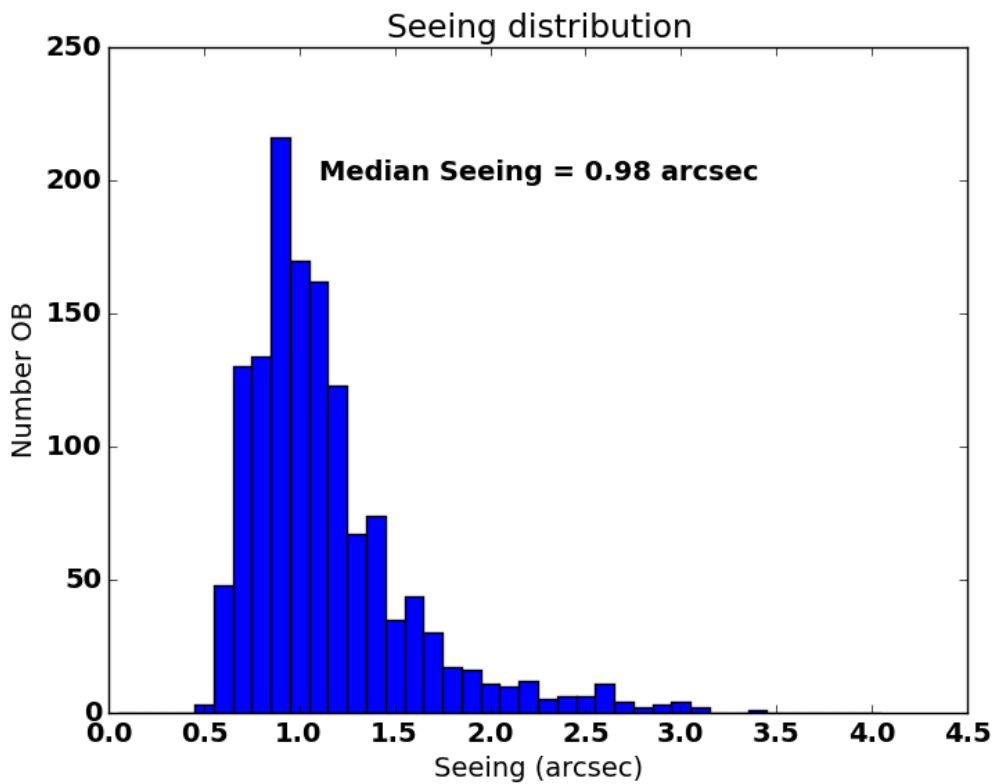


Figure 2: Histogram of seeing per exposure for MW, SD and CL fields and for all observations obtained up to July 2014.

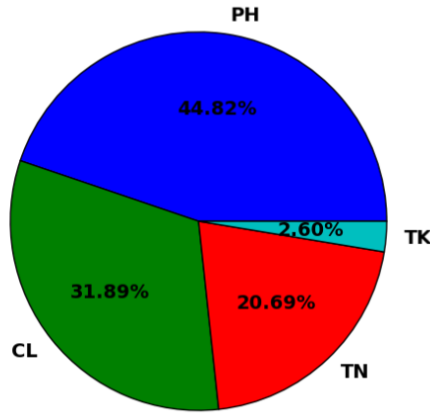


Figure 3: Percentages of observing conditions experienced during the observations up until July 2014: PH = Photometric, CL = Clear, CY = Cloudy, TN = Thin cirrus cloud, TK = Thick cirrus cloud.

The primary source catalogue for the Milky Way field stars is VISTA imaging, ensuring excellent recent astrometry, and adding maximal value to the VISTA surveys.

Photometry for the open clusters mainly comes from the literature and 2MASS. Astrometry is from 2MASS.

## Release Content

The present data release includes spectra, as well as a catalogue containing photometry and advanced products for a fraction of the stars for which spectra are delivered.

### 1. THE SPECTRA

The observations in this data release are summarised in Table 3. The total number of submitted data files is 44210 (size: 13 GB uncompressed) comprising spectra of 25533 unique targets. Included in this sample are 2342 ESO archive spectra, reduced and analysed within GES by the GES pipelines. Figure 4 presents the histograms of the  $J$  magnitudes of the targets included in this release.

### 2. THE CATALOGUE

For the stars for which spectra are delivered, advanced products are also released. These may include one or more of the following quantities: radial velocities (for 96% of stars), astrophysical parameters (APs: effective temperature (for 76%), surface gravity (for 47%), metallicity [Fe/H] (for 57%)), lithium I 6707.8 Å equivalent width, H $\alpha$  emission information, a gravity index ( $\gamma$ , see Damiani et al. 2014, A&A, 566, 50) and individual abundances for a number of elements (with abundances delivered for from 1% (for N) to 45% (for Li) of stars). See Table 3 for a full list of the advanced products delivered in this release.

Parameters that passed the quality thresholds discussed below are included in the table. When a star has been observed with more than one setting and/or with multiple exposures, more than one spectrum is delivered per star (i.e., HR10 and HR21, or HR15N and UVES580). In such cases only one recommended set of parameters (one row of data) is written to the catalogue.

Table 2: Summary of observations and spectra for GES iDR4 (31/12/2011-19/07/2014.)

Field Type <sup>6</sup>	Instrument	Grating	Spectral Range (Å)	Resolving Power (R)	No. Objects	No. Spectra	Median SNR	No. Objects per field type
MW	Giraffe	HR10	5339-5488	19800	6186	6186	30	
		HR21	8484-9001	16200	6303	6303	60	
	UVES	580	4771-6785	47000	1570	3140	53	7870

<sup>6</sup> See Table 1 for a more detailed breakdown of the field types, denoted by the 'GES\_TYPE' keyword.

SD	Giraffe	HR10	5339-5619	19800	2471	2783	49	
		HR15N	6470-6750	17000	2051	2317	70	
		HR21	8484-9001	16200	2471	2749	93	
		HR9B	5143 - 5356	25900	13	79	170	
	UVES	520	4140-6210	47000	82	710	76	
	UVES	580	4771-6785	47000	445	1670	97	3710
CL	Giraffe	HR15N	6470-6790	17000	12379	12396	57	
		HR9B	5143 - 5356	25900	1108	1124	35	
		HR3	4033-4201	24800	688	688	18	
		HR5A	4540-4587	18470	690	690	30	
		HR6	4538-4759	20350	689	689	23	
		HR14A	6308-6701	17740	684	684	39	
	UVES	520	4140-6210	47000	177	354	145	
	UVES	580	4771-6785	47000	824	1648	87	14309

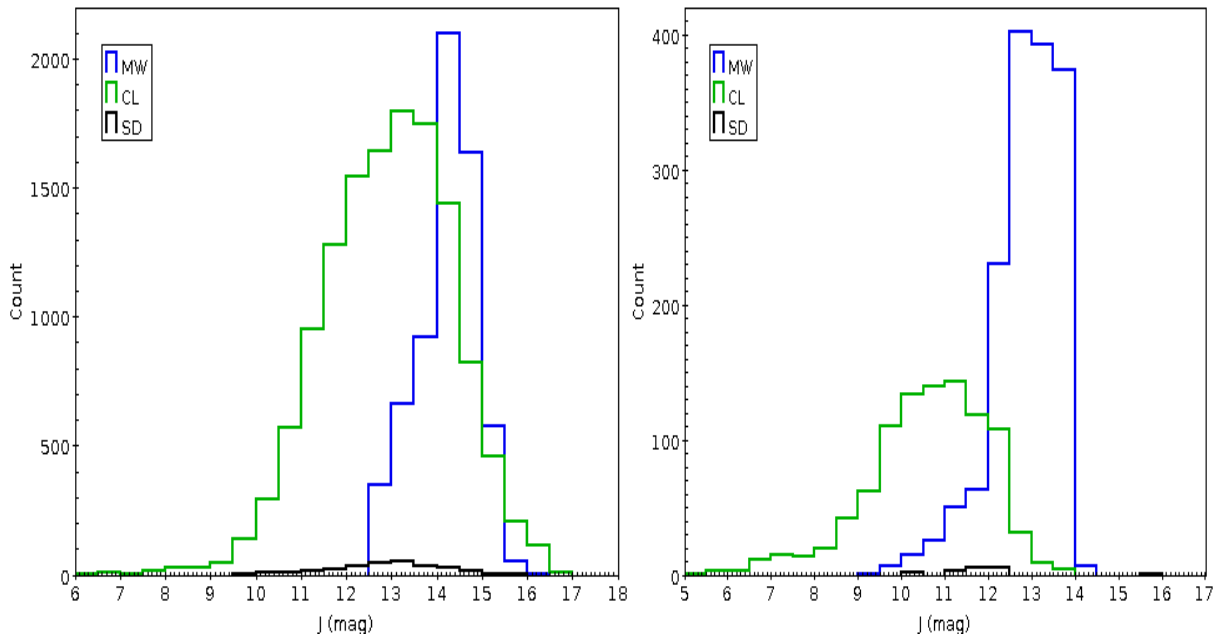


Figure 4: Histogram of magnitudes of targets ( $J_{VISTA}$  for MW fields,  $J_{2MASS}$  for clusters.) Left-hand panel: GIRAFFE; right-hand panel: UVES.

## Release Notes

### Data Reduction and Calibration

The standard Gaia-ESO observing procedure is to divide each observing block into three exposures (except for Giraffe HR21 which is normally divided into two). Two of these are long exposures which are then co-added to eliminate residual cosmic rays, and one is a short exposure (of a few seconds), which is taken for the purpose of obtaining a simultaneous arc lamp spectrum

(SIMCAL) with Giraffe for the wavelength calibration. Spectra from the short exposures are not co-added when creating the final spectra.

Departures from this observing pattern exist – in the case, for example, of periods of poor seeing when additional exposures of a field have been obtained with the aim of increasing the signal-to-noise ratio (SNR). Conversely, during occasional periods of exceptionally good seeing, only one exposure of a field may be taken.

Multi-epoch exposures are defined to be those composed of individual exposures originating from more than one night. A night is defined as the 24-hour period from noon-noon local time (16:00-16:00 UT).

### **Reduction Pipeline: Giraffe**

The Giraffe spectra were reduced by a pipeline that was specially written at the Cambridge Astronomical Survey Unit (CASU). It performs all of the following steps:

- *Bias correction and 2D flat fielding.* The latter is done using test dome flats that are taken periodically as part of the instrumental health checks. Although these are not dispersed flat fields and, of course, flat fields are wavelength dependent, using these does take out a large amount of the pixel-to-pixel variation. (Unfortunately dispersed flat fields without the fibre feed in the light path are not available);
- *Localisation and tracing of the fibre spectra using fibre flat field images.* The optimal extraction profile fits are also done at this point;
- *Extraction of arc spectra, identification of arc lines and wavelength-solution calculation;*
- *Removal of scattered light, extraction and wavelength calibration of object spectra.* The spectra are wavelength calibrated using the arc solution and also shifted to the solar rest frame. For all but the HR21 setting the SIMCAL lamp spectra are used to define a correction to the wavelength solution that is also applied here. For HR21 a similar correction is applied using a subset of well-studied night-sky lines;
- *Sky correction using combined sky fibres from the field.* For all but HR21 the combined sky spectrum is used as is and is subtracted from each object spectrum. For HR21, the sky spectrum is scaled by the relative fluxes of the sky lines to ensure cleaner sky removal;
- *Repeat exposures of the same objects are stacked and cosmic rays are removed.* These are then normalised by the fibre flat field to remove the large-scale wavelength-dependent variation in each fibre.

### **Reduction Pipeline: UVES**

The UVES data were reduced at INAF-Arcetri, using the public ESO FLAMES-UVES pipeline (version 4.9.8 or later for GES iDR4 spectra) for the standard steps of the data reduction process (e.g., bias subtraction, flat-fielding and wavelength calibration) and a pipeline written at INAF-Arcetri for the sky-subtraction, barycentric correction, co-addition. Details of the reduction process can be found in Sacco et al. 2014, A&A 565, 113). The main steps are summarized below.

The reduction is performed in a semi-automatic way, following a reduction cascade. Relevant raw data, including both calibration and science frames, are selected and inserted into the reduction path.

All acquired data are pipeline-reduced using the best possible master calibration products, which are produced starting from the best available day-time calibration frames. After quality checks, these are applied to the reduction of science data. The standard reduction steps followed are:

- *Bias subtraction;*
- *Flat-fielding;*
- *Tracing of the spectral order position;*
- *Wavelength calibration;*
- *Optimal extraction of science spectra* (spectra are de-convolved for fibre cross talk and intra-order background is subtracted);
- *Spectra are corrected for differences in fibre transparency;*
- *The orders are merged;*
- *The sky spectrum from the fibre allocated to the sky is subtracted from the target spectra.* This step is performed both on the individual orders, and on the merged spectra. When more than one fibre is allocated to the sky, the median of the sky spectra is subtracted;

- Both single order and merged spectra are shifted to a Heliocentric reference system;
- Both single order and merged spectra of the same target are co-added;
- A median SNR ratio across the whole spectrum is calculated, for both CCDs;
- All co-added spectra are flagged for binarity;
- Final quality checks are performed on the spectra (see **Data Quality** section);

### **Post-processing**

The normalisation applied to the spectra depends on the particular science goal of the analysis. The choice of continuum level in particular is an individual one which is left as a scientific choice for the end user. As we did in the first release, we deliver here non-normalised spectra to ensure that no valuable information is lost from the spectra.

For the UVES echelle spectra, we have merged the spectral orders and deliver only the merged spectra.

### **Radial Velocity Pipeline: UVES**

Radial velocities (RVs) are derived by cross-correlating each spectrum with a grid of synthetic template spectra. The grid is composed of 36 spectra convolved at the FLAMES-UVES spectral resolution. It covers seven effective temperatures ( $T_{\text{eff}} = 3100, 4000, 5000, 6000, 7000, 8000$  K), three surface gravities ( $\log(g) = 2.5, 4.0, 5.0$ ), and two values of metallicity ( $[\text{Fe}/\text{H}] = 0.0, -1.0$ ).

Each spectrum is cross-correlated with all the spectra of the grid, using the IRAF task FXCOR masking the Balmer lines ( $\text{H}\alpha$  and  $\text{H}\beta$ ) and regions of the spectra with strong telluric lines. To derive the RV, the cross-correlation function (CCF) with the highest peak is selected and the peak is fitted with a Gaussian function to derive its centroid. This procedure fails for very early-type stars with an effective temperature above the highest temperature of our grid, which are characterised by the presence of no, or very few, absorption lines other than the Balmer lines. Radial velocities for these stars are not included in the present release.

To estimate the precision of the RVs, we used the differences between RVs measured from the lower ( $RV_L$ ) and upper ( $RV_U$ ) spectra, which are measured independently by the pipeline. Assuming identical uncertainties on RVs from the two wavelength ranges, and since there is no systematic offset between lower and upper spectra ( $\text{median}(RV_U - RV_L) = 0.007 \text{ km s}^{-1}$ ), the empirical error on the RVs derived by our pipeline is ; the statistical error on RV is equal to the 68th percentile rank of the distribution of these empirical errors, after outliers have been removed. ( $\sigma = 0.18 \text{ km s}^{-1}$ ).

Since the upper and the lower spectrum are calibrated using the same arc lamp, our approach for the error estimate does not take into account the error due to the variations of the zero point of the wavelength calibration. In order to estimate this source of uncertainty, we used spectra of targets observed multiple times in different epochs. Similarly to the above case, the empirical error is estimated as, where  $|\Delta RV|$  is the difference between two observations of the same target performed in different nights. The distribution of this empirical error is much wider than the distribution of the errors  $\sigma_{UL}$ ; the 68th percentiles are  $\sigma_U = 0.38 \text{ km s}^{-1}$  and  $\sigma_L = 0.40 \text{ km s}^{-1}$  for the lower and upper ranges, respectively; this proves that the variations of the zero point of the wavelength calibration are the main source of uncertainty. Therefore, we adopt  $\sigma \sim 0.4 \text{ km s}^{-1}$  as the typical error for the RVs derived from the FLAMES-UVES spectra of the *Gaia*-ESO Survey.

### **Radial Velocity Pipeline: GIRAFFE**

All spectra are iteratively matched against a range of templates to identify the most suitable object-specific templates, thus determining the output RV, and its probability distribution function. Errors are estimated from the curvature of the chi-square surface around the minimum and then empirically corrected to reflect the systematic error floor limit different for each instrument setup as further described in Koposov et al. 2011, ApJ,736,146.

Thanks to the observations of radial velocity standard stars, the radial velocities for all of the setups observed could be shifted to a common zero point. The corrections applied to the radial velocities are described in Hourihane et al. 2016 (to be submitted).

### **Spectrum analysis**

Five working groups (WGs) share this task, focusing on Giraffe and UVES spectra of FGK normal stars, of cool pre-main sequence stars, of OBA-type stars, and on unusual objects, respectively. Within each WG several nodes participate in the analyses. An early lesson from working with

many analysis teams was the critical need to have a well-understood, common, suitable line-list for the analyses, a common set of model atmospheres, a common grid of synthetic spectra, and a common approach to data formats and standards. All of these have been made available to the analysis groups and are regularly updated thanks to the efforts of dedicated teams. Once the node analysis within the different WGs has been completed, WG recommended parameters are derived using the calibrators (in particular the Gaia benchmark stars) to evaluate and weight node performances. Depending on the particular WG and on the number of nodes, the final recommended parameters are computed as weighted medians or mean values after outlier rejection and, when needed, offset correction. After this stage, parameter and abundance homogenisation across WGs is performed. This step involves putting the parameters and abundances derived by the different WGs for the different types of stars on the same scale. It is carried out based on common targets and calibrators analysed by all the spectrum analysis nodes and WGs (see Hourihane et al., 2016, in preparation).

The different node analyses are based on several complementary standard, as well as special-purpose, spectrum analysis methodologies. The structure of the WGs provides close coordination between the teams, ensuring the optimum range of analyses are applied to the various stellar and data types as appropriate. The methodologies are all established, all publicly well-documented, forming the basis of most modern spectrum analyses in the literature. Below we provide a general description of the strategy and methods followed by the spectrum analysis WGs. For the details, node value combination, and error estimate we refer to Smiljanic et al. (2014, A&A 570, 122), Lanzafame et al. (2015, A&A, 576, 80L), Recio-Blanco et al. (2016, in preparation), Blomme et al. (2016, in preparation).

For the analysis of the spectra and the determination of the advanced products contained in this release, the five WGs in charge of the spectrum analysis follow a similar approach, summarized below:

- The data analysis has been duplicated among the nodes contributing to each WG. Specifically, more than one group has normally analysed and produced results for (nearly) all relevant Survey targets. This duplication of different methods has allowed, given performance comparison of the results, production of a set of recommended parameters. Also, through rigorous quality control, it has provided a quantitative estimate of both random and method-dependent uncertainties. When discordant results are produced for a specific star, individual checks have been conducted. These include checks against the benchmark stars and calibrating clusters. The methods of combining the node results within a WG detailed in the publications above include the calculation of weighted means and transformations composed of linear combinations of the parameters to transform parameters from one setup onto the scale of another setup. For the advanced products included in this release (GES iDR4), the method used by Smiljanic et al. (above) in WG11 for earlier releases has been replaced by a robust ensemble method, using MCMC, that assumes the node values are normally distributed and attempts to infer the uncertainty for each node, allowing for biases and trends within the node. The predictive power of the model performs well against the benchmarks.
- Depending on the star's spectral type and characteristics, appropriate optimal tools, software, and model atmospheres have been used; however, some methodologies in common to all WGs have been identified. As mentioned, a common line list has been implemented (Heiter et al., 2015, PhyS, 90, 4010); likewise, common model atmospheres, covering the range from A to M spectral type (MARCS models), as well as a grid of synthetic spectra based on those models, have been used by most nodes, with exception of those analysing warm stars;
- The methods used to derive APs can be roughly divided in two broad categories. The first one, based on comparisons between observed spectra and a grid of templates (other synthetic or observed ones), includes the main types of parameterisation methodology, such as exhaustive search algorithms, global optimization methods, projection algorithms, pattern-recognition methods, and Bayesian parameterisation approaches (like e.g., MyGisFos, SME, ROTFIT, FASTWIND, etc.) ; the second one consists of more classical approaches, based on measurements of equivalent widths (EWs) of absorption lines and inversion codes (like e.g., MOOG) , or use of curves of growths (COGs) for particular lines/elements (e.g., Li). In these cases EWs are measured with (semi-) automatic codes by fitting Gaussian profiles to the lines. The available codes include: DAOSPEC, ARES, and SPECTRE.



- Additional methods to derive APs have been used in special subsets of the sample (e.g., H $\alpha$  wings, line-depth-ratios). In most cases the codes are automatic, and proven to be able to handle Gaia-ESO scale data. An alternative approach for GIRAFFE/HR15N, proposed by Damiani\_et al. (2014), is also used. This is based on spectral indices in different wavelength ranges of the spectrum. The derived spectral indices are calibrated against stars with known parameters, yielding quantitative estimates of APs;
- The width of H $\alpha$  at 10% and H $\alpha$  equivalent widths have been measured on the continuum-normalised spectra, using a semi-automatic procedure. After manually defining the wavelength range and level of continuum, the equivalent width is calculated by a direct integration of the flux above the continuum, while the width at 10% is derived by considering the level corresponding to 10% of the maximum flux above the continuum in the selected wavelength range. All measurements are visually checked and repeated in case of miscalculation (e.g. due to the presence of multiple peaks). Uncertainties are estimated using multi-epoch observations of the same star;
- Lithium equivalent widths released here (young cluster targets only) have been measured with three independent methods and then combined, after careful comparison (see Lanzafame et al. 2015). Namely, IRAF-plot (Gaussian fitting), DAOSPEC (Gaussian fitting), and a semi-automatic direct integration procedure in IDL, specifically developed for the Gaia-ESO Survey. As a conservative uncertainty estimate on the recommended equivalent width, we adopted the larger of the standard deviation and the mean of the individual method uncertainties;
- Finally, a special object-by-object analysis process has been applied to spectra that are not consistent with any of the stellar classes (e.g., binaries, carbon stars, etc.). A set of flags has been implemented and the dictionary is delivered in this release (Table 4).
- The parameters returned by the various WGs at the end of the first round of analysis (parameter determination) are homogenised by the homogenisation Working Group, WG15., to produce a set of recommended parameters for the Survey. The first step in the homogenisation is the checking of parameters with respect to literature values for the well-studied sample of Gaia FGK benchmark stars, where these are within the parameter range of the specific analysis group. Other calibrators employed include well-studied open and globular clusters, which provide another important inter-setup calibration set, and a sample to compare cool vs hot stars and main sequence versus evolved stars. The clusters are also important calibrators for the WGs whose analysis sample falls outside of the FGK range covered by the benchmarks (e.g. the hot star group). Checks of results from each instrumental setup are performed and if necessary, offsets or transformations are applied to put the data onto the scale defined by the calibrating samples.
- Elemental abundances have been derived based on recommended parameters. The methods used to derive elemental abundances can be broadly categorised into spectrum synthesis or equivalent width methods, using the same codes employed for the parameter determination. Depending on the working group, element, number of lines, abundances from the different nodes have been combined either on a average node abundance basis, or on a line-by-line node abundance basis.
- A similar procedure is employed to homogenise the abundances returned by the WGs at the end of the second round of analysis (abundance determination). The abundances are again checked against the benchmark abundances (alpha and iron-peak elements; Jofre et al. 2015), where available, and inter-setup corrections are additionally applied based on comparisons within clusters (see Data Quality section). This process produces a recommended set of elemental abundances per star, which may combine results from several WGs where the setups analysed cover different spectral ranges and allow the determination of different elemental abundances.

## Data Quality

### Spectra - general

The quality array ('QUAL') delivered along with the spectra in the data files codes data values as good quality (0) or bad quality (1). These code values are derived from weight maps where a value of '1' represents a bad pixel.

The error arrays in the Giraffe spectra in this release (reduced with the iDR4 – internal data release 4 - version of the Gaia-ESO Survey Giraffe pipeline) were calculated incorrectly which led to their being underestimated, which gave falsely high SNR ratios. Hence the SNR values from the

latest version of the pipeline, that used in GES iDR5, have been used for selection of the spectra (which were reduced with the iDR4 pipelines) for this release. The distribution of these SNR values is shown in Figure 5. (Updated versions of the spectra, reduced with the iDR5 pipelines, are being made available in the ESO archive as a supplement to this release). For consistency, the UVES spectra are also selected on the GES iDR5 SNR.

Further quality control that is applied to the spectra is described below.

**Quality Control: UVES spectra**

Quality control (QC) on the UVES data is performed in three steps:

- Check on the quality of the calibration frame by comparing the QC parameters, which are given as output by the ESO pipeline, with the typical values published on the ESO website. This approach allows us to verify the instrument stability (e.g. the stability of the bias frame or the precision of the wavelength calibration);
- Visual inspection of the final spectra aimed at discovering artifacts or other anomalies (e.g., in the wavelength calibration). If this analysis identifies anomalies in one or more spectra, the whole workflow, since fibre allocation, is investigated. Once the problem is identified, the reduction is performed again to improve the quality of the spectra;
- Selection of SNR thresholds. (The SNR is the median value and is quoted per pixel.)

**Quality Control: Giraffe spectra**

QC on the Giraffe data is carried out as part of the spectral template fitting which then assigns a basic classification.

The main criteria defining this classification are:

- the  $\chi^2$  of the fit;
- the SNR (calculated per pixel);
- the  $\chi^2$  of the pure continuum fit; and
- the distance to the best fit template.

The spectrum is marked as UNKNOWN instead of STAR when the continuum-only fit is better or almost as good as the template fit. The  $\chi^2$  and/or distance to the best-fit template are higher than a certain SNR-dependent threshold.

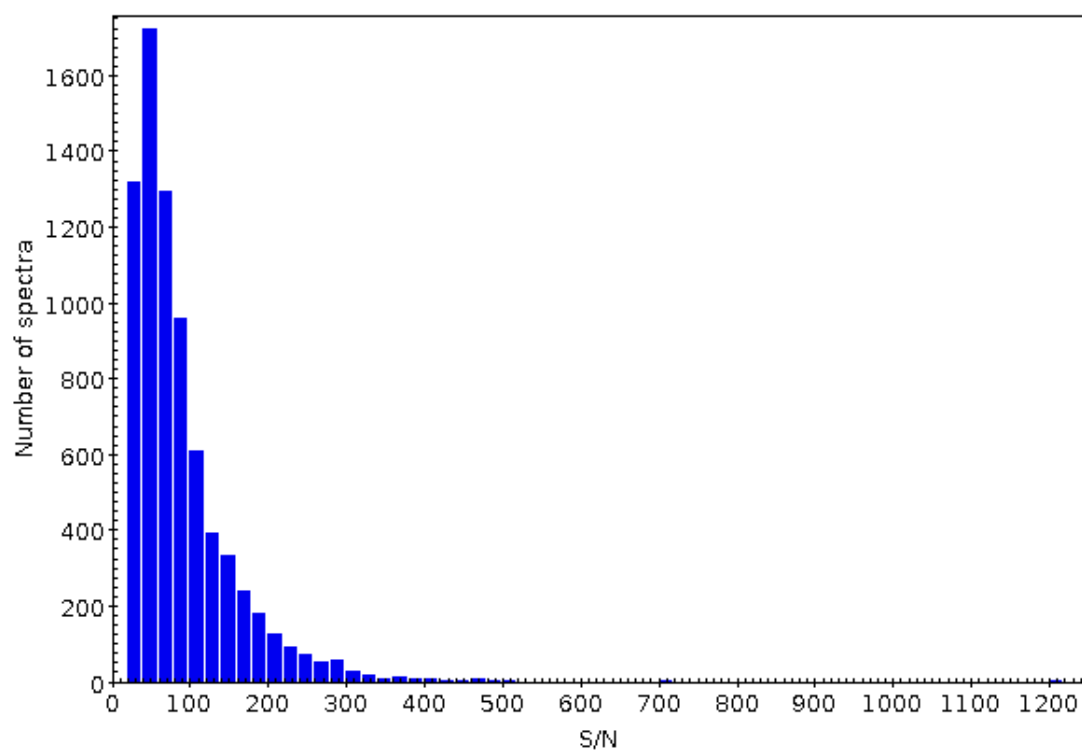
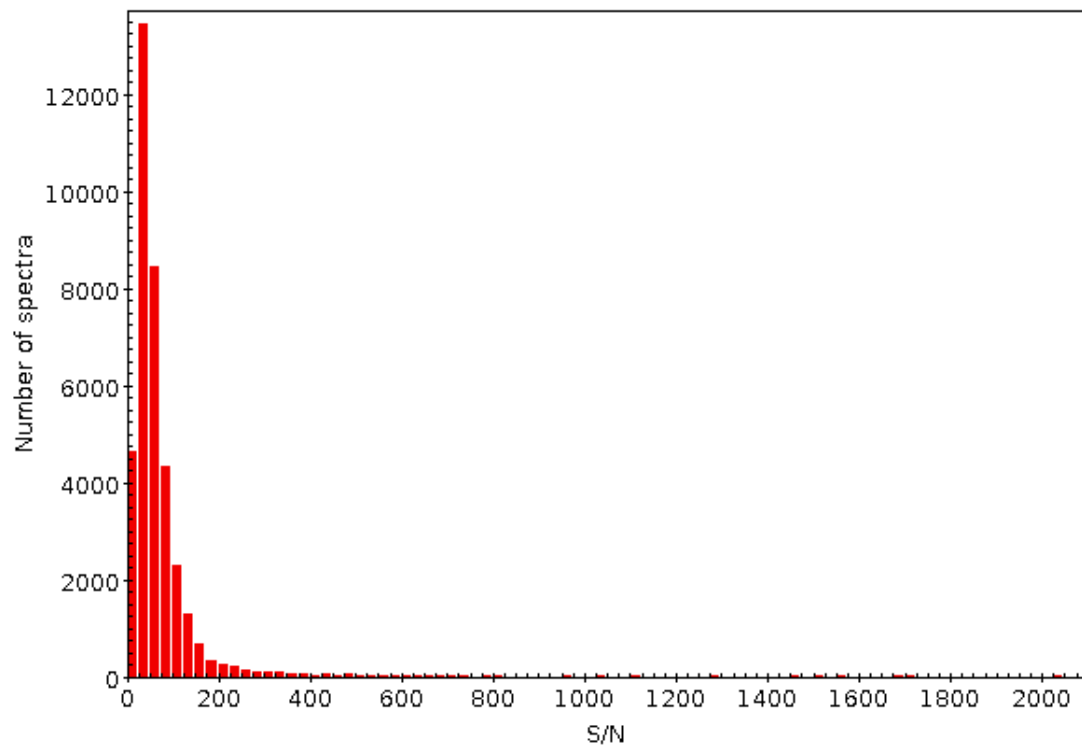


Figure 5: SNR ratio distributions for the released spectra; GIRAFFE and UVES are shown in the upper and lower panels, respectively. The values of SNR shown here are from the subsequent reduction, corresponding to version 5 of the Survey pipeline reductions.

### Quality control: Radial velocities

The following set of figures shows selected quality diagnostics for the radial velocities (RVs). As part of the radial velocity homogenisation, different instrumental setups were compared, and offsets were applied to bring the radial velocities on to the same scale. The RVs from the HR10 setup were used as the zeropoint of the scale due to their good agreement with the literature values of the Gaia Radial Velocity Standards (Soubiran et al.). Figure 6 shows an example of the cross-match between different instrumental setups and the offset found and applied during homogenisation.

The radial velocity precision being achieved for the majority of survey targets is well within the required 1 km/s. Indeed it reaches below to 0.3 km/s (our requirement) for cool stars in clusters (see Figure 7 and Jackson et al. 2015). Similarly, very good precision is obtained for UVES (see Sacco et al. 2014).

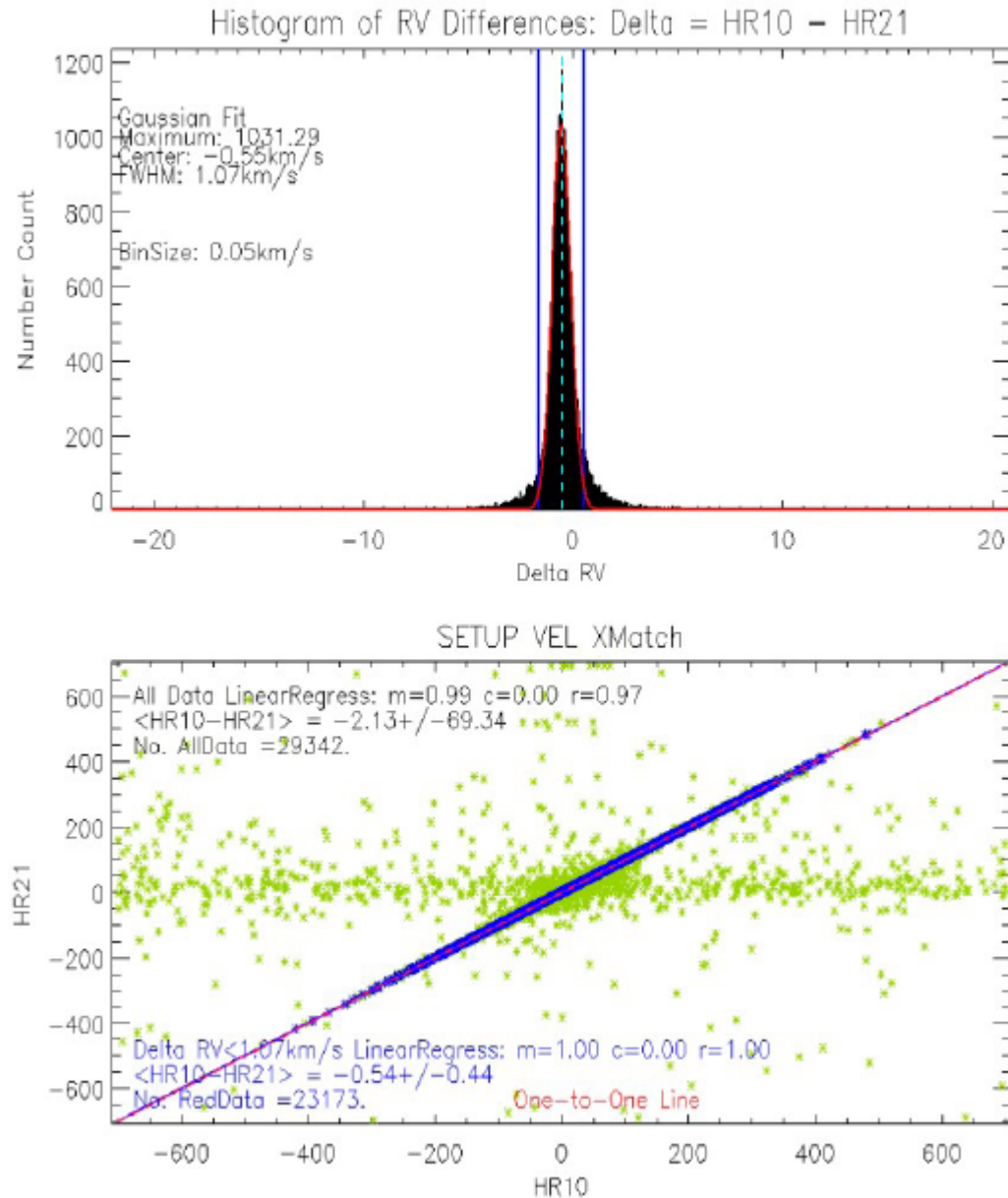


Figure 6: comparison of the radial velocities obtained from the HR10 and HR21 instrumental setups, showing the known offset of 0.5 km/s.

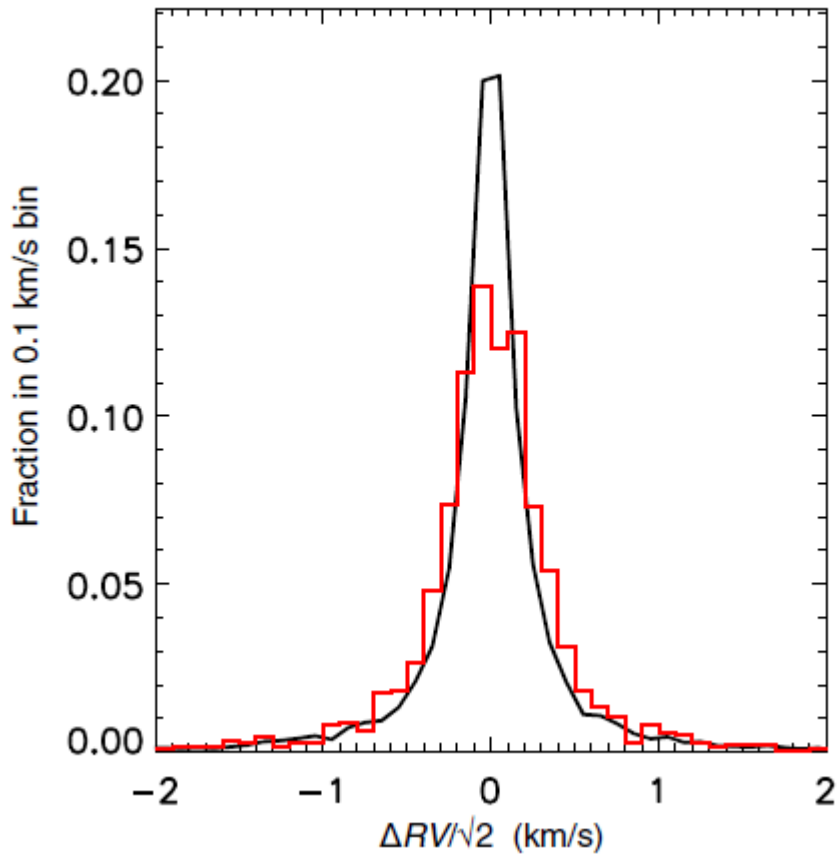


Figure 7. Comparison of the probability density of the observed precision in RV using long (different OBs) and short term (same OB) repeats in the setup HR15N. The black line shows results for short-term repeats (i.e. pairs of observations within the same OB). The red histogram shows results for long-term repeats (i.e. spectra of the same targets but taken from different Obs where individual targets are allocated to different fibres). (Jackson et al. 2015)

#### Quality control: astrophysical parameters

A variety of quality checks have been performed on the astrophysical parameters (see Smiljanic et al. 2014; Lanzafame et al. 2015; Recio Blanco et al. 2016, in preparation; Hourihane et al. 2016, in preparation); namely:

- the final recommended set of parameters for each benchmark star is compared with the literature values for these stars;
- HR diagrams of both field stars and clusters are produced. For the clusters comparison with the isochrones is performed (see Figures 8-10);
- Within the clusters checks for the homogeneity of metallicity and lack of trends with parameters are performed;
- Within clusters, checks against isochrones are also performed;
- For the MW fields we checked that the expected (from original selection) temperature and gravity distributions are recovered

#### Quality control: elemental abundances

The checks performed for the elemental abundances include:

- the final recommended set of abundances for each benchmark star is compared with the literature values for these stars (Jofre et al. 2015; Gaia FGK benchmarks)
- for Milky Way stars, the abundances have additionally been checked against the high-resolution dataset of Bensby et al. 2014 (714 FG solar neighbourhood stars)

- the compactness of the cluster element (El) vs.  $[Fe/H]$  plots (see Figure 11) is investigated
- the behaviour of the  $[El/Fe]$  vs.  $[Fe/H]$  distributions (based on which we have discarded some elements) is investigated (see Figure 12)

These checks showed good agreement and behaviour for the homogenised parameters and abundances.

A few quality plots are included below.

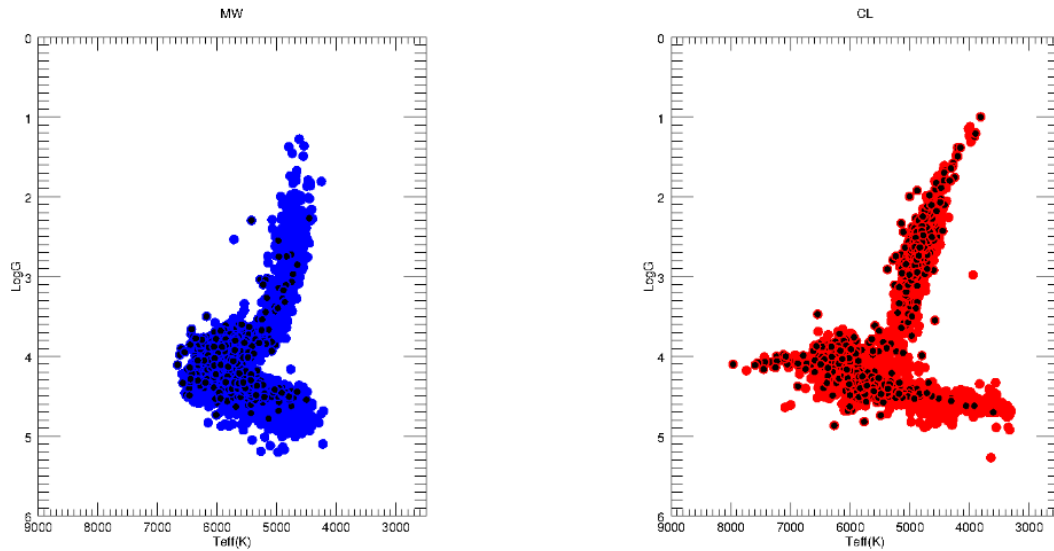


Figure 8: HR diagrams of the MW fields (left-hand panel) and cluster fields (all clusters together; right-hand panel). Stars observed with UVES are overlaid in black on the Giraffe results (blue and red).

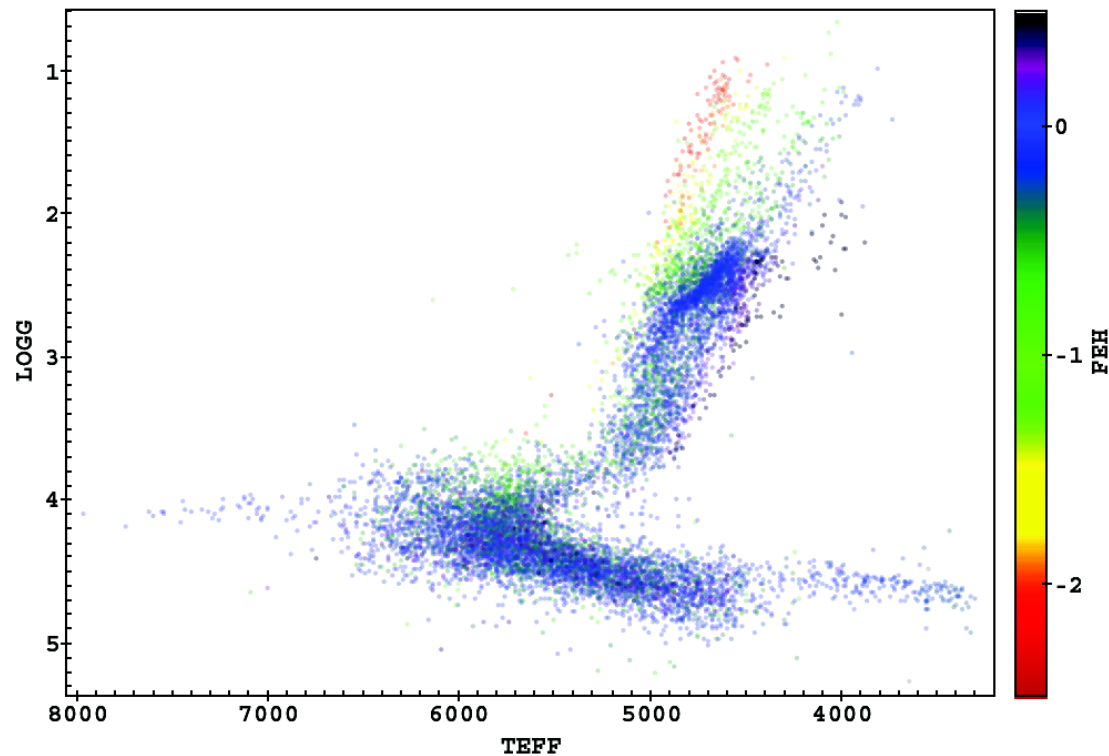


Figure 9: HR diagram of MW and cluster fields, Giraffe and UVES together, colour-coded by  $[Fe/H]$

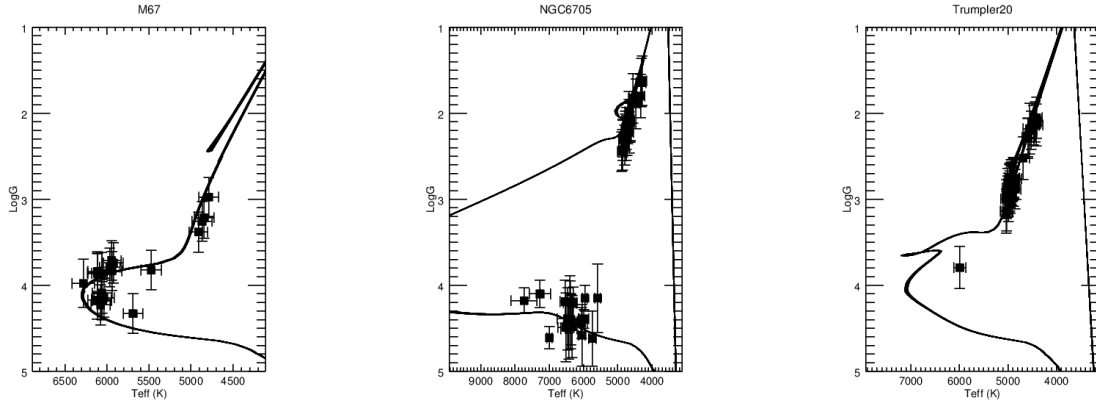


Figure 10: HR diagrams of three calibrator clusters (M67, NGC6705 and Trumpler 20) with isochrones overimposed. Note that spectra for M67 were retrieved from the ESO archive.

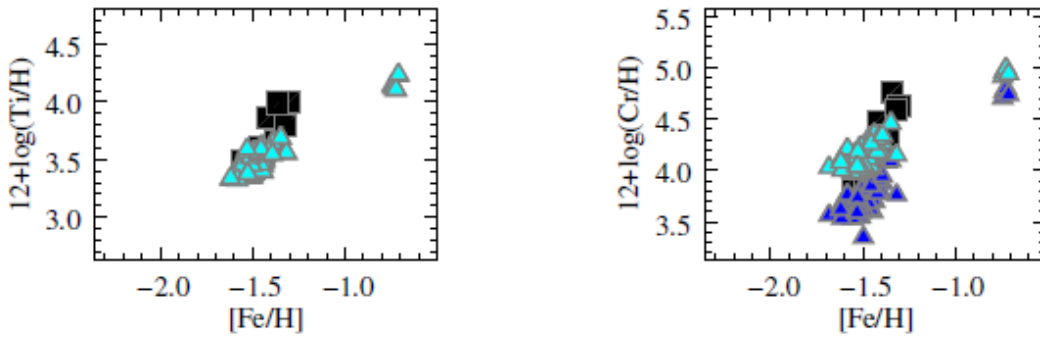


Figure 11: Abundances for the cluster M2, coming from WG11/UVES (black squares) and WG10/GIRAFFE (triangles). The blue triangles are the Giraffe results before correction and the cyan are after correction, which was applied to the Cr abundances coming from GIRAFFE.

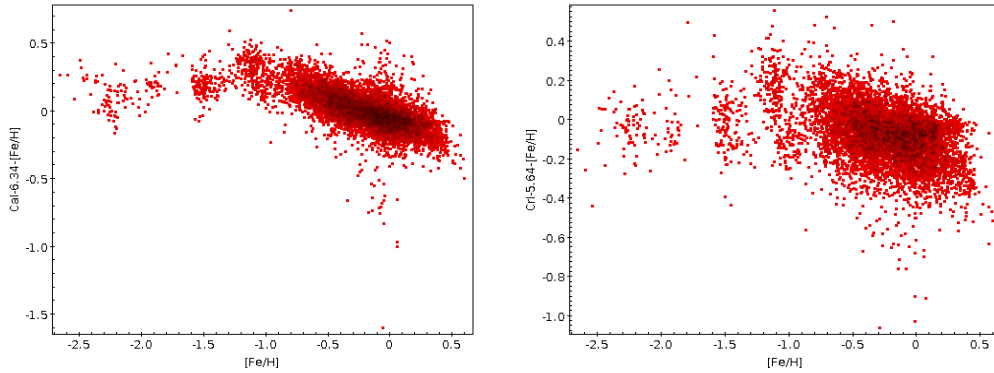


Figure 12:  $[\text{El}/\text{Fe}]$  vs  $[\text{Fe}/\text{H}]$  plots for Ca I (left-hand panel) and Cr I (right-hand panel) showing the expected trend with  $[\text{Fe}/\text{H}]$ .

### Quality Criteria:

The spectra included in this data release satisfy the following key criteria:

1. Observed prior to 20 July 2014;
2. Target observations are complete for the survey;
3. They are not characterized by obvious reduction issues
4. They have all undergone a cycle of analysis (the GES internal iDR4), even if the resulting parameters are not released, because they do not match our selection criteria (see below)

5. Above SNR limit:
  1. **UVES 520 and 580:** SNR > 20 in both lower and upper CCDs
  2. **Giraffe:**
    - a. **MW and SD:** Giraffe spectra exist for the target in both HR21 and HR10, achieving HR21 SNR>30 and HR10 SNR >20. MW bulge stars have a single setup (HR10 or HR21), and thus the requirement of spectra in both HR21 and HR10 is relaxed. HR15N spectra have SNR > 20 (standard fields, including calibrating open clusters, CoRoT and benchmarks).
    - b. **CLUSTERS:** Giraffe HR15N and HR9B spectra have SNR > 20. No SNR thresholds are imposed on the bluer gratings, since the current SNR ratios values for those spectra (with which warm stars are observed) may be underestimated. Since a great fraction of those spectra have been analysed, we believe they are of sufficient quality.
6. Observations are complete on all cluster targets, for all UVES and Giraffe settings during the required epoch;

As noted in the **Data Quality** section (**Spectra – General**), the GES version 5 (iDR5) SNR values have been used for selection of the spectra.

**The radial velocities included in this data release satisfy the following criteria:**

1. For the clusters:  $-100 \text{ km/s} < RV < 150 \text{ km/s}$
2. For the clusters: The error must have been determined
3. For the clusters: The error must be below 2 km/s
4. No cuts were applied for the MW

**Based on the error distributions, the parameters included in this release satisfy the following criteria:**

1. All parameters and abundances from WG13 (hot stars) are discarded
2. All parameters and abundances coming from the HR9B setup are discarded, due to large errors in  $\log g$  and/or inconsistent position in the HR diagram
3. All abundances (except Li) derived from WG12 are discarded since they did not pass our quality criteria for inclusion in this release
4. In the other cases, the parameters included in this release satisfy the following criteria:
  - i.  $T_{\text{eff}}$ : error available and relative error  $e_{T_{\text{eff}}}/T_{\text{eff}} \leq 5\%$
  - ii.  $\log g$ : error available and  $e_{\log g} \leq 0.4$  dex for the HR15N setup,  $e_{\log g} \leq 0.3$  dex for the other setups. For open clusters in HR15N, also consistency with the HR diagram and isochrones has been checked.
  - iii.  $[\text{Fe}/\text{H}]$ : error available and  $e_{[\text{Fe}/\text{H}]} \leq 0.2$  dex

For HR10, HR21 and UVES, all the three criteria must be verified (e.g. if  $e_{T_{\text{eff}}}/T_{\text{eff}} \leq 5\%$  and  $e_{\log g} > 0.3$  dex all parameters and abundances are discarded and only radial velocities are included). For HR15N, if the  $T_{\text{eff}}$  criterium is satisfied, both  $T_{\text{eff}}$  and the Li abundance are always retained even if  $\log g$  and/or  $[\text{Fe}/\text{H}]$  errors are above the threshold.

5. No cuts on the individual abundances were performed. The following elements are included: C1, Li1, N\_CN, O1, Na1, Mg1, Al1, S1, Ca1, Sc2, Ti1, Ti2, V1, Cr1, Co1, Ni1, Zn1, Y2, Zr2, Ba2, La2, Ce2, Nd2, Eu2 in addition to Li1, as specified above. Abundances (except Li1) are discarded when one or more parameters do not satisfy the criteria.
6. In the case of the cluster Br44, parameters and abundances from U580 are retained only if  $T_{\text{eff}} \leq 5000 \text{ K}$
7. No cuts were applied for H $\alpha$ , lithium equivalent width, and gravity index



## Known issues

There is a known issue with the error arrays of the Giraffe spectra in this release. For this reason, updated versions of the spectra are also being made available as a supplemental release. See the section on Data Quality/Spectra – General for further details.

## Previous Releases

The previous release was number 2. The changes in the present release are as follows:

1. A larger set of spectra is included. In release #2 a subsample of the spectra obtained up to 31 December 2013 were delivered. Here we submit a fraction of the spectra obtained up to 20 July 2014. The spectra selection criteria have changed, but not significantly (see Quality Criteria, above). The change in selection criteria result in some spectra from the previous release not making the selection thresholds for this release and so not having an updated version present in this release. The spectra released here have been reduced with updated versions of the GIRAFFE and UVES pipelines compared to the previous release. This difference in pipeline version may also cause some change in the spectra making the selection criteria.
2. A catalogue is delivered, including ancillary information (i.e., target photometry), radial velocities, astrophysical parameters, H $\alpha$  emission information, lithium equivalent widths, elemental abundances and flags (see Table 3). The number of columns in the catalogue is greater than in the previous release (100 vs 38 columns) and elemental abundances are released for the first time.

## Data Format

### Files Types

The files provided for this release are in the format as specified in issue 5 of document GEN-SPE-ESO-33000-5335. This consists of a FITS file with a primary header unit containing no data and a binary FITS table extension containing the data. The header cards in the header unit of each extension contain the information requested in the above document.<sup>7</sup> The wavelength array (WAVE), spectrum (FLUX), error array (ERR) and quality array (QUAL) are each provided in a single cell of the one row contained in the binary table.

The objects in each file have a name which is derived from the object's equatorial coordinates. This is formed by splicing the RA (in hours, minutes and seconds to two decimal places) and Declination (in degrees, minutes and seconds to one decimal place) as integers with the declination sign in the middle. Thus an object at 3h40m21.767s and -31°20'32.71" will have the name 03402177-3120327 (the cname).

The name of the file is of the form <prefix>\_<cname>\_<expmode>\_<version>.fits, or <prefix>\_<cname>\_<expmode>\_<index>\_<version>.fits. The value of <prefix> is either 'gir3' (Giraffe), 'uvl3' (UVES lower), or 'uvu3' (UVES upper) for GES-observed spectra. For archival spectra, the prefix is 'gar3' (Giraffe), 'url3' (UVES lower), or 'uru3' (UVES upper). The value of <expmode> is derived from the central wavelength and grating for the instrument, e.g. H875.7. The value of the <index> suffix is an integer assigned to distinguish each individual exposure spectrum for the unstacked benchmark spectra (all spectra without an <index> are stacked from the available spectra). The value of <version> denotes the GES internal data release version of the spectrum.

The catalogue consists of a FITS file with a primary header unit containing no data and two binary FITS table extensions containing the catalogue data and the provenance file information. The catalogue column GES\_TYPE specifies the type of the Gaia-ESO Survey field (values are listed in Table 1 in the **Overview of Observations** Section). GES\_TYPE is intended to provide useful supplementary information on the field for the user.

---

<sup>7</sup> Please note that the SPEC\_RES keyword in the primary header denotes the spectral resolving power,  $\lambda/\Delta\lambda$ , rather than the FWHM resolution.

## Catalogue Columns

The catalogue comprises 100 columns with 25533 rows of data. The columns are described in the Table 3, and the flags used in columns PECULI, REMARK and TECH are defined in Table 4.

Table 3: Column names, data format, description and units contained within the Gaia-ESO Survey iDR4 Catalogue.

Column	Data Format	Description	Units
CNAME	16A	GES object name from coordinates, corresponds to OBJECT in header of FITS spectrum	
GES_FLD	23A	GES field name from CASU	
OBJECT	74A	GES object name from OB	
GES_TYPE	20A	GES Classification System of Target Programmes <sup>8</sup>	
SETUP	23A	Grating setups used for analysis (for stellar parameters) <sup>9</sup>	
RA	D	Object Right Ascension	Deg
DECLINATION	D	Object Declination	Deg
VRAD	E	Radial Velocity	km/s
E_VRAD	E	Error on VRAD	km/s
TEFF	E	Effective Temperature	K
E_TEFF	E	Error on TEFF	K
LOGG	E	Log Surface Gravity (gravity in $\text{cms}^{-2}$ )	dex
E_LOGG	E	Error on LOGG	dex
FEH	E	Metallicity	dex
E_FEH	E	Error on FEH	dex
XI	E	Microturbulence	km/s
E_XI	E	Error on XI	km/s
EW_LI	E	Li(6708A) equivalent width	mÅ
LIM_EW_LI	I	Flag on EW_LI (0=detection, 1=upper limit; NULL=-1)	
E_EW_LI	E	Error on EW_LI	mÅ
EWC_LI	E	Blend-corrected Li(6708A) equivalent width	mÅ
LIM_EWC_LI	I	Flag on EWC_LI (0=detection, 1=upper limit)	
E_EWC_LI	E	Error on EWC_LI	mÅ
EW_HA_ACC	E	Halpha EW: accretion	Å
E_EW_HA_ACC	E	Error on EW_HA_ACC	Å
HA10	E	Halpha EW at 10% of peak - accretion	km/s
E_HA10	E	Error on HA10	km/s

<sup>8</sup> The GES\_TYPE values are taken from the results of the internal data release four (iDR4) of GES, and include the GES\_TYPES of all of the spectra which were analysed. The originating spectra for the stellar parameters in the catalogue are in all cases included as part of this ESO Phase III release, but not all of the additional spectra that were analysed meet the selection thresholds for inclusion in this release and are therefore not included in all cases. Please see Table 1 for a full list of GES\_TYPES.

<sup>9</sup> Specifically, the grating setup of the originating spectrum/spectra from which the stellar atmospheric parameters were derived. May be NULL if no stellar parameters are provided. HR10|HR21 denotes parameters were derived from a joint analysis in the GIRAFFE setups HR10 and HR21.

GAMMA	E	Gravity sensitive spectral index	
E_GAMMA	E	Error on GAMMA	
PECULI	29A	Peculiarity Flag(s): WG14 Dict.1000-2999 (see Table 3)	
REMARK	16A	Spec. Class. Flags(s): WG14 Dict.3000-8999 (see Table 3)	
TECH	87A	Technical Flag(s): WG14 Dict.9000-15000 (see Table 3)	
LI1	E	Neutral Lithium Abundance	dex
LIM_LI1	I	Flag on LI1 (0=detection, 1=upper limit)	
E_LI1	E	Error on LI1	dex
C1	E	Neutral Carbon Abundance	dex
E_C1	E	Error on C1	dex
N_CN	E	Nitrogen Abundance from CN Molecular Bands	dex
E_N_CN	E	Error on N CN	dex
O1	E	Neutral Oxygen Abundance	dex
E_O1	E	Error on O1	dex
NA1	E	Neutral Sodium Abundance	dex
E_NA1	E	Error on NA1	dex
MG1	E	Neutral Magnesium Abundance	dex
E_MG1	E	Error on MG1	dex
AL1	E	Neutral Aluminium Abundance	dex
E_AL1	E	Error on AL1	dex
S1	E	Neutral Sulfur Abundance	dex
E_S1	E	Error on S1	dex
CA1	E	Neutral Calcium Abundance	dex
E_CA1	E	Error on CA1	dex
SC2	E	Ionised Scandium Abundance	dex
E_SC2	E	Error on SC2	dex
TI1	E	Neutral Titanium Abundance	dex
E_T1	E	Error on TI1	dex
TI2	E	Ionised Titanium Abundance	dex
E_T2	E	Error on TI2	dex
V1	E	Neutral Vanadium Abundance	dex
E_V1	E	Error on V1	dex
CR1	E	Neutral Chromium Abundance	dex
E_CR1	E	Error on CR1	dex
CO1	E	Neutral Cobalt Abundance	dex
E_CO1	E	Error on CO1	dex
NI1	E	Neutral Nickel Abundance	dex
E_NI1	E	Error on NI1	dex
ZN1	E	Neutral Zinc Abundance	dex
E_ZN1	E	Error on ZN1	dex
Y2	E	Ionised Yttrium Abundance	dex
E_Y2	E	Error on Y2	dex
ZR2	E	Ionised Zirconium Abundance	dex
E_ZR2	E	Error on ZR2	dex
BA2	E	Ionised Barium Abundance	dex

E_BA2	E	Error on BA2	dex
LA2	E	Ionised Lanthanum Abundance	dex
E_LA2	E	Error on LA2	dex
CE2	E	Ionised Cerium Abundance	dex
E_CE2	E	Error on CE2	dex
ND2	E	Ionised Neodymium Abundance	dex
E_ND2	E	Error on ND2	dex
EU2	E	Ionised Europium Abundance	dex
E_EU2	E	Error on EU2	dex
j_vista	D	J Band 4th aperture magnitude of VHS	mag
j_vista_err	D	J Band 4th aperture magnitude error of VHS	mag
h_vista	D	H Band 4th aperture magnitude of VHS	mag
h_vista_err	D	H Band 4th aperture magnitude error of VHS	mag
k_vista	D	K Band 4th aperture magnitude of VHS	mag
k_vista_err	D	K Band 4th aperture magnitude error of VHS	mag
dist_vista	D	Distance to VHS co-ordinate match	arcsec
j_2mass	E	J band magnitude of 2MASS	mag
j_2mass_err	E	J band magnitude error of 2MASS	mag
h_2mass	E	H band magnitude of 2MASS	mag
h_2mass_err	E	H band magnitude error of 2MASS	mag
k_2mass	E	K band magnitude of 2MASS	mag
k_2mass_err	E	K band magnitude error of 2MASS	mag
dist_2MASS	E	Distance to 2MASS co-ordinate match	arcsec
UMAG	E	U Band magnitude from Cluster photometry compilation	mag
BMAG	E	B Band magnitude from Cluster photometry compilation	mag
VMAG	E	V Band magnitude from Cluster photometry compilation	mag
RMAG	E	R Band magnitude from Cluster photometry compilation	mag
IMAG	E	I Band magnitude from Cluster photometry compilation	mag

Table 4: Definitions for the flags included in the PECULI, REMARK and TECH columns of the catalogue as stated in the GES WG14 Dictionary. Confidence flags that can be attached to a flag are defined as: A=Probable, B=Possible, C=Tentative

Flag	WG14 Dictionary Definition
<i>PECULI Flags</i>	
1010	Suspicion that (H $\alpha$ ) emission lines are extrinsic rather than intrinsic (e.g. HII region in the line of sight)
1011	Emission line detection (Hydrogen (01), ionization level)
1013	Emission line detection (Balmer H $\alpha$ )
2005	Stars with large radial velocity variations, indicating either large jitter or binary motion
2010	SB1 (Stars with radial velocity variations larger than expected jitter for its type, indicating probable binary motion)
2020	SBn,n $\geq$ 2 (Spectroscopic Binary, Ncomponents $\geq$ 2)
2030	SBn,n $\geq$ 3 (Spectroscopic Binary, Ncomponents $\geq$ 3)
2040	SBn,n $\geq$ 4 (Spectroscopic Binary, Ncomponents $\geq$ 4)
2070	Composite spectrum
2100	Abnormal rotators

2302	Abnormally strong absorption lines enhanced He
2422	Abnormally strong molecular bands enhanced TiO
2472	Abnormally strong molecular bands enhanced 12C12C
<i>REMARK Flags</i>	
3010	C/O < 1 in star, classified as M star
3030	C/O < 1 in star, classified as Ba star
4130	T Tauri
4140	weak T Tauri
4400	Pre Main Sequence
4410	Main sequence
4415	Horizontal Branch (HB) stars
4420	Giant or supergiant
15101	No recommended FEH available (WG11)
15102	No recommended XI available (WG11)
15103	Recommended XI computed with calibration (WG11: For stars with $4000 < T_{\text{eff}} \text{ (K)} < 7000$ , $0 < \log g < 5$ , and $-4.5 < [\text{Fe}/\text{H}] < +1$ and previously XI=NaN)
<i>TECH Flags</i>	
9020	Radial velocity determination problem
9021	Reported Radial Velocity Error > 10km/s
9030	Data reduction issues
9050	Bad CCF: cross-correlation function looks distorted or is missing
14101	Single component emission - one intrinsic H $\alpha$ emission component
14102	Single component emission - one intrinsic H $\alpha$ emission component, additional nebular emission component
14201	below-zero or oversubtracted spectrum
14300	At least one absorption component of H $\alpha$ is found in one of the exposures of the object CNAME
14301	At least one intrinsic emission component of H $\alpha$ is found in one of the exposures of the object CNAME
14302	At least one nebular emission component of H $\alpha$ is found in one of the exposures of the object CNAME
15100	Out of grid boundaries: outside the range of $2500 < T_{\text{EFF}} < 8000$ or $0 < \text{LOGG} < 5.5$
15110	Out of grid boundaries: outside the range $2500 < T_{\text{EFF}} < 8000$ or $0 < \text{LOGG} < 5.5$ or $-4 < \text{FEH} < 0.5$ :
15130	Out of grid boundaries: outside the range $5.5 > \text{LOGG} > 0$ :

## Acknowledgements

Please use the following statement in your articles when using these data:  
Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 188.B-3002.