M-subdwarf Research. I. Identification, Modified Classification System, and Sample Construction

Shuo Zhang, A-Li Luo, Georges Comte, John Gizis, Rui Wang, Yinbi Li, Li Qin, Xiao Kong, Bai Yu, Zhenping Yi

M SUBDWARF RESEARCH. I. IDENTIFICATION, MODIFIED CLASSIFICATION SYSTEM, AND SAMPLE CONSTRUCTION

Shuo Zhang,[1,2] A-Li Luo,[1,2] Georges Comte,[3] John E. Gizis,[4] Rui Wang,[1,2] Yinbi Li,[1] Li Qin,[2] Xiao Kong,[1,2] Yu Bai,[1] and Zhenping Yi[5]

[1] Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] Aix-Marseille Univ, CNRS, CNES, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France

[4] Department of Physics and Astronomy, University of Delaware, Newark, DE 19716, USA

[5] School of Mechanical, Electrical and Information Engineering, Shandong University at Weihai, Weihai 264209, China

ABSTRACT

We propose a revision of the system developed by Lépine et al. (2007) for spectroscopic M subdwarf classification. Based on an analysis of subdwarf spectra and templates from Savcheva et al. (2014), we show thatthe CaH1 feature originally proposed by Gizis (1997) is important in selecting reliable cool subdwarf spectra. This index should be used in combination with the [TiO5, CaH2+CaH3] relation provided by Lépine et al. (2007) to avoid misclassification results.

In the new system, the dwarf-subdwarf separators are first derived from a sample of more than 80,000 M dwarfs and a "labeled" subdwarf subsample,these objects being all visually identified from their optical spectra. Based on these two samples, we re-fit the initial [TiO5, CaH1] relation, and propose a new [CaOH, CaH1] relation supplementing the [TiO5, CaH1] relation to reduce the impact of uncertainty in flux calibration on classification accuracy. In addition, we recalibrate the $\zeta_{TiO/CaH}$ parameter defined in Lépine et al. (2007) to enable its successful application to LAMOST spectra.

Using this new system, we select candidates from LAMOST Data Release 4 and finally identify a set of 2791 new M subdwarf stars, covering the spectral sequence from type M0 to M7. This sample contains a large number of objects located at low Galactic latitudes, especially in the Galactic anti-center direction, expanding beyond previously published halo- and thick disk-dominated samples. Besides, we detect magnetic activity in 141 objects.

We present a catalog for this M subdwarf sample, including radial velocities, spectral indices and errors, activity flags, with a compilation of external data (photometric and GAIA DR2 astrometric parameters). The catalog is provided on-line, and the spectra can be retrieved from the LAMOST Data Release web portal.

Keywords: stars: late-type stars: low-mass stars: statistics: subdwarfs

Corresponding author: A-Li Luo
lal@nao.cas.cn

1. INTRODUCTION

M subdwarfs are Galactic fossils with lifetimes much longer than the Hubble time (Laughlin et al. 1997). These faint low mass stars were originally discovered because they combined a large proper motion and a low luminosity (Kuiper 1939), and were subsequently found to share similar kinematics as the inner halo and thick disk stellar populations (Kirkpatrick et al. 1991; Bochanski et al. 2013). M subdwarfs are rare in the solar neighborhood, but they are supposed to be the largest stellar component of the Milky Way's halo (Bochanski et al. 2013), while their cousin M main-sequence dwarfs are the most numerous stellar inhabitants of the Milky Way disk (Reid et al. 2002; Chabrier 2003; Bochanski et al. 2010). Studies of M subdwarf spectra show that they are metal-poor objects compared with the common M dwarfs of near-solar metallicity (Gizis 1997; Lépine et al. 2007), thus making them crucial touchstones of the star formation and metal enrichment histories of the Milky Way (Jao et al. 2016).

Initially, cool subdwarfs were selected from high proper motions catalogs rather than by their spectroscopic features, because the efficiency of photographic spectrographs for faint object is poor in the red spectral range, where their molecular absorption bands are most prominent. Rare examples of spectroscopically identified metal-poor M-subdwarfs were provided by Joy (1947). Then several works suggested that subdwarfs are metal-poor stars associated with halo kinematic population (Mould 1976a; Mould & McElroy 1978). Ake & Greenstein (1980) conducted a spectroscopic survey of high velocity stars and published the spectra of four targets which appeared to have "extreme metal deficiency" compared with the "normal" M-subdwarfs. Similar stars were identified spectroscopically in searches for nearby white dwarfs (Liebert et al. 1979), for Population II halo stars (Hartwick et al. 1984), and in a survey of cool M-dwarfs (Bessell 1982), all focused on faint stars having high proper motions.Until the end of the 20th century, the term "subdwarf" encompassed any number of characteristics including photometric, spectral, or kinematic properties. Meanwhile, a star, whether it shows only one or more of these characteristics, might still be considered as a genuine "subdwarf" (Kirkpatrick et al. 1991).

Spectroscopic analysis is fundamental to access the physical parameters of stars (effective temperature, surface gravity, and surface chemical composition) and their kinematics (radial velocity component). However, the local scarcity of subdwarfs and their intrinsic faintness have long-term made their spectra difficult to obtain and limited to low or moderate resolution. Thus spectral classification is an important step for understanding the physics of M subdwarfs from low or moderate resolution spectra.

The first systematic analysis and classification of M subdwarf spectra was proposed by Gizis (1997), hereafter G97, who divided M dwarfs into three metallicity subclasses ("ordinary" dM, "metal-poor subdwarf" sdM and "extreme-subdwarf" esdM) based on the differences between TiO and CaH molecular absorption bands. Empirical relations were drawn from the spectra of a sample of large proper-motion or low-luminosity targets. Then Lépine et al. (2007), hereafter L07, proposed a revision of the system from more than 400 subdwarfs having halo-like kinematics.L07 defined a metallicity indicator $\zeta_{TiO/CaH}$, which provided a numerical estimate of how the TiO to CaH ratio in a star compares to the value measured in solar-metallicity objects. Besides, L07 also introduced an additional metallicity subclass of "ultrasubdwarf" (usdM) to distinguish the most metal-poor objects and make the classification results consistent with the study of common binaries. In this system, subdwarfs are classified into sdM/esdM/usdM based on their ζ values.

The classification systems have been used in large spectroscopic data sets from modern spectroscopic surveys, which have fulfilled the task of observing multiple targets simultaneously through mature multi-fiber technology, such as the Sloan Digital Sky Survey (SDSS, York et al. 2000), and the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST, Cui et al. 2012). The number of spectroscopically identified subdwarfs is thus increasing rapidly (e.g., Lépine & Scholz 2008; Savcheva et al. 2014; Zhang et al. 2016; Bai et al. 2016). The current largest sample of M subdwarf spectra (Savcheva et al. 2014), hereafter S14, contains 3517 targets from SDSS Data Release 7 (DR7).

However, we made a comparison of the G97 system and the L07 system with S14 subdwarf sample, and the result shows that the L07 system is not identical with the G97 one. Three relations on the [CaH1,TiO5] [CaH2,TiO5] and [CaH3,TiO5] diagrams were first defined in G97 as separators of dM/sdM/esdM, and L07 revised the schemes to a single [CaH2+CaH3, TiO5] relation. Based on a more detailed analysis (which can be seen in Section 2.3), we suggest that the CaH1 index abandoned by L07 is important when searching for genuine optical spectroscopic subdwarfs (see details in Section 2.3).

Up to now, the search for M subdwarfs has been limited to high Galactic latitude areas, favouring the detection of halo and thick disk objects. Whether or not a cool subdwarf population could be present in the Galac-

tic thin disk, mixed (or not) with the immense M dwarf population, and having similar (or not) kinematic and dynamic properties remains an open question. LAMOST data provide an excellent source to address this problem, since this instrument is conducting the most complete spectroscopic survey of the Galactic disk (Luo et al. 2015) especially targeted to the regions of the Galactic anti-center. Therefore, the need for a reliable classification system well adapted to LAMOST products was felt when searching for spectroscopic M subdwarfs in LAMOST data set.

In the present work, in order to revise the classification standards, we use more than 90,000 visually identified M dwarf and subdwarf spectra. Then we apply the modified system to the entire LAMOST Data Release 4 (DR4) to select subdwarf candidates. Ultimately, we obtain a sample of 2791 M subdwarfs validated via visual inspection, covering the spectral sequence from M0 to M7. The catalog is available at http://paperdata.china-vo.org/szhang/DR4_Subdwarfs.csv. Figure 1 shows the distribution of these objects in the Galactic coordinate system.

The present work is limited to spectroscopic identification, classification of candidates, and sample selection in LAMOST data. The paper is organized as follows: we review the previous works and propose solutions to existing problems in Section 2. In Section 3, we introduce the LAMOST data set and our radial velocity (RV) measurement method. In Section 4, we define the revised classification system, present the final subdwarf sample, and briefly summarize the catalog. Finally, in Section 5 we give a summary and conclusion.

2. IDENTIFICATION AND SPECTRAL CLASSIFICATION OF M SUBDWARF STARS: A REVIEW OF EXISTING PROBLEMS

2.1. Basic Characteristics of M Subdwarf Spectra At Low Resolution

M subdwarf stars, as cool stars of low effective temperature (4000~2400K, Rajpurohit et al. 2013) have a spectrum dominated by molecular absorption bands, which show considerable overlapping across the whole visible range. The continuum being weak in the blue, the best observing range with standard optical spectrographic instruments is the red and the deep red, where the spectrum is covered by a dense forest of molecular lines, hiding or blending most of the atomic lines used in usual spectral analysis and diagnostic. At low resolution, molecular absorption bands from metal oxides and hydrides dominate the red range, such as titanium oxide (TiO), vanadium oxide (VO), CaH and H_2O. Eggen & Greenstein (1965) tried to identify subdwarfs with MgH

and TiO bands, then Mould (1976b) proposed a similar method using CaH and TiO bands: because cool subdwarf stars are metal-deficient with respect to normal dwarfs, the metal oxides are depleted with respect to the hydrides at a given effective temperature in their atmosphere, hence the intensity ratio of molecular absorption between an oxide and an hydride should be useful as a separator. The most popular set of indices was defined by Reid et al. (1995) and expanded by Lépine et al. (2003b). We list this classical set of indices in Table 1. Figure 2 shows these wavebands on a template spectrum.

The value of a spectral index is given by:

$$Index = \frac{F_{fea}}{F_{cont}} \qquad (1)$$

where F_{fea} is the mean flux in the molecular absorption feature band, and F_{cont} is a pseudo-continuum flux, i.e., the mean flux in the waveband from λ_{Begin} to λ_{End} (for CaH1 index, the pseudo-continuum flux is the average flux of two bands). For example, the CaOH index is measured using the mean flux within fea_CaOH divided by the one within cont_CaOH.

2.2. Review of the Selection Schemes and Classification Systems

2.2.1. G97 system

The spectroscopic classification system of M subdwarfs was initially defined by G97, from spectra of 79 targets mostly selected from large-proper-motion catalogs such as LHS (Luyten 1979) and the Lowell Proper Motion Survey (Giclas et al. 1971). Based on quantitative measures of TiO and CaH features (CaH1, CaH2, CaH3 and TiO5), G97 defined a set of empirical relations on the CaHn−TiO5 (n=1,2,3) diagrams that classified M dwarf spectra into three subclasses corresponding to increasing metal-poor levels: dwarfs of solar metallicity (dM), normal subdwarfs (sdM), and extreme-subdwarfs (esdM).

The initial condition proposed by G97 to identify a spectroscopic subdwarf was the [CaH1, TiO5] relation:

$$CaH1 < 0.695 \times TiO5^3 \ - \ 0.818 \times TiO5^2 \\ + \ 0.413 \times TiO5 \ + \ 0.651 \qquad (2)$$

G97 suggested it to be the primary dwarf / subdwarf separator, and added a [CaH2, TiO5] relation to separate classic subdwarfs from extreme subdwarfs. Finally, G97 assigned the spectral subclass (linked to effective temperature) using the CaH3 index. For the coolest stars with TiO5<0.49, G97 also suggested that the CaH2 or CaH3 index must be used to avoid inaccuracy caused by the "saturation" of the [CaH1, TiO5] relation.
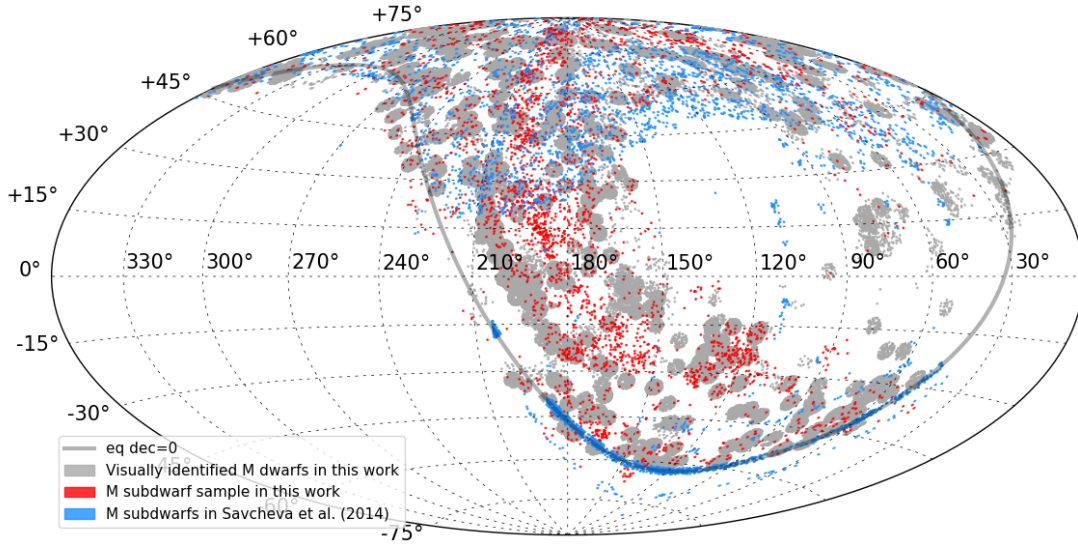
Figure 1. The LAMOST DR4 M dwarfs and subdwarfs in the Galactic coordinate system

Table 1. The features used for classification

No.	Feature	$\lambda_{Begin}(\mathring{A})$	$\lambda_{End}(\mathring{A})$
1	fea_CaOH	6230.0	6240.0
2	fea_CaH1	6380.0	6390.0
3	fea_CaH2	6814.0	6846.0
4	fea_CaH3	6960.0	6990.0
5	fea_TiO5	7126.0	7135.0
6	cont_CaOH	6345.0	6355.0
7	cont_CaH1_1	6345.0	6355.0
8	cont_CaH1_2	6410.0	6420.0
9	cont_CaH2	7042.0	7046.0
10	cont_CaH3	7042.0	7046.0
11	cont_TiO5	7042.0	7046.0

NOTE—The useful quantity is the mean flux across the band, and the bands are in air wavelength.

2.2.2. *L07 system*

Lépine et al. (2003a) later suggested that the [CaH1, TiO5] relation could be abandoned due to its shorter dynamic range and the lower signal-to-noise of CaH1 band in very cool stars. They proposed its replacement by a [CaH2+CaH3, TiO5] relation, as producing a result almost equivalent to the former classification for the same stars, according to Lépine et al. (2003b). Along this way, Burgasser & Kirkpatrick (2006) went one step further, defined two separators on the [CaH2+CaH3,TiO5] diagram to differentiate between dM/sdM/esdM:

$$sdM : (CaH2 + CaH3) < 1.31 \times TiO5^3$$
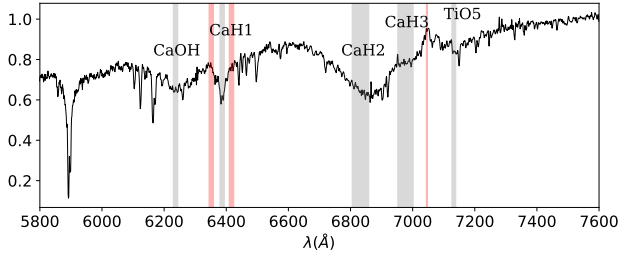$$- 2.37 \times TiO5^2 + 2.66 \times TiO5 - 0.20 \quad (3)$$

Figure 2. A template esdM spectrum in the rest frame. The wavebands used to calculate the spectral indices are marked by the translucent gray bands, corresponding to the "fea_" features in Table 2, and the bands marked in red show the "cont_" features used as pseudo-continuum.

$$esdM : (CaH2 + CaH3) < 3.54 \times TiO5^3$$
$$- 5.94 \times TiO5^2 + 5.18 \times TiO5 - 1.03 \quad (4)$$

From a much larger published sample of M type spectra selected from the high proper motion catalog LSPM-north (Lépine & Shara 2005), L07 proposed an update of this classification system based on an empirical calibration of the TiO/CaH ratio for stars of near-solar metallicity. For this, they introduced a parameter $\zeta_{TiO/CaH}$. which quantifies the weakening of the TiO band-strength due to metallicity effect, with values ranging from $\zeta = 1$ for stars of near-solar metallicity to $\zeta = 0$ for the most metal-poor (and TiO depleted) subdwarfs. The ζ parameter is defined as:

$$\zeta_{CaH/TiO} = \frac{1 - TiO5}{1 - [TiO5]_{Z_\odot}} \quad (5)$$

the $[TiO5]_{Z_\odot}$ is a third order polynomial fit of the TiO5 spectral index as a function of the CaH2+CaH3 index. In L07, these indices are measured in spectra of a kinematical selection of local disk dwarfs of roughly solar metallicity, giving:

$$[TiO5]_{Z_\odot} = -0.05$$
$$- 0.118 \times (CaH2 + CaH3)$$
$$+ 0.670 \times (CaH2 + CaH3)^2 \quad (6)$$
$$- 0.164 \times (CaH2 + CaH3)^3$$

L07 suggested that the separation between subdwarfs and ordinary dwarfs could be effective when applying the single condition: $\zeta < 0.825$. L07 also refined the scheme by introducing an additional class of ultra-subdwarfs (usdM) corresponding to the most metal-poor ones.

Since then, L07 system has been widely used in the identification, classification, and subtype determination of the cool subdwarf spectra (e.g., Lépine & Scholz 2008; Bochanski et al. 2013; Savcheva et al. 2014; Bai et al. 2016; Zhang et al. 2017).

2.2.3. ζ and the metallicity

There is a major problem with the sdM/esdM/usdM subclasses in the classification systems above. The subclass defined by CaH and TiO indices were originally used to infer metallicity levels, although the separators do not seem to run parallel to the lines of iso-metallicity drawn from synthetic model grids such as the NextGen grid used by L07. Jao et al. (2008) then compared GAIA stellar atmosphere model grid with 88 subdwarf spectra and pointed out that these indices are affected in complicated ways by combinations of the temperatures, metallicities, and gravities of subdwarfs. Lépine et al. (2013) also compared observed spectra with BT-Settl stellar atmosphere model grids, and concluded that the TiO/CaH ratio is not primarily sensitive to the classical metallicity value [Fe/H], but rather depends on the [α/H] because O, Ca, and Ti are all α-elements. Variations in [α/Fe] would thus weaken the correlation between ζ and [Fe/H]. More information is necessary to assess stellar atmospheric parameters of a subdwarf, such as a binary membership, or detailed model fitting of a high-resolution spectrum (Rajpurohit et al. 2016).

Yet even so, the separators defined by L07 in the [CaH2+CaH3, TiO5] space can still indicate a rough ordering in metallicity, as they were verified on a sample of resolved subdwarf binaries, and in fact the metallicity subclasses are more associated with the kinematics of the objects as they were first determined by subdwarfs with different kinematic properties. To separate subdwarf spectra from normal main-sequence dwarfs, the parameter $\zeta_{TiO/CaH}$ still appears to be to be a robust indicator.

2.3. Problems in the Dwarf/Subdwarf Separation and Subclass Limits

2.3.1. Specific Problems with S14 Sample Selection

To search for subdwarfs, S14 adopted L07 system as filter to select candidates from M type spectra in SDSS Data Release 7, and finally identified 3517 M subdwarfs with optical spectra. However, two problems may be identified in this subdwarf sample:

1)As shown in Figure 3, a large fraction (1447) of objects in the sample do not seem to satisfy equation (2), which means the L07 system is not identical to the original G97 system. In fact, these targets do not exhibit classical "subdwarf" characteristics such as large proper motions. This leads to the suspect that there is a group of stellar objects having subdwarf-like [CaH2+CaH3]-TiO5 spectral features ratios but which might not be
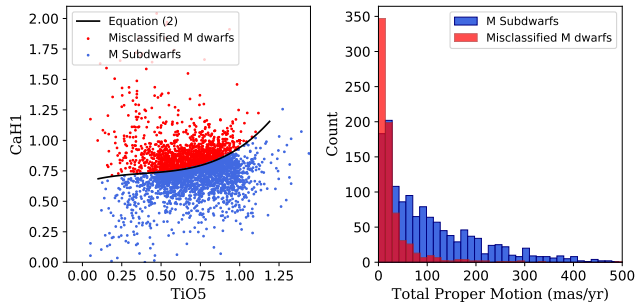
Figure 3. Distribution of the subdwarfs from Savcheva et al. (2014) catalog in the [CaH1, TiO5] index diagram and histogram of their proper motions. The subdwarfs were selected by a [CaH2+CaH3]-TiO5 relation defined in Lépine et al. (2007) from SDSS Data Release 7. The black curve on the left diagram is the original subdwarf selection criterion, i.e. the CaH1-TiO5 relation defined in Gizis (1997) and abandoned by Lépine et al. (2007). There are 1447 targets (red dots as group I) in this "subdwarf" sample not fulfilling the CaH1-TiO5 relation which thus might be mis-classified M dwarfs. The other ones, which could be considered as reliable subdwarfs are figured as blue dots (group II). The right panel shows that most group I stars do not exhibit large proper-motions.

classified as genuine "subdwarfs". Hence, neglecting the primary condition defined by G97 –CaH1 feature– leads to the misclassification results of some stellar objects when using only L07 system.

2) In the work of S14, the online catalog provided CaH, TiO indices and ζ values for the M subdwarf sample, a large part of which were directly taken from West et al. (2011). These measurements were performed on spectra reduced to rest-frame using radial velocity (RV) measurements by cross-correlation with M dwarf templates. As mentioned above, S14 adopted $\zeta < 0.825$ as the subdwarf selection criterion. However, S14 also provided a value of RV for each target derived from cross-correlation with a set of M subdwarf template spectra built by the authors. We have recomputed the CaH, TiO spectral indices and ζ index using these last RV values: the results plotted in Figure 4 show that there are 744 objects with $\zeta > 0.825$ which should be classified as dwarfs rather than subdwarfs according to this selection criterion.

These discrepancies clearly arise from the accuracy of the rest-frame reduction of spectra used to measure indices. The whole process is highly sensitive to the template spectra used in the cross-correlation. Therefore, a method independent of templates to measure RVs of the candidates is mandatory for us to ensure the accuracy of the final sample, since the LAMOST pipeline does not contain subdwarf templates.

2.3.2. *The Need to Recalibrate $\zeta_{TiO/CaH}$*

Lépine et al. (2013) made a comparison of spectra from the same stars obtained at different observatories, which revealed that spectral band index measurements are dependent on spectral resolution, spectrophotometric calibration, and other instrumental factors. Different spectral resolution, spectrophotometric calibration, and other instrumental factors would all lead to different spectral index values for a same target spectrum (Lépine et al. 2013), and different data sets may also have their own biases due to selection effects. The authors thus suggested that a consistent classification scheme requires that spectral indices be calibrated and corrected for each observatory/instrument combination used. For example, Dhital et al. (2012) recalibrated the definition of ζ with a sample of SDSS binary systems to correct the bias in ζ observed in early-type M dwarfs. Further on, Lépine et al. (2013) recalibrated the ζ parameter with corrected spectral index values from their new larger dwarf sample. Table 3 lists the coefficients of the polynomial fit of the [CaH2+CaH3] vs TiO5 mean relation for ordinary dwarfs obtained by these authors.

Hence, we have decided to derive a new calibration of the ζ index based on LAMOST data.

2.3.3. *Difficulties in Using LAMOST Spectra*

Due to the low luminosity of low-mass stars and efficiency limits of LAMOST instrument, a large fraction of the M type stars don't have high-quality spectra. In addition, since the spectral indices are defined as flux ratios, the accuracy of them used for selection and classification may also be affected by continuum slope deformation from flux calibration problems (Du et al. 2016). For example, the waveband of CaH1 feature is ~ 700Å away from TiO5, hence the inaccuracy of continuum slope may affect the comparison result of the CaH1 index with the TiO5 index.

Bai et al. (2016) made an effort to search for M subdwarfs in the 2nd data release of LAMOST survey (which contains more than 200,000 M type spectra). The authors used visual inspection as the final identification tool to avoid mistakes caused by the factors mentioned above. As a result, they finally verified 108 objects from the candidates after a prior selection by G97/L07 schemes. This number very probably underestimates the subdwarf fraction potentially contained in DR2 because the classification standards were defined on much smaller samples and should not be directly applied to LAMOST data.

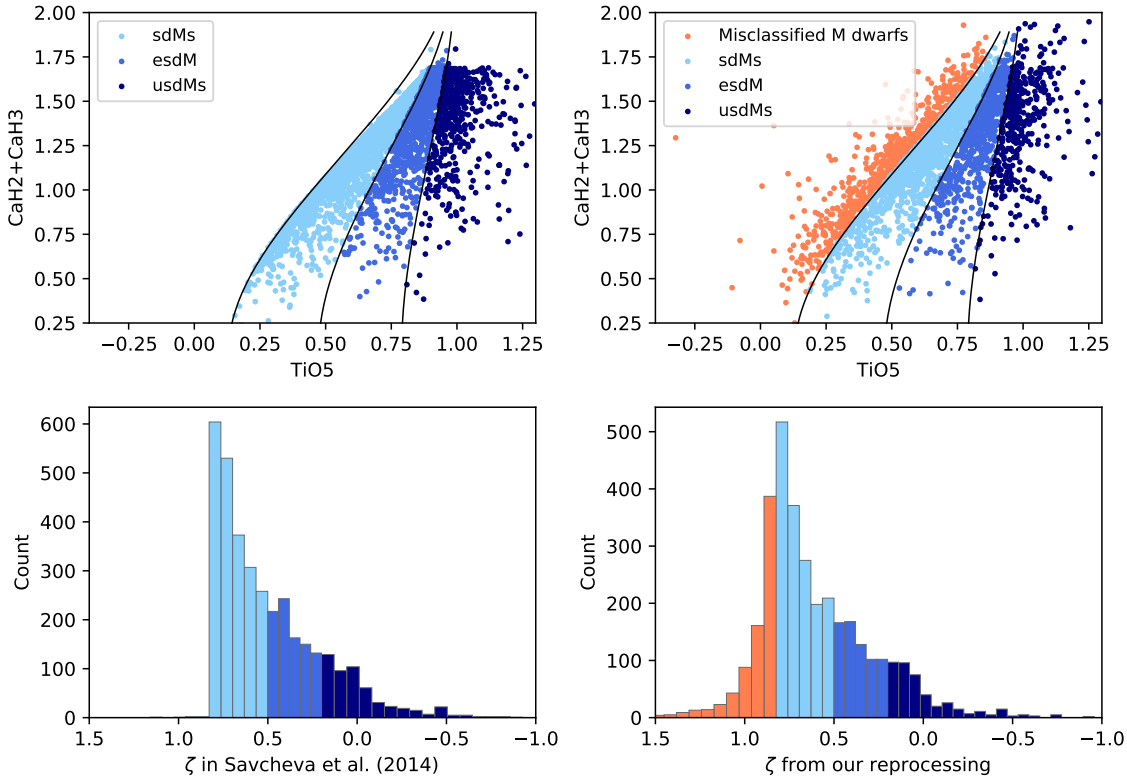### 2.4. *A Short Introduction of Proposed Solutions*

Figure 4. Sensitivity of the spectral classification indicators to the spectrum rest-frame reduction accuracy. The two left panels show the distribution of spectral indices and ζ parameter taken from the large subdwarf catalog of Savcheva et al. (2014) (S14). They are based on spectra reduced with radial velocities (RVs) from West et al. (2011), obtained by cross-correlation with ordinary M dwarf templates. The subdwarfs are selected from dwarfs using $\zeta < 0.825$ and further divided into sdM (light blue), esdM (blue) and usdM (dark blue) based on decreasing ζ values. The right panels show our reprocessing of the S14 sample using the set of RVs (also provided by S14, but not used by them) based on cross-correlation with purposely built subdwarf templates. 744 objects (orange dots) escape the selection criterion and have $\zeta > 0.825$, which makes them probable misclassified ordinary dwarfs.

Taking factors above into account, we conclude that a reliable M dwarf sample and a carefully verified M subdwarf sample from LAMOST data must first be built in order to define the separators between subdwarfs and dwarfs. An accurate rest-frame reduction, based on a RV measurement method independent of template spectra is necessary to get correct spectral indices. The dwarfs can be used to calibrate the ζ parameter, and the CaH1 index should be taken into consideration. In addition, the CaOH index, is much closer to CaH1 (~ 150Å) than TiO5, and its comparison with CaH1 is as useful as TiO5 in separating subdwarfs from dwarfs. Then, combining revised G97 and L07 schemes, an automated search for subdwarf candidates in the entire LAMOST DR4 can be made. Finally, visual inspection will be adopted as the ultimate tool for confirmation of subdwarf identification.

3. LAMOST OBSERVATION AND DATA PREPARATION

3.1. *LAMOST Data Release 4*

LAMOST is an all-reflective Schmidt-type telescope located in Xinglong Station of National Astronomical Observatory, China (105° E, 40° N). It has a 6.67 m spherical primary mirror, which, combined with the corrector, provides an effective aperture of 4 m. It offers both a large field of view (5°) and a large aperture ratio. 4000 fibers mounted on the focal plane, feeding 16 double channel spectrographs, allow a high spectral acquisition rate (Cui et al. 2012). As it is dedicated to a spectral survey of celestial objects over the entire available northern sky, both Galactic and extragalactic surveys are conducted. The Galactic one, LEGUE, focuses on the Galactic Anti-center direction, the disk at selected longitudes away from the Galactic Anti-center, and the halo (Luo et al. 2015).

Standard techniques are used in the analysis of the spectra. On the original CCD image, 4K pixels are used to record blue and red wavelength regions across 3700-

5900Å and 5700-9000Å, respectively. The blue and red channels are combined and each spectrum is re-binned to calculate the radial velocity. A combined spectrum is re-sampled at a scale of 69 km s^{-1} per pixel, i.e., the difference between two adjacent points in wavelength is $\Delta\log(\lambda) = 0.0001$. The raw CCD data are reduced by the LAMOST data reduction software named LAMOST 2D pipeline. The 2D pipeline performs the tasks of subtraction of dark and bias, correction of flat field, extraction of spectra, subtraction of skylight, calibration of wavelength, merging of sub-exposures with cosmic ray elimination and combination of blue and red wavelength ranges (Luo et al. 2015).

The Data Release 4 of LAMOST regular survey contains 7,620,612 spectra covering the entire optical band ($\sim$ 3700 $-$ 9000Å) at resolution R$\sim$1800, including 6,944,971 stellar spectra, 153,348 extragalactic spectra (galaxies and QSO), and 522,293 spectra classified as Unknown (most of the latter because of insufficient signal-to-noise).

3.2. Radial Velocity Measurement

Heliocentric radial velocity (RV) is the projection of the star's 3D velocity along the line of sight and can be measured from the Doppler shift of spectral lines. A conventional and accurate way to determine RV is by cross-correlating the target spectrum with a rest-frame template spectrum of similar type (Tonry & Davis 1979). LAMOST 1D pipeline provides a redshift measurement "redshift z" for each object by matching with a template spectrum. However, up to now, M subdwarf templates are not yet available for the 1D pipeline, and thus the RVs measured using ordinary M dwarf templates for subdwarf candidates would be of insufficient accuracy.

To overcome this problem, we have adopted an alternate RV determination method independent of templates. The natural way is to fit single Gaussian profiles to a set of absorption lines of neutral metals (Savcheva et al. 2014; Bai et al. 2016). However, M subdwarfs are basically faint red objects whose spectra often have a low or very low signal-to-noise ratio, and the useful metal lines at low resolution often don't exhibit nice Gaussian profiles. In many cases, the profiles are constructed by only 3$\sim$5 flux points and not even symmetric. Therefore, we designed the following process to estimate RV from the Doppler shift of eight prominent absorption lines listed in Table 3.2. The method is based on the fact that the wavelengths of the 8 lines would shift simultaneously due to the Doppler effect caused by radial motion, hence the offsets of the line centers would be identical in logarithmic wavelength. The details of this method are described below:

Table 2. Absorption Lines Used For RV Measurment

Line	Center (Å)
K I	7667.0089
K I	7701.0825
Na I	8185.5054
Na I	8197.0766
Ti I	8437.2600
Ca II	8500.3600
Ca II	8544.4400
Ca II	8644.5200

NOTE—The line centers adopted from NIST are in vacuum wavelength.

1. The wavelengths of a LAMOST spectrum are recorded in exponential form in the fits file header as

$$\lambda = 10^{(head+n*step)}$$

where *head* corresponds to the COEFF0 field which records the decimal logarithm of the central wavelength of the first pixel, (usually around 3.5682), $n = 0, 1, 2, 3, ...$, and $step \equiv 0.0001$. Therefore the observational accuracy of RV when the shift of a line is defined to 1 step unit i.e., when $\Delta n = \pm 1$, is $c*(1-10^{head+\Delta n*0.0001}) \approx \pm 69$ km/s. We recall that this corresponds to the sampling step of the pipeline reconstructed spectrum: the actual physical spectral resolution (somewhat variable along the spectral range) is of the order of 2.5 sampling steps, but naturally, radial velocity measurement methods actually allow final accuracies of a fraction of a spectrum pixel.

2. Since the LAMOST subdwarfs are within 1 kpc of the Sun (from apparent magnitude considerations), the Sun's escape velocity—550 km/s (Kafle et al. 2014)—can be used as a proxy for the escape velocity of the entire sample. When $\Delta n = \pm 10$ the corresponding RV is ± 690 km/s, hence we suppose that the shift of a line would be less than 10 step units.

3. Without any prior knowledge of the type of spectrum (late K star, M star, QSO or Unknown), for each line in Table 3.2, we get the wavelength step index n_0 in the spectrum corresponding to its rest-frame wavelength, and then find the Δn of the lowest flux point in the range $[\lambda_{n_0-10}, \lambda_{n_0+10}]$. We insist that this is an entirely automated process that could provide positions of "true" absorption lines as well as simple accidental deeps in the object's continuum, but its virtue is that the "line" profiles themselves are not required to be gaussian.

4. We count the number of "lines" affected by each $\Delta n \in [-10, 10]$. The Δn corresponding to the largest number of lines is adopted as our target $\Delta n*$. Note that the number of lines should be at least 3 to avoid a large random error, otherwise we'll flag the target as unable to be measured. And at later visual inspection stage of the selection process, we also filter out the objects with more than 3 "lines" which are merely local flux minima instead of real absorption lines.

5. The "lines" affected by the shift $\Delta n*$ are then used to measure the RV: this is achieved on an interpolated spectrum with artificially 15 times increased resolution to acquire a better precision (Bochanski et al. 2007). The mean value of the RVs obtained from these lines is our final RV, and the standard deviation is the error.

Our RV measurements have a median achieved typical uncertainty of 13.9 km s^{-1}. In Figure 5 we compare the RVs of 1926 M dwarfs from our M dwarf sample (see Section 4.1 below) with the RVs provided by GAIA DR2 on the same stars. Our measurements show a fair agreement with GAIA RVs with a offset of 5.50 km s^{-1} and a dispersion of 13.5 km s^{-1}. Similar systematic offsets are also found in other comparisons: 4.54 km s^{-1} between LAMOST and APOGEE (Anguiano et al. 2018), and 3.8 km s^{-1} between LAMOST and MMT+Hectospec (Huang et al. 2015). Note that four of the eight absorption lines (the K I doublet and the Na I doublet) are located within telluric absorption bands and sometimes the line centers are altered by insufficiently corrected telluric absorption by O_2 and H_2O. To solve this problem, we assign a same weight to each of the eight lines of Table 3.2 and impose that three "lines" at least must be detected to provide a valuable result. In addition, we also inspect the 7000-9000Å range of each spectrum in our final subdwarf sample and remove the objects which do not show identifiable absorption lines by eye-check to improve the reliability of the final results.
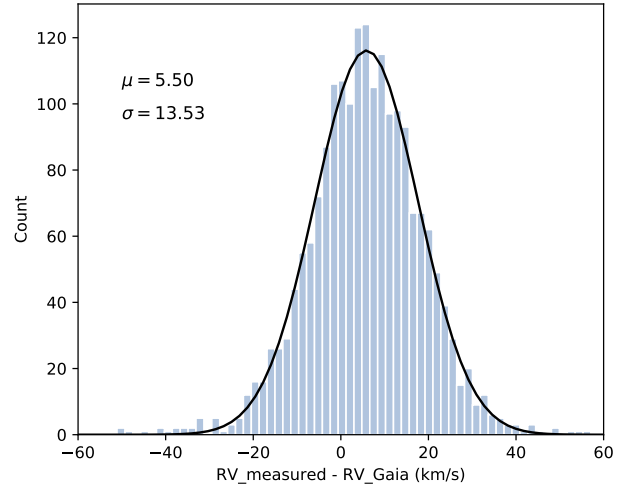

Figure 5. Comparison of RVs measured in this work (Subsection 3.2) on 1926 stars in the M dwarf comparison sample with RVs provided by GAIA radial velocity survey. The GAIA spectrometer collects medium resolution (R $\sim$ 11,700) spectra across the wavelength range 845-872 nm centered on the Calcium triplet region (Cropper et al. 2018).

4. SEARCH FOR M SUBDWARFS IN LAMOST DR4

In order to search for M subdwarf spectra in the entire LAMOST Data Release 4, we modify the L07 system and apply its new version onto the data set. The implementation process to obtain the final subdwarf sample is thus as follows:

1. Selection of M dwarf comparison sample and of the subdwarf "labeled" sample. We visually inspect the entire M-type stellar spectra collection (99,671) from LAMOST 4th year regular survey. Among this sample, 325 subdwarfs and 90,018 ordinary M dwarfs have been identified. Adding the subdwarfs from Bai et al. (2016), the identified M subdwarf sample contains more than 400 objects. We call this subsample the subdwarf "labeled" sample. The inspection process is discussed in Section 4.1.

2. Calibration of ζ. Based on the M dwarf sample, we recalibrate the ζ index as initially defined in L07 with the fitting of the [TiO5, CaH2+CaH3] relation. We adopt as new maximal ζ values 0.75, 0.5, and 0.2 to separate M type subdwarf stars into sdM/esdM/usdM respectively. The details are described in Section 4.2.

3. Definition of CaH1 separators. We re-fit the original G97 [TiO5, CaH1] relation using LAMOST subdwarf "labeled" sample and derive equation 8. We also define a new [CaOH, CaH1] relation

(Equation 9) for supplement. The equations are shown in Figure 9. The process is detailed in Section 4.3.

4. Determination of spectral subtype using Equation 10.

5. Searching for M subdwarfs in LAMOST DR4. Applying the constraints above on 1,271,426 spectra from DR4 dataset, we obtain nearly 10,000 candidates. Finally, 2791 targets among these candidates are confirmed as M subdwarfs by visual inspection. Section 4.5 gives the details and Section 4.6 introduces the subdwarf catalog and the additional external data.

4.1. The Subdwarf "Labeled" Sample and the M Dwarf Comparison Sample

The initial step consists in assembling two reference samples from LAMOST data: one of ordinary M dwarfs, the other of carefully validated M subdwarfs. These allow the revision of the various selection and classification tools of G97 and L07 in order to apply them to LAMOST spectra. Since the existing standard M subdwarf spectral templates are not perfectly appropriate for LAMOST spectra, we choose the most reliable identification/confirmation method underlined in all the former works —visual inspection— as our first identification tool to assemble a reliable subdwarf sample from LAMOST data. Although Bai et al. (2016) provided 108 verified M subdwarfs from LAMOST DR2, the candidates submitted to visual inspection were provided by screening through the G97/L07 systems. Therefore, to obtain a larger subdwarf sample and meanwhile select a comparison sample of pure ordinary M dwarfs, the visual inspection was targeted at the entire M type spectra dataset (auto-classified by LAMOST 1D pipeline) from the 4th year of LAMOST regular survey, containing nearly 100,000 spectra. In this process, potential contaminants such as M giants, double stars and unrecognizable objects can be removed, and the spectral subtype of each M dwarf spectrum can also be verified based on the eyecheck result.

To achieve this arduous task, we use the manual "eyecheck" mode of Hammer spectral typing facility (Covey et al. 2007), an IDL code which uses the relative strength of a series of spectral features in the range 4000∼9100Å for classifying stellar spectra (Lee et al. 2008; West et al. 2011; Dhital et al. 2012; Savcheva et al. 2014). Yi et al. (2014) have modified the Hammer code to better adapt it to LAMOST M spectra. The manual mode of Hammer allows the user to compare a target spectrum with a collection of M dwarf template spectra and to assign a label to each target.

Referring to the aspect of single subdwarf spectra and/or templates provided in previous works such as G97, L07, Jao et al. (2008), S14, Bai et al. (2016), we base our initial assignment of a target spectrum into the "giant", "dwarf" or "subdwarf" category on the following basic recipes:

1) The K I doublet around 7700Å and Na I doublet around 8200Å are highly gravity-sensitive, thus they are almost invisible in giants, but prominent both in dwarfs and subdwarfs. Comparing with dwarfs, the CaH1 absorption band of an M giant spectrum of same spectral type is almost invisible, as is the CaH3 band, while its TiO5 minimum is as deep as CaH2 minimum.

2) The most obvious differences between M subdwarfs and dwarfs are the features around 6200-6400Å and 6800-7200Å. CaH1 absorption band is much more prominent in subdwarf spectra than in dwarfs, almost as deep as CaOH and even deeper respectively for esdM and usdM. Regarding subdwarfs, TiO5 minimum is less deep than CaH2 minimum for sdM, even less deep than CaH3 minimum for esdM, and the entire TiO absorption band is almost invisible for usdM.

Note that the reddest part of TiO5 is blended with a strong atomic line 7148Å of CaI which is very close to the third (and deepest in normal conditions) "tooth"-like absorption of the TiO, which makes it easy to misidentify the minimum of TiO5 on low-resolution noisy spectra.

In this step, we label each target with a spectral subtype (M0, M1,...M9) and a type information such as "dwarf", "giant", "subdwarf", "double star", or "odd". As a result, 325 subdwarfs are selected and validated. Combining with the 108 M subdwarfs from Bai et al. (2016), we now have a "labeled" subdwarf sample including 433 visually identified spectra in total.

Besides, 90,018 ordinary M dwarfs are identified, of which 83,213 have measurable RVs: these define our M dwarf comparison sample. This comparison sample is also used for the calibration of ζ.

4.2. Calibration of $\zeta_{TiO/CaH}$

We measure the spectral indices defined in Section 2.1 and Table 1 on all spectra of the "labeled" subdwarf sample and on all spectra of the M dwarf comparison sample. The M dwarf results are used to build the TiO5−[CaH2+CaH3] diagram shown in Figure 6. The barycentric line across the data cloud provides directly
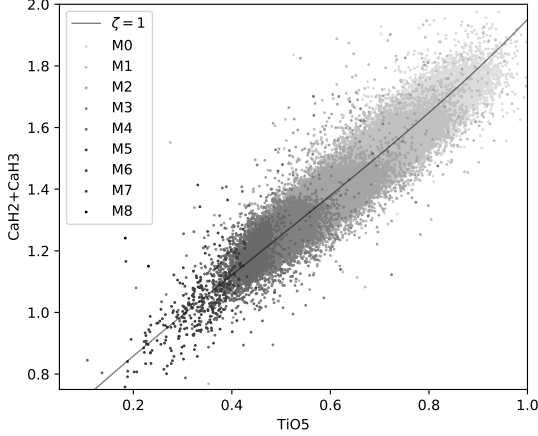
Figure 6. Recalibration of the $\zeta_{TiO/CaH}$ parameter (denoted ζ for short) based on a sample of 83,213 M dwarf spectra from LAMOST regular survey. ζ is a combination of the TiO5, CaH2, and CaH3 spectral indices (see definition in equation 5) and has been shown to be correlated with metallicity. The black curve is an iso-ζ line with $\zeta = 1$, providing an estimate of TiO to CaH ratio in an object of solar-metallicity. Smaller ζ values correspond to stars of lower metallicity. The index was first introduced in Lépine et al. (2007), and then recalibrated in Dhital et al. (2012) and Lépine et al. (2013). Table 3 compares these calibrations.

the recalibration of the $\zeta_{TiO/CaH} = 1$ curve for average solar metallicity objects. We get:

$$
\begin{aligned}
[TiO5]_{Z_\odot} = & -1.326 \\
& + 2.417 \times (CaH2 + CaH3) \\
& - 1.083 \times (CaH2 + CaH3)^2 \\
& + 0.2465 \times (CaH2 + CaH3)^3
\end{aligned}
\tag{7}
$$

Based on this updated calibration of ζ, we find that the optimal separator curve used to select subdwarfs from dwarfs in the [TiO5, [CaH2+CaH3]] space can be defined as $\zeta < 0.75$ as shown in Figure 7: it exactly separates almost all our subdwarfs belonging to the "labeled" subsample. Furthermore, $\zeta < 0.75$ excludes a sufficient fraction, in statistical sense, of ordinary dwarfs (98%). This ζ value is smaller than the canonical 0.825 recommended by L07, Dhital et al. (2012), and Lépine et al. (2013), which reflects the differences in instrumental response to the real spectra between LAMOST and other setups. The separators between sdM/esdM/usdM are still adopted as $\zeta < 0.5$ and $\zeta < 0.2$. Figure 7 shows the comparison between these several systems.

4.3. *CaH1 - based Dwarf/Subdwarf Separators*

The characteristic of prominent CaH1 absorption in subdwarfs can be clearly seen on Figure 8, where we

compare the index values of M dwarf and subdwarf templates built by Bochanski et al. (2007) and S14 respectively. The result shows excellent discrimination between dwarfs and subdwarfs across the entire spectral sequence when plotting on the [CaH1, TiO5] diagram. Therefore, we suggest that the CaH1 feature should be taken into consideration to select spectroscopic subdwarfs, especially for early-M type subdwarfs of moderate metal deficiency in which TiO5 absorption is not very strong and ζ loses efficiency as separator.

In addition, as the middle diagram of Figure 8 shows, the CaOH index and the TiO5 index show similar monotonic trend along the effective temperature sequence. The result is to be expected, because the CaOH band depth, as well as the TiO5 one, heavily relies on the atmospheric abundance of the alpha element O. Due to the CaOH feature being much closer to the CaH1 band than the TiO5 feature, the ratio of CaOH index to CaH1 index maybe less affected by errors arising from the distortion of continuum affected by various instrumental factors or flux calibration problems (Du et al. 2016).

Using the visually identified dwarfs and "labeled" subdwarfs, we derive two equations to define separator curves on the CaH1 vs. TiO5 index diagram and CaH1 vs. CaOH index diagram respectively, yielding:

$$
CaH1 < 0.6522 \times TiO5^2 - 0.577 \times TiO5 + 0.8937 \tag{8}
$$

$$
\begin{aligned}
CaH1 < & \ 0.4562 \times CaOH^3 - 0.1977 \times CaOH^2 \\
& - 0.01899 \times CaOH - 0.7631
\end{aligned}
\tag{9}
$$

Equation (8) is shown in the left panel of Figure 9 with a comparison with the original one from G97, i.e., equation (2). Equation (9) is shown in the right panel of Figure 9.

4.4. *Spectral Subtyping*

The determination of spectral subtypes (SpT) for cool subdwarfs can be made by several methods, such as using the depth of the CaH molecular bands (G97; L07) or the shape of relatively opacity-free region within 8200-9000Å (Jao et al. 2008). In this work we follow the definition in L07 in which the "metallicity" subclass assignment is dependent on the combined value of the CaH2 and CaH3 indices, as:

$$
Sp = 1.4 \times (CaH2+CaH3)^2 - 10 \times (CaH2+CaH3) + 12.4 \tag{10}
$$

It will thus be easy to compare our sample with previous works on the spectral sequence. The derived value of SpT is rounded up to the nearest integer, giving the spectral subtype value from 0 for M0 to 7 for M7.
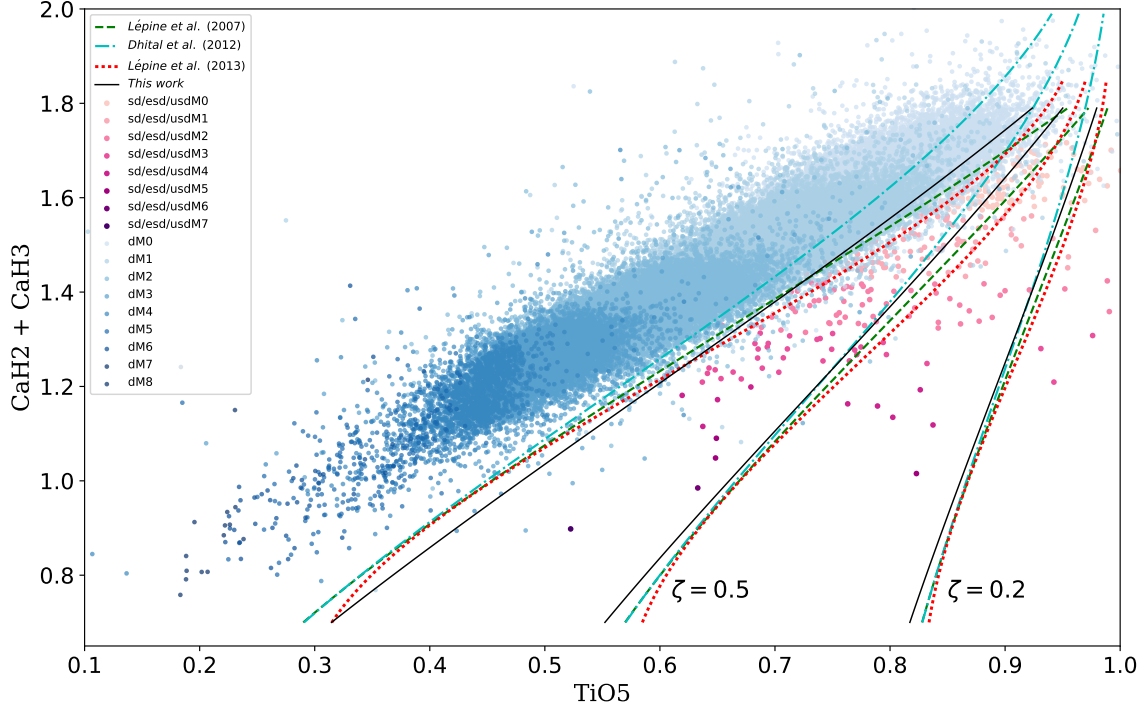
Figure 7. New calibration of the ζ parameter used for subdwarf selection and classification. We plot the ordinary M dwarf sample (83,213 objects) as blue dots and the "labeled" subdwarf sample (433 objects) as pink dots (see Section 4.1) in a [TiO5, CaH2+CaH3] index diagram. The final condition we adopt to classify a target as a "subdwarf" is $\zeta < 0.75$ (plotted as the solid black curve) because it can exactly cover most of our "labeled" subdwarf sample and excludes more than 98% of the M dwarfs. The iso-ζ contours from the earlier calibrations of Lépine et al. (2007), Dhital et al. (2012) and Lépine et al. (2013) are also shown for comparison as dashed, dash-dotted, and dotted lines, respectively. The iso-ζ curves used to separate subdwarfs from dwarfs in the three former systems correspond to $\zeta < 0.825$, as shown in the figure, (all different from each other due to sample selection bias). We formally adopt $\zeta < 0.5$ and $\zeta < 0.2$ as separators between sdM, esdM and usdM, following the former systems.
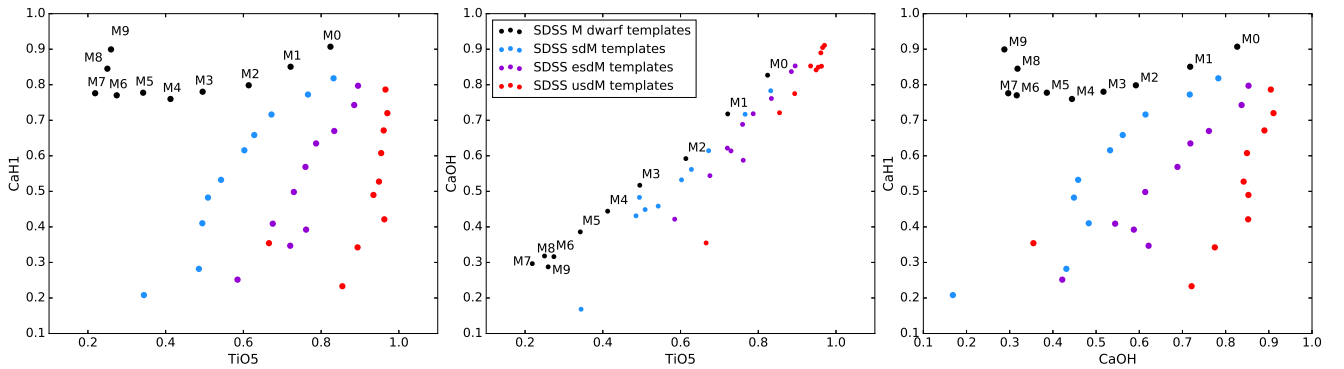

Figure 8. The three panels compare the M dwarf template spectra built by Bochanski et al. (2007) from the Sloan Digital Sky Survey (SDSS) with M subdwarf template spectra (sdM/esdM/usdM) built by Savcheva et al. (2014), also from SDSS, on the [TiO5, CaH1], [TiO5, CaOH], and [CaOH, CaH1] index diagrams respectively. The left panel shows that M subdwarfs are clearly separated from dwarfs by the CaH1-TiO5 index couple. The central panel shows that the CaOH index variation is consistent with that of TiO5 index, leading to similar results between CaH1 vs. TiO5 (left panel) and CaH1 vs. CaOH (right panel) distributions. A CaOH-CaH1 relation can thus be defined on the [CaOH, CaH1] index diagram to separate subdwarfs from dwarfs.

Table 3. Coefficients for different calibrations of $[TiO5]_{Z_\odot}$ (See equation 6 and 7)

Coefficients	Lépine et al. (2007)	Dhital et al. (2012)	Lépine et al. (2013)	This work
c0	-0.05	-0.047	0.622	-1.326
c1	-0.118	-0.127	-1.906	2.417
c2	0.67	0.694	2.211	-1.083
c3	-0.164	-0.183	-0.588	0.2465
c4	...	-0.005	...	...
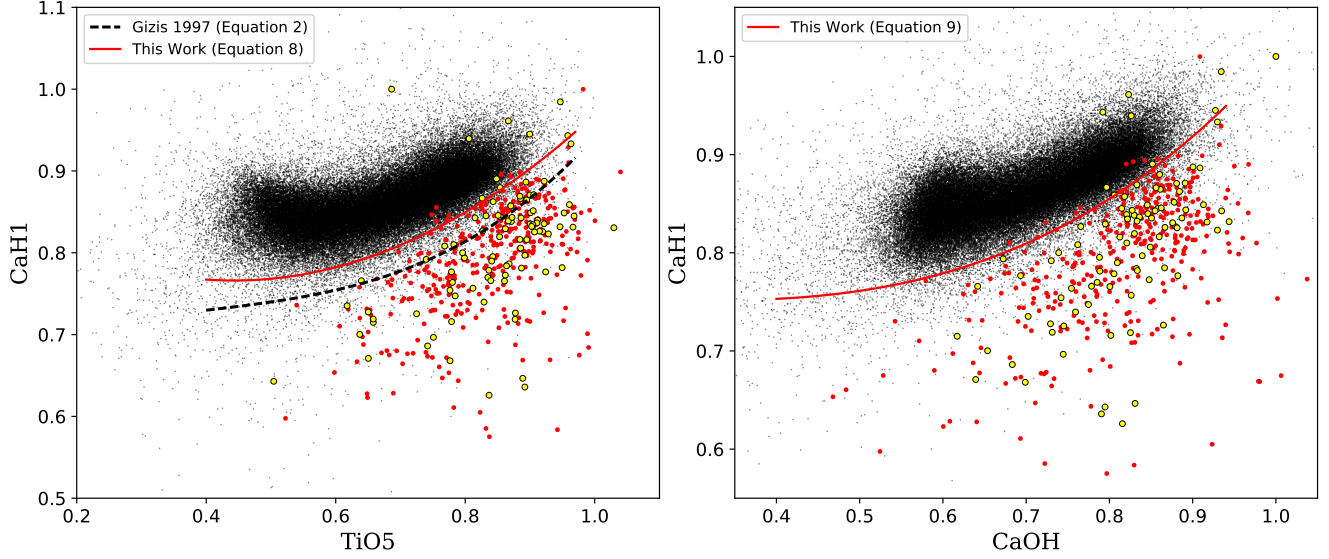
Figure 9. To separate M subdwarfs from M dwarfs on the [TiO5, CaH1] and [CaOH, CaH1] index diagrams, we plot the visually identified M dwarfs as black dots and the "labeled" subdwarf sample as red dots (see Section 4.1), the subdwarfs identified by Bai et al. (2016) are figured as yellow dots. We define two separator lines (i.e., equation 8 and 9) as red solid curves. The dashed black curve on the left panel is the dwarf-subdwarf separator relation originally proposed by Gizis (1997), which does not appear suitable for LAMOST data.

4.5. *The LAMOST DR4 Subdwarf Final Sample*

To build the final sample, we search by automatic processing LAMOST DR4 using the constraints detailed above. A spectrum is adopted as a candidate if it meets the following three conditions:
1) It has an available radial velocity measurement;
2) $\zeta < 0.75$;
3) It satisfies the equation (8) *or* equation (9).

Ultimately, visual examination of the spectrum is utilized to confirm (or not) the nature of the candidate.

To avoid missing targets as much as possible, we search for candidate subdwarfs among late K type stars, M type stars, QSOs and Unknowns (as classified by the LAMOST 1D automated pipeline), 1,271,426 spectra in total. Adding QSOs to the sample to be explored was decided because strong molecular absorption by TiO (and VO in the coolest objects) give a very choppy appearance to the spectra of late-type dwarf stars, quite similar to continuum breaks exhibited by QSOs in specific redshift ranges. As a matter of fact, cool M-type stars and QSO are often contaminating each other in surveys dedicated to each class (Kirkpatrick et al. 1997).

Our final subdwarf sample is constructed via visual inspection of each candidate spectrum. It consists of 2791 objects and covers the spectral sequence from M0 to M7, 291 objects had been investigated by previous researches. Figure 10 summarizes the distribution of their spectral subtypes. Most objects are earlier than M4, a result not unexpected since the metal deficiency has a tendency to decrease atmospheric opacity and to move the spectral types towards higher effective temperatures, on one hand, and on the other hand, instrumental selection by the red channel spectrograph performance of LAMOST also plays against detection of very cool faint objects. According to our newly recalibrated ζ parameter, the present sample can also be divided into the sdM/esdM/usdM classes with $\zeta < 0.75$, $\zeta < 0.5$ and $\zeta < 0.2$, respectively. There are 2386 sdMs, 295 esdMs, and 110 usdMs in total.

4.6. *The LAMOST DR4 Subdwarf Catalog and External Data*

We provide an M subdwarf catalog containing 2791 objects, the spectra of which can be accessed from the LAMOST Data Release web portal[1] and the catalog can be obtained on-line as well as by contacting the corresponding author. We give the values and errors of each spectral index (CaOH, CaH1, CaH2, CaH3, and TiO5), compute the ζ value, and provide an Hα activity indica-

tor for each target. The radial velocities are measured with the method described in Section 3.2. The spectral subtypes for subdwarfs are computed based on Equation 10 (Lépine et al. 2007).

Additional data from external sources are compiled in this catalog: photometric magnitudes measured in various bandpasses: five optical and near-infrared bands (g, r, i, z, y Tonry et al. 2012) from the Pan-STARRS1 (PS1) 3π survey (Kaiser et al. 2010), which is a systematic imaging survey of 3/4 of the sky north of -30°; the near-infrared J (1.25 μm), H (1.65 μm), and Ks (2.16 μm) bands from 2MASS; and W1, W2, W3, W4 bands centered at wavelengths of 3.4, 4.6, 12, and 22 μm respectively are provided by the AllWISE Data Release (Cutri et al. 2014), which has produced a new Source Catalog and Image Atlas with enhanced sensitivity and accuracy compared with earlier WISE (Wright et al. 2010) data releases (Kirkpatrick et al. 2014, 2016). G, BP and RP magnitudes in the new GAIA photometric system (Gaia Collaboration et al. 2018)are also provided.

The astrometric parameters from GAIA Data Release 2 are also provided: proper motions and parallaxes (Gaia Collaboration et al. 2018). We also give estimated distances recorded in the catalog from Bailer-Jones et al. (2018). An important point must be underlined: the distances provided by Bailer-Jones et al. (2018) are probabilistic estimates based on an elaborate statistical processing of GAIA DR2 data, using specific prior modeling (in particular stellar density distribution along Galactic lines of sight) to avoid physical errors resulting from "blind" simple inversions of measured parallaxes, as discussed e.g. in Luri et al. (2018).

Table 4 in Appendix give the description of each column of the catalog. A partial catalog extract is also shown in Appendix as Table 5. The complete catalog is available in CSV format in the online journal [2].

4.7. *A By-product: Active Objects Detection from Hα Emission Line*

Chromospheric activity is common among ordinary M dwarfs and has been largely studied by several authors, in particular from the large spectroscopic data base of the SDSS, see e.g. West et al. (2004), West et al. (2011) and it has also been found in many M subdwarfs (Bochanski et al. (2007), Savcheva et al. (2014)). At the spectral resolution of survey instruments like the SDSS combination or LAMOST, the brightest and most easily detectable emission line is Hα, accompanied in

---

the most active objects by CaII emission (both in H and K lines and in the far-red triplet) and sometimes by higher Balmer series lines. According to the criteria for designating a star as active defined by West et al. (2011), we measure the equivalent widths (EWs) of Hα emission line for each object, and remove the unreliable ones via visual inspection. Finally, a total of 141 active subdwarfs with prominent Hα emission lines are found, including 120 sdMs, 18 esdMs and 3 usdMs, amounting to more than 5 percent of the total sample. The Hα activity flag for each object is listed in the catalog. The flag is set as 1 for the active ones, 0 for the ones without emission features, and -9999 for the ones with S/N<3 in r or i band.

As Figure 10 shows, the activity fraction of M dwarfs is increasing in later types a result consistent with former studies (see e.g. (West et al. 2011)). Unfortunately, due to the observational selection against subdwarfs cooler than M3-M4 in our sample, it's difficult to derive a reliable estimate of the fraction of active subdwarfs. Previous studies have shown that among M subdwarfs, chromospheric activity is significantly less frequent than among ordinary M dwarfs, but the spectral type at which the maximum rate of activity is observed does not appear to differ. This reduction is claimed to be an argument to confirm that the chromospheric activity is a proxy for the age of the subdwarfs, whose bulk of observationally selected population, associated with the Galactic halo and thick disk, should be of substantially older formation than that of the disk dwarfs.

5. SUMMARY AND CONCLUSION

In order to search for spectroscopic cool subdwarfs from the fourth data release (DR4) of LAMOST regular survey, we review the history of current classification system for M subdwarfs developed by Lépine et al. (2007) and suggested to make some modifications based on the analysis of subdwarf spectra and templates from Savcheva et al. (2014). We suggest that the [CaH1, TiO5] index relation originally proposed by Gizis (1997) should be kept as a necessary condition for safe spectroscopic identification of subdwarfs, in combination with the classical [TiO5, CaH2+CaH3] index relation.

To adapt the molecular index relations to LAMOST data as separators between dwarfs and subdwarfs, we built a sample of more than 80,000 M dwarf spectra from LAMOST regular survey verified by means of the manual "eyecheck" mode of Hammer spectral typing utility. Meanwhile, a "labeled" subdwarf sample consists of 325 targets from this procedure and 108 subdwarfs from LAMOST DR2 in Bai et al. (2016). Using the dwarf sample and "labeled" subdwarf sample, we revise
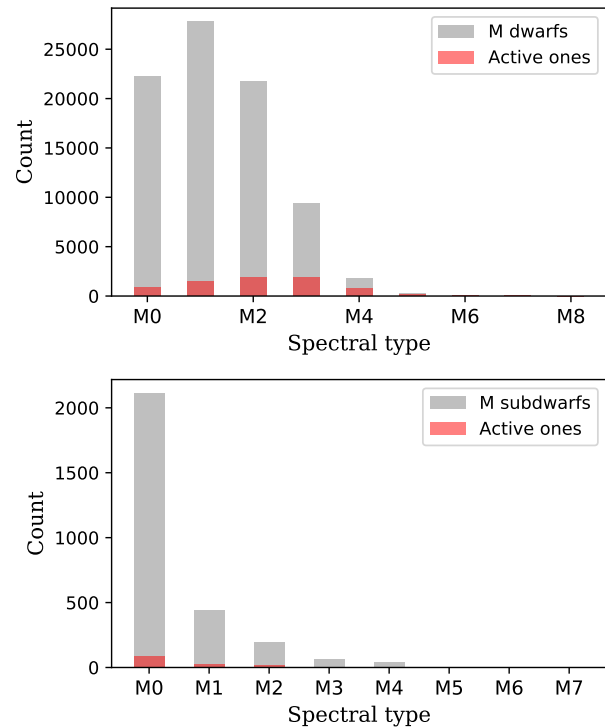

Figure 10. Spectral type (SpT) distribution of visually identified M dwarfs (83,213 in total) (upper panel) compared with SpT distribution of subdwarfs (lower panel). The magnetic activity fraction for each spectral subtype, detected by Hα emission following the precepts of West et al. (2011), is shown as red bars. 7075 active M dwarfs and 141 M subdwarfs were found. The figure shows that magnetic activity fraction is a function of spectral type for M dwarfs, which is consistent with former studies (West et al. 2011). For subdwarfs, the active subsample remains too small to draw a reliable conclusion.

the separator relation on the [TiO5 vs. CaH1] diagram and define a new separator on the [CaOH vs. CaH1] diagram. Further, the ζ index is recalibrated from the [TiO5, CaH2+CaH3] relation obeyed by LAMOST M dwarfs. The best new separator ζ value between dwarfs and subdwarfs is found to be 0.75. The spectral subtypes of subdwarfs are determined by the compound CaH2+CaH3 index.

Using these revised schemes as filters and visual inspection as the final verification tool, we identify 2791 M subdwarfs from the entire LAMOST DR4, out of more than a million of spectra submitted to the screening. In this sample, 141 have detectable Hα emissions.

For a more convenient use of this sample, the catalog also provides the astrometric information from GAIA DR2, distance estimated by Bailer-Jones et al. (2018) and a variety of magnitudes including photometric magnitudes from Pan-STARRS1, 2MASS, AllWISE and

GAIA. The catalog is provided on-line and the spectra can be accessed from the LAMOST Data Release web portal.

The spectroscopic selection and classification systems were originally built from objects with "old population" kinematic properties. However, the reciprocal statement, i.e., whether the large "subdwarfs" samples selected and classified using spectral indices all exhibit these kinematic properties, is still an open question to explore. Moreover, the subdwarfs appearing in directions corresponding to the Galactic thin disk are especially interesting, because subdwarfs are believed to evolve extremely slowly and most of those already identified are residents of the halo and the thick disk. Is this a pure projection effect or have these objects migrated or been trapped in the thin disk? Therefore, the correlations between their metallicity, true spatial location, and kinematics might reveal crucial information on the formation and evolution history of our Milky Way Galaxy. Our M subdwarf sample contains a large number of objects located at low Galactic latitudes especially in the Galactic anti-center direction, allowing exploration of differential kinematic properties with respect to previously published halo- and thick disk-dominated samples (in preparation). In addition, this sample also supplies candidates for high-resolution observations in the future.

Software: Hammer (Covey et al. 2007), Astropy (Astropy Collaboration et al. 2013), Matplotlib (Barrett et al. 2005), Pandas (http://pandas.pydata.org), Numpy (http://www.numpy.org)

APPENDIX

Table 4. Data description of the subdwarf catalog

Column Name	Unit	Description
designation	-	Target Designation
fitsname	-	Spectrum name
mjd	-	Day
planid	-	Plan Name
spid	-	Spectrograph ID
fiberid	-	Fiber ID
ra	degrees	Right Ascension
dec	degrees	Declination

Table 4. Data description of the subdwarf catalog

Column Name	Unit	Description
l	degrees	Galactic longitude
b	degrees	Galactic latitude
snrr	-	Average spectrum signal to noise Ratio in r band
snri	-	Average spectrum signal to noise Ratio in i band
rv	km s^{-1}	Heliocentric radial velocity
rv_err	km s^{-1}	Standard error of radial velocity
CaOH	-	Spectral index of CaOH band
CaOH_err	-	The error of spectral index of CaOH band
CaH1	-	Spectral index of CaH1 band
CaH1_err	-	The error of spectral index of CaH1 band
CaH2	-	Spectral index of CaH2 band
CaH2_err	-	The error of spectral index of CaH2 band
CaH3	-	Spectral index of CaH3 band
CaH3_err	-	The error of spectral index of CaH3 band
TiO5	-	Spectral index of TiO5 band
TiO5_err	-	The error of spectral index of TiO5 band
Zeta	-	The value of ζ parameter
spt	-	Spectral subtype
activity	-	Hα emission tag
flag_caution	-	A flag for spectrum quality from visual inspection
parallax	mas	GAIA DR2 Parallax
parallax_error	mas	Standard error of GAIA DR2 parallax
pmra	mas yr^{-1}	GAIA DR2 proper motion in right ascension direction
pmra_error	mas yr^{-1}	Standard error of GAIA DR2 proper motion in right ascension direction
pmdec	mas yr^{-1}	GAIA DR2 proper motion in declination direction
pmdec_error	mas yr^{-1}	Standard error of GAIA DR2 proper motion in declination direction
phot_g_mean_mag	mag	GAIA DR2 G-band mean magnitude
phot_bp_mean_mag	mag	GAIA DR2 integrated BP mean magnitude
phot_rp_mean_mag	mag	GAIA DR2 integrated RP mean magnitude
bp_rp	mag	GAIA DR2 BP-RP colour.
rest	pc	Estimated distance from Bailer-Jones et al. (2018), based on GAIA DR2
b_rest	pc	Lower bound on the confidence interval of the estimated distance
B_rest	pc	Upper bound on the confidence interval of the estimated distance
rlen	pc	Length scale used in the prior for the distance estimation
gmag	AB magnitudes	Pan-STARRS1 mean PSF magnitude from g filter detections
e_gmag	AB magnitudes	Error in Pan-STARRS1 g magnitude detections
rmag	AB magnitudes	Pan-STARRS1 mean PSF magnitude from r filter detections
e_rmag	AB magnitudes	Error in Pan-STARRS1 r magnitude detections

Table 4. Data description of the subdwarf catalog

Column Name	Unit	Description
imag	AB magnitudes	Pan-STARRS1 mean PSF magnitude from i filter detections
e_imag	AB magnitudes	Error in Pan-STARRS1 i magnitude detections
zmag	AB magnitudes	Pan-STARRS1 mean PSF magnitude from z filter detections
e_zmag	AB magnitudes	Error in Pan-STARRS1 z magnitude detections
ymag	AB magnitudes	Pan-STARRS1 mean PSF magnitude from y filter detections
e_ymag	AB magnitudes	Error in Pan-STARRS1 y magnitude detections
Jmag	mag	2MASS default J-band magnitude
Hmag	mag	2MASS default H-band magnitude
Kmag	mag	2MASS default Ks-band magnitude
e_Jmag	mag	Photometric uncertainty for the 2MASS J-band magnitude
e_Hmag	mag	Photometric uncertainty for the 2MASS H-band magnitude
e_Kmag	mag	Photometric uncertainty for the 2MASS Ks-band magnitude
W1mag	mag	AllWISE W1 standard aperture magnitude
W2mag	mag	AllWISE W2 standard aperture magnitude
W3mag	mag	AllWISE W3 standard aperture magnitude
W4mag	mag	AllWISE W4 standard aperture magnitude
e_W1mag	mag	Uncertainty in the AllWISE W1 standard aperture magnitude
e_W2mag	mag	Uncertainty in the AllWISE W2 standard aperture magnitude
e_W3mag	mag	Uncertainty in the AllWISE W3 standard aperture magnitude
e_W4mag	mag	Uncertainty in the AllWISE W4 standard aperture magnitude

NOTE—The complete LAMOST DR4 M subdwarf catalog can be downloaded from the on-line journal. Columns 2 to 12 record the information provided by LAMOST data base and the spectrum FITS files headers. Radial velocities, spectral indices and subtypes are measured in this work. The activity flag is used to indicate the magnetic activity of an object, which is measured by $H\alpha$ emission line. A flag_caution is provided based on the visual inspection procedure: the objects with low quality spectra or inaccurate flux calibration are set as "*" to warn that the data on these objects should be used carefully.

Table 5. A partial extract of the LAMOST DR4 M subdwarf catalog

designation	fitsname	mjd	planid	spid	fiberid	ra	dec	l	b
	snrr	snri	rv	rv-err	CaOH	CaOH-err	CaH1	CaH1_err	CaH2
	CaH2_err	CaH3	CaH3_err	TiO5	TiO5_err	Zeta	spt	activity	flag_caution
J002531.37+355526.4	spec-56952-M31005N35M1_sp09-206	56952	M31005N35M1	9	206	6.380737	35.924017	117.0639812	-26.65228601
	14.9	32.2	37	36	0.729	0.029	0.833	0.022	0.699
	0.015	0.875	0.018	0.828	0.02	0.677	0	0	
J003921.57+345754.2	spec-55910-M31_007N34_B2_sp10-129	55910	M31_007N34_B2	10	129	9.839904	34.965069	120.1336293	-27.84104438
	16	40.6	52	8	0.777	0.023	0.844	0.02	0.693
	0.012	0.851	0.014	0.798	0.015	0.733	0	0	
J004613.83+335010.3	spec-55910-M31_007N34_B2_sp05-222	55910	M31_007N34_B2	5	222	11.55766	33.836202	121.6953138	-29.02304224
	11.8	32.6	48	28	0.91	0.038	0.774	0.026	0.567
	0.012	0.757	0.014	0.842	0.018	0.357	1	0	
J005303.21+271507.1	spec-56551-VB010N27V2_sp06-142	56551	VB010N27V2	6	142	13.263398	27.251989	123.3736522	-35.61837284
	9.9	24.2	34	12	0.701	0.04	0.776	0.034	0.474
	0.014	0.706	0.018	0.695	0.022	0.55	2	0	
J011701.85+335911.3	spec-57373-M31019N31B1_sp11-217	57373	M31019N31B1	11	217	19.25772	33.986477	128.9722096	-28.58492645
	147.6	90.7	29	11	0.787	0.004	0.857	0.004	0.719
	0.003	0.859	0.003	0.816	0.004	0.735	0	0	
J012534.63-041817.6	spec-56591-EG012606S021203F01_sp01-034	56591	EG012606S021203F01	1	34	21.394302	-4.304911	144.0657935	-65.7660619
	10.1	17.4	90	11	0.802	0.042	0.78	0.036	0.674
	0.03	0.816	0.034	0.845	0.038	0.492	0	0	
J012615.42+352030.6	spec-56657-M31020N36M1_sp08-104	56657	M31020N36M1	8	104	21.564258	35.341851	130.8954916	-26.99090458
	12.5	28.4	-105	11	0.812	0.034	0.773	0.024	0.655
	0.016	0.836	0.018	0.843	0.02	0.498	0	0	
J012706.90-032656.1	spec-56233-EG012606S021203B03_sp05-009	56233	EG012606S021203B03	5	9	21.778777	-3.44894	144.2726261	-64.83205927
	7.3	28.1	27	11	0.641	0.046	0.815	0.04	0.596
	0.019	0.81	0.023	0.72	0.026	0.738	1	0	*
J013745.91+321343.0	spec-56650-M31021N31M1_sp13-050	56650	M31021N31M1	13	50	24.441311	32.228621	134.1987802	-29.62755946
	12.2	23.2	-47	14	0.702	0.038	0.842	0.026	0.678
	0.023	0.824	0.027	0.785	0.028	0.702	0	0	*
J022329.68+034930.6	spec-56299-EG023131N032619F04_sp14-192	56299	EG023131N032619F04	14	192	35.873708	3.825187	162.1417055	-51.89742013
	16.1	35.9	-29	22	0.815	0.028	0.873	0.02	0.756
	0.014	0.891	0.015	0.872	0.018	0.641	0	0	

Note—An extract of the 28 first columns of catalog records on 10 objects. The full catalog can be accessed on-line.

REFERENCES

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543-558

Anguiano, B., Majewski, S. R., Allende-Prieto, C., et al. 2018, arXiv:1807.07625

Ake, T. B., Greenstein, J. L. 1980, ApJ, 240,859

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

Bai, Y., Luo, A.-L., Comte, G., et al. 2016, Research in Astronomy and Astrophysics, 16, 107

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58

Barrett, P., Hunter, J., Miller, J. T., Hsu, J.-C., & Greenfield, P. 2005, Astronomical Data Analysis Software and Systems XIV, 347, 91

Bessell, M. S. 1982, PASA, 4,417

Bochanski, J. J., West, A. A., Hawley, S. L., et al. 2007, AJ, 133,531

Bochanski, J. J., Hawley, S. L., Covey, K. R. 2010, AJ, 139,2679

Bochanski, J. J., Savcheva, A., West, A. A. 2013, AJ, 145,40 (B13)

Burgasser, A. J., & Kirkpatrick, J. D. 2006, ApJ, 645, 1485

Chabrier, G. 2003, ApJ, 586,133

Covey, K. R., Ivezic, Z., Schlegel, D., et al. 2007, AJ, 134,2398

Cropper, M., Katz, D., Sartoretti, P., et al. 2018, A&A, 616, A5

Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, Research in Astronomy and Astrophysics, 12, 1197

Cutri, R. M., Wright, E. L., Conrow, T., et al. 2014, http://wise2.ipac.caltech.edu/docs/release/allwise/expsup/

Dhital, S., West, A. A., Stassun, K. G., et al. 2012, AJ, 143, 67

Du, B., Luo, A.-L., Kong, X., et al. 2016, ApJS, 227, 27

Eggen, O. J. & Greenstein, J. L. 1965, ApJ, 142, 925

Finkbeiner, D. P., Schlafly, E. F., Schlegel, D. J., et al. 2016, ApJ, 822, 66

Fuchs, B., Dettbarn, C., Rix, H.-W., et al. 2009, AJ, 137, 4149

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1

Giclas, H. L., Burnham, R., & Thomas, N. G. 1971, Flagstaff, Arizona: Lowell Observatory, 1971,

Gizis, J. E. 1997, AJ, 113,806 (G97)

Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2015, ApJ, 810, 25

Guo, Y.-X., Yi, Z.-P., Luo, A.-L., et al. 2015, Research in Astronomy and Astrophysics, 15, 1182

Hartwick, F. D. A., Cowley, A. P., Mould, J. R. 1984, ApJ, 286,269

Huang, Y., Liu, X.-W., Yuan, H.-B., et al. 2015, MNRAS, 449, 162

Jao, W.-C., Henry, T. J., Beaulieu, T. D., & Subasavage, J. P. 2008, AJ, 136, 840

Jao, W.-C. 2011, 16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, 448, 907

Jao, W. C., Nelan, E. P., Henry, T. J. 2016, AJ, 152,153

Joy, A. H. 1947, ApJ, 105,96

Kafle, P. R., Sharma, S., Lewis, G. F., & Bland-Hawthorn, J. 2014, ApJ, 794, 59

Kaiser, N., Burgett, W., Chambers, K., et al. 2010, Proc. SPIE, 7733, 77330E

Kirkpatrick, J. D., Henry, T. J., McCarthy, D. W. J. 1991, ApJS, 77,417

Kirkpatrick, J. D., Henry, T. J., Irwin, M. J. 1997, AJ, 113,1421

Kirkpatrick, J. D., Schneider, A., Fajardo-Acosta, S., et al. 2014, ApJ, 783, 122

Kirkpatrick, J. D., Kellogg, K., Schneider, A. C., et al. 2016, ApJS, 224, 36

Kuiper, G. P. 1939, ApJ, 89, 548

Laughlin, G., Adams, F. C. 1997, ApJ, 491,51

Lépine, S., Rich, R. M., Shara, M. M. 2003, AJ, 125, 1598

Lépine, S., Shara, M. M., Rich, R. M. 2003, ApJ, 585, 69

Lépine, S., & Shara, M. M. 2005, AJ, 129, 1483

Lépine, S., Rich, R. M., Shara, M. M. 2007, ApJ, 669,1235 (L07)

Lépine, S., & Scholz, R.-D. 2008, ApJL, 681, L33

Lépine, S., Hilton, E. J., Mann, A. W., et al. 2013, AJ, 145, 102

Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008, AJ, 136, 2022

Liebert, J., Dahn, C. C., Gresham, M., Strittmatter, P. A., et al. 1979, ApJ, 233,226

Luo, A.-L., Zhao, Y.-H., Zhao, G., et al. 2015, Research in Astronomy and Astrophysics, 15, 1095

Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9

Luyten, W. J. 1979, Minneapolis: University of Minnesota, 1979, 2nd ed.,

Mould, J. R. 1976, ApJ, 207,535

Mould, J. R. 1976, ApJ, 210,402

Mould, J. R.,& McElroy, D. B. 1978, ApJ, 220,935

Olmsted, C. M. 1908, Contributions from the Mount Wilson Observatory / Carnegie Institution of Washington, 21, 1

Öhman, Y. 1934, ApJ, 80, 171

Rajpurohit, A. S., Reyl?(C), C., Allard, F., et al. 2013, A&A, 556,15

Rajpurohit, A. S., Reylé, C., Allard, F., et al. 2016, A&A, 596, A33

Reid, I. N., Hawley, S. L., Gizis, J. E. 1995, AJ, 110, 1838

Reid, I. N., Gizis, J. E., Hawley, S. L. 2002, AJ, 124, 2721

Roeser, S., Demleitner, M., & Schilbach, E. 2010, AJ, 139, 2440

Savcheva, A. S., West, A. A., Bochanski, J. J., et al. 2014, ApJ, 794,145 (S14)

Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131,1163

Tonry, J., & Davis, M. 1979, AJ, 84, 1511

Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, ApJ, 750, 99

West, A. A., Hawley, S. L., Walkowicz, L. M., et al. 2004, AJ, 128, 426

West, A. A., Morgan, D. P., Bochanski, J. J., et al. 2011, AJ, 141, 97

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868

Yi, Z., Luo, A., Song, Y., et al. 2014, AJ, 147, 33

Yi, Z.-P., Luo, A.-L., Zhao, J.-K., et al. 2015, Research in Astronomy and Astrophysics, 15, 860

Yuan, H. B., Liu, X. W., & Xiang, M. S. 2013, MNRAS, 430, 2188

York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579

Zhang, L., Pi, Q., Han, X. L., et al. 2016, NewA, 44, 66

Zhang, Z. H., Pinfield, D. J., Gálvez-Ortiz, M. C., et al. 2017, MNRAS, 464, 3040