TARGET SELECTION FOR THE SDSS-III MARVELS SURVEY

MARTIN PAEGERT1, KEIVAN G. STASSUN1,2, NATHAN DE LEE3,9, JOSHUA PEPPER4, SCOTT W. FLEMING5,6,9, THIRUPATHI SIVARANI7,9, SUVRATH MAHADEVAN9,10,11, CLAUDE E. MACK III1, SAURAV DHITAL12, LESLIE HEBB8, and JIAN GE9

1 Vanderbilt University, Nashville, TN 37235, USA
2 Fisk University, Nashville, TN 37208, USA
3 Northern Kentucky University, Highland Heights, KY 41099, USA
4 Lehigh University, Bethlehem, PA 18015, USA
5 Computer Sciences Corporation, 3700 San Martin Dr, Baltimore, MD 21218, USA
6 Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21218, USA
7 Indian Institute of Astrophysics, Koramangala, Bangalore-560034, India
8 Hobart & William Smith Colleges, Ithaca, NY, USA
9 University of Florida, 2014 Turfing Hall, Gainesville, FL 32611, USA
10 Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
11 Center for Exoplanets and Habitable Worlds, The Pennsylvania State University, University Park, PA 16802, USA
12 Department of Physical Sciences, Embry-Riddle Aeronautical University, 600 South Clyde Morris Blvd., Daytona Beach, FL 32114, USA

Received 2015 January 2; accepted 2015 March 27; published 2015 May 18

ABSTRACT

We present the target selection process for the Multi-object APO Radial Velocity Exoplanets Large-area Survey (MARVELS), which is part of the Sloan Digital Sky Survey (SDSS) III. MARVELS is a medium-resolution (R ~ 11,000) multi-fiber spectrograph capable of obtaining radial velocities for 60 objects at a time in order to find brown dwarfs and giant planets. The survey was configured to target dwarf stars with effective temperatures approximately between 4500 and 6250 K. For the first 2 years MARVELS relied on low-resolution spectroscopic pre-observations to estimate the effective temperature and log(g) for candidate stars and then selected suitable dwarf stars from this pool. Ultimately, the pre-observation spectra proved ineffective at filtering out giant stars; many giants were incorrectly classified as dwarfs, resulting in a giant contamination rate of ~30% for the first phase of the MARVELS survey. Thereafter, the survey instead applied a reduced proper motion cut to eliminate giants and used the Infrared Flux Method to estimate effective temperatures, using only extant photometric and proper-motion catalog information. The target selection method introduced here may be useful for other surveys that need to rely on extant catalog data for selection of specific stellar populations.

Key words: catalogs – planets and satellites: detection – stars: general – surveys: techniques: radial velocities

1. INTRODUCTION

Target selection is a crucial step for most astronomical surveys, one that may have a significant impact on the result even before the first image is taken. This is especially true for exoplanet surveys. A common method is to pre-select stars according to brightness, then derive stellar parameters from reconnaissance observations and compile them into an Input Catalog from which the final set of targets is drawn. The NASA Kepler mission is one of the most prominent projects following this process. However, such pre-observations require telescope time and extensive effort to process and evaluate the reconnaissance data. Therefore, for Multi-object APO Radial Velocity Exoplanet Large-area Survey (MARVELS; Ge et al. 2009) we opted to devise a technique to find the stellar populations suitable for the survey using only existing catalog data, thus saving the time and effort that would otherwise go into pre-observations, and streamlining the target selection process significantly. We hope that our method will be useful for future surveys like the upcoming Transiting Exoplanet Survey Satellite, or indeed any effort to select a particular population of stars from existing catalog data.

MARVELS is part of the Sloan Digital Sky Survey (SDSS)—III program (Eisenstein et al. 2011a) and uses a specially built 60-fiber spectrograph to obtain medium-resolution (R = 11,000) spectra to derive the precision radial velocities needed to find exoplanets and brown dwarfs orbiting main sequence stars. The MARVELS instrument itself is described in J. Ge et al. (2015a, 2015b, in preparation), the data reduction pipeline is described in N. Thomas et al. (2015, in preparation), and the final DR12 data release is described in S. Alam et al. (2015, in preparation). In this paper we focus on describing the MARVELS target selection process.

For each target field with a circular field of view of 7 square degrees, 56 stars are selected for observation and assigned to a fixed fiber which then is plugged to a hole in a metal plate placed in the focal plane of the 2.5 m SDSS telescope (Gunn et al. 2006). Four fibers are reserved for guide stars, which are chosen after the science targets are known. The plugs require a minimal distance of 75 arcsec and thus define the required minimal distance between target stars. Between 2008 October and 2012 July, MARVELS made 1565 observations of 92 fields collecting multi-epoch data for 5520 stars, more than 90% of them with enough epochs to be processed through the pipeline and yield sufficient RV observations to search for companions, including stellar companions, brown dwarfs, and giant planets.

Due to technical and administrative changes in 2011 January—change of fibers, joint observation with the the APOGEE SDSS-III survey (the APO Galactic Evolution Experiment, Allende Prieto et al. 2008; Eisenstein et al. 2011b)—the observation is divided into two different phases: before and after 2011 January, hereafter referred to as “initial” (Years 1–2) and “final” (Years 3–4) phases.

Section 2 first describes the final target selection process used for fields observed after 2011 January. It then describes
the initial process that was used prior to 2011 January as well as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with very few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.

2. TARGET SELECTION METHODS

MARVELS observed 5520 stars over four years, observing 54 science targets per field at a time. MARVELS was designed to achieve a radial velocity precision of <30 m s\(^{-1}\) for stars as faint as \(V = 12\) mag in order to discover brown dwarfs and giant planets of a homogenous sample of targets with only few, well-understood biases. Prime targets for MARVELS are FGK dwarfs, limiting the effective temperatures as the lessons learned and why the initial process was abandoned. Section 3 presents a summary of the properties of the targets observed. We conclude with a brief summary in Section 4.
values from the NASA Exoplanet Archive. As we show below, the results using the updated KIC are not substantially different from that used in our actual target selection process.

We define giants as having \( \log(g) \leq 3.5 \), dwarfs having \( \log(g) \geq 4.1 \), and subgiants as those with \( \log(g) \) between these values. Figure 1 shows the RPM\(_J\) diagram for the MARVELS-Kepler overlap stars (using original KIC values in the left panel, revised values in the right panel). The solid black line marks the border between RPM\(_J\)-dwarfs below and RPM\(_J\)-giants above the line as defined by Equation (3). Green symbols above the line are \( \log(g) \)-dwarfs that are mis-identified as giants by the RPM\(_J\) cut. Red symbols below the line are in turn \( \log(g) \)-giants mis-identified as dwarfs. The MARVELS “region of interest” is \( 0.3 < (J - H) < 0.54 \) and below the RPM\(_J\) cut, which translates to dwarfs or subgiants of spectral types F9 to K3.

A few stars on the left panel in Figure 1 are missing in the right panel—most notably those with \( J - H > 0.7 \). These are evolved giant stars that were observed by Kepler in quarter 0 only then dropped and are thus not part of the revised values published in the NASA Exoplanet Archive for stars observed by Kepler in quarters 1–16. Nonetheless, the right panel with the revised values is more populated and confirms the ability of the RPM\(_J\) cut to distinguish between dwarfs and giants—although again subgiants are not well discriminated.

Using the original KIC values we find that six stars (1.7\%) are false negatives—\( \log(g) \) dwarfs according to KIC, but giants according to the RPM\(_J\) cut. Another six stars (1.7\%) are false positives—they are \( \log(g) \) giants, falsely identified as dwarfs by the RPM\(_J\) method. Most of the sub-giants are below the line and thus in the “dwarf”-region. Collier Cameron et al. (2007) included only stars with \( \log(g) < 3.0 \) or \( \log(g) > 4.0 \), thus excluding subgiants. In our analysis we included the missing \( \log(g) \) interval and conclude that the RPM\(_J\) cut does not seem able to distinguish sub-giants from dwarfs. For MARVELS this is not a problem, because sub-giants are valid target stars, but it should be considered for any future statistics derived from the MARVELS data set, as subgiants will be included in the “dwarf” sample.

### Table 1

| Step | Criterion | Reason |
|------|-----------|--------|
| 1    | Keep stars with \( 7.6 \leq V \leq 13.0 \)     | Include only stars in MARVELS magnitude limits |
| 2    | Keep stars with \( J - K_s \geq 0.29 \)       | Exclude stars that are clearly too hot |
| 3    | Keep stars with known proper motions           | Allows exact positioning of fibers and permits use of reduced proper motion cut |
| 4    | Ensure that positional coordinates, after correcting for proper motion, indicate that the star is in the field for at least 2 years from projected start of observations | Exclude stars that might wander off-plate |
| 5    | If two stars are closer than 75°, keep the brighter star | Prefer bright stars for good SNR |
| 6    | Closest star with \( V < 9 \) must be more than 5° away | Prevent flux contamination of target star |
| 7    | Exclude star if too close to APOGEE targets    | Prevent fiber collision between APOGEE and MARVELS |
| 8    | Limit the results to the 1000 brightest stars in \( V \) | Build large enough pool for subsequent steps |
| 9    | Apply reduced proper motion cut to filter out all but the six brightest giants | Only six giants wanted |
| 10   | Exclude hot stars \( T_{\text{eff}} > 6250 \text{ K} \) according to Infrared Flux Method | Exclude hot stars |
| 11   | Limit F stars (those with \( 5800 \text{ K} \leq T_{\text{eff}} \leq 6250 \text{ K} \)) to 40% of all MARVELS targets in the field | Guarantee 50% GK stars |
| 12   | Limit the total number to 100 per field        | 60 plugged, 40 as “reserve” in case of collisions |
| 13   | Check the six selected giants in Simbad        | Verify for \( T_{\text{eff}} \), exclude close binaries and known variables |

Note. \( T_{\text{eff}} \) from Casagrande relations.
Note that the KIC stars span a magnitude range from $V = 9$ to 11.5, thus excluding the bright and faint end of the MARVELS magnitude range (7.6–13.0). Therefore, one concern is that the good performance of the RPM$_J$ cut for the KIC stars might not apply to the full range of MARVELS targets. In addition, when considering stars in other parts of the sky, one would necessarily need to use proper motions from heterogeneous catalog sources, thus potentially introducing systematic errors. Thus for the MARVELS target stars we checked how much of an influence a change from the modified GSC to UCAC as source catalog for proper motions would have on our selection. Generally GSC and UCAC are in good agreement, with a mean difference in total proper motion of $\Delta \mu = 0.37 \pm 4.75$ mas yr$^{-1}$. Thus on average we do not expect dramatic changes. We also checked the rate of stars switching from giant to dwarf classification according to the RPM$_J$ cut when changing from GSC to UCAC as the proper motion source catalog. We found this rate to be 1.75%, thus low enough to not cause concern.

Encouraged by these comparisons with the KIC we next tested the RPM$_J$ cut all-sky with the *Hipparcos* catalog (ESA 1997; Perryman et al. 1997). We combined the *Hipparcos* catalog with isochrones to derive a log($g$) determination based upon location within the HR diagram. We started with the Padova isochrones (Marigo et al. 2008; Girardi et al. 2010) from color–magnitude diagram 2.3 (http://stev.oapd.inaf.it/cmd). The isochrones were chosen to be solar metallicity ($Z = 0.019$) and range to in $\log_{10}$(Age) (yr) from 6.60 to 10.10 (inclusive) in steps of 0.05. The intrinsic luminosities from the isochrones were transformed into the Johnson–Cousins filters using Maíz Apellániz (2006) and Bessell (1990). The HR diagram is a phase space of absolute $V$ magnitude and $B − V$ color. Each isochrone occupies a particular region of this phase space, but they fall in such a way that some regions have multiple isochrones overlapping and some regions have no isochrones. In order to quantify this, the phase space was separated into 40 bins in $B − V$ and 80 bins in the $M_V$. The minimum value, maximum value, and step size in $B − V$ were −0.23, 3.12, and 0.0837, respectively. For $M_V$ the minimum value, maximum value, and step size were −6.14, 13.00, and 0.239, respectively. The isochrones are made up of a series of three coordinate data points ($M_V$, $B − V$, and log($g$)). Each isochrone data point is put into its appropriate bin in color–magnitude phase space. Once this was done, the median log($g$) of the data points in each color–magnitude bin was assigned as the log($g$) of that bin as shown in Figure 2. The isochrones do not completely cover the HR diagram, so there are bins that do not have a log($g$) value. The next step is to associate each *Hipparcos* star with a bin and throw out any stars that do not fall within 0.5 mag (in both $M_V$ or $B − V$) of a bin center that had a log($g$) value. If a *Hipparcos* star has more than 1 bin within the 0.5 mag bin radius, then the star was associated with the closest bin center and was then assigned the log($g$) of that bin. This results in a table of *Hipparcos* stars with log($g$) that can then be used to test the RPM$_J$ method; see Figure 2. Note that our use here of 0.5 mag, and of a linear interpolation in log age, are simplified and arbitrary choices, however they suffice for the purposes of the RPM$_J$ method to perform the broad performance of the RPM$_J$ method to distinguish dwarfs from giants.

To compare the RPM$_J$ results with directly measured log($g$) values, we chose the RAVE catalog, data release 2 (Zwitter et al. 2008) and 3 (Siebert et al. 2011). Both RAVE releases give comparable results, except that the fraction of log($g$) giants which would be classified as RPM$_J$-dwarfs is nearly three times higher with data release 2 (15% instead of 5.8%). We ascribe this to the improved RAVE pipeline used for data release 3.

The results for *Hipparcos* and RAVE DR3 are summarized in Table 2. If the RPM$_J$ cut is the only information used, 92.5% of stars flagged as giants would be true giants according to the log($g$) value derived for *Hipparcos* stars. 2.6% would be subgiants and 4.9% would be dwarfs. Furthermore 2.6% of the stars flagged as dwarfs would be giants, 39.9% subgiants, and 57.5% dwarfs.

The results for all stars in RAVE DR3 are similar with a notable shift from sub-giants toward dwarfs and giants. While...
the ratio of correctly identified giants decreases by 9%, the ratio of correctly identified dwarfs improves by 17%. If we limit the RAVE stars to valid MARVELS targets—those matching the MARVELS magnitude and color cut \( 7.6 \leq V < 13 \), respectively \( (J - K_S \geq 0.29) \)—the results improve slightly, but not significantly, as shown in the bottom part of Table 2.

For MARVELS we conclude that using the RPM\(_J\) cut as the only method for selecting dwarfs will result in a giant contamination rate of about 4%, which is much better than the rate we experienced from spectroscopic pre-observations (see Section 2.2). Importantly, however, subgiants comprise a large fraction of the “dwarf” sample. Therefore while the target selection procedure described above is highly effective at removing evolved red giants, subgiants are unavoidably mixed in with the dwarfs at the level of 20%–40% (see Table 2).

### 2.1.3. Effective Temperature from the Infrared Flux Method

We compute the effective temperature using color-metallicity—temperature relations based on the IRFM as described in Casagrande et al. (2010). For stars with \( 0.78 < x = V - K < 3.15 \) and \(-5.0 < [\text{Fe/H}] < 0.4 \), and defining \( x = V - K \), and \( T_{\text{eff}} = 5040.0/\theta_{\text{eff}} \), Casagrande et al. (2010) gives the relation

\[
\theta_{\text{eff}} = 0.5057 + 0.2600x - 0.0146x^2 \\
- 0.0131x[\text{Fe/H}] + 0.0288[\text{Fe/H}] \\
+ 0.0016[\text{Fe/H}]^2.
\]  

Using instead the \( J - K \) colors, for stars with \( 0.07 \leq J - K_S \leq 0.80 \) and the same metallicity restrictions, and now defining \( x = J - K_S \):

\[
\theta_{\text{eff}} = 0.6393 + 0.6104x + 0.0920x^2 \\
- 0.0330x[\text{Fe/H}] + 0.0291[\text{Fe/H}] \\
+ 0.0020[\text{Fe/H}]^2.
\]  

Without a measured value for [Fe/H], we assume solar metallicity for all our target stars. The additional error induced by this assumption does not exceed 80 K at the extreme ends of the color range and for [Fe/H] = ±0.4. Comparing \( T_{\text{eff}} \) from the Casagrande relations with the original \( T_{\text{eff}} \) values in KIC, we estimated an error of 105 K for dwarfs and 165 K for giants. According to Table 8 and Figure 18 in Pinsonneault et al. (2012), the KIC \( T_{\text{eff}} \) values are approximately 200 K systematically too low for dwarfs and giants alike, so in reality the \( T_{\text{eff}} \) errors for the MARVELS target selection is likely closer to this value of 200 K.

This error on \( T_{\text{eff}} \) is larger than the 1-σ error derived from benchmark grade stars given in Casagrande et al. (2010). One reason is the assumed solar metallicity of our target stars, as noted above. A second reason is the fact that the Casagrande relations are calibrated for dwarfs and subgiants, but not for giants. For MARVELS target selection this is not important because we exclude all but the six brightest giants which are vetted manually using Vizier (see above). While \( V - K \) as a temperature estimator is more sensitive to reddening than, e.g., \( J - K \), it is less sensitive to metallicity errors. As reddening does not play a major role for the nearby dwarf stars that dominate the MARVELS target sample, we consider \( V - K \) to be a valid and optimal choice for our purposes (see Section 3.2 for an estimation of reddening and extinction). In cases where the limits of the Casagrande relations for \( J - K \) and \( V - K \) allow us to compute a \( T_{\text{eff}} \) from both relations, we use the mean value.

We test the Casagrande relations with RAVE DR3, the result is shown in Figure 4. Running a least-squares fit over all stars we find an offset of \( 100 \pm 10 \) K between IRFM and RAVE temperatures. This offset matches the errors we estimate using KIC and the 85 ± 14 K the RAVE team reports when comparing their data to high-resolution external results. As reported by the RAVE team, the data show a wide spread of temperatures and a noticeable trend to yield higher temperatures especially for dwarf stars. However, the shift is about one MK subclass—say G5 instead of G4—and not significant for the MARVELS target selection.

While the agreement between IRFM based temperatures and the RAVE spectroscopic temperatures is within the error margin for dwarfs with \( T_{\text{eff}} \leq 5000 \) K, it degrades quickly for hotter dwarfs. Notably the agreement between RAVE and IRFM-derived temperatures for giants and subgiants is not worse than the agreement for hot dwarfs. Thus we conclude that the choice to use the Casagrande relations for giants and subgiants, although they are calibrated for dwarfs, is reasonable.

### 2.2. Initial Target Selection Process

#### 2.2.1. Basic Input Catalog Construction

Field names, center coordinates and number of observations for the year 1 and 2 fields are listed in Table 5. The basis for the target selection is the modified Guide Star Catalog 2.3 (Lasker et al. 2008) as described in Section 2.1.1.

Unlike in the final target selection, for the 1000 brightest stars matching the brightness and color cut of \( J - K_S \geq 0.29 \), a spectroscopic snapshot was taken by the SDSS double spectrograph, mainly used for SEGUE (Yanny et al. 2009). The instrument has a resolution of \( R \approx 2000 \) and is described in greater detail in Section 2 of Smee et al. (2013). The double spectrograph saturates at \( V = 9 \), thus brighter stars needed special treatment during the initial phase. The stellar parameters
were derived using a modified version of the SEGUE Stellar Parameter pipeline (SSPP Lee et al. 2008). The target selection was a two-step process.

1. Select up to 1000 stars for stellar characterization with SDSS spectrograph.
2. Using the characterization from step 1 select 100 stars for drilling, 60 of them will get observed with MARVELS.

For every star in a given field the steps in Table 3 were applied. The distance parameter in step 6 of the selection process (62") is different from the 75" reported in step 5 of Table 1. The reason is that MARVELS switched to wider fibers in order to maximize the throughput in years 3 and 4. In addition the effective temperatures used here are derived from pre-observations using the modified SSPP pipeline instead of based on the Casagrande relations as in Table 1. In order to allow a consistent comparison, we computed $T_{\text{eff}}$ using the Casagrande relations for all stars targeted during the initial phase.

The observations took place during twilight at the Apache Point Observatory. The spectra were evaluated for $T_{\text{eff}}$, log($g$) and [Fe/H] using the adapted SSPP. The stars were then split in a bright ($7.6 \leq V < 9.0$) and a faint ($9.0 < V < 13.0$) sample. Both samples were split between main sequence stars (log($g$) > 3.0) and giants (log($g$) < 3.0). Although this split is different from the classification introduced with the final target selection (log($g$) < 3.5 for giants) it does not play a role in the large giant contamination of the initial phase. Only 10 stars from all stars observed in this phase have $3.0 \leq \log g \leq 3.5$ from the SSPP-pipeline and are flagged as RPM$_{J}$-giants, thus this shift is not responsible for the high contamination by giants in the initial phase.

The bright stars were checked against SIMBAD and usually rejected—allowing for special targets—if any of the following conditions were met.

Table 3

| Step | Criterion | Reason, Comments |
|------|-----------|-----------------|
| 1    | Keep stars with $7.6 \leq V < 13.0$ | Include only stars in MARVELS magnitude limits |
| 2    | Keep $J - K_s \geq 0.29$ | Exclude stars that are clearly too hot |
| 3    | Ensure that positional coordinates, after correcting for proper motion, indicate that the star is in the field for at least 2 years from projected start of observations | Exclude stars that might wander off-plate |
| 4    | Closest star with $V < 9$ must be more than 5" away | Prevent flux contamination of target star |
| 5    | If two stars are closer than 62", keep the brighter star | Prefer bright stars for good SNR |
| 6    | Keep the brightest 1000 stars | Limit the number of stars to number of SDSS spectrograph fibers |
| 7    | Stars must stay on plate for 2 years | Re-check because date of observation might have changed |
| 8    | Keep the six brightest giants | Only six giants wanted |
| 9    | Keep stars with $T_{\text{eff}} < 6250$ K | Exclude hot stars, $T_{\text{eff}}$ from SSPP pipeline |
| 10   | Limit F stars (those with $5800$ K $\leq T_{\text{eff}} < 6250$ K) to 40% of all MARVELS targets in the field | Guarantee 50% GK stars, $T_{\text{eff}}$ from SSPP pipeline |
| 11   | If two stars are closer than 75", keep the brighter star | Prefer bright stars for good SNR |
| 12   | Limit the total number to 100 per field | 60 plugged, 40 as “reserve” in case of collision with guide stars |

Note. $T_{\text{eff}}$ from the modified SSPP Pipeline.

$T_{\text{eff}}$, log($g$) and [Fe/H] were derived using a modified version of the SEGUE Stellar Parameter pipeline (SSPP Lee et al. 2008). The target selection was a two-step process.

1. Select up to 1000 stars for stellar characterization with SDSS spectrograph.
2. Using the characterization from step 1 select 100 stars for drilling, 60 of them will get observed with MARVELS.
1. The spectral type was not between late F and early K for Main Sequence Stars or between mid G and early K for giants.
2. They are known variable stars.
3. They are in a visual binary with a companion less than 5″ away.
4. They are known exoplanet hosts (except for benchmark stars).
5. Any anomalies were found making it unlikely that MARVELS could detect a substellar companion.

Bright stars passing these tests were combined with the faint star sample and steps 7 to 12 from Table 3 were applied. While it may appear that step 7—the first step after pre-selection—is redundant with step 3, several months may have elapsed between pre-selection and step 7. In this time the planned observations may have been delayed to a later date, thus necessitating a new check that the target stays on plate even with the new, later observation start date.

Although the spectrograph has only 64 fibers, we keep 100 stars in order to have a “reserve” if it turns out that a star cannot be plugged because it is too close to a guiding star or for other technical reasons.

2.2.2. Giant Contamination In Initial Target Selection

It was initially assumed that in this process 10% of the selected dwarfs would actually be giants due to errors in the log(g) determinations from the SSPP, yielding a final giant fraction of about 15% in the final sample. Instead, the contamination rate was about 35%, as determined later by the RPM_f method. Some MARVELS fields overlapped with the Kepler field, and we compared the stellar characteristics obtained from SDSS spectra with those in the KIC. Figure 5 shows the log(g) values from SDSS spectra versus values in the KIC. At $T_{\text{eff}} < 5000$ K results diverge rapidly and there is no agreement at all for $T_{\text{eff}} < 4500$ K. Moreover, for all stars with KIC log(g) < 3 the values disagree strongly. Given the fact that the original KIC values for log(g) are up to 1.0 dex too high (Casagrande et al. 2014), the true contamination by giants in the SSPP selected sample is even higher than suggested by Figure 5. Figure 6 shows the HR diagram (effective temperature and log(g)) for the same set of stars from KIC and the SSPP-pipeline modified for MARVELS, again showing the strong discrepancies in the SSPP estimated log(g) values for cool giants. We conclude that cool giants are misidentified as dwarfs by the modified SSPP pipeline. We therefore abandoned the spectroscopic pre-observations in favor of the streamlined target selection process described in Section 2.1.

3. RESULTS

3.1. Summary of Selected Stars

The initial phase was significantly longer than the final phase—26 versus 15 months. The number of stars selected for observation reflects this asymmetry: 4130 stars in the initial phase, and 2900 stars for the final phase—adding to 7030 stars designed for observation out of which 5520 actually got observed. Figure 7 shows the distribution on the sky in galactic coordinates. The field centers along the galactic plane are from the final phase and located at galactic latitudes of $-8^\circ, -4^\circ, 0^\circ, 4^\circ$ and $8^\circ$, and thus appear to blend into each other.

Figure 8 shows the magnitude distribution in the V-band, for the initial phase at the left, for the final phase at the right. Aside from the different total numbers mentioned above, the most pronounced difference is a shift of the maximum by 0.5 mag—from around 11.25 mag for the initial to 11.55 mag for the final phase. The reason is that coordination with APOGEE placed the fields outside of the galactic plane. Since MARVELS and APOGEE could not observe the same stars in those sparse fields, the available stars were fainter.

Figure 9 is a stacked histogram for effective temperatures of RPM_f-dwarfs and RPM_f-giants during the initial phase (left) and the final phase (right). For this comparison we computed effective temperatures for all stars from the initial phase using the Casagrande relations in order to allow a direct comparison to the temperatures estimated in the final phase. In the initial phase giants are overrepresented for $T_{\text{eff}} < 5000$ K, indicating again that the log(g) values from the modified SSPP-pipeline were unreliable for cooler stars. Out of the 4130 stars selected for observation during the initial phase, 1414 stars are flagged as giants by the RPM_f method. This is 34% of the sample.

To estimate the fraction of giants in our sample we take the rates for MARVELS-selected RAVE stars from Table 2. About 14% of the RPM_f-giants are false positives, and thus are either dwarfs or sub-giants. On the other hand 3% of the RPM_f-dwarfs are false
Figure 7. MARVELS footprint in galactic coordinates. Blue circles: target fields during the initial phase; green crosses: target fields during the final phase. The field centers along the galactic plane are located at latitudes of $-8^\circ$, $-4^\circ$, $0^\circ$, $4^\circ$ and $8^\circ$ and thus appear to blend into each other.

Figure 8. $V$ magnitude distribution for the initial phase (left) and final phase (right).

Figure 9. Distribution of effective temperatures (IRFM method). Left: initial phase, right: final phase.
positives, and therefore giants. The estimated rate for the initial phase is then $34 - 0.14 \times 34 + 0.03 \times 66 = 31\%$. For the final phase we manually checked six giants per field (10\%). To this we add the 4\% error for RPM\textsubscript{J}-dwarfs and thus end with 14\% giants in stars selected for the final phase. Given that we do want 10\% of the stars to be giants, the difference between estimated and wanted giants gives the contamination rate. The results are summarized in Table 4.

3.2. Effects of Reddening and Extinction on Inferred Stellar Properties

Most of the target fields of year 3 and 4 are located near the galactic plane ($-8 \leq b \leq 8$), so reddening and extinction might have to be taken into account. Using the RPM\textsubscript{J} cut to distinguish dwarfs and subgiants from giants, extinction moves stars down in the RPM\textsubscript{J} diagram, reddening moves them to the right. There are three possible effects.

1. Giants are pushed downwards over the cut by extinction, contaminating our sample.
2. Dwarfs are shifted out of the region of interest by reddening and are lost.
3. Hot stars are moved downward and to the right into our region of interest, polluting the sample.

| Phase          | Initial | Final |
|----------------|---------|-------|
| RPM\textsubscript{J}-giants | 34\%   | 10\%  |
| RPM\textsubscript{J}-dwarfs  | 66\%   | 90\%  |
| Est. giants     | 31\%   | 14\%  |
| Wanted          | 10\%   | 10\%  |
| Contamination   | 21\%   | 4\%   |

Figure 10. Reddening and extinction for the Kepler field. Left panel: stars with $V = 10$; Right panel: stars with $V = 13$. Gray shade: MARVELS region of interest. Box and whiskers: galactic model position of stars with given magnitude and spectral type. Numbers on the lower legend are the typical distances of dwarfs (top, green) and giants (bottom, red).

Figure 11. Reddening and extinction for the Galactic Center. Left panel: stars with $V = 10$; Right panel: stars with $V = 13$. Gray shade: MARVELS region of interest. Box and whiskers: galactic model position of stars with given magnitude and spectral type. Numbers on the lower legend are the typical distances of dwarfs (top, green) and giants (bottom, red).
Table 5
Name, Center Coordinates, and Number of Observations for Fields from the Initial Phase

| Name   | R.A.  | Decl. | l    | b     | Obs |
|--------|-------|-------|------|-------|-----|
| 47UMA  | 164.86| 40.43 | 175.78| 63.3645| 29  |
| 51PEG  | 344.37| 20.77 | 90.0669| −34.7279| 36  |
| FIELD1068| 264.3 | 30   | 54.18431| 28.24962| 23  |
| FIELD1110| 100.8 | 33   | 182.19045| 12.7955| 23  |
| FIELD1348| 262.86| 42   | 67.41378| 32.0944| 23  |
| FIELD1349| 268.57 | 42  | 68.33689| 27.92775| 25  |
| FIELD1572| 321.7 | 54   | 95.6606 | 2.3394 | 31  |
| FIELD1631| 101.54| 60   | 155.71982| 22.6218| 34  |
| GJ176  | 70.73 | 18.96 | 180.01238| −17.42718| 30  |
| GJ436  | 175.55| 26.71 | 210.53025| 74.75275| 29  |
| GL237  | 111.85| 5.23  | 212.33978| 10.37355| 28  |
| HAT-P-1| 344.45| 38.67 | 99.7928 | −19.04242| 36  |
| HAT-P-3| 206.09| 47.97 | 100.09623| 66.74412| 24  |
| HAT-P-4| 229.99| 36.13 | 58.29627| 57.34666| 6    |
| HD 118203| 203.51 | 53.73 | 109.3461| 62.26016| 27  |
| HD 17092| 41.59  | 49.65 | 141.3032| −9.07813| 30  |
| HD 17156| 42.44  | 71.75 | 131.99263| 10.99141| 38  |
| HD 219828| 349.69 | 18.65 | 94.26312| −39.00768| 36  |
| HD 37605| 85.01  | 6.06  | 199.10635| −12.89319| 28  |
| HD 4203| 11.17  | 20.45 | 120.78828| −42.39362| 36  |
| HD 43691| 89.49  | 41.09 | 172.65002| 11.92868| 31  |
| HD 46375| 98.3   | 5.4   | 206.04629| −1.57714| 35  |
| HD 49674| 102.88 | 40.87 | 175.33733| 17.37424| 26  |
| HD 68988| 124.59 | 61.46 | 155.26411| 33.91797| 40  |
| HD 80606| 140.60 | 50.54 | 167.51619| 44.3251 | 29  |
| HD 88133| 152.53 | 18.12 | 217.91036| 51.87233| 25  |
| HD 89307| 154.39 | 12.56 | 227.39155| 51.36809| 29  |
| HD 89744| 155.54 | 41.17 | 178.49952| 56.40393| 35  |
| HD 9407| 23.64   | 58.95 | 126.80841| 6.40343| 29  |
| HIP 14810| 47.81  | 21.1  | 161.54971| −31.09206| 31  |
| K10    | 294.12 | 46.01 | 78.80125| 11.96774| 23  |
| K14    | 299.64 | 44.87 | 79.67739| 8.02455 | 23  |
| K15    | 296.12 | 43.53 | 77.23814| 9.55879 | 20  |
| K20    | 294.71 | 39.63 | 73.25579| 8.6313  | 23  |
| K21    | 291.58 | 38.15 | 70.78675| 10.11117| 19  |
| K4     | 295.69 | 49.9  | 82.83844| 12.79648| 19  |
| K5     | 291.93 | 48.45 | 80.38966| 14.37021| 20  |
| K7     | 285.05 | 45.2  | 75.36922| 14.72520| 25  |
| K8     | 281.91 | 43.44 | 72.80583| 18.91956| 26  |
| KEPLER3-TRES2| 285.9  | 49.2 | 79.51202| 18.32046| 11  |
| KEPLER4| 282.52 | 47.46 | 76.97688| 19.84864| 23  |
| WASP-1 | 3.17   | 31.99 | 115.36902| −30.42966| 36  |
| XO-1   | 240.55 | 28.09 | 45.73607| 48.00637| 20  |
| XO-2   | 117.03 | 50.16 | 168.36412| 29.32105| 31  |

We estimate the effect on a typical field for the first 2 years of observation (Kepler field) and—as worst case scenario—when observing directly toward the galactic center (assuming that reddening and extinction both increase toward the galactic center).

Taking the absolute magnitudes for dwarfs and giants from Allen (2001) and extending to bluer colors using Schaefer et al. (1982), we computed the spectroscopic distances for dwarfs and giants of different spectral types with apparent magnitudes of $V = 10$ and 13, representing the bright and faint end of our targets. We then computed the typical proper motions these stars would have according to the galactic model of Dhital et al. (2010) and placed them into a $PM_2$ diagram.

Figure 10 shows the result for the Kepler field with stars with $V = 10$ in the left and stars with $V = 13$ in the right panel. The MARVELS region of interest is shaded. For illustration we plotted the overlap stars of Kepler and MARVELS with colored symbols in the left panel. In order not to overcrowd the right panel showing the faint end of MARVELS magnitude range, we do not overplot the MARVELS-Kepler overlap stars, thus keeping the reddening/extinction vector more visible. The typical position of stars of a given spectral type and magnitude according to the galactic model from Dhital et al. (2010) are marked as box and whiskers. Each box represents 50% of all stars, the whiskers the upper and lower 27%—leaving 3%...
outliers apart. The spectroscopic distances of dwarfs and giants are given at the bottom. For each of the boxes we computed a reddening-extinction vector. We multiplied an assumed mean density of \( n_H = 1 \text{ atom/ccm} \) with the spectroscopic distance, yielding a column density of \( n_H = N_H d \). This column density we converted to \( A_V \) and further to \( E(B-V) \), adopting the relations

\[
A_V = n_H/2.30 \times 10^{21}, \quad E(B-V) = A_V/3.1. \quad (6)
\]

Taking \( A_f/A_V = 0.282 \) and \( A_H/A_V = 0.190 \) from Cardelli et al. (1989) we converted \( E(B-V) \) to \( E(J-H) \) which completes the vector \( (E(J-H), A_J) \). For dwarfs we plotted this vector at each box. For \( V = 10 \) there is no noticeable shift in and out of the region of interest. For \( V = 13 \) the brightest stars \((A0)\) can be shifted in the region of interest. However, they are a very small fraction of the stellar population and thus will not significantly pollute the sample.

We repeated the same analysis for observations toward the Galactic Center, with the results shown in Figure 11. Compared to the Kepler field there is a slight but insignificant shift in the position of the boxes. For bright stars \((V = 10, \text{ left panel})\) reddening and extinction do not play a significant role. For faint stars \((V = 13, \text{ right panel})\) A0 stars get shifted into the region of interest. We might see pollution by late F-dwarfs; A5–F5 are not reddened enough. As even in the worst case of observing toward the Galactic Center we only would see a light pollution by late F-dwarfs, we concluded that correcting for reddening and extinction is not necessary for the MARVELS target selection.

Reddening and extinction get stronger if we assume a higher density than average 1 atom per ccm. The spectroscopic distance of giants with an apparent magnitude of \( V = 13 \) is \( 4.8 \text{ kpc} \). As long as we do not hit a denser region within this distance, our estimation of reddening and extinction holds.

4. SUMMARY

In this paper we have discussed the target selection methodology for creating an input catalog for the MARVELS radial velocity survey. The MARVELS survey was interested in looking for radial velocity companions to stars of the FGK spectral type primarily focusing on dwarf stars. To achieve this goal, a target selection criteria of 10% giant stars 90% dwarf stars was set. Target selection for MARVELS is broken down into two distinct phases, the initial phase which found targets for the first half of the survey, and the final phase which found targets for the second half of the survey. The initial target selection method used low-resolution \((R \sim 2000)\) spectra from the SDSS spectrographs processed by a modified SSP pipeline. This method ultimately proved to be inadequate for removing giant star contamination primarily on the cool end \((T_{eff} < 4500)\). This result was not entirely unexpected because the MARVELS stars are significantly brighter than what the SSP was designed to work with. As a result, the giant contamination rate for the initial phase was 31%. Given the results of the initial phase, the final phase of target selection used a different method. Instead of low-resolution spectra, a RPM\(_f\) method was employed. This method did a much better job meeting our criteria providing a giant contamination rate of just 13%.

This investigation also revealed two other notable results. First is that interstellar reddening is not a major factor in influencing the stars selected for MARVELS. This is due to primarily to the relatively short distances to the MARVELS stars and the way in which the reddening vector points in the RPM\(_f\) diagram. Second is that the RPM\(_f\) method, although being quite useful for separating dwarfs and giants, is not able to adequately separate dwarfs and subgiants (subgiant contamination of the dwarf sample is on order 30%). For the MARVELS scientific goals this was not important. However, future surveys or missions need to be aware of this fact when designing their respective input catalogs.

We acknowledge the generous support from the W. M. Keck Foundation for developing the MARVELS survey instruments. The MARVELS project was also supported from NSF with grant AST-0705139, NASA with grant NNX07AP14G and the University of Florida.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is \url{http://www.sdss3.org/}.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

APPENDIX

Table 5 provides field names, coordinates, and the number of observations during the initial phase. Table 6 provides the same information for the final phase.

REFERENCES

Allen, 2001, Allen’s Astrophysical Quantities (Berlin: Springer)
Allende Prieto, C., Majewski, S. R., Schiavon, R., et al. 2008, AN, 329, 1018
Bessell, M. S. 1990, PASP, 102, 1181
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Casagrande, L., Ramírez, I., Meléndez, J., Bessell, M., & Asplund, M. 2010, A&A, 512, A54
Casagrande, L., Silva Aguirre, V., Stello, D., et al. 2014, ApJ, 787, 110
Collier Cameron, A., Wilson, D. M., West, R. G., et al. 2007, MNRAS, 380, 1230
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of Point Sources, \url{http://irsa.ipac.caltech.edu/applications/Gator}
Dhillon, S., West, A. A., Stassun, K. G., & Bochanski, J. J. 2010, AJ, 139, 2566
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011a, AJ, 142, 72
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011b, AJ, 142, 72
ESA 1997, The HIPPARCOS and TYCHO Catalogues. Astrometric and Photometric Star Catalogues Derived From the ESA HIPPARCOS Space Astrometry Mission (ESA Special Publication)
Ge, J., Lee, B., de Lee, N., et al. 2009, Proc. SPIE, 7440, 0
Girardi, L., Williams, B. F., Gilbert, K. M., et al. 2010, ApJ, 724, 1030
Gunn, J. E., Siegmund, W. A., Munnery, E. J., et al. 2006, AJ, 131, 2332
Kepler Mission Team 2009, yCat, 5133, 0
Lasker, B. M., Lattanzi, M. G., McLean, B. J., et al. 2008, AJ, 136, 735
Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008, AJ, 136, 2022
Maíz Apellániz, J. 2006, AJ, 131, 1184
Marigo, P., Girardi, L., Bressan, A., et al. 2008, A&A, 482, 883
Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. 1997, A&A, 323, L49
Pinsonneault, M. H., An, D., Molenda-Zakowicz, J., et al. 2012, ApJS, 199, 30
Schaeifers, K., Voigt, H. H., Landolt, H., Boernstein, R., & Hellwege, K. H. 1982, Astronomy and Astrophysics, B: Stars and Star Clusters (Berlin: Springer)
Siebert, A., Williams, M. E. K., Siviero, A., et al. 2011, AJ, 141, 187
Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377
Zacharias, N., Finch, C., Girard, T., et al. 2010, AJ, 139, 2184
Zacharias, N., Monet, D. G., Levine, S. E., et al. 2004, in American Astronomical Society Meeting Abstracts, 1418
Zwitter, T., Siebert, A., Munari, U., et al. 2008, AJ, 136, 421