Astron. Astrophys. 363, 947–957 (2000)

ASTRONOMY
AND
ASTROPHYSICS

The G dwarf problem

Analysis of a new data set

Bjarne Rosenkilde Jørgensen⋆

Copenhagen University Observatory, Niels Bohr Institute, Juliane Maries vej 30, 2100 Copenhagen Ø, Denmark

Received 27 July 2000 / Accepted 26 September 2000

Abstract. The G dwarf metallicity distribution of a mass limited sample of stars within 40 pc and south of $\delta = -26°$ and for which Strömgren $uvby$ and H_β photometry is available is studied. Special care is taken to ensure that the sample consists of stars with a main sequence lifetime longer than the age of the Galactic disk. Furthermore, binary stars, subgiants, active stars, etc. are removed to ensure that the sample is as clean as possible. The sample was limited to the mass interval $0.7 - 1.0 M_\odot$ using masses determined by interpolation between theoretical evolutionary tracks. A mass interval was used as opposed to a color interval, as it is more directly comparable with theoretical models. The final sample contains 253 stars. Even though the sample shows a remarkably small low metallicity tail a new scale height weighting method was developed, and used. The resulting G dwarf distribution cannot be reconciled with the simple model for chemical evolution, in fact this dataset shows that the G dwarf problem is even larger than earlier results indicate.

Key words: Galaxy: evolution – Galaxy: solar neighbourhood

1. Introduction

The G dwarf problem is one of the more persistent problems in Galactic astronomy. The problem is the discrepancy between the prediction of the simple model for the evolution of the solar neighbourhood and observations. Basically the simple model for galactic chemical evolution predicts too many low metallicity stars among the long lived stars. It was first examined by van den Bergh (1962) and Schmidt (1963).

Many theoretical models have been created to explain the G dwarf problem (e.g. Larson 1972; Lynden-Bell 1975; Clayton 1985; Pagel 1989; Truran & Cameron 1971 and Pagel & Tautvaišienė 1995). Most of these models fit the present observations fairly well.

Unfortunately the observations are not of sufficient quality to discriminate between these models. The problem has, in general, been to remove biases from the sample. e.g. kinematical bias, or bias resulting from the problem of isolating long-lived stars.

⋆ *Present address:* Lund Observatory, Box 43, SE-22100 Lund, Sweden (bjarne@astro.lu.se)

In this paper an investigation of a large volume limited sample of F and G dwarfs in the solar neighbourhood is presented. The sample is subjected to several tests to ensure that it is as pure as possible. The metallicity distribution of a final mass limited sample is then compared with several other observational and theoretical distributions.

This paper is organized as follows: In Sect. 2 the sample used is briefly described. In Sect. 3 the steps taken to make the sample as clean as possible is described. In Sect. 4 the corrections applied to the sample is described. In Sect. 5 the resulting G dwarf distribution is discussed. In Sect. 6 it is compared with observational and theoretical distributions. Finally in Sect. 7 the conclusions that can be drawn from this investigation are presented.

2. The sample

2.1. Observations

This work is based on observations obtained with the Danish 50cm and 1.54m telescopes on La Silla. All stars from the Michigan Spectral Catalogue (Houk & Cowley 1975; Houk 1978, 1982) in the spectral range F to K2, and below $\delta = -26°$ were included. This spectral range was used to ensure that all stars in the G dwarf mass interval were included, as different age and metallicity spread stars of a given mass over a wide spectral range.

The sample has varying magnitude limits based on color. That can be seen in Table 1. $M_{\rm pg}$ is the absolute photographic magnitude and $m_{\rm pg,lim}$ is the established apparent photometric limit in the corresponding color interval. For the F stars the limit was $m_{v,\rm lim} = 8.3$ which is more than adequate.

All stars of luminosity class V, IV or undertermined class are included. The sample consists of 5561 stars.

Strömgren four color photometry was obtained for all stars. The H_β index was measured for F and early G-type stars. The observations are more fully described by Olsen (1983, 1993, 1994).

The calibrations used are split in two parts; the G dwarfs and the F dwarfs, for which separate calibrations were used. In the zone where these groups overlap the F dwarf calibration was used. The F dwarf group used the Crawford (1975) calibration for M_v, the Nissen (1981) calibration for [Fe/H], the Magain

Table 1. The sample limits as a function of spectral type.

Sp. T	b-y	M_v(ZAMS)	$M_{\rm pg}$(ZAMS)	$m_{\rm pg,lim}$
G0	0.368	4.44	5.18	8.6
G1	0.378	4.54	5.30	8.7
G2	0.387	4.65	5.44	8.8
G3	0.402	4.85	5.67	9.1
G5	0.420	5.11	5.98	9.5
G6	0.440	5.42	6.34	9.8
G8	0.456	5.64	6.61	10.1
K0	0.478	5.88	6.90	10.4
K1	0.498	6.06	7.13	10.6
K2	0.518	6.23	7.34	10.8

(1987) calibration for T_e and the Olsen (1988) calibration for $(b-y)_0$. In the G dwarf group the calibration by Olsen (1984) was used.

2.2. Distance limit

To ensure that the sample is reasonably volume complete a distance limit must be established. This is done by comparison with a homogeneous distribution. For $b-y < 0.520$ the sample is reasonably complete to $40\rm pc$. For the F stars in the sample the limit is beyond $100\rm pc$, causing a lot of them to be removed when the sample is reduced to stars within $40\rm pc$. The distance limit leaves 1141 stars in the sample.

2.3. Mass derivation

The masses of the stars in the sample were determined by linear interpolation in the evolutionary models by VandenBerg et al. (2000). These models are brand new, and include an improved equation of state with non-ideal corrections, updated nuclear reactions, neutrino cooling rates, and modern opacites. The new models are so superior compared to the older models (VandenBerg 1985) that the older models should be considered obsolete. Even so the old models were used for comparison to gauge the effect of different models on the resulting masses. The models agree within 0.05 to $0.1 M_\odot$, with a small trend, where the old models give slightly lower masses for low metallicity stars.

In referring to these models, the following definitions are used: A point is the data describing a single model star of a given mass and metallicity at one age in its evolution. A track is the collection of points describing the evolution of a model star of a given mass and metallicity. A set is the collection of tracks describing stars of a given metallicity.

This is an interpolation in three dimensions. The dimensions involved are: M_v, $T_{\rm eff}$ and [Fe/H]. The interpolation is split in three parts to match these dimensions. For each of the two sets nearest to the star in [Fe/H], the two tracks nearest to the star is found. Temporary points are then found on these tracks by interpolating along the track, so that the temporary points have the same M_v as the star. It is then possible to find a temporary mass for the star in each set. A final interpolation (between the two sets) gives the mass of the star. As a test this mass is compared with the four masses used to derive it. If the resulting mass is within $0.05 M_\odot$ of any of these points the result is accepted. This leaves 749 stars in the sample. If the sample is restricted to stars where no extrapolation is used to derive their mass, the number of stars is reduced to 550. In both cases this is still a fairly large sample.

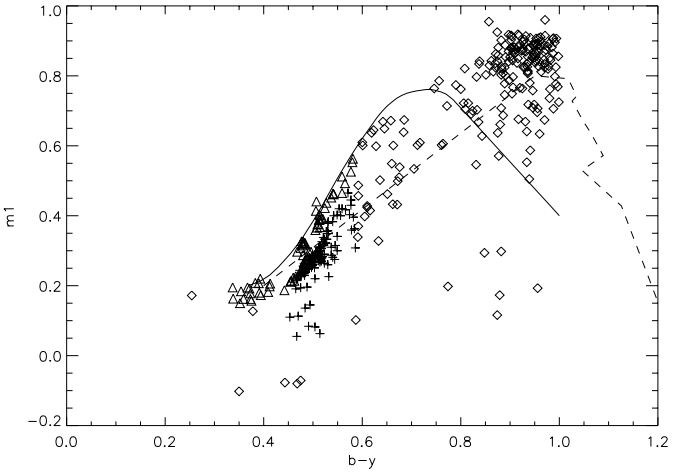

Fig. 1. $(b-y)-m_1$ diagram for the discarded stars. The diamonds are stars outside the calibration ranges of Schuster & Nissen (1989), the crosses are stars with conflicting [Fe/H] and the triangles are the rest. The solid line is the preliminary standard relation for dwarfs of Hyades metallicity. The dashed line is for giants.

A total of 392 stars were discarded by the interpolation program. It is of course very interesting to examine these stars. The Olsen (1984) calibration is only valid for luminosity class V stars. In general, subgiants will get a too low [Fe/H]. To test for subgiants, the Olsen [Fe/H] was compared to the calibrations of Schuster & Nissen (1989), as their calibration does include subgiants. Of the 392 rejected stars, 206 were outside the calibration limits of Schuster & Nissen. 116 had $[Fe/H]_{\rm SN} > [Fe/H]_{\rm Olsen} + 0.5$. This indicates that they are subgiants. Looking at the $(b-y)-m_1$ diagram (Fig. 1) it is obvious that the 116 stars with conflicting [Fe/H] lie in the area where subgiants normally are found (see Olsen 1984). It is also obvious that the stars outside the calibration range are mostly giants and stars on the lower main sequence (K-dwarfs etc.).

Of the remaining 70 discarded stars, 14 were close binaries according to the Hipparcos catalogue (ESA 1997). In addition, one was found to be chromospherically active according to the EMSS survey (Einstein Observatory Extended Medium Sensitivity Survey , see Sect. 3.3). It is not clear why the remaining stars are discarded, but their small number justifies simply accepting that they did not fit in some way.

2.4. The mass interval

The range of colors included in the sample corresponds to a mass interval, within which the sample is reasonably complete. The lower mass limit is established by comparion to the lowest temperature for each metallicity in the theoretical models

Table 2. Main sequence lifetimes at different [Fe/H].

Mass ($M_\odot$)	Metallicity ([Fe/H])	t(MS)
1.0	–1.01	6.8 Gyr
1.0	–0.83	7.6 Gyr
0.9	–1.01	10.3 Gyr
0.9	–0.83	11.5 Gyr
0.8	–1.01	16.3 Gyr
0.8	–0.83	18.1 Gyr
0.7	–1.01	27.0 Gyr
0.7	–0.83	30.0 Gyr

by VandenBerg et al. (2000). This establishes the lower mass limit to between $0.7M_\odot$ and $0.8M_\odot$. As stars with high metallicity has lower $\log T_e$ than comparable low metallicity stars, a limit at $0.7M_\odot$ will lead to high metallicity stars being underrepresented. The upper mass limit must be established so that all stars below the mass limit have main sequence lifetimes longer than the age of the disk. As this age is still controversial this limit is subject to some uncertainty. The limit is established by comparing the age of a model star 0.02 dex (in $\log T_e$) past the bluest point of the evolutionary track. This extra distance is chosen so that a $1.0M_\odot$ star with solar metallicity has a main sequence lifetime of 10.7 Gyr, and to allow for the observational scatter. The resulting ages are shown in Table 2.

As the region below [Fe/H] = -1.0 is only marginally relevant for the purposes of this paper, as there are very few stars in the sample below this limit, it is safe to assume that the limit is somewhere between 0.9 and 1.0 $M_\odot$. This does represent a rather narrow range when compared with the color limit.

3. Steps towards a clean sample

A number of corrections was applied to the sample to ensure that it is as free of contaminants as possible.

3.1. Accuracy & errors

It is of prime importance to study the accuracy of the calculated masses. Nissen (1994) made a study of the calibration of the Strömgren system for F and G stars. In this study he finds among other things the standard scatter in $T_{\rm eff}$, M_v and [Fe/H] determined from Strömgren photometry. These results can then be used to estimate the accuracy of the mass as found by the method described above. Nissen finds that the error in photometric distance is about 15%, which corresponds to 0.3 mag. For disk stars the error in effective temperature lies around 100 K. The error in [Fe/H] is estimated to 0.15 dex from Olsen (1984). This is a conservative value as for most stars (those above [Fe/H] =-0.6) the scatter in the calibriation was 0.13, furthermore for the F stars the scatter is between 0.08 and 0.11 dex.

A grid of calculated masses running from [Fe/H]=0.00 to [Fe/H]=-1.05 in steps of alternately 0.07 and 0.08, from M_v = 1.5 to M_v = 9.9 in steps of 0.15, and from $\log T_{\rm eff} = 3.950$ (8000 K) to 3.477 (3000 K) in steps of 50 K, was constructed.

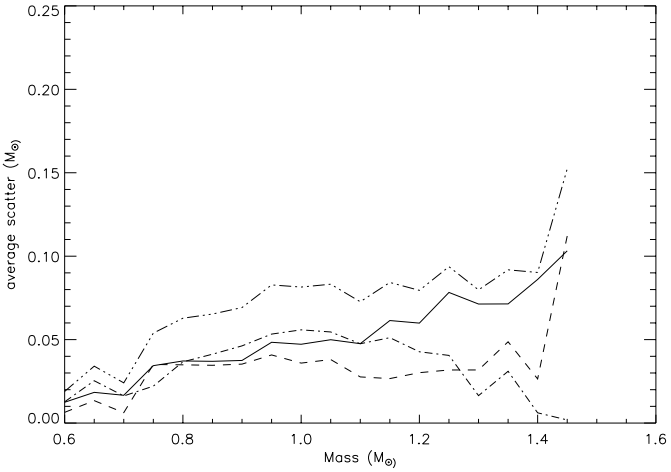

Fig. 2. Average scatter in the mass determinations. The solid line is the scatter from M_v, the dashed line is the scatter from $\log T_{\rm eff}$, the dash-dotted line is the scatter from [Fe/H] and the dash-dot-dot-dot line is the combined scatter. Only stars within $0.05M_\odot$ of an interpolation point have been used and invalid neighbours have been removed.

This is then a 'double resolution' grid, i.e. the step length is half the scatter. Using the next nearest neighbour it is possible to estimate the average scatter as a function of mass. It is obvious that the grid points outside the theoretical sets give nonsensical results. To avoid using these points in the estimation of the scatter, any points with a calculated mass below $0.5M_\odot$ or above $1.5M_\odot$ were not used. In addition only calculated masses within $0.05M_\odot$ of any of the interpolation points were used. The results are shown in Fig. 2.

The mass is determined with an average scatter of $\approx 0.08M_\odot$. There is a slight increase with increasing mass, but this is explainable. This increase is mostly derived from M_v, and as the distance in M_v between tracks is larger on the lower main sequence, a "step" in mass up or down in the HR diagram will influence the calculated mass less. It is interesting to note that the average scatter is comparable to the difference between using the old and the new models.

3.2. Binaries & subgiants

To correct for any possible contamination of the sample by multiple stars or subgiants, the sample has to be examined closely to remove any such stars. To do this the remaining stars were all checked in several catalogues. The catalogues used were:

- The Washington Double Star Catalogue (Worley & Douglass 1996) This catalog contains 78100 known visual double stars.
- The SIMBAD online database was checked to find MK-classifications for the stars in the sample. It was also used as a source to find multiple stars, generally spectroscopic binaries.
- The Hipparcos catalogue (ESA 1997) was used to check the photometric parallaxes. In general the photometric paral-

laxes was verified, although a few subgiants were discovered in this way.

Using these catalogues, the stars were classified into several categories:

- Visual binaries. The visual binaries in the sample were split in two. The systems with a separation greater than $10''$ were regarded as single stars as the photometry of the individual components should be unafected by the other component. For the remaining visual binaries there are three possibilities. If the magnitude difference, Δm, is larger than 5.0 the primary star dominates, and the system is regarded as a single star. If Δm is below 0.2 the two stars are regarded as identical, and the system is counted twice if the photometric parallax places the system within 29 pc to account for the extra luminosity. If Δm is between 0.2 and 5.0 the photometry is regarded as corrupted and the stars are discarded.
- Spectroscopic binaries. There were eight stars that SIMBAD categorized as spectroscopic binaries. These stars were removed as their photometry may be unreliable.
- Variable stars. These stars were removed from the sample.
- Eclipsing binaries. The four eclipsing binaries in the sample were removed.
- Alpha Centauri was in the sample, but it was removed as the photometry was of both α Cen A & B combined, and as their $\Delta m = 1.4$, they were discarded.
- Giants. A few giant stars were still in the sample. They were removed.
- Subgiants. Although a lot of the subgiants in the sample were already removed by the interpolation program several remain. These stars were removed here. The removal process was stepwise. First the MK classifications from the SIMBAD database were examined. If there were more than three sources that all agreed on the luminosity class, the star was either accepted or rejected on that basis. If that was not enough, the position of the individual star in the $(b-y)-c_1$, $(b-y)-m_1$ and c_1-m_1 diagrams were examined. Finally, the distances to all the remaining stars were checked with Hipparcos data. Here a conservative limit of 60 pc (corresponding to 40 pc + 3σ) was adopted. Stars with larger Hipparcos distances were discarded.
- CORAVEL spectroscopic binaries. CORAVEL data was used to remove 15 spectroscopic binaries.
- Halo stars. One halo star (HD 113083) was found from the velocity data described below.
- Single main sequence stars. These stars were accepted and used.

The sample should now be reasonably free of disturbing double stars and subgiants, and it now contains 497 stars.

3.3. Chromospheric activity

Chromospheric activity may affect the photometric metallicities, through an influence on the Strömgren m_1 index (e.g. Giampapa et al. 1979; Morale et al. 1996). To examine whether this has any effect on the current sample, the Henry et al. (1996) survey of HK line emission in solar-like stars (between G0 and K2) south of $\delta = -25^\circ$ and with HD magnitude < 9.0 was used. For a G2V star this corresponds to a distance of $68\mathrm{pc}$. This sample contains 650 stars. Comparing with stellar densities they expect their sample to contain about 50% of the stars present within $50\mathrm{pc}$.

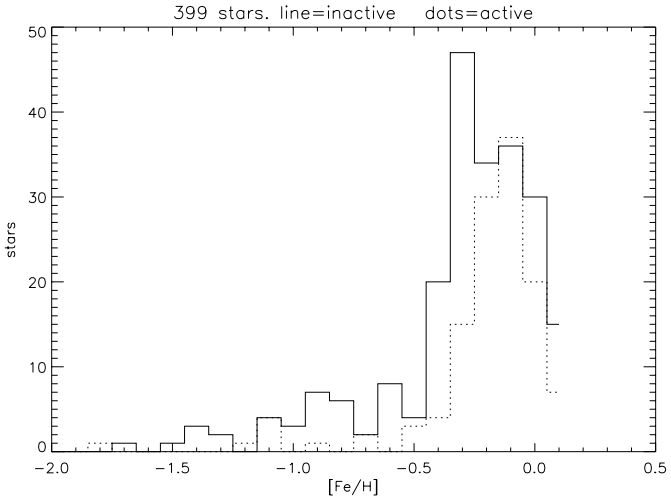

Fig. 3. Comparison of the metallicity distribution of 399 stars classified as active and inactive by Henry et al. (1996). The solid line is inactive stars, the dotted line is active stars.

Of the 1141 stars in the current sample within $40\mathrm{pc}$, 399 were also present in the Henry et al. sample. This is perhaps a bit low as Henry et al.'s sample extends to $50\mathrm{pc}$, thus having almost twice the volume, but the current sample also includes a fair amount of F stars and giants etc. This is a sufficiently large sample that an investigation of the effect on the photometric [Fe/H] can be made. Using Henry et al.'s limit of $\log R'_{HK} = -4.75$ to divide the sample into active and inactive stars, two subsamples are made which can then be compared. The effect suggested by Morale et al. and Giampapa et al. should lead to quite different metallicity distributions, with the high activity part considerably more metal poor than the low metallicity part. But as can be seen in Fig. 3, no real difference can be seen.

Even when limited to the stars that are accepted after all tests there are still 235 stars left, enough for these investigations. When the sample is limited to these stars, the result does not change, as seen in Fig. 4. This leaves a significant discrepancy between these results and the predictions of others. In an attempt to understand this phenomenon, the distribution of [Fe/H] versus $\log R'_{HK}$ (Fig. 5) was examined.

As can be seen, there is little indication that low metallicity stars are seriously polluted by active stars. There is a clear relation between metallicity and activity to around $\log R'_{HK} \approx -4.9$, as expected from the activity - age - metallicity connection. Around solar metallicity there is a range of activity levels. Only those stars categorized as very active by Henry et al. (1996) ($\log R'_{HK} > -4.2$) show a tendency to have underestimated metallicities. These stars form a very small part (2.6%

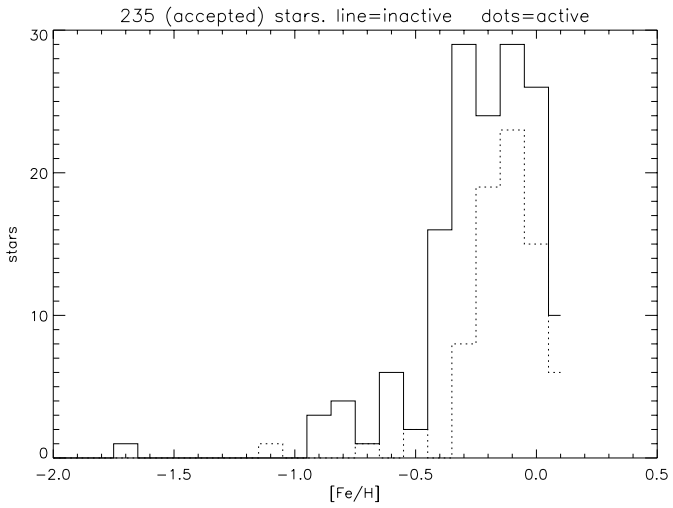

Fig. 4. Comparison of the metallicity distribution of stars classified as active and inactive by Henry et al. (1996), using only the 235 stars accepted after all tests. Again the solid line is inactive stars, the dotted line is active stars.

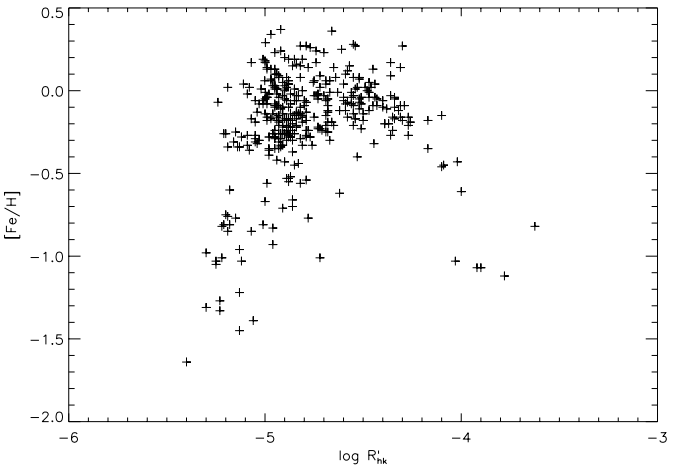

Fig. 5. [Fe/H] versus $\log R'_{HK}$ for the 399 stars from Fig. 3. Active stars, with $\log R'_{HK} > -4.75$ are predicted to have lower photometric metallicities.

according to Henry et al.) of the G dwarfs, so even without any attempt to remove such stars from a sample they would not affect it greatly. Among the active stars there are a few stars below [Fe/H] $= -0.5$, but the great majority lies around solar [Fe/H]. Even among the few highly active stars with low [Fe/H] only one was present after all the other tests. It was removed.

4. Corrections

A large program has been underway for some years now, to determine the space velocities of nearby stars. the author was kindly lent some data from this project by Andersen & Nordström (private communication). For all but 33 of the stars left in the sample there are velocity data and calculated orbits (described in Fux 1997). These data are based on the same absolute magnitudes as the results presented here.

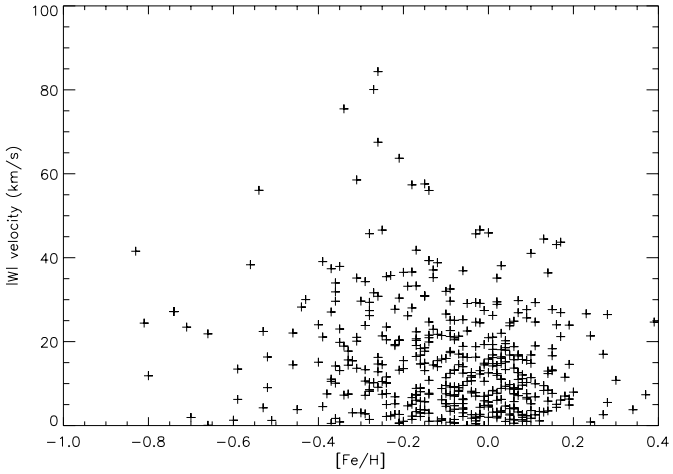

Fig. 6. Velocities perpendicular to the galactic plane versus metallicity.

With these data it is possible to make an in-depth study of the scale height weighting problem. As theoretical models usually concern a cylinder around the sun and the observed sample is confined to the solar neighbourhood, some correction must be made. This is usually done by weighting the sample according to the average $|W|$ velocity in each velocity bin to correct for the different scale hights. According to Mayor et al. (1977) the U and V velocities does not produce a noticeable effect (meaning that the radial metallicity gradient is low), thus only $|W|$ velocities are usually considered.

From Fig. 6 it is clear that most of the stars have $|W|$ velocities below $50\mathrm{km/s}$. There are 10 stars with $|W|$ velocities above $50\mathrm{km/s}$. These stars have $z_{\max}$ between 1 and 2 kpc, and could be either halo stars or the end of the disk velocity distribution. They do not have exceptional metallicities. The remaining stars are approximately gaussian distributed. A fitted exponential distribution would predict no stars with $|W|$ velocities above $\approx 50\mathrm{km/s}$, which suggest that this might be a way to separate the possible halo stars from the sample.

Another result evident from this figure is that many stars have very low $|W|$ velocities. This can cause a problem when using conventional corrections such as weighting the stars by their $|W|$ velocities; When one star has a $|W|$ velocity of e.g. $1.5\mathrm{km/s}$ and another has a $|W|$ velocity of e.g. $9.0\mathrm{km/s}$, this will cause the second star to have six times the weight of the first star, while in reality none of them spend much, if any, time more than 40 pc from the galactic plane. As seen in Fig. 6 there is only a very general trend for $|W|$ to depend on [Fe/H]. This makes it somewhat unreliable to use average values, but they will at least give the general trend. Averaging the velocities in 0.1 [Fe/H] bins gives Fig. 7. In the figure a linear fit is shown (with a dashed line). Note though, that it suffers from the bad statistics in the low metallicity part with $2-4$ stars per bin, as opposed to ≈ 70 in the more metal rich part.

Another possible correction is to use the f method from Sommer-Larsen (1991). Using this method would ignore the velocity data, which is not optimal as this discards useful information.

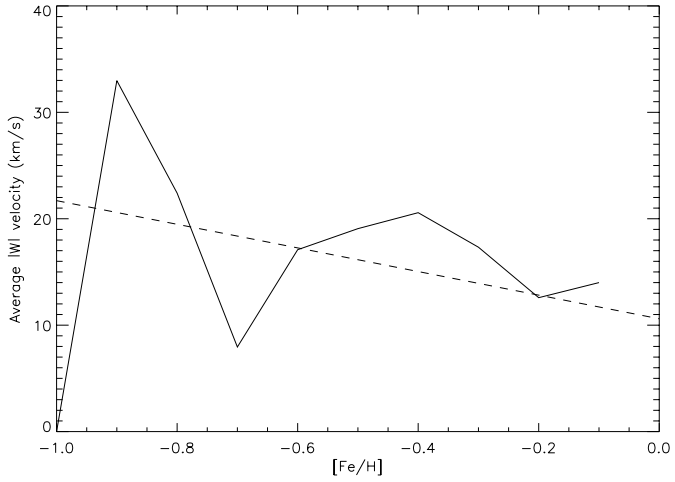

Fig. 7. Average velocities perpendicular to the galactic plane versus metallicity.

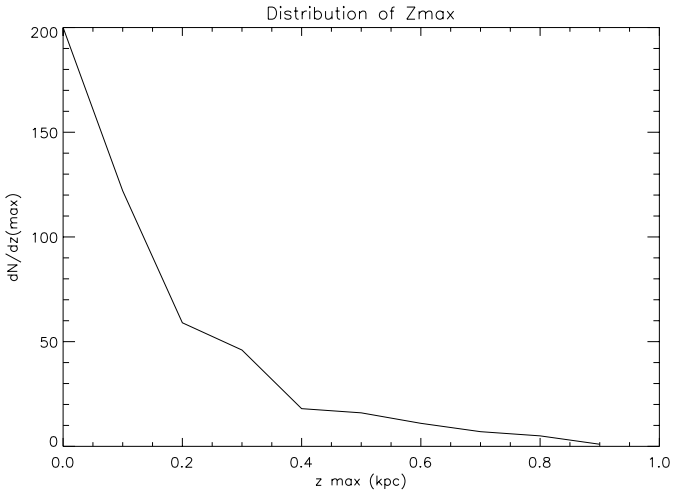

Fig. 8. Distribution of $z_{\max}$.

Another possibility is to use the calculated maximum height above the plane. The distribution of $z_{\max}$ is shown in Fig. 8. There seem to be an exponential decay (a rough guesstimate suggests a scale height of around 300 pc). The important feature of the plot is that the maximum height of a considerable part of the sample is comparable to the distance limit (40 pc). Thus $|W|$ is not a good indicator of the time the star spends inside our volume, as some stars inside the sample will be near their $z_{\max}$ and will therefore have low $|W|$ velocities. These low velocities (e.g. $2\mathrm{km/s}$) will give extremely low weights with a standard correction method.

The $z_{\max}$ correction is done by assuming that the stars move vertically in a harmonic oscillator potential. A harmonic oscillator potential is a fair approximation, as long as z is small compared to the scale height of the disk, and this is true for most of the sample (it is not the method used to calculate $z_{\max}$, but for these purposes it is acceptable). This method also assumes that the Sun is located in the Galactic plane. While a more complete analysis would remove this assumption is is used as it greatly

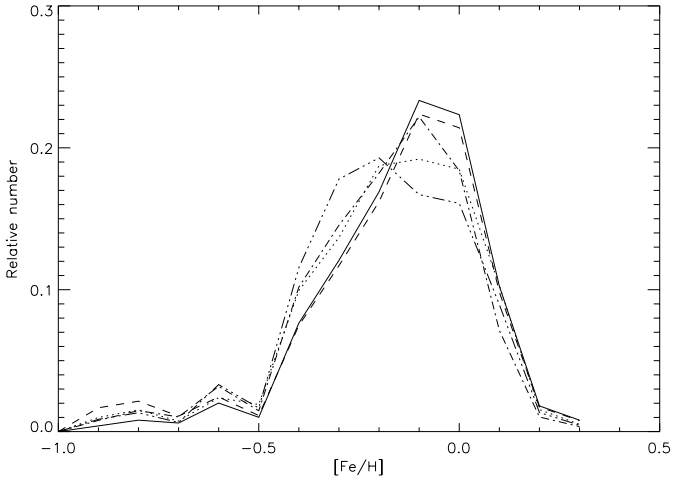

Fig. 9. Metallicity distribution of all remaining stars, with different scale height corrections applied. The solid line is the uncorrected sample, the dotted line is the sample corrected with the $z_{\max}$ method and stars with $|W| > 50\mathrm{km/s}$ removed, the dashed line is using the Sommer-Larsen correction, the dot-dashed line is with a linear correction and the dot-dot-dot-dashed line is with the $z_{\max}$ method using all stars.

simplifies the problem. With these assumptions it is then easy to find the proportion of time spent within 40 pc of the Galactic plane knowing the maximum distance that the star reaches. Each star is then weighted with the reciprocal time:

$$w = \frac{1}{1 - (\arccos(40\mathrm{pc}/z_{\max})2/\pi)} \tag{1}$$

Eq. 1 is used, if the star reaches further than 40 pc away from the Galactic plane. If it does not, its weight is set to one. For the 33 stars where no velocity data were available, the correction from the average values of $|W|$ was used, normalized to the weights from the method described above.

Unfortunately this method is also prone to distortions from a few stars. e.g. some stars gain weights of ≈ 75, compared to weight$= 1$ for a star that does not move more than 40 pc from the Galactic plane. This is of the same order of magnitude as the sample size and can be very disruptive. This problem is compounded by the fact that the harmonic oscillator approximation breaks down at large distances from the plane. Another problem is that with the large scatter in velocities and the very few stars in the low metallicity bins, this effect will primarily affect the bins with many stars, as the bins with few stars will probably not have any extreme stars. To rectify this problem all stars with $|W|$ velocities above $50\mathrm{km/s}$ were removed.

The different metallicity distributions that result from the different corrections are shown in Fig. 9. The most interesting point of this plot is that the distribution below [Fe/H]$= -0.5$ does not change much, although the Sommer-Larsen correction is a bit above the others. Another interesting point is the position of the main "bulge". The bulge shifts position with changing correction method. The difference between the $z_{\max}$ correction and the truncated $z_{\max}$ correction indicates that the cut-off at $50\mathrm{km/s}$ does remove some of the weighting, but as the distribu-

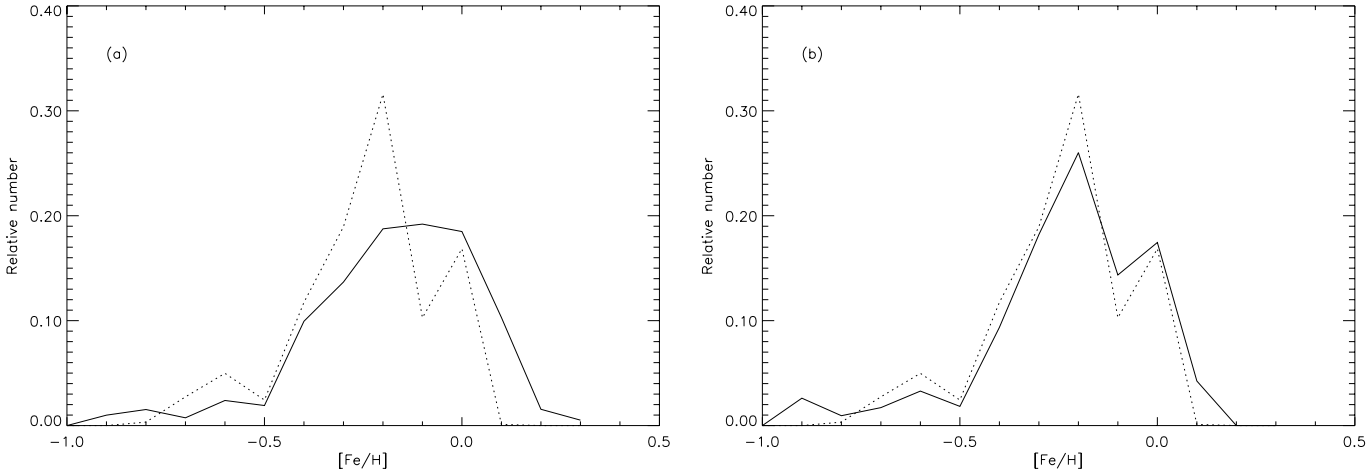

Fig. 10a and b. Comparison of (**a**) all stars (A, solid line) and the conservative sample (B, dotted line) and (**b**) the $0.7 - 0.9M_{\odot}$ sample (C, solid line) with the conservative sample (B, dotted line). See Table 3.

Table 3. The definition of the different subsamples.

Subsample	**ID**	**# of stars**
All stars still in the sample.	A	497
All stars between 0.8 and 0.9 $M_{\odot}$.	B	117
All stars between 0.7 and 0.9 $M_{\odot}$.	C	177
All stars below 0.9 $M_{\odot}$.	D	216
All stars between 0.8 and 1.0 $M_{\odot}$.	E	193
All stars above 0.8 $M_{\odot}$.	F	398
All stars above 1.0 $M_{\odot}$.	G	205
All stars between 0.7 and 1.0 $M_{\odot}$.	H	253

Table 4. The final metallicity distribution.

Metallicity bin	# of stars	weighted
-1.0 to -0.9	0	0.000
-0.9 to -0.8	2	0.018
-0.8 to -0.7	2	0.006
-0.7 to -0.6	3	0.013
-0.6 to -0.5	7	0.035
-0.5 to -0.4	2	0.013
-0.4 to -0.3	17	0.090
-0.3 to -0.2	38	0.172
-0.2 to -0.1	56	0.219
-0.1 to 0.0	48	0.144
0.0 to 0.1	59	0.198
0.1 to 0.2	19	0.091

tions are comparable at lower metallicities, the cut-off does not adversely affect the distribution. As the distribution with cut-off is comparable to the other corrections, it will be used.

5. Results

After these corrections it is possible to review the mass limits of the sample. To facilitate this investigation, a number of subsamples are defined in Table 3.

A comparison of the metallicity distribution of all stars left in the sample and the conservative ($0.8-0.9M_{\odot}$) sample is shown in Fig. 10a. Fig. 10b shows the conservative sample compared with the $0.7-0.9M_{\odot}$ sample. There is little difference between the conservative sample and the $0.7 - 0.9M_{\odot}$ sample, and it is thus possible to extend the sample down to $0.7M_{\odot}$. As seen in Fig. 11a the sample is affected if the range is extended all the way down to the lowest mass, the main peak is lower and wider, which is not the expected result. The main peak shows no influence from the lack of high-metallicity low-mass stars predicted from the color limit. This can be explained in two ways: The first explanation is that the interpolation program forms a natural lower mass limit, e.g. if very few stars below a certain mass, say 0.7 $M_{\odot}$, are accepted because of the location of the tracks in the HR diagram. The second explanation is that the conservative sample is also affected. In that case the low metallicity tail would be over-represented in this study. This should be taken into consideration when evaluating the results from this study. The next step is to examine the upper limit. The first thing to notice here is that the low metallicity tails in Fig. 10b are comparable. This indicates that the upper mass limit is beyond $0.9M_{\odot}$. Fig. 11b shows the effect of extending the sample upward to $1.0M_{\odot}$. There is little change, although the low metallicity tail is slightly less prominent. This indicates that the upper mass limit can safely be set to $1.0M_{\odot}$.

In Fig. 12a the upper mass limit is removed. This has an effect on the low metallicity tail, which is nearly halved. The main bulge is also slightly broader, and it is thus not prudent to extend the sample to higher masses. This is obvious in Fig. 12b where only the stars above $1.0M_{\odot}$ is shown. This is caused by the short mainsequence lifetimes of high mass - low metallicity stars. In conclusion it is possible to use the interval from $0.7M_{\odot}$ to $1.0M_{\odot}$. The distribution of the stars in this interval is shown in Fig. 13. This distribution is also tabulated in Table 4. The weighted column indicates the normalized weight of the bin after scaleheight correction.

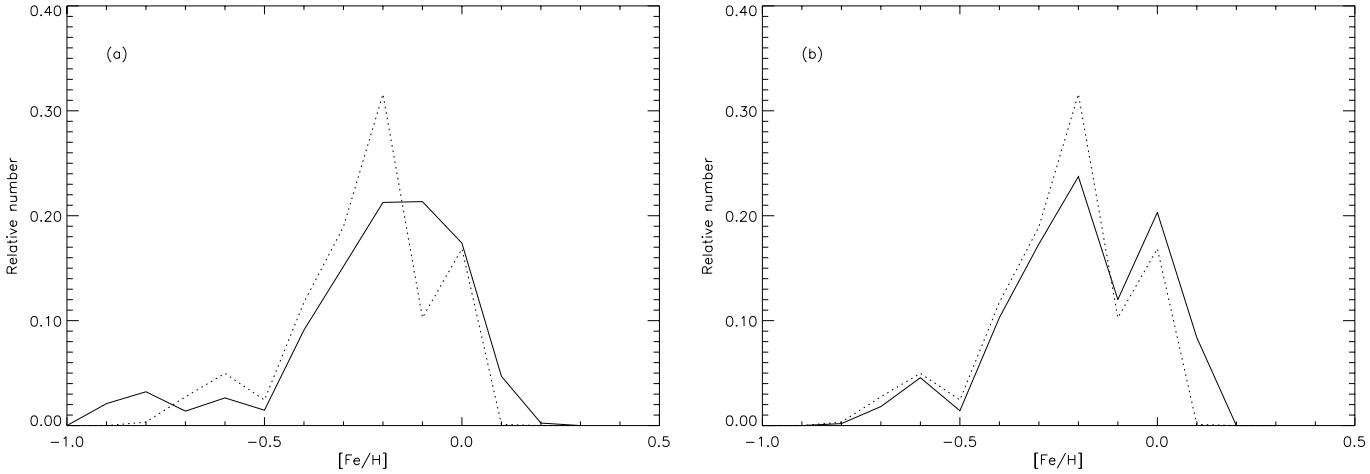

Fig. 11a and b. Comparison of (**a**) all stars below $0.9M_{\odot}$ (D, solid line and the conservative sample (B, dotted line) and (**b**) the $0.8 - 1.0M_{\odot}$ sample (E, solid line) with the conservative sample (B, dotted line).

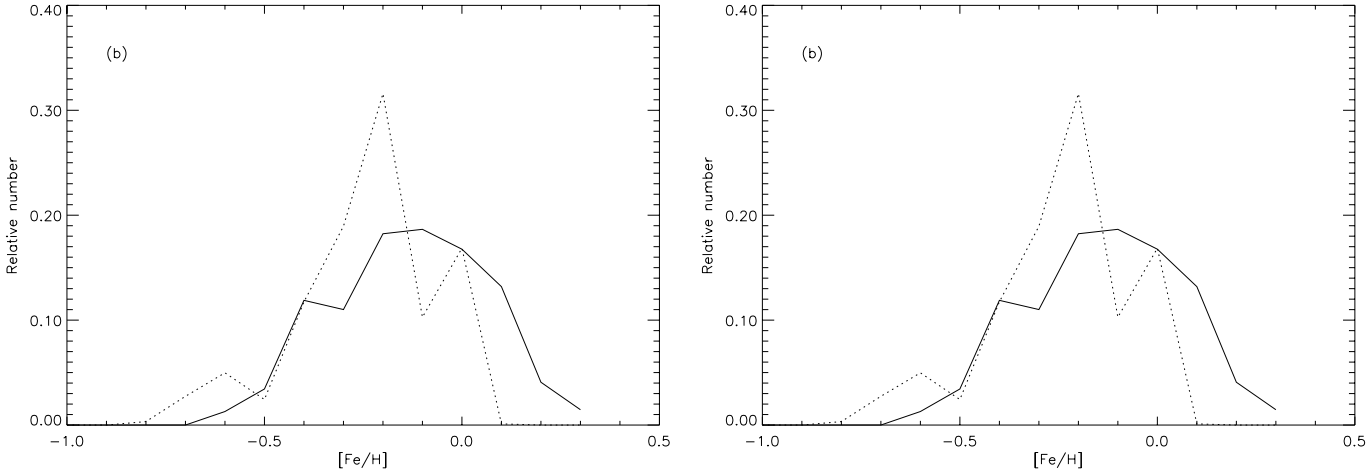

Fig. 12. Comparison of (**a**) all stars above $0.8M_{\odot}$ (F, solid line) and the conservative sample (B, dotted line) and (**b**) all stars above $1.0M_{\odot}$ (G, solid line) with the conservative sample (B, dotted line).

5.1. Scatter estimation

It is of great interest to have an estimate of the error associated with this distribution. Such an estimate is obtained in two steps. First a smoothed distribution is found, and then a Monte Carlo simulation of the error in [Fe/H] is made using this smoothed distribution. This ignores errors arising from the mass limits, the model uncertainies etc., but will give an estimate of the scatter that is present.

The observational scatter is taken to be 0.15 dex, although for stars above [Fe/H] =-0.6, 0.13 dex is more appropriate. The intrinsic scatter is taken to be the intrinsic scatter in the age-metallicity relation for stars at solar age and galactocentric distance. This is $\approx$ 0.15 dex from Edvardsson et al. (1993). Deconvolving the final distribution with a normal distribution with a standard deviation of $\sqrt{0.15^2 + 0.15^2} = 0.21$ dex, and using the optimum noise filter, results in a filtered distribution. From the filtered distribution a reconstructed sample is made by a convolution of the filtered distribution and a normal distribution ($\sigma = 0.21$).

The [Fe/H] of all stars in the sample was shifted with a normally-distributed random function with a standard deviation of 0.15 dex. The mass was shifted with a standard deviation of $0.1M_{\odot}$. Error bars were constructed so that 67% of all distributions fall inside them. The result of this simulation is shown in Fig. 14. The error bars in this figure do match the extended sample and the reconstructed distribution very well, with the only problems located near the "edges" of the main peak. It should also be noted that while there are large uncertainties in the distribution, the low metallicity tail can not be explained away, as the error bars are established based on the presence of stars in the low metallicity tail, and do not in any way touch on the possibility that they should not be in the sample e.g. mis-identified subdwarfs or RS CVn stars. It is however highly unlikely that this is the case, as the tail consists of several stars neatly distributed into the bins between [Fe/H]=-1.0 and [Fe/H]=-0.5. However, a more detailed observational investigation of these stars might be interesting.

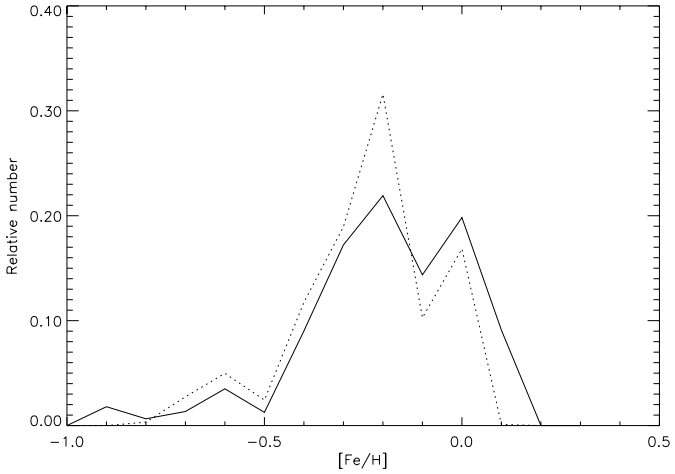

Fig. 13. Comparison of the extended sample (H, solid line) with the conservative sample (B, dotted line).

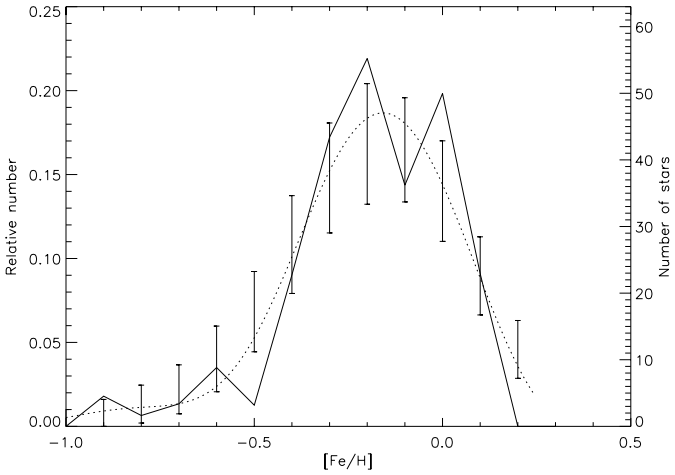

Fig. 14. Error bars from a Monte Carlo simulation based on the filtered distribution compared with the extended sample (H) and the reconstructed distribution (dotted line).

6. Comparisons

It is interesting to compare the distribution of the extended sample with distributions obtained by other authors. In Fig. 15 the distributions from Pagel (1989) and Rocha-Pinto & Maciel (1996) are compared with the distribution of the extended sample.

The Pagel (1989) distribution (containing 132 stars) does not compare well with the extended sample distribution. It is much broader, and the low metallicity tail begins at a much lower [Fe/H] (≈ -0.9). This suggests that the discrepancy might be caused by the different metallicity estimators (Strömgren δm_1 vs Johnson UV excess δ_{U-B}) and callibrations, i.e. [Fe/H] vs the [O/H] used in Pagel (1989). This makes it hard to compare these distributions.

The Rocha-Pinto & Maciel distribution displayed in Fig. 15 contains 287 stars, and is corrected with the Sommer-Larsen f method (mentioned in Sect. 4). It has the same pronounced peak at [Fe/H] ≈ -0.2 as the extended sample distribution, but the

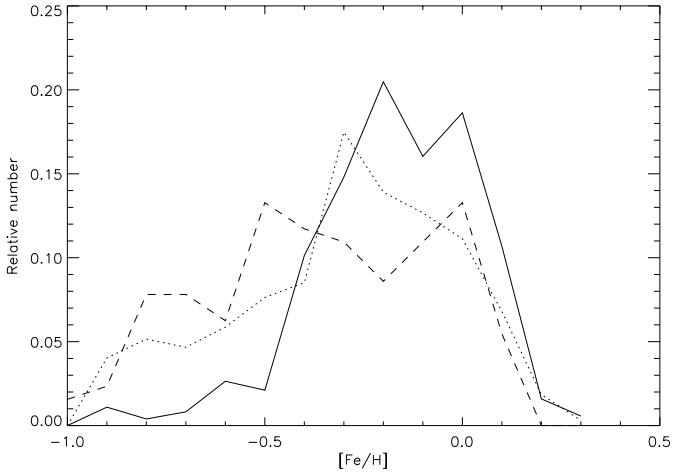

Fig. 15. Comparison with the Pagel (1989) (dashed) and the Rocha-Pinto & Maciel (1996) (dotted) distributions. The solid line is the extended sample (H).

tail is not as small. While these two distributions are af samples having comparable size (287 vs. 253 stars), the more thorough methods used in obtaining the extended sample implies that more weight should be placed on the extended sample.

6.1. Comparison with theoretical models

When comparing with theoretical models it is practical to use the recontructed sample, as that removes some effects from observational scatter. However, the observational and intrinsic scatter are still present. To compensate for this the theoretical models are convolved with a normal distribution with $\sigma = 0.20$. A variation of σ for the normal distribution has the greatest effect on steep distributions. A high σ will make any steep distribution considerably wider.

The extended sample was compared with the simple model (e.g. Pagel 1997), the Prompt Initial Enrichment model (Truran & Cameron 1971) and Larson's model (Larson 1972). None of these models fits the data in any acceptable way.

In Fig. 16 the extended sample is compared with Pagel & Tautvaišienė (1995). Their model fits reasonably well. The adopted parameters were: $u_1 = 0.14$, $u_0 = 1.3$ and $p_1 = 0.8Z_\odot$. Considering the scatter inherent in the distribution this is a good fit. While the convolution does degrade the fit, it is still not unreasonable.

The Lynden-Bell model is also a good fit (see Fig. 17). The Lynden-Bell model has $M_\infty = 10$ and a yield of $1.0Z_\odot$. This value of M_∞ is not in any way unreasonable (in Pagel 1989 a value twice this is used). The reconstructed sample lies more or less directly between the convolved and the unconvolved models. This indicates that if the intrinsic and observational scatter are less than anticipated an even better fit can be made. In any case, the infall models fits the distribution rather well, and are perfectly capable of explaining the G dwarf problem considering the scatter in the observed distribution. The only problem with these models is around [Fe/H] =-0.6. Here the models

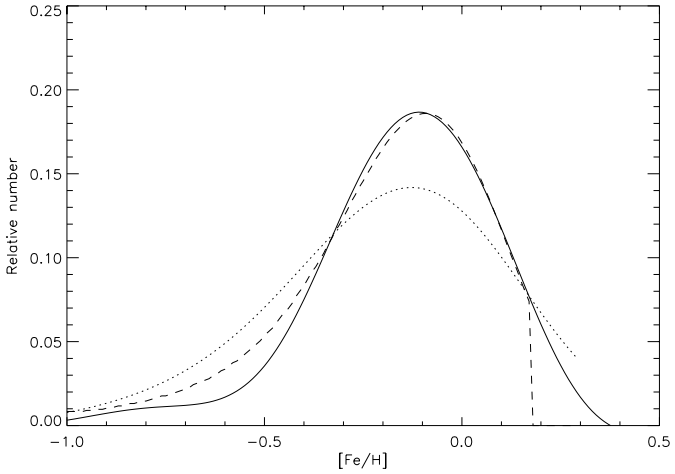

Fig. 16. Comparison with the Pagel & Tautvaišienė (1995) model. The solid line is the reconstructed sample. The dashed line is the Pagel & Tautvaišienė (1995) model, while the dotted line is including the convolution.

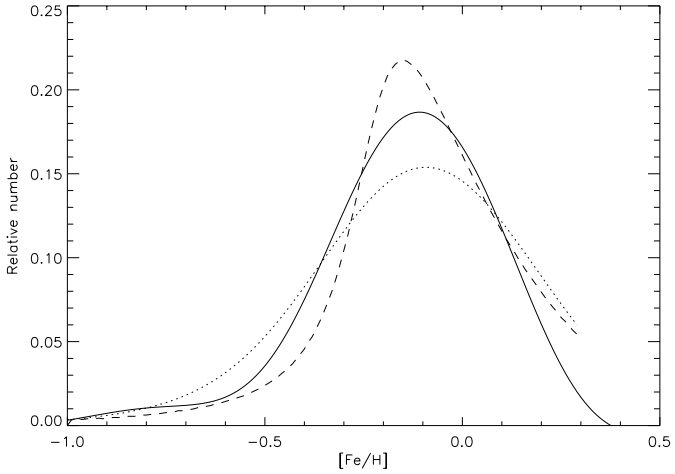

Fig. 17. Comparison with the Lynden-Bell (1975) model. The solid line is the reconstructed sample. The dashed line is the Lynden-Bell (1975) model with $M_\infty = 10$, while the dotted line is including the convolution.

are a little too broad, although this disappears if the scatter is overestimated. This indicates that models should perhaps try to relax the perfect mixing assumption, which could counter the intrinsic scatter making models more directly comparable to observational data.

7. Conclusion

A large sample of possible G dwarfs was examined. The sample was limited so as to be reasonably volume complete within 40 pc and south of $\delta = -26°$. Masses were calculated for the stars in the sample, using models of stellar evolution by VandenBerg et al. (2000). The masses calculated from the interpolation are in general accurate within 10% (as estimated in Sect. 3.1). Binary stars, subgiants and other pollutants were removed from the sample. The sample was examined for any influence from chromospheric activity, but the sample did not appear to be affected. A new $|W|$ velocity correction method was developed, using the space velocities of the individual stars. This method did not produce a result that was markedly different from other methods (e.g. Sommer-Larsen 1991). After some investigations of the consequences, the sample was limited to stars with masses between 0.7 and 1.0 $M_\odot$. Even after limiting the sample in these ways it is still significantly larger than most published until now.

The strict procedures undertaken to make sure that the sample used was free of any effects that could affect the resulting distribution (e.g. by using a mass interval and not a color interval, evolution effects were largely avoided; by removing multiple stars, their faulty photometric metallicities were not used etc.) makes it an even stronger result. It is the belief of the author, that even though these procedures have removed more than 75% of the stars from the raw sample, the end result is probably a more correct representation of the G dwarf metallicity distribution of the solar cylinder than any previously published.

When compared with prior observational work, this sample has a much smaller low metallicity tail. In Sect. 2 it is mentioned that the low metallicity tail might be over-represented in this study, which strenghens this point. It makes any reconciliation with the simple model impossible. The derived distribution can be fitted with infall models, such as the models by Pagel & Tautvaišienė (1995) and Lynden-Bell (1975). The careful selection procedure used in this work makes it even more unlikely that the G dwarf problem can be explained by selection effects.

That infall models are capable of explaining the G dwarf problem (perhaps with some delayed recycling and a relaxation of perfect mixing assumption) is a powerful result. Although it can not be ruled out that other effects are important, for now it appears to be sufficient to use an infall model to resolve the G dwarf problem. Thus the G dwarf metallicity distribution presented here is a solid constraint on numerical models taking into account the interaction of the different processes in the Galaxy throughout the lifetime of the Galaxy. This will hopefully one day lead to a better understanding of the formation and evolution of the Milky Way Galaxy.

Acknowledgements. The author would like to thank Erik Heyn Olsen for providing the data that this work is based on, and for general help and guidance. The author would also like to thank Johannes Andersen and Birgitta Nordström for making their velocity data available to me. The SIMBAD database available at CDS was used in preparing this article.

References

Clayton D.D., 1985, In: Arnett W.D., Truran J.W. (eds.) Nucleosynthesis, Challenges and New Developments. Chicago University Press
Crawford D.L., 1975, AJ 80, 955
Edvardsson B., Andersen J., Gustafsson B., et al., 1993, A&A 275, 101
ESA, 1997, The Hipparcos and Tycho Catalogues, ESA SP-1200
Fux R., 1997, Ph.D. Thesis
Giampapa M.S., Worden S.P., Gilliam L.B., 1979, ApJ 229, 1143
Henry T.J., Soderblom D.R., 1996, AJ 111, 439
Houk N., 1978, Michigan Catalogue of Two-Dimensional Spectral Types for the HD stars. Vol. 2, University of Michigan, Ann Arbor, Michigan

Houk N., 1982, Michigan Catalogue of Two-Dimensional Spectral Types for the HD stars. Vol. 3, University of Michigan, Ann Arbor, Michigan
Houk N., Cowley A.P., 1975, University of Michigan Catalogue of Two-Dimensional Spectral Types for the HD stars. Vol. 1, University of Michigan, Ann Arbor, Michigan
Larson R.B., 1972, Nat. Physical Science 236, 7
Lynden-Bell D., 1975, Vistas in Astronomy 19, 299
Magain P., 1987, A&A 181, 323
Mayor M., Turon Lacarrieu C., Martinet L., 1977, A&A 61, 433
Morale F., Micela G., Favata F., Sciortino S., 1996, A&AS 119, 403
Nissen P.E., 1981, A&A 97, 145
Nissen P.E., 1994, Rev. Mex. Astron. Astrofis. 29, 129
Olsen E.H., 1983, A&AS 54, 55
Olsen E.H., 1984, A&AS 57, 443
Olsen E.H., 1988, A&A 189, 173
Olsen E.H., 1993, A&AS 102, 89
Olsen E.H., 1994, A&AS 104, 429
Pagel B.E.J., 1989, In: Beckman J.E., Pagel B.E.J. (eds.) Evolutionary Phenomena in Galaxies. Cambridge University Press
Pagel B.E.J., 1997, Nucleosynthesis and Chemical Evolution of Galaxies. Cambridge University Press
Pagel B.E.J., Tautvaišienė G., 1995, MNRAS 276, 505
Rocha-Pinto H.J., Maciel W.J., 1996, MNRAS 279, 447
Schmidt M., 1963, ApJ 137, 758
Schuster W.J., Nissen P.E., 1989, A&A 221, 65
Sommer-Larsen J., 1991, MNRAS 249,368
Truran J.W., Cameron A.G.W., 1971, Ap&SS 14, 179
VandenBerg D.A., 1985, ApJS 58, 711
VandenBerg D.A., Swenson F.J., Rogers, F.J., Iglesias C.A., Alexander D.R., 2000, ApJ 532, 430
van den Bergh S., 1962, AJ 67, 486
Worley C.E., Douglass G.G., 1996, A&AS 125, 523