Astron. Astrophys. 363, 223–227 (2000)

ASTRONOMY AND ASTROPHYSICS

On the problem of confirming ZAMS stars in the field

R. Wichmann

Hamburger Sternwarte, University of Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany (rwichmann@hs.uni-hamburg.de)

Received 3 April 2000 / Accepted 8 August 2000

Abstract. The problem of confirming the youth of presumed zero-age main sequence (ZAMS) stars among field stars, i.e. outside of clusters and associations, is discussed within a statistical framework. A generic method to evaluate the completeness and contamination of samples is presented, criteria for the design of appropriate tests are outlined, and the performance of X-ray selection and the so-called 'lithium test' is estimated. It is shown that for any test applied to this problem, the error of the second kind is far more important than the error of the first kind.

Key words: stars: pre-main sequence – X-rays: stars – methods: data analysis

1. Introduction

The current star forming rate (SFR) in the solar neighbourhood is estimated to about 3.5–5 $M_\odot$ pc^{-2} Gy^{-1} (see Rana 1991 and references therein). Therefore we should expect several thousand stars with ages up to the Pleiades age within hundred parsec around the Sun. Presumably only a few per cent of these are born in open clusters (Wielen 1971), while the majority might be born in less tightly bound systems like T-Tauri and OB associations. However, even open clusters typically dissolve into the field on timescales of 30 My to few 100 My (c.f. Wielen 1971, van den Bergh 1981), i.e. their stars will not remain 'locked up' for a long time.

Therefore one has to assume that a small fraction of field stars are young stars on the zero-age main sequence (ZAMS), which raises the question whether they can be found, and how reliable – in terms of completeness and contamination – the employed method(s) work.

The age of open clusters is usually determined by main-sequence fitting, a method that works well for a large sample of coeval stars, but cannot be used on individual objects.

For isolated stars, one has to fall back on correlations between the age of a star and other properties like e.g. X-ray luminosity, lithium surface abundance, or kinematics. Some of these correlations, e.g. kinematics, are statistical only, while others, e.g. X-ray luminosity or Li abundance, are deterministic, but show a large spread in coeval samples, thus rendering their application to individual objects problematic.

With the availability of large X-ray surveys, in particular the ROSAT All-Sky Survey (RASS), it has become possible to define candidate samples of X-ray active, hence presumably young stars. The first surveys were restricted to the surroundings of nearby star forming regions (c.f. Wichmann et al. 1996; Alcalá et al. 1995; Alcalá et al. 1996; Krautter et al. 1997; Wichmann et al. 1997; Magazzu et al. 1997). However, recent work by Guillout et al. (1998a, 1998b) has demonstrated that the Gould Belt star forming complex shows up prominently in the RASS, and one also might expect a large population of somewhat older, less X-ray active and more dispersed, young stars (Briceño et al. 1997).

The confirmation of the youth of the objects found in the above surveys has been problematic, and uncertainties remain with respect to the trustworthyness of the method(s) employed, and the constrains that can be put on the stellar ages. I will try to resolve some of these uncertainties in this paper.

2. Interpreting the evidence

Whenever some observation O yields new evidence for the truth of a hypothesis H, the probability $P_{\rm H}$ that H is true can be updated as

$$P_{\rm H,new} = \frac{P_{\rm H,prior} P_{\rm O|H}}{P_{\rm H,prior} P_{\rm O|H} + (1 - P_{\rm H,prior}) P_{\rm O|nonH}}. \quad (1)$$

Here, $P_{\rm H,prior}$ is the *prior* probability that H is true. $P_{\rm O|H}$ is the *true-positive* probability of observing O if H is true (the *sensitivity* of O), while $P_{\rm O|nonH}$ is the *false-positive* probability of observing O if H is false ($1 - P_{\rm O|nonH}$ is called the *specificity* of O).

Alternatively, $(1 - P_{\rm O|H})$ is often called the *error of the first kind*, and $P_{\rm O|nonH}$ the *error of the second kind*. $P_{\rm H,prior}$ is the *prior* probability that H is true. In Fig. 1, Eq. (1) is visualized as a flow chart.

Eq. (1) follows from the sum and product rules for probabilities and is known as *Bayes theorem* (Bayes 1763, 1764). Note that there is substantial controversy on the application of Bayes theorem in cases where $P_{\rm H,prior}$ is unknown. This controversy is irrelevant here, because only cases with known $P_{\rm H,prior}$ will be discussed.

The following points should be noted: (i) If $P_{\rm O|H} \leq P_{\rm O|nonH}$ then $P_{\rm H,new} \leq P_{\rm H,prior}$. In other words, the trust in the truth of H can only be enhanced if both $P_{\rm O|H}$ and $P_{\rm O|nonH}$ are known, and $P_{\rm O|H} > P_{\rm O|nonH}$ can be proved.

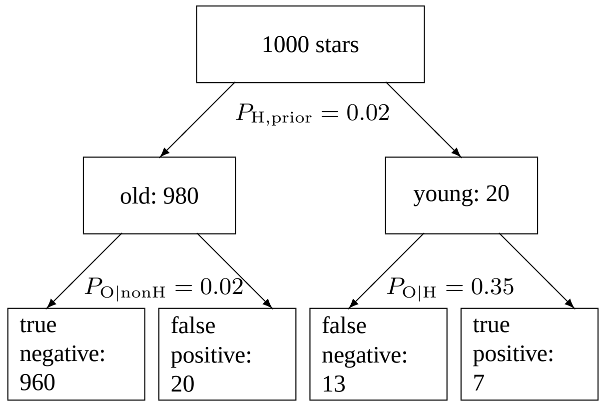

Fig. 1. Flow chart visualization of Bayes theorem. For the age determination of field stars. $P_{\rm H,prior}$ is estimated to 0.02 (see Sect. 3.1), and $P_{\rm O|nonH}$ and $P_{\rm O|H}$ are set here to 0.02 and 0.35, respectively (see Sect. 3.3). In this example, $20 + 7 = 27$ stars would be accepted as young, of which 20 are falsely accepted old stars, corresponding to $P_{\rm H,new} = 0.26$. This example represents the more pessimistic one of the two cases for Li selection summarized in Sect. 3.3.3.

(ii) Usually, one can trade false positives for true positives (i.e. improving the sensitivity $P_{\rm O|H}$ at the cost of lowering the specificity $(1 - P_{\rm O|nonH})$. This raises the problem of finding the optimum balance between sensitivity and specificity.

(iii) This optimum balance obviously depends on the prior probability $P_{\rm H,prior}$, if the fraction of false positives is of interest.

3. Age determination of field stars

In order to make use of Eq. (1), one must first specify the hypothesis H, and evaluate $P_{\rm H,prior}$. In a one-sided age determination, $\rm H = (Age_{star} < Age_{limit})$.

3.1. The prior probability

$P_{\rm H,prior}$ is the expected fraction of stars with ages below the limit, and can be evaluated from the stellar mass density and the current star forming rate. For a limit of 100 My, a star forming rate of 4 $\rm M_\odot\ pc^{-2}\ Gy^{-1}$ and a stellar density of 25 $\rm M_\odot\ pc^{-2}$ (see Rana 1991 and references therein), one may estimate $P_{\rm H,prior} \simeq 0.02$ (assuming an average mass of 0.8 $\rm M_\odot$ and equal scale heights for stars of all ages).

3.2. Sensitivity versus specificity

Fig. 1 demonstrates that because of the very low value of $P_{\rm H,prior}$, the specificity of any applied test is the crucial issue. It probably will not hurt if e.g. 20 per cent of young stars are missed, but it certainly will render the result almost worthless, if 20 per cent of the old stars were included as false positives. It is therefore very important to assess the specificity of any test used to confirm the youth of a field star, and to aim for a high specificity rather than for high sensitivity, or equivalently to aim for a small error of the second kind rather than the first kind.

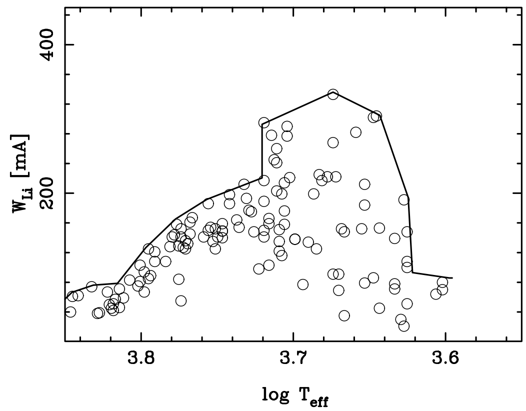

Fig. 2. Scatter plot of Li I $\lambda 6708$ equivalent widths of Pleiades stars vs. $\rm T_{eff}$, and their upper envelope. Data are from Soderblom et al. (1993).

3.3. The lithium test

This test is based on the fact that at the bottom of the convective zone of low-mass stars lithium may be burned by nuclear reactions, and thus will be depleted. Li depletion, as measured by the strength of the Li I $\lambda 6708$ doublet line, depends not only on age and stellar mass, but also on additional and poorly understood parameters that may cause a significant spread for coeval stars of equal masses (c.f. Soderblom et al. 1993).

To circumvent this complication, the test is carried out by plotting a scatter plot of Li I equivalent widths EW(Li) vs. effective temperatures for an appropriate open cluster used as benchmark (most often the Pleiades), defining the upper envelope for EW(Li) (see Fig. 2), and verifying whether a star falls above or below this upper envelope. In this form, it has probably first been used by Maggazu et al. (1997), although a similar plot appears in Briceño et al. (1997). Usually EW(Li) rather than the abundances N(Li) are used to avoid introducing additional uncertainties from model calculations.

To study the specificity and sensitivity of the test, it is necessary to compare distribution functions of the *EW(Li)-offsets* with respect to the upper envelope of the benchmark cluster for several different samples of different ages. Fig. 2 suggest that the pattern of Li depletion may be different for hot stars and cool stars, as the observed spread appears to be much smaller for the hot stars. The 'break' temperature appears to be about 5300 K, but may vary somewhat between clusters (see Jones et al. 1999, Randich et al. 1997 for a comparison of different clusters). For the present discussion, a value of 5300 K is adopted, and distribution functions for hot and cool stars are computed seperately.

Cumulative distribution functions (c.d.f) for the EW(Li)-offsets of open clusters, with respect to the Pleiades upper envelope, were calculated for M67 (4.5 Gy, data from Jones et al. 1999), the Hyades (600 My, data from Thorburn et al. 1993), the Pleiades (100 My, data from Soderblom et al. 1993), and IC 2602 (30 My, data from Randich et al. 1997). In addition, the c.d.f for T Tauri stars was determined from data given by Magazzu et al. (1992), Basri et al. (1991), and Martin et al. (1994). The resulting c.d.f are shown in Figs. 3 and 4. In computing

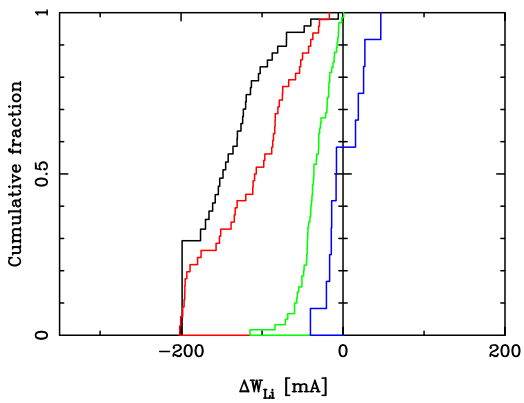

Fig. 3. Cumulative distribution functions of EW(Li)-offsets with respect to the Pleiades upper envelope, for stars hotter than 5300 K. From left to right: M67, Hyades, Pleiades, IC 2602.

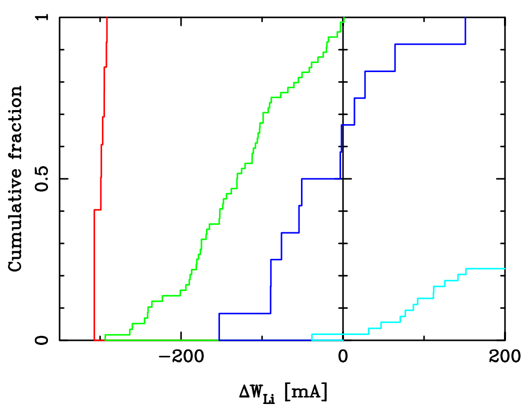

Fig. 4. Cumulative distribution functions of EW(Li)-offsets with respect to the Pleiades upper envelope, for stars cooler than 5300 K. From left to right: Hyades, Pleiades, IC 2602, T Tauri stars. The c.d.f of the T Tauri stars extends up to 850 mÅ. There are no data for cool stars in M67.

these c.d.f, upper limits have been taken into account, i.e. they represent the Kaplan-Meier estimator of the c.d.f (the Kaplan-Meier estimator is the maximum-likelihood estimator for the sample c.d.f in the presence of censored data).

3.3.1. Specificity

To evaluate the specificity of the lithium test, one must consider the fraction of objects that would falsely be accepted as young, i.e. older stars above the Pleiades upper envelope.

From Figs. 3 and 4 one can see that stars above the Pleiades upper envelope cannot be found in any older cluster. Therefore, the specificity of the Li test is very high. However, due to the limited sample size, the observed value of $P_{\rm O|nonH} = 0.0$ has some formal error, which may be evaluated from the binominal probability to observe zero events if $P_{\rm O|nonH} = x$, with $x > 0.0$.

For hot stars, a 1 σ confidence limit of $P_{\rm O|nonH} < 0.02$ can be given. (There are 48 Hyades data points, and the binominal probability of observing 0 events in 48 trials is 0.32 ($= 1\,\sigma$) for a success probability of 0.02).

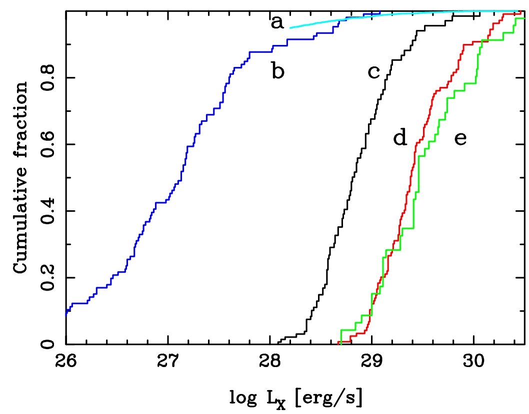

Fig. 5. Cumulative distribution functions of X-ray luminosity $\log \mathrm{L}_X$ (erg s^{-1}). for stars of spectral type G and cooler. From left to right: (a) Field stars within 25 pc (only the high-luminosity tail where the RASS is comple is shown), (b) field stars within 7 pc (only K and M stars), (c) Hyades, (d) Pleiades, (e) IC 2602.

For cool stars, there are only 15 Hyades data points, yielding a 1 σ confidence limit of $P_{\rm O|nonH} < 0.07$.

3.3.2. Sensitivity

The sensitivity of the lithium test is given by the fraction of younger stars accepted by the lithium test. By definition, for Pleiades stars the sensitivity is 0.0. For stars with an age of about 30 My, about 58 (cool) to 67 (hot) per cent of the stars would be falsely rejected.

On the other hand, for cool T Tauri stars the sensitivity is very high (0.98 ± 0.02).

3.3.3. Result

If we approximate the sensitivity by a piecewise linear function, increasing from 0.0 to 0.4 between 100 and 30 Myr, and from 0.4 to 1.0 between 30 and 0 Myr, we obtain a total sensitivity $P_{\rm O|H} \simeq 0.35$. Using a value of $P_{\rm O|nonH} = 0.02$ then yields $P_{\rm H,new} = 0.26$, while $P_{\rm O|nonH} = 0.0$ yields $P_{\rm H,new} = 1.0$. Clearly, the small uncertainty on $P_{\rm O|nonH}$ makes a big difference.

3.4. X-ray selection

As the candidate samples for surveys for young stars are often X-ray selected, it is worthwile to study what effect an X-ray pre-selection has.

Fig. 5 shows the X-ray c.d.f for field stars within 25 pc (only high-luminosity tail down to the completeness limit of the RASS, data from Hünsch et al. 1999), field stars of spectral type K and M within 7 pc, (complete, data from Schmitt et al. 1995), Hyades (data from Stern et al. 1995), Pleiades (data from Stauffer et al. 1994), and IC 2602 (data from Randich et al. 1995). As for Li c.d.f, upper limits have been taken into account.

From Fig. 5 one can infer that for a threshold of (e.g.) $\log \mathrm{L}_X = 29.0$, we have a sensitivity $P_{\rm O|H} \simeq 0.9$, which may be

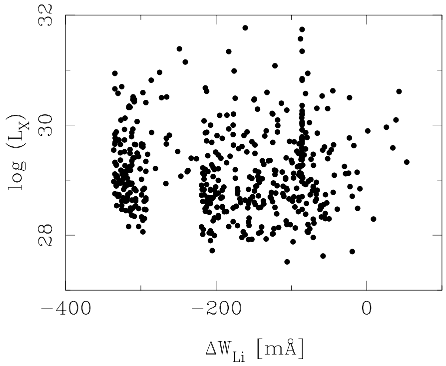

Fig. 6. Distribution of EW(Li)-offsets vs X-ray luminosity $\log \mathrm{L}_X$ (erg s^{-1}) for field stars.

constant for stars younger than the Pleiades. $P_{\mathrm{O|nonH}}$ decreases from 0.9 for Pleiades-age stars to 0.3 at the age of the Hyades. For nearby field stars, $P_{\mathrm{O|nonH}} = 0.02$; the ages of these stars are unknown, but we may assume an average solar-like age of about 4.5 Gyr.

3.4.1. Result

With the numbers mentioned above, we can estimate $P_{\mathrm{O|H}} \simeq 0.9$ and $P_{\mathrm{O|nonH}} \simeq 0.18$ for X-ray selection of stars younger than the Pleiades and a selection threshold of $\log \mathrm{L}_X = 29.0$. The value of $P_{\mathrm{O|nonH}} \simeq 0.18$ (from piecewise linear interpolation between Pleiades–Hyades–4.5 Gyr) is a conservative estimate, as the field stars represent a mix of ages. Because L_X is correlated with age, we may assume that only the youngest field stars contribute to the important high-luminosity tail. However, in the absence of data on clusters older than the Hyades, it is problematic to speculate about the behaviour of the c.d.f of L_X with age.

This probabilities estimated above would result in $P_{\mathrm{H,new}} = 0.09$, i.e. a fourfold increase in the fraction of young stars, as compared with an unselected sample. The problem with this approach is that it requires some knowledge of the distance of the stars. However, candidates obtained from cross-correlating the RASS and the Tycho catalogue are appropriate for such an analysis.

3.5. Lithium and X-ray combined

Combining two methods is a straightforward procedure. First, $P_{\mathrm{O|H}}$ and $P_{\mathrm{O|nonH}}$ for X-ray selection, as well as Eq. (1), are used to update the prior probability from stellar densities (see Sect. 3.1). Then the new probability $P_{\mathrm{H,new}}$ thus obtained will become the prior probability for any subsequent test (e.g. the lithium test).

Of course this approach is only valid if $\log \mathrm{L}_X$ is uncorrelated with the subsequently used test criterium. Fig. 6, which is based on preliminary data from a survey of field stars (Wichmann & Schmitt 2000), indicates that $\log \mathrm{L}_X$ is not correlated with EW(Li)-offsets.

Using the results from the discussion above, we now have $P_{\mathrm{H,prior}} = 0.09$ (from the X-ray selection) as prior probability for the lithium test. Depending on the $P_{\mathrm{O|nonH}}$ for the lithium test, we now obtain $P_{\mathrm{H,new}} = 0.6...1.0$.

3.6. Measurement errors

So far, the effect of measurement errors has not been discussed. The observed distribution functions for samples of known age already include measurement errors. As long as these errors are random, they will just broaden the observed distributions, thus making the test less conclusive.

Measurement errors in the candidate stars will modify the probabilities $P_{\mathrm{O|H}}$ and $P_{\mathrm{O|nonH}}$, i.e. these do not depend only on the distribution functions of the comparison samples anymore.

E.g. if O ist the observation that a star is above some threshold T, measurement errors will scatter some stars below T, while other stars will be scattered above T. Thus, in Eq. (1), $P_{\mathrm{O|H}}$ may be replaced by

$$P^E_{\mathrm{O|H}} = P_{\mathrm{O|H}} P^{\mathrm{in}} + P_{\mathrm{nonO|H}} P^{\mathrm{out}}, \tag{2}$$

where P^{in} is the probability that star truly is above T, while P^{out} is the probability that the star is below T. P^{in} and $P^{\mathrm{out}} = (1 - P^{\mathrm{in}})$ can be calculated from the probability density function of the measurement error. Similarly, $P_{\mathrm{nonO|H}}$ should be replaced by $P^E_{\mathrm{nonO|H}}$.

If the errors are much larger than the effect to be studied, evidently $P^E_{\mathrm{O|H}} \simeq P^E_{\mathrm{nonO|H}}$, and therefore $P_{\mathrm{H,new}} \simeq P_{\mathrm{H,prior}}$.

4. Conclusions

It has been demonstrated that even false-positive probabilities as low as a few per cent will lead to severe contamination with old stars, when one tries to select young (ZAMS) stars in the field. Therefore, in the design of observational tests, one must aim to maximize the specificity of the selection criterium, even if this results in severely incomplete samples. Such incompleteness might be problematic in further studies of the samples thus obtained, because it is not random, and thus represents a selection bias.

For discriminating stars above and below Pleiades age in the field, the available data indicate that the lithium test can reach the required sensitivity, at the cost of an incompleteness on the order of 60 per cent for 30 My old stars. Pre-selection by X-ray luminosity is useful, especially in view of the fact that the lithium test is rather ‘expensive’, as it requires high-quality spectra. The main problem appears to be the uncertainty on the specificity of the lithium test that stems from the count statistic, because even a small, non-zero $P_{\mathrm{O|nonH}}$ has a major impact on the result.

Both tests considered, and in fact all threshold-based tests, have the property that $P_{\mathrm{O|H}} = P_{\mathrm{O|nonH}}$ at the age of the benchmark cluster. As discussed in Sect. 2, this renders the test useless unless $P_{\mathrm{O|nonH}}$ decreases above this age and/or $P_{\mathrm{O|H}}$ increases below this age. Thus the usefulness of the test is really determined by the behaviour of both in the neighbourhood of the limiting age.

Ideally, one would like to have a very steep increase in $P_{\mathrm{O|H}}$ and a steep decrease in $P_{\mathrm{O|nonH}}$, corresponding to a sharp transition of the selection property across the limiting age. While the available data are somewhat sparse, it seems that by combining lithium abundance and X-ray activity, this requirement can be fulfilled satisfactorily, and field ZAMS stars can be found in an economic way.

In the present discussion, we have assumed a constant and homogeneous SFR rate. Estimates of the current SFR rate are usually based on the present-day mass function and the age distribution of stars, where the former is only complete in the immediate solar neighbourhood, and the latter usually binned in large bins of 10^9 yr typically (see Rana 1991 and references therein).

On the other hand, the Sun is located within a relatively young and large star forming complex – the Gould Belt. The Gould Belt is old enough (~ 70 Myr) for many of its stars to have dispersed into the field already, while it is far enough (the nearby edge is at ~ 150 pc, the far edge at ~ 800 pc) to be missed in the low-mass star counts, and the large age bins in SFR studies will further smooth out any trace of the Gould Belt.

However, the low-mass Gould Belt stars represent a significant enhancement in the number of young dispersed field stars, with respect to the expectations from the current SFR, as evidenced by the results of Guillout et al. (1998a, 1998a). In the regions of the sky covered by the Gould Belt, the age distribution of stars will be skewed towards ages less than the Gould Belt age. While it is difficult to put firm numbers on the effect, the selection relieability for young stars as discussed in this work obviously will be improved.

Acknowledgements. It is a great pleasure to thank J.H.M.M. Schmitt for critical reading of the manuscript. This project has been supported by grants from the Deutsche Forschungsgemeinschaft (DFG Schwerpunktprogramm 'Physics of star formation').

References

Alcalá J.M., Krautter J., Schmitt J.H.M.M., et al., 1995, A&AS 114, 109
Alcalá J.M., Terranegra L., Wichmann R., et al., 1996, A&AS 119, 7
Basri G., Martin E.L., Bertout C., 1991, A&A 252, 625
Bayes T., 1763, Philosophical Transactions 53, 376 Reprinted in Biometrica 45 (1958), 296
Bayes T., 1764, Philosophical Transactions 54, 298 Reprinted in Biometrica 45 (1958), 296
Briceño C., Hartmann L.W., Stauffer J.R., Gagne M., Stern R.A., 1997, AJ 113, 740
Guillout P., Sterzik, M. F., Schmitt J.H.M.M., et al., 1998a, A&A 334, 540
Guillout P., Sterzik, M. F., Schmitt J.H.M.M., Motch C., Neuhäuser R., 1998b, A&A 337, 113
Hünsch M., Schmitt J.H.M.M., Sterzik M.F., Voges W., 1999, A&AS 135, 319
Jones B.F., Fischer D., Soderblom D.R., 1999, AJ 117, 330
Krautter J., Wichmann R., Schmitt J.H.M.M., et al., 1997, A&AS 123, 329
Magazzu A., Rebolo R., Pavlenko Y.V., 1992, ApJ 392, 159
Magazzu A., Martin E.L., Sterzik M.F., et al., 1997, A&AS 124, 449
Martin E.L., Rebolo R., Magazzu A., Pavlenko Y.V., 1994, A&A 282, 503
Rana N.C., 1991, ARA&A 29, 129
Randich S., Schmitt J.H.M.M., Prosser C.F., Stauffer J.R., 1995, A&A 300, 134
Randich S., Aharpour N., Pallavicini R., Prosser C.F., Stauffer J.R., 1997, A&A 323, 86
Schmitt J.H.M.M., Fleming T.A., Giampapa M.S., 1995, ApJ 450, 392
Soderblom D.R., Jones B.F., Balachandran S., et al., 1993, AJ 106, 1059
Stauffer J.R., Caillault J.-P., Gagne M., Prosser C.F., Hartmann L.W., 1994, ApJS 91, 625
Stern R.A., Schmitt J.H.M.M., Kahabka P.T., 1995, ApJ 448, 683
Thorburn J.A., Hobbs L.M., Deliyannis C.P., Pinsonneault M.H., 1993, ApJ 415, 150
van den Bergh S., 1981, PASP 93, 712
Wichmann R., Krautter J., Schmitt J.H.M.M., et al., 1996, A&A 312, 439
Wichmann R., Sterzik M., Krautter J., Metanomski A., Voges W., 1997, A&A 326, 211
Wichmann R., Schmitt J.H.M.M., 2000, In: García López R.J., Rebolo R., Zapatero Oserio M.R. (eds.) Proceedings of the 11th Cool Star Workshop. ASP Conf. Series, in press
Wielen R., 1971, A&A 13, 309