Mon. Not. R. Astron. Soc. **365,** 590–594 (2006) doi:10.1111/j.1365-2966.2005.09739.x

On the upper limit on stellar masses in the Large Magellanic Cloud cluster R136

Chris Koen★
Department of Statistics, University of the Western Cape, Private Bag X17, Bellville, 7535 Cape, South Africa

Accepted 2005 October 11. Received 2005 October 7; in original form 2005 May 26

ABSTRACT

A truncated power-law distribution is fitted to the 29 largest stellar masses known in R136. Two different statistical techniques are used, with comparable results. An upper limit to the mass distribution of the order of 140–160 $M_{\odot}$ is derived, while the power-law exponent is in the approximate range 0.9–1.7. A power-law distribution with no upper limit on the mass can be rejected with considerable confidence. It is recommended that the calculations be repeated when more reliable mass estimates are available.

Key words: methods: statistical – stars: early-type – stars: luminosity function, mass function – Magellanic Clouds – galaxies: stellar content.

1 INTRODUCTION

A power-law form for the 'original mass function', or 'initial mass function' (IMF), was introduced by Salpeter (1955). The IMF of, for example, a star cluster can be thought of as the mass distribution of stars on its zero-age main sequence (ZAMS). The Salpeter IMF has the specific form

$$f(m) = \beta m^{-(a+1)}, \qquad (1)$$

where β is a constant and the exponent $a \approx 1.3$. The usual interpretation of (1) is that it specifies the number of stars with masses in a small interval centred on m. Equivalently, it may be seen as a probability density function (pdf), specifying the probability that the mass of a randomly selected star lies in a small interval around m. In this paper the second interpretation will implicitly be assumed, but it is stressed that it is equivalent to the first, provided that (1) is seen as a statistical, rather than a deterministic, relation.

There are good reasons for the continued interest in the IMF, such as the potential information about star formation – see e.g. the review by Kroupa (2002). It is remarkable that, at least for stars with $M \gtrsim 1\,M_{\odot}$, quite similar values of the exponent a are found in many physical environments, including the Large Magellanic Cloud (LMC) (Kroupa 2001). At lower masses there is some evidence for changes in the form of (1) with decreasing mass (Kroupa 2001).

The IMF for massive stars ($M > 15\,M_{\odot}$) was reviewed by Massey (1998). He concluded that (1), with $a \approx 1.3$, applies to clusters and associations in the Milky Way, LMC and Small Magellanic Cloud (SMC), despite substantial variations in metallicities and star densities. The paucity of stars with masses above about 120 $M_{\odot}$ he ascribed to the very low formation probability of such objects, as predicted by (1). This contrasts with the possibility that there is a stellar upper mass limit, i.e. that very massive stars simply do not exist. The present paper addresses the two issues of a possible upper mass limit, and the behaviour of the exponent a in (1) as m increases. For these purposes, use is made of determinations of masses in the LMC cluster R136.

R136 is a good source of data for investigating the behaviour of the IMF at high masses, since it is richer in O stars than any other known cluster. Furthermore, the massive stars are young (ages of the order of a Myr – Massey 1998), which means that the highest mass stars have not yet expired.

Massey & Hunter (1998) obtained *Hubble Space Telescope* (*HST*) spectra of a number of very hot, luminous stars in R136. Effective temperatures were then deduced from spectral classifications, using calibrations by Vacca, Garmany & Shull (1996) and Chlebowski & Garmany (1991) respectively; these will be referred to as VGS96 and CG91 in the rest of the paper. Bolometric corrections followed from the effective temperatures, and after de-reddening photometric magnitudes, absolute bolometric magnitudes could be calculated. The masses of the stars were then found by comparison of the derived ($T_{\rm eff}$, $M_{\rm bol}$) values to evolutionary tracks taken from Schaerer et al. (1993). Luminosities of the brightest stars exceeded the upper limits for which evolutionary tracks were available, and their masses therefore had to be estimated by extrapolation.

A table listing the 29 highest masses in their sample, estimated using both VGS96 and CG91 effective temperature calibrations, was presented by Massey & Hunter (1998). The CG91 temperature scale leads to lower effective temperatures, hence to lower luminosities, and therefore smaller mass estimates. Both sets of masses will be used below, since a comparison of the results gives further insight into the level of uncertainty in the analysis. The authors concluded that the distribution of high stellar masses in R136 was consistent with the known distribution at lower masses, i.e. $a \approx 1.3$–1.4, and that there was no evidence for an upper mass limit.

By contrast, Weidner & Kroupa (2004) and Oey & Clarke (2005) deduced from the same data that there is an upper limit to the

★E-mail: ckoen@uwc.ac.za

stellar mass distribution in R136. Weidner & Kroupa (2004) essentially assumed that (1) is a deterministic relation, from which the precise numbers of stars in a given mass interval can be derived once the cluster mass is specified. For a cluster of the size of R136, and with $a = 1.25$, the authors found that stars with masses in excess of 750 $M_\odot$ should exist, inconsistent with the highest observed mass $\sim 150\,M_\odot$. Weidner & Kroupa (2004) considered various modifications of (1) and concluded that the most likely is that there is a fundamental upper stellar mass limit of about about 150 $M_\odot$.

The treatment in Oey & Clarke (2005), on the other hand, was probabilistic rather than deterministic: given that some differences in the mass distributions of different clusters of the same age and metallicity are to be expected, this seems reasonable. The authors derived an expression for the expected value (ensemble average) of the maximum mass in a sample of N stars, assuming a power-law mass distribution of the form (1). If $a = 1.35$, then for values of N as low as 100 the expected value of the maximum mass far exceeds 150 $M_\odot$ unless there is an upper limit to the mass distribution. The authors also calculated the probability of finding a largest mass below 200 $M_\odot$ in R136, and found it to be less than 10^{-5} unless the upper mass limit is itself below 1000 $M_\odot$.

A point investigated by both Weidner & Kroupa (2004) and Oey & Clarke (2005) is the consequence of a change in the exponent a in (1) in the distribution of large masses. This is taken up in Section 2, in which a truncated form of (1) is fitted to the 29 largest masses estimated by Massey & Hunter (1998). The aims are to investigate whether a law of the form (1) is appropriate for these masses, and to estimate relevant parameters such as the mass range over which it applies and the value of the exponent a. In Section 3, likelihood ratio statistics are used to compare models with and without upper mass limits. Conclusions are presented in Section 4.

2 DISTRIBUTION FITTING

We consider a pdf of the form (1), defined on the mass interval $[L, U]$ where the upper limit U may be infinite. Weidner & Kroupa (2004) worked with $L = 1\,M_\odot$, while numerical values quoted by Oey & Clarke (2005) were based on $L = 10\ M_\odot$.

The cumulative distribution function (cdf) corresponding to (1) is

$$F(m) = \int_L^m f(x)\,\mathrm{d}x = \frac{\beta}{a}(L^{-a} - m^{-a}). \tag{2}$$

The cdf is the probability that the mass lies in the interval $[L, m]$. The normalization $F(U) = 1$ requires that

$$\beta = \frac{a}{L^{-a} - U^{-a}}. \tag{3}$$

The distribution of the maximum mass in a random sample of size N satisfying (1)–(2) is of interest. The notation $m_{(j)}$ will be used for the jth largest mass in the sample, i.e. the largest mass is $m_{(N)}$. The cdf of this so-called 'maximum order statistic' is well known to be

$$G\left[m_{(N)}\right] = \left\{F\left[m_{(N)}\right]\right\}^N$$

(e.g. David 1970). Oey & Clarke (2005) calculated the expected value, or ensemble average, of the mass $m_{(N)}$ numerically. An analytical result is available in the statistics literature:

$$\mathrm{E}m_{(N)} = L\mathcal{H}[a^{-1}, N; N+1; 1 - (U/L)^{-a}]$$

(Khurana & Jha 1991), where $\mathcal{H}$ is the hypergeometric function.

There is a straightforward graphical technique which can be used to gain insight into the question of whether the upper mass limit in (1) and (2) is infinite. From (2) and (3),

$$F(m) = \frac{1 - (m/L)^{-a}}{1 - (U/L)^{-a}}, \tag{4}$$

which can be rewritten as

$$\log[1 - \kappa F(m)] = -a \log m + a \log L, \tag{5}$$

where $0 < \kappa = 1 - (U/L)^{-a} \leqslant 1$. In the special case $U \to \infty$, $\kappa = 1$ and (5) reduces to

$$\log[1 - F(m)] = -a \log m + a \log L. \tag{6}$$

A standard way of dealing with equations of the form (6) is to order the masses from small to large, and use the 'empirical cdf'

$$\widehat{F}\left[m_{(j)}\right] \equiv \frac{j}{N+1} \tag{7}$$

(e.g. Rice 1988). A plot of $\log\{1 - \widehat{F}[m_{(j)}]\}$ against $\log[m_{(j)}]$ should then be approximately linear if (6) holds. On the other hand, for a truncated distribution (finite U, $\kappa < 1$) the plot will curve downward for large $m_{(j)}$ (see e.g. Aban, Meerschaert & Panorska 2005): effectively an estimate of $\log[1 - F(m)]$ is plotted against $C_1 + C_2 \log[1 - \kappa F(m)]$ (with C_1, C_2 and κ constants); it is then not difficult to verify that the graph will have downward curvature for $0 < \kappa < 1$.

Both Figs 1 and 2, for the two different sets of mass determinations given by Massey & Hunter (1998), show very obvious downward curvature with increasing m. This then serves as strong evidence that U is finite. Proceeding on the assumption that (1) is appropriate, the mass limits L and U and the exponent a are now estimated.

Using (5) to estimate U, L and a by least-squares fitting looks inviting. However, the optimal solution has $a = 0$ (since this satisfies the identity $0 = 0$), which is not useful. Instead, we use

$$\log m = \log L - \frac{1}{a}\log[1 - \kappa F(m)], \tag{8}$$

replacing m by $m_{(j)}$ and $F(m)$ by $\widehat{F}[m_{(j)}]$ from (7). Fitting is easily performed by non-linear least squares. Details of the solutions are given in Table 1. Plots of the data are repeated in Figs 3 and 4; also shown (solid lines) are the functions

$$\log[1 - F(m)] = \log\left[1 - \frac{1 - (m/L)^{-a}}{\kappa}\right]$$

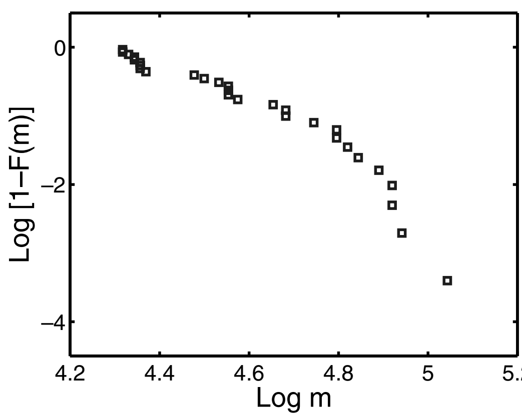

Figure 1. The logarithm of the estimated complementary cdf $[1 - \widehat{F}(m)]$ plotted against the logarithm of the VGS96 masses. If the mass distribution obeyed a power law with no mass limit, the points would have delineated a straight line to a reasonable approximation.

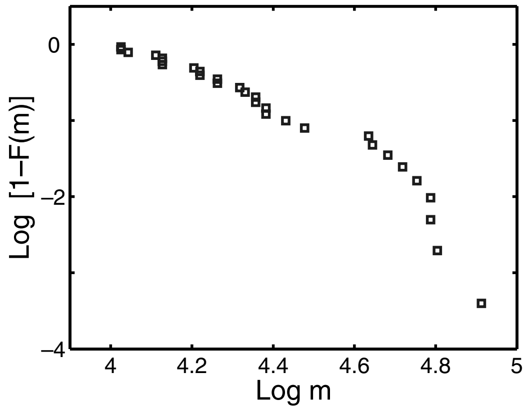

Figure 2. As for Fig. 1, but for the CG91 masses.

Table 1. Parameter estimates for the two sets of masses, using least-squares and maximum likelihood techniques. The exponents a' were calculated under the assumption of infinite U, and are shown for purposes of comparison.

	Least-squares method				Maximum likelihood method			
Data set	L	U	a	a'	L	U	a	a'
VGS96	70.4	158.1	0.94	4.0	75	155	1.74	3.5
CG91	54.5	143.9	1.10	3.3	56	136	1.11	2.7

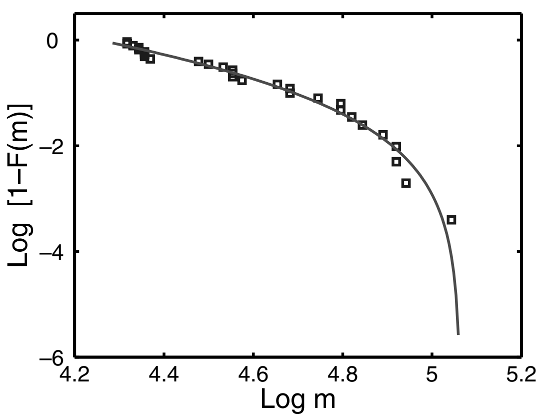

Figure 3. The fit of a truncated power-law distribution to the VGS96 masses.

[see equation (4)] calculated using the estimated parameter values. Clearly the derived values of L, a and κ give good fits to the observed mass distribution.

It is perhaps worth pointing out that the solutions for L are close to (more precisely, slightly less than) the lower limits (75 and 56 $M_\odot$ respectively) of the 29 masses taken from Massey & Hunter (1998). This is plausible, since, given only these data, the probability associated with yet lower masses is negligible, or, put differently, the sample does not allow pronouncements on the distribution of masses below L.

Table 1 also contains estimates for L, U and a derived by a second method, namely maximum likelihood. The necessary formulae were recently derived by Aban et al. (2005). The estimators for L and U are simply the lower and upper order statistics of the sample, i.e.

$$\widehat{L} = m_{(1)}, \qquad \widehat{U} = m_{(N)}, \tag{9}$$

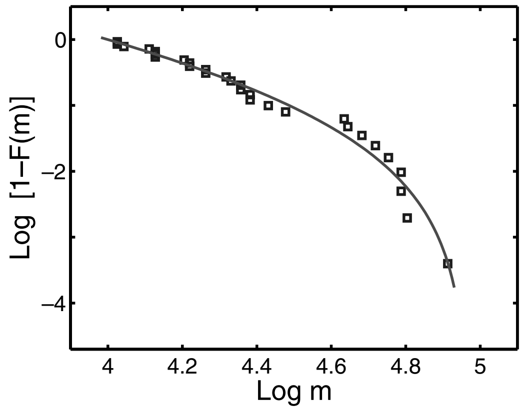

Figure 4. The fit of a truncated power-law distribution to the CG91 masses.

whereas $\widehat{a}$ is given by the solution of

$$\frac{N}{\alpha} + \frac{N r^\alpha \log r}{1 - r^\alpha} - \sum_{j=1}^{N} \log\left[m_{(j)}/m_{(1)}\right] = 0, \quad r \equiv m_{(1)}/m_{(N)} \tag{10}$$

for the unknown α.

The statistical significance levels of the finite upper mass limits given in Table 1 are discussed in Section 3, where it is shown that they are highly meaningful. It is also noted in passing that power-law exponents (denoted a' in Table 1) estimated for the case $U \rightarrow \infty$ are substantially larger ($2.7 \leqslant a' \leqslant 4.0$) than their counterparts for finite U.

The lower limits on the masses tabulated by Massey & Hunter (1998) are somewhat arbitrary, and it is therefore of some interest to see what parameter estimates result if different values are chosen. This point is easily explored by deleting the lowest masses from the data set and re-estimating the parameters. Of course, the reduction in the sizes of the data sets means that the reliability of the estimates decreases with increasing lower mass limit: for this reason the minimum sample size allowed was taken to be $N = 15$. The ranges of the minimum masses were then $75 \leqslant m_{(1)} \leqslant 95\,M_\odot$ and $56 \leqslant m_{(1)} \leqslant 78\,M_\odot$ for the VGS96 and CG91 calibrations respectively.

The results show very little variation in the estimates of U: for the VGS96 masses, $152 \leqslant U \leqslant 158\,M_\odot$; for the CG91 masses, $139 \leqslant U \leqslant 144\,M_\odot$. The lower limits L followed the values of $m_{(1)}$, in the mean being about $3\,M_\odot$ smaller than $m_{(1)}$. The results for the estimated exponents a can be seen in Figs 5 and 6, which are plotted on the same scales. Exponents estimated from the CG91 masses do not vary as much as those calculated from the VGS96 masses.

3 STATISTICAL SIGNIFICANCE LEVELS OF THE FINITE UPPER MASS LIMIT

A standard method can be used for a rough comparison of the models in Table 1 with the best fits obtainable under the assumption that $U \rightarrow \infty$, i.e. $\kappa = 1$, namely evaluation of the likelihood ratio statistic (e.g. Mood, Graybill & Boes 1974).

3.1 The least-squares method

Denote the residuals associated with the non-linear regression (8) by

$$\epsilon_j = \log m_j - \log L + \frac{1}{a}\log[1 - \kappa F(m_j)], \quad j = 1, 2, \ldots, N.$$

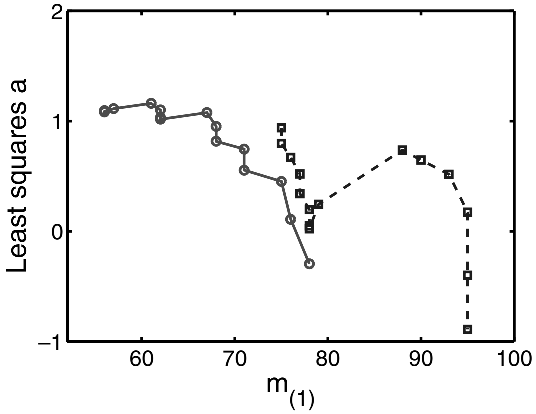

Figure 5. The least-squares estimated exponents for different values of the minimum sample mass $m_{(1)}$. Circles (connected by solid lines) and squares (connected by dashed lines) denote results using CG91 and VGS96 masses respectively.

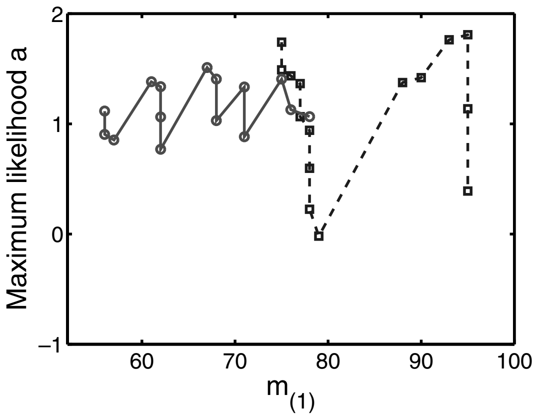

Figure 6. The maximum likelihood estimated exponents for different values of the minimum sample mass $m_{(1)}$. Circles (connected by solid lines) and squares (connected by dashed lines) denote results using CG91 and VGS96 masses respectively.

The Gaussian log likelihood is then given by

$$\mathcal{L} = -\frac{1}{2}\left[N\log 2\pi + N\log\sigma^2 + \sum_j\left(\frac{\epsilon_j}{\sigma}\right)^2\right], \tag{11}$$

where σ^2 is the variance of the ϵ_j. Using the usual estimate

$$\widehat{\sigma}^2 = \frac{1}{N}\sum_j \epsilon_j^2,$$

(11) reduces to

$$\mathcal{L} \approx -\frac{N}{2}[\log 2\pi + \log\widehat{\sigma}^2 + 1]. \tag{12}$$

Models with $\kappa = 1$ and $\kappa \neq 1$ can be compared by calculating the likelihood ratio statistic

$$\Lambda = 2[\mathcal{L}(L, a, U) - \mathcal{L}(L, a)], \tag{13}$$

where the log likelihoods are maximized with respect to their respective arguments, and κ is fixed at unity in the maximization of $\mathcal{L}(L, a)$. Substitution of (12) into (13) then gives

$$\Lambda \approx N\left(\log\widehat{\sigma}_1^2 - \log\widehat{\sigma}_\kappa^2\right), \tag{14}$$

where subscripts 1 and κ refer to the models with $\kappa = 1$ and $\kappa \neq 1$ respectively.

Since the value $\kappa = 1$ is at the boundary of its parameter space, the statistic Λ has the non-standard asymptotic distribution

$$\text{Prob}(\Lambda > x) = \frac{1}{2}\text{Prob}\left(\chi_1^2 > x\right), \tag{15}$$

where χ_1^2 is the chi-squared distribution with one degree of freedom (Chernoff 1954).

The values of Λ for the VGS96 and CG91 data sets are 56.4 and 50.8 respectively; given that $\text{Prob}(\chi_1^2 > 10.8) = 0.001$, both values are significant at levels far better than 0.05 per cent. The interpretation is that under the null hypothesis ($U \to \infty$), such large values of Λ are extremely unlikely (probability less than 0.0005).

3.2 The maximum likelihood method

From (1) the likelihood of the set of mass values is given by

$$Lik = \prod_j \beta m_j^{-(a+1)},$$

and hence

$$\mathcal{L} = \log Lik = N\log a - N\log(L^{-a} - U^{-a}) - (a+1)\sum_j \log m_j,$$

where β has been taken from (3). The likelihood ratio statistic is again given by (13), with $U^{-a} = 0$ in the expression for $\mathcal{L}(L, a)$.

Estimators maximizing $\mathcal{L}(L, a, U)$ are given in (9) and (10). In the case of infinite U, $\widehat{L} = m_{(1)}$ as before, while

$$\widehat{a} = \frac{N}{\sum_j \log\left[m_j/m_{(1)}\right]}.$$

The large sample distribution of Λ is again given by (15).

The likelihood ratio statistic is 8.2 and 9.8 for the VGS96 and CG91 masses respectively; these values are significant at better than the 0.2 per cent level.

4 CONCLUSIONS

The maximum likelihood values of a estimated from the CG91 masses are in reasonable agreement with the accepted Salpeter IMF value of 1.35, as are the least-squares values for $m_{(1)} \leqslant 67\,\mathrm{M}_\odot$. Subject to the validity of (1) for high masses – which certainly is plausible according to Figs 3 and 4 – an upper stellar mass limit of the order of 140–160 $\mathrm{M}_\odot$ is predicted.

Although the data sets analysed are particularly useful for investigation of the upper reaches of the mass distribution, it would also be interesting to apply the methods to a more extensive collection of masses, i.e. with smaller $m_{(1)}$. Provided that (1) holds over the entire range studied, this would allow more accurate estimates. Of course, the smaller $m_{(1)}$, the more important issues of sample completeness are bound to become.

Since publication of the Massey & Hunter (1998) mass estimates used in this paper, a huge amount of work has been done in order to understand O-type stars better. Recent non-local thermodynamic equilibrium (NLTE) models include all or some of the effects of the hydrodynamics of stellar winds, line blanketing by metals, wind blanketing (backscattering of radiation) and spherical extension of the atmosphere (e.g. Hillier & Miller 1998; Pauldrach, Hoffmann & Lennon 2001; Lanz & Hubeny 2003; Puls et al. 2005). This work has shown that the VGS96 calibration generally overestimates the effective temperatures of the hottest stars by several thousands of degrees even for Magellanic Cloud stars, where the effects are smaller than for Milky Way O stars (e.g. Martins, Schaerer & Hillier 2005; Massey et al. 2005). As a result the mass estimates of the stars are sharply decreased: this can be seen in the re-analysis of a few R136 stars by Massey et al. (2004, 2005).

The mass determinations are bound to continue changing over the next few years as there are a number of relevant issues which remain unresolved. To name but two of these: there are substantial discrepancies between spectroscopic and evolutionary masses for the hottest stars (Massey et al. 2005), and there are differences in bolometric corrections found by different groups (Martins et al. 2005; Massey et al. 2005).

ACKNOWLEDGMENTS

Comments from a referee led to a considerable improvement of the original version of the paper.

REFERENCES

Aban I. B., Meerschaert M. M., Panorska A., 2005, J. Am. Stat. Assoc., in press
Chernoff H., 1954, Ann. Math. Stat., 25, 573
Chlebowski T., Garmany C. D., 1991, ApJ, 368, 241 (CG91)
David H. A., 1970, Order Statistics. John Wiley, New York
Hillier D., Miller D. L., 1998, ApJ, 496, 407
Khurana A. P., Jha V. D., 1991, Sankhya Ser. B, 53, 11
Kroupa P., 2001, MNRAS, 322, 231
Kroupa P., 2002, Sci, 295, 82
Lanz T., Hubeny I., 2003, ApJS, 146, 417
Martins F., Schaerer D., Hillier D. J., 2005, A&A, 436, 1049
Massey P., 1998, in Gilmore G., Howell D., eds, ASP Conf. Ser. Vol. 142, The Stellar Initial Mass Function. Astron. Soc. Pac., San Francisco, p. 17
Massey P., Hunter D. A., 1998, ApJ, 493, 180
Massey P., Bresolin F., Kudritzki R. P., Puls J., Paudrach A. W. A., 2004, ApJ, 608, 1001
Massey P., Puls J., Paudrach A. W. A., Bresolin F., Kudritzki R. P., Simon T., 2005, ApJ, 627, 477
Mood A. M., Graybill F. A., Boes D. C., 1974, Introduction to the Theory of Statistics, 3rd edn. McGraw-Hill, New York
Oey M. S., Clarke C. J., 2005, ApJ, 620, L43
Pauldrach A. W. A., Hoffmann T. L., Lennon M., 2001, A&A, 375, 161
Puls J., Urbaneja M., Venero R., Repolust T., Springmann U., Jokothy A., Mokiem M. R., 2005, A&A, 435, 669
Rice J. A., 1988, Mathematical Statistics and Data Analysis. Wadsworth & Brooks, Pacific Grove
Salpeter E. E., 1955, ApJ, 121, 161
Schaerer D., Meynet G., Maeder A., Schaller G., 1993, A&AS, 98, 523
Vacca W. D., Garmany C. D., Shull J. M., 1996, ApJ, 460, 914 (VGS96)
Weidner C., Kroupa P., 2004, MNRAS, 348, 187

This paper has been typeset from a TEX/LATEX file prepared by the author.