

# ON THE PRECISION OF THE MK SPECTRAL CLASSIFICATION SYSTEM

C. JASCHEK and M. JASCHEK

*Observatorio Astronómico, La Plata, Argentina*

**Abstract.** It is shown that a large scatter exists in all color-color and color-spectral type diagrams. This scatter can not be neglected and its importance is shown in several examples. A new approach is then developed for the study of the precision attainable in spectral classification. It is found that for A-type stars MK classification is about as accurate as photometry, whereas for B-type dwarfs photometry is at least two times more precise than MK classification. A precision of  $\pm 0.6$  tenths of a spectral class seems attainable in both groups.

The purpose of this note is to communicate some reflections concerning the accuracy of the MK spectral classification system. In principle one can either compare the results of the MK system with those of other spectral classification systems and assess the accuracy through the intercomparison, or one can compare the results of the MK systems with those of other classification systems, for instance photometric ones.

The first procedure was applied in a previous paper (Jaschek and Jaschek, 1964), where it was found that

$$\sigma(S) = \pm 0.6$$

( $S$  = spectral class in tenths of a spectral type). Gliese (1971) has confirmed this result for nearby dwarfs, mostly of types later than F.

The present paper will deal with the second of the procedures mentioned above. The great advantage of the method is that it permits to evaluate the accuracy of all classification systems involved, both spectroscopic and photometric. We shall apply specifically the method to a material of B- and A-type dwarfs.

Let us start the analysis with the examination of the relation between MK types and UBV photometry. The mean relation between these parameters has been examined many times and the different results are in general agreement. The few discrepancies can be attributed to statistical fluctuations and are generally equal to or less than  $0^m03$ . In what follows we have adopted Johnson's values (1966) for B0–B9 dwarfs, whereas for A0–A9 dwarfs we have preferred those of Cowley *et al.* (1969), because these authors give finer subdivisions of spectral type than Johnson.

The use of such average relations has become so familiar that it has been almost forgotten that they are statistical in origin, implying also a dispersion around the average. The existence of this dispersion has been so much ignored that we have been unable to find an analysis of it. We have therefore calculated the dispersions in a material of field A-type dwarfs classified in a homogeneous way (Cowley *et al.*, 1969) and in a material of field B-type dwarfs. We have excluded from our material all

stars fainter than 6<sup>m</sup>5 and stars which had companions closer than  $d=15''$  and/or  $\Delta m < 4^m$ . In the case of the A-type dwarfs we have examined the existence of reddening corrections by forming average colors for dwarfs both brighter and fainter than 5<sup>m</sup>0. If reddening were statistically significant, one would expect both averages to differ systematically, the average for the fainter group being redder than the average of the brighter stars. Since such trend did not show up, we omitted reddening corrections for the A-type dwarfs. For B-stars this is clearly not permissible and it was decided therefore to correct the  $U-B$  index by means of

$$E(U - B) = 0.71 \cdot E(B - V).$$

Because of this correction, only the dispersion in  $U-B$  can be used in the B dwarfs.

The  $UBV$  colors of the A-type dwarfs were taken from Cowley *et al.* (1969), whilst those of the B-type dwarfs were taken from Blanco *et al.* (1968).

Tables I and II summarize conveniently the results. In conclusion it can be said that the average value of the dispersion for

B-type dwarfs is  $\pm 0^m07$

A-type dwarfs is  $\pm 0^m04$ .

Both values are very large and imply that the relation is a band rather than a curve. This is a very obvious fact, which is, however, very seldomly mentioned.

TABLE I  
Dispersion of  $U-B$  indices in B-type dwarfs

	Hiltner (1956)	Slettebak (1954, 1955)	Hiltner <i>et al.</i> (1969)	Cowley <i>et al.</i> (1969)	Adopted
B0	0 <sup>m</sup> 08 13				0 <sup>m</sup> 08
B0.5	07 30				07
B1	07 60				07
B2	08 21	0 <sup>m</sup> 11 14	0 <sup>m</sup> 05 25		07
B2.5			04 15		
B3	07 11	10 22	04 18		07
B4			03 12		
B5		04 13	05 16		05
B8		06 15		0 <sup>m</sup> 09 18	07
B9		06 9		08 36	08
B9.5				05 37	05

The first column gives the dispersion, the second column gives the number of data.

The easiest objection one can make to these high dispersions is that they contain the dispersion of the spectral types. The next step is therefore to calculate the dispersions without intervention of the spectral types. This can be done by calculating the dispersion in  $U-B$  colors from all stars having  $B-V$  indices within prefixed intervals. From the data assembled in Table III it can be seen that the dispersion is practically

TABLE II  
Dispersion of  $U-B$  and  $B-V$  indices in A-type dwarfs

	$N$	$\sigma(U-B)$	$\sigma(B-V)$
A0	101	0 <sup>m</sup> 07	0 <sup>m</sup> 03
A1	101	5	3
A2	96	4	4
A3	73	3	3
A4	20	2	4
A5	20	3	3
A7	17	4	2
A8	10	7	5
F0	18	6	3

$N$  = number of stars. The colors were taken from a variety of sources. See Cowley *et al.* (1969).

TABLE III  
Dispersion of  $U-B$  indices in A type stars  
for prefixed values of  $B-V$

$B-V$	$N$	$\sigma(U-B)$
0 <sup>m</sup> 00 to 0 <sup>m</sup> 05	141	$\pm$ 0 <sup>m</sup> 054
05	10	046
10	15	034
15	20	034
20	25	036
25	30	044

constant and of the order of  $\pm 0^m04$ . In conclusion, the dispersion in spectral type does not influence the dispersion in  $U-B$ . This can be explained through the slope of the relation spectral type vs  $U-B$ , which is very small for the range A0-A9.

One possible contribution to the dispersion could originate from systematic errors between observers. Since our colors were averaged from series of different observers, there exists the possibility that this heterogeneity increases considerably the dispersion. To check this possibility we derived the dispersions anew from the material gathered by Johnson *et al.* (1966). The observations contained in this paper were all made by Johnson and collaborators, in the most homogeneous fashion possible. Since the list provides not only  $UBV$  but also  $RI$  measurements, we calculated the dispersions also for these additional indices. The results are given in Table IV. Since also here the average dispersion  $\sigma(U-B) = \pm 0^m04$ , it seems clear that systematic differences between observers do not constitute a large source of error.

A second conclusion from this table is that in no color the dispersions are less than 0<sup>m</sup>03. Therefore one can say quite generally that broad band two color diagrams give 'bands' and not 'curves'. The dispersion around the average curve is thus an essential feature of the two color diagrams. This dispersion is caused by a number of sources,

TABLE IV  
Dispersion of multicolor indices in A-type stars

Group	<i>N</i>	$\sigma(U-V)$	$\sigma(B-V)$	$\sigma(V-R)$	$\sigma(V-I)$
B9.5 V	12	0 <sup>m</sup> 06	0 <sup>m</sup> 03	0 <sup>m</sup> 03	0 <sup>m</sup> 05
A0	18	9	4	3	5
A1	19	6	3	2	4
A2	31	5	3	3	5
A3	22	4	2	3	5
A5	9	2	2	4	6
A7	7	(2)	(2)	(2)	(2)
F0	5	(2)	(4)	(3)	(5)
$\langle\sigma\rangle$		0 <sup>m</sup> 05	0 <sup>m</sup> 03	0 <sup>m</sup> 03	0 <sup>m</sup> 05

The colors were taken exclusively from Johnson *et al.* (1966).

as for instance observational errors, interstellar reddening, chemical composition effects and rotational effects. We will not deal here with the disentanglement of these different effects, but we will rather show some consequences of the existence of a scatter in the two color diagrams.

Let us take the case of the interstellar extinction corrections. Usually stars are 'de-reddened' by moving them back to the 'mean curve'. But if this relation is a 'band' the reddening is indeterminate by a amount of the order of

$$\Delta A_v = \pm 3 \cdot \sigma(B - V).$$

Since the dispersion is independent of the number of objects, the  $\Delta A_v$  cannot be diminished by enlarging the number of stars. It should be added that this criticism was first raised by Becker (1966).

A somewhat similar situation arises if spectral types are used to determine intrinsic colors. Such a method only works if the scatter around the average relation is negligible although this basic assumption is never explicitly mentioned. In order to see what happens, we have given in Table V the correlation between  $U-B$  colors and spectral types for bright southern stars. The spectral types were taken from Hiltner *et al.* (1969) and the colors from Blanco *et al.* (1968). It can be seen that no unique intrinsic color

TABLE V  
Distribution of  $U-B$  indices and spectral types in early nearby B-type dwarfs

$U-B$ :	-0 <sup>m</sup> 85	80	75	70	65	60	55	50	45	40	35	<i>N</i>
B2	6	10	4	5	2	2						29
B2.5	2	-	3	7	2	1						15
B3			1	4	8	4	-	1	1	1		19
B4					3	3	2	1				9
B5					1	8	3	2	1	1		16
B6							1	3	1	2		7

exists for each spectral type, the range being of the order of  $0^m2$  to  $0^m3$ . This is in line with dispersions in  $U-B$  of the order of  $0^m07$ , but shows that intrinsic colors derived via spectral types are imprecise. In general one should avoid such a procedure, and use spectral classification only for selecting normal dwarfs.

One objection which can be raised to the previous procedures is that the scatter is large because broad band photometry is used, which is less precise than narrow band photometry. To examine this argument we have taken several intermediate band and narrow band indices and have analysed the scatter in these indices. Of the many available indices we have selected three, namely the  $H\beta$  index of Crawford *et al.* (1966, 1970) the K-index of Henry (1969, 1971) and the ' $b-y$ ' index of Strömrgren (Strömrgren and Perry, unpubl.). The first two were chosen because Balmer lines and the K-line are also used in spectral classification of A-type stars, which permits an easy comparison of the different techniques. The ' $b-y$ ' index on the other side was included essentially for comparison with the broad band index ' $B-V$ '.

We have calculated the scatter of these indices in a sample of or A-types dwarfs, the sample being selected according to principles described in more detail below. The dispersions were calculated from the relation of each parameter with  $B-V$ , and the results are given in Table VI. As one can see, the dispersions are very large even with narrow band indices. The exception is ' $b-y$ ' which correlates very strongly with ' $B-V$ ', and has accordingly a very small scatter.

TABLE VI  
Dispersion of photometric parameters

Parameter	Dispersion
K index	$\pm 0^m06$
$H\beta$ index	04
$b-y$	01

Let us now examine the question of the relative precision of these indices. To the three parameters quoted above we will add the ' $B-V$ ' index and the spectral type. Luminosity classes can be left out because we deal only with dwarfs, defined as such by the usual spectroscopic criteria. We have thus five different parameters at our disposal and we would like to examine which parameter gives the least dispersion.

This problem becomes mathematically treatable if it can be assumed that the relation between any pair of the five parameters is linear. This hypothesis can be checked easily and is valid probably in all cases if one uses only small parts of the main sequence.

If  $X$  and  $Z$  are two of the indices, we have

$$Z = \alpha X + \alpha_0.$$

Now  $Z$  and  $X$  have both errors which we will denote by  $\varepsilon$  and  $\delta$ . We assume further that  $\langle \varepsilon \rangle = 0$  and  $\langle \delta \rangle = 0$ . The dispersion of the errors of  $Z$  is  $\sigma(\varepsilon)$  and that of  $X$ ,

$\sigma(\delta)$ . The  $\varepsilon$  and  $\delta$  combine both the observational errors and the intrinsic errors of each variable, because at the moment it does not seem profitable to split them into components. Statistical theory (Kendall and Stuart, 1967; Deeming, 1968) shows that it is possible to formulate a system of three equations involving the second moments of  $X$  and  $Z$  and the first mixed moment, with four unknowns. This system is obviously not solvable. But if instead of having two linearly related variables  $X$  and  $Z$ , one has three –  $X$ ,  $Z$  and  $Y$  – then the problem can be solved. The appropriate formulae are given in the mathematical appendix.

Proceeding along these lines we have studied a sample of A-type dwarfs selected according to the following criteria.

- (1) All stars should be dwarfs between spectral types B9 and F0.
- (2) All stars should be nearby, to avoid reddening corrections,  $m < 6^m5$ .
- (3) All stars should have measures of the K index, the  $H\beta$  index,  $B-V$ ,  $b-y$  and estimates of the spectral type in the MK system.

We found finally a sample of 87 stars which satisfy all these criteria. The indices were taken from Henry (1971), Crawford *et al.* (1966, 1970), Blanco *et al.* (1968), and Cowley *et al.* (1969). We applied then the technique described above and in more detail in the appendix. Since with five indices there exist five possible combinations of the three indices needed to apply our technique, we decided to combine in turn  $b-y$  and  $B-V$  with each one of the remaining indices. From the results given in Table VII it can be seen that some indices have a very large scatter, both in absolute value as when compared with its range of variability.

TABLE VII  
Intrinsic dispersion of different parameters, from a sample of 87 A-type dwarfs

Index	Dispersion
$b-y$	$\pm 0^m011$
$B-V$	$\pm 0^m013$
K	$\pm 0^m073$
$H\beta$	$\pm 0^m041$
$S$	0.62

Attention is also called to the value of 0.6 obtained for the spectral type, which agrees very well with the one found by Gliese (1971).

A sample of 76 B type stars was analysed in the same way. Here we used the parameters  $U-B$ ,  $c_1$ ,  $H\beta$  and  $S$ . It turned however out that  $S$  is not linearly related to the other variables, a fact shown by the imaginary values for  $\sigma(S)$  obtained when solving the equations. For the other variables the results are those given in Table VIII.

We can compare now the dispersions obtained for each parameter with its range of variation. The results are given in Table IX. In A-type dwarfs it is certainly no surprise that  $B-V$  and  $b-y$  do have the same efficiency, defined as the ratio of range and dispersion. Spectral class turns out to be approximately as precise as  $b-y$  or

TABLE VIII

Intrinsic dispersion of different parameters, from a sample of 76 B-type dwarfs

Index	Dispersion
$c_1$	$\pm 0^m020$
$U-B$	$\pm 032$
$H\beta$	$\pm 017$

TABLE IX

Dispersion and range of classification parameters

	Parameter	Range	$\sigma$	$R/\sigma$
A-type stars:	$b-y$	$0^m24$	0.011	22
	$B-V$	$0^m37$	0.015	24
	$S$	11	0.62	18
	$K$	$0^m97$	0.083	12
	$H\beta$	$0^m20$	0.041	5
B-type stars:	$c_1$	$1^m08$	0.020	54
	$U-B$	$0^m97$	0.032	30
	$H\beta$	0.25	0.017	15
	$S$	12	0.7	17

$B-V$  and to be definitely more precise than both the  $K$  and  $H\beta$  indices. This is understandable because of the large dispersions found in Table VI. It implies that for A-type dwarfs spectral classification is about as accurate as photometry.

For B-type stars the situation is different. Here Table IX shows beyond doubt that photometry is at least twice as precise as spectral classification. Spectral classification is not competitive here, but remains essential even in this case because of the need to eliminate non-dwarfs.

### Acknowledgements

Thanks are due to Mr B. Kucewicz who carried out the calculations.

### Mathematical Appendix

Let  $X_i$  and  $Y_i$  be two mathematical variables affected by observational errors so that

$$x_i = X_i + \delta_i \quad (1)$$

$$y_i = Y_i + \varepsilon_i \quad (2)$$

and assume that  $X_i$  and  $Y_i$  are related linearly

$$Y_i = \alpha X_i + \alpha_1 \quad (3)$$

Assume further that  $\langle \varepsilon_i \rangle = \langle \delta_i \rangle = 0$  and that the dispersions of  $\varepsilon_i$  and  $\delta_i$  are  $\sigma_\varepsilon$  and  $\sigma_\delta$ . Then by substitution

$$y_i - \varepsilon_i = \alpha(x_i - \delta_i) + \alpha_1 \quad (4)$$

and averaging

$$\langle y \rangle = \alpha \langle x \rangle + \alpha_1. \quad (5)$$

By calculating dispersions of (1) one gets

$$\sigma_x^2 = \sigma_X^2 + \sigma_\delta^2 \quad (6)$$

and by calculating the dispersion of (4), after some algebra

$$\sigma_y^2 = \alpha^2 \sigma_x^2 + \sigma_\varepsilon^2 \quad (7)$$

provided

$$\langle X_i \delta_i \rangle = \langle Y_i \varepsilon_i \rangle = 0.$$

One can also calculate the covariance of  $x$  and  $y$  and obtains

$$\text{cov}(x, y) = \frac{1}{n} \sum (x_i - \langle x \rangle)(y_i - \langle y \rangle) = \alpha \cdot \sigma_x^2 \quad (8)$$

provided  $\langle \varepsilon_i \delta_i \rangle = \langle \varepsilon_i X_i \rangle = 0$ . The condition on  $\langle \varepsilon_i \delta_i \rangle$  implies that the errors are uncorrelated, the same as  $\langle \varepsilon_i X_i \rangle = 0$ .

Under these hypothesis, equations (6), (7) and (8) form a system of three equations with the four unknowns  $\sigma_x$   $\sigma_\varepsilon$   $\sigma_\delta$   $\alpha$

$$\begin{aligned} \sigma_x^2 &= \sigma_X^2 + \sigma_\delta^2 \\ \sigma_y^2 &= \alpha^2 \sigma_x^2 + \sigma_\varepsilon^2 \\ \text{cov}(x, y) &= \alpha \cdot \sigma_x^2. \end{aligned} \quad (9)$$

A small reflection shows that if one wants  $\sigma_\delta > 0$  and  $\sigma_\varepsilon > 0$  then

$$\begin{aligned} \sigma_x^2 &> \sigma_X^2 \\ \sigma_y^2 &> \alpha^2 \sigma_x^2 \\ \text{cov}(x, y) &= \alpha \cdot \sigma_x^2 \end{aligned} \quad (10)$$

implying that one must have

$$\text{cov}(x, y) < \sigma_x \cdot \sigma_y. \quad (11)$$

This condition proves to be very helpful.

Obviously the system (8) cannot be solved. Assume now that a third variable  $Z_i$  exists with

$$Z_i = z_i + \zeta_i \quad (12)$$



connected with  $X_i$  and  $Y_i$  linearly, like

$$\begin{aligned} Z_i &= \beta Y_i + \beta_1 \\ X_i &= \gamma Z_i + \gamma_1. \end{aligned} \tag{13}$$

Then one can write down the system (8) for each pair of variables, yielding nine equations with twelve unknowns. By dropping those equations and unknowns which repeat themselves because of the cyclic arrangement (13) one gets a system of seven equations with seven unknowns, which becomes solvable.

It was found convenient in practice not to solve directly the system of seven equations, but to solve for instance the systems  $(X, Y)$  and  $(X, Z)$  by introducing assumed values of  $\sigma_X$  and then to select that triple of values  $\sigma_X$ ,  $\sigma_Y$  and  $\sigma_Z$  which satisfies simultaneously all equations.

### References

- Becker, W.: 1966, *Z. Astrophys.* **64**, 77.  
 Blanco, V., Demers, S., Douglass, G. G., and Fitzgerald, M. P.: 1968, *Publ. U.S. Naval Obs.*, 2nd Ser. **XXI**.  
 Cowley, A., Cowley, C., Jaschek, M., and Jaschek, C.: 1969, *Astron. J.* **74**, 375.  
 Crawford, D. L., Barnes, J. V., Faure, B. Q., Golson, J. C., and Perry, C. L.: 1966, *Astron. J.* **71**, 709.  
 Crawford, D. L., Barnes, J. V., and Golson, J. C.: 1970, *Astron. J.* **75**, 624.  
 Deeming, T. J.: 1968, *Vistas in Astronomy* **10**, 125.  
 Gliese, W.: 1971, *Veröffentl. Astron. Rechen-Inst. Heidelberg*, No. 24.  
 Henry, R. C.: 1969, *Astrophys. J. Suppl.* **18**, 47.  
 Henry, R. C.: 1971, *Astrophys. J. Suppl.* (in press).  
 Hiltner, W. A.: 1956, *Astrophys. J. Suppl.* **2**, 389.  
 Hiltner, W. A., Garrison, R. F., and Schild, R. C.: 1969, *Astrophys. J.* **157**, 313.  
 Jaschek, C. and Jaschek, M.: 1966, in K. Lodén, L. O. Lodén, and U. Sinnerstad (eds.), 'Spectral Classification and Multicolor Photometry', *IAU Symp.* **24**, 6.  
 Johnson, H. L., Mitchell, R. I., Iriarte, B., and Wisniewski, W. Z.: 1966, *Commun. Lunar Planetary Lab.*, No. 63.  
 Johnson, H. L.: 1966, *Ann. Rev. Astron. Astrophys.* **4**, 193.  
 Kendall, M. G. and Stuart, A.: 1967, *The Advanced Theory of Statistics II*, Charles Griffin and Co., London.  
 Slettebak, A.: 1954, *Astrophys. J.* **119**, 146.  
 Slettebak, A.: 1955, *Astrophys. J.* **121**, 102.