Evaluation of Computed K-Indices with Orthogonalized S_R Method

Shinji Kadokura

Abstract

'Orthogonalized S_R Method' (OSM), one of the algorithms to derive K-indices, has been tested following the decision during IUGG (Vancouver) and IAGA (Exeter); by histograms of differences between hand-scaled and computed K-indices for seven observatories (Memambetsu, Hermanus, Crozet, Kerguelen, Nurmijärvi, Ottawa, and Sodankylä) from March, 1985 to February 1986. The OSM achieves good results except histograms for Hermanus (HER); (1) total agreements are 60 − 75%, (2) the occurrence of the differences larger than one unit is less than 2%. The wrong results for HER suggest wrong hand-scaling. It was revealed by analysis of multi-observatories data set that OSM, which is the optimal in the linear algorithms on the basis of statistical principle, is adequate for the algorithm to derive K-indices.

1. Introduction

The K-index has been internationally adopted as an indicator of the geomagnetic activity since Bartels et al. introduced it in 1939. Mayaud (1967) established morphological rules as guidelines to estimate the non-K variations, that is S_R (Solar Regular Variation). Since the patterns of S_R's are variable for each day, it is essential and difficult to estimate S_R on deriving K-index, so that K-indices must be hand-scaled by a well-trained observer. There are sometimes discrepancies between the K-indices determined by two well-trained observers originated from their subjectivity (Mayaud and Menvielle, (1979)). Therefore, it is desirable to find an algorithm for an automatic determination of the K-index. As recent advances in electronics and computer enable the geomagnetic observatories to adopt automatic operation including the digitization of the data, it is worth while developing a method for an automatic scaling of K-indices by using a computer. On the other hand, the computer derivation of K-indices is required for computation and circulation of IAGA K-derived indices, such as Km, Kp etc., within short time (Menvielle (1991)). Many algorithms that claim to derive K-indices have been proposed, and tested with the different data set (e.g. Van Wijk and Nagtegaal (1977), Hopgood (1986), Walker (1987), Wilson (1987)). Thus, it was decided during the IUGG (Vancouver) and IAGA (Exeter) meeting to organize a comparison between proposed algorithms for computer determination of K-indices with the same data set and formats (see Appendix).

All of previous algorithms for automatic K-index represent S_R with some parameters (e.g. Fourier coefficients in harmonic analysis, values obtained by filtering, etc.). Any parameter should be constrained (e.g. to be kept between the maximum and minimum values) on the basis of the data on quiet days in the past to separate S_R from disturbances. If the parameters correlate each other, the correlation should be represented in the constraint formulae. This representation causes the complicatedness of the formulae. Therefore, it is essential and necessary for 'real automation' that

(1) the number of parameters should be as small as possible,

(2) the parameters are independent of each other so that each formula of constraint does not include more than one parameter.

'Orthogonalized S_R Method' (OSM, Kadokura (1988)) based on the principal component analysis seeks the optimal solution on the basis of the statistical principle.

This report summarizes the comparison between the hand-scaled indices and those computed by OSM following the decision of the meeting.

2. Data Analysis

The analysis by OSM consists of two parts; the preparation process and the K-derivation process. In the preparation process, the data of quiet days are selected on the basis of K-index, and orthogonalized to obtain orthonormal basis of S_R (ONS), eigenvalues and statistical parameters (means and standard deviations) of coefficients. These results are recorded in magnetic disk media as 'the prepared data' and used repeatedly, so that this process should be executed only once. Since the pattern of S_R depends on season, 'the prepared data' are obtained for each month of the year with quiet data of relevant month. To derive S_R and K-indices, the one-minutes data of the relevant day and 'the prepared data' are used in the derivation process. In this process, S_R is represented as a linear combination of the ONS, and K-indices are derived from the range of the difference between the one-minutes data and derived S_R.

For some observatories, quiet data could be collected insufficiently in amount. To apply OSM to such observatories, OSM is modified in two point on the present analysis. The first is that quiet days are selected on the basis of Kp-indices instead of K-indices at each station. The second is that the quiet data of neighboring months are used for preparation process: e.g. quiet data on December, January and February are used to obtain ONS for January. Nevertheless, these modifications will cause no significant change.

The data are analyzed following the decision of the meeting described in Appendix, with three differences;

(1) since OSM requires the data of 20 quiet days for each month of year at least, the data of AIA, BEL, CNB, HAD and NEW, of which data could have been collected insufficiently in amount, were not analyzed,

(2) all of data are converted into 'quasi-local-time' (QLT: UT + 3 × n hours, where n is the integer to make QLT closest to the local time) base before analysis,

(3) the presentation of selected 30 days (histograms of (a5) and drawings) is abbreviated.

The purpose of the second difference is as follows; the variation pattern of S_R is active only at daytime, and it changes day by day; if the center of analysis unit is nighttime, S_R patterns before and after night will happen to be different; therefore, more likely S_R is obtained by setting daytime in the center of analysis unit; for the convenience to determine K-indices, QLT is used as the time base on analysis instead of the local time.

3. Results and discussion

Fig. 1−7 show the histograms of the differences between the hand-scaled (K(hand)) and computer-derived (K(comp)) K-values of seven observatories: MMB, HER, CZ, KG, NUR, OTT and SOD. The information of the seven observatories is tabulated in Table 1. Three panels of these figures correspond to (a1)−(a4), (b1) and (c) of the instruction of the K-comparison (Appendix): (A) is histograms of all results and of results for three seasons; (B) is histograms of all results as a function of the time of the day; (C) is histograms of all results as a function of the K(hand).

Memambetsu (MMB): The histograms for MMB are shown in Fig. 1. The panel (A) indicates the total agreement between K(hand) and K(comp) is 65%. The occurrence of the differences larger than one unit is less than 2%. The symmetric distribution indicates that the means of K(hand) and K(comp) are almost identical. This symmetry is strongly expected on replacement of K(hand) by K(comp). Considering that the noon of QLT (Quasi-Local Noon, abbreviated to QLN) is 3 UT, the panel (B) of Fig. 1 shows that the agreement is better in the nighttime than in the daytime. In the nighttime, the range of S_R is small, so that variety of S_R is small. Therefore, it is easier to estimate S_R for the nighttime than to estimate for the daytime. K(comp) is underestimated at K(hand) < 2, and is overestimated at K(hand) > 2, as shown in the panel (C). The feature similar to MMB, that K(comp) is underestimated on small values of K(hand) and is overestimated on large values of K(hand), is seen in the histograms of all stations.

Table 1 Observatories used in the present analysis.

observatory	abbr.	Geographic lat.	Geographic long.	K=9	DT	quiet data base
Crozet	CZ	−46.26	51.52	500	3	1984, 1987
Hermanus	HER	−34.25	19.14	300	0	1984, 1987, 1988
Kerguelen	KG	−49.21	70.12	750	6	1984, 1987
Memambetsu	MMB	43.54	114.12	350	9	1985–1989
Nurmijärvi	NUR	60.52	24.65	750	3	1983, 1984, 1988, 1989
Ottawa	OTT	45.24	284.27	750	−6	1984, 1987
Sodankylä	SOD	67.37	26.63	1500	3	1983, 1984, 1988, 1989

abbr. : abbreviation, DT : QLT − UT

Hermanus (*HER*): Fig. 2 shows the histograms for HER. Although HER is in a low latitude similarly to MMB, the histograms have characteristics quite different from those of MMB: (1) the total agreement is 48%, distinctly less than that of MMB; (2) the mean of the difference, $\Delta K = K(comp) - K(hand)$, is about 0.5 unit; (3) the occurrence of $|\Delta K| \geq 2$ is 7%, not a negligible ratio. Since the same characteristics are seen in the histograms obtained by other algorithms (e.g. FMI method, Sucksdorff et al. (1991)), I believe that the wrong results are caused by wrong hand-scaling.

Crozet (*CZ*): Fig. 3 shows the histograms for CZ. The total agreement is 61%, wronger in the daytime (QLN=9UT) than in the nighttime, and occurrence of $|\Delta K| \geq 2$ is 1%. The week asymmetry in the total distribution is a slight problem, because the mean of ΔK, 0.2, is small enough.

Kerguelen (*KG*) *and Nurmijärvi* (*NUR*): Fig. 4 and Fig. 5 show the histograms for KG and NUR, respectively. They are quite similar: (1) the total agreements between K(hand) and K(comp) are about 70%, (2) the occurrence of $|\Delta K| \geq 2$ is 1%, (3) although there is an asymmetry in the total distribution, the mean of ΔK is less than 0.2 unit, small enough to be neglected, (4) better agreements are achieved for the nighttime than for the daytime.

Ottawa (*OTT*): The histograms for OTT, Fig. 6, indicate the same features as (1)−(4) of KG and NUR. However, the asymmetry of the histogram is opposite sense, i.e. the means of K(hand) are larger than those of K(comp). This tendency is distinct in the daytime as shown in panel (B). The panel (C) indicates the mean of K(hand) are lager than K(comp) when K(hand) > 0. Since the similar characteristics are seen in the histogram obtained by FMI algorithm (Sucksdorff et al. (1991)), this asymmetry suggests that the hand-scaled values are wrong rather than computed values. The hand-scaling is probably based on an principle that the pattern of S_R changes not so much; this principle causes overestimation of K-variation amplitude, especially in daytime when the variety of S_R is large.

Sodankylä (*SOD*): The histograms for SOD, Fig. 7, achieve the best agreement among those for the seven observatories: the total agreement is 76%, not decreasing in the vicinity of QLN, furthermore the histograms keep symmetry. SOD is an observatory in a high latitude, so that the range for K=9 is the largest in those of the seven observatories, 1500 nT. On the other hand, the amplitude and variety of S_R are not so different from those of the observatories in low latitudes. Therefore, errors in S_R estimated on hand-scaling or computing affect K-values less than those of low-latitudes observatories affect.

4. Conclusion

OSM, one of the algorithms to derive K-values, seeks the optimal solution on the basis of the statistical principle. It has been evaluated with the data of seven observatories, following the decision of the meeting of IAGA/IUGG. It achieves good agreements with hand-scaled K-values; total agreements are about 60−75% with some exceptions.

There was appreciable discrepancy between hand-scaled and computed K-values for HER and OTT. This discrepancy is probably suggestion that hand-scaled values are wrong rather than computed values are. It may be a new problem of magnetic observatories.

The comparison reveals that OSM is proper to derive K-indices.

Acknowledgement: The author wishs to acknowledge technical staffs of each observatories and 'World Data Center A for Solar-Terrestrial Physics' for sending him geomagnetic one-minute data. The author also acknowledges Professor Sugiura (Tokai University) for valuable discussion.

references

Bartels, J., N. H. Heck and H. F. Johnston, The three-hour-range index measuring geomagnetic activity, Terr. Magn. Atmos. Elec., **44**, 411, 1939.

Hopgood, P. A., On the computer generation of geomagnetic K-indices from digital data, J. Geomag. Geoelectr., **38**, 861-871, 1986.

Kadokura, S., A new method to determine K-indices by using geomagnetic one-minute data (in Japanese), Mem. Kakioka Mag. Obs., **22**, 45-61, 1988.

Mayaud, P. N., Atlas des Indices K, IAGA Bull. **21**, IUGG Publication Office, Paris, 1967.

Mayaud, P. N. and M. Menvielle, A report on Km observatories visit, IAGA Bull. **32i**, 113-135, IUGG Publication Office, Paris, 1979.

Menvielle, M., Evaluation of algorithms for computer production of K-indices, Geophys. Tranzaction, **36**, 313-320, Eötvös Lorand Geophysical Institute of Hungary, Budapest, 1991.

Sucksdorff, C., R. Pirjola and L. Häkkinen, Computer production of K-indices based on linear elimination, Geophys. Tranzaction, **36**, 333-345, Eötvös Lorand Geophysical Institute of Hungary, Budapest, 1991.

Van Wijk, A. M. and T. Nagtegaal, K measurements by computer, J. Atmos. Terr. Phys., **39**, 1447-1450, 1977.

Walker, J. K., Adaptive separation of regular and irregular magnetic activety for K indices, J. Atmos. Terr. Phys., **49**, 1017-1025, 1987.

Wilson, L. R., An evaluation of digitally derived K-indices, J. Geomag. Geoelectr. **39**, 97-109, 1987.

Appendix

This is the instruction on comparison of computer methods of K-index production decided by the meeting held in Exeter (U.K.) on July 28, 1989.

period: from March 1985 to February 1986 (1 year)

observatories: Argentine Island (AIA), Belsk (BEL), Canberra (CAN), Crozet (CZ), Hartland (HAD), Hermanus (HER), Kerguelen (KG), Memambetsu (MMB), Newport (NEW), Nurmijärvi (NUR), Ottawa (OTT), Sodankylä (SOD),

histograms: computed K-values should be presented as histograms of the differences, $\Delta K = K(\text{computed}) - K(\text{hand-scaled})$, in the following 'formats':

(a1) histogram of all results (Mar. 85 ... Feb. 86)

(a2) histogram of Nov. 85 ... Feb. 86 (Winter)

(a3) histogram of May 85 ... Aug. 85 (Summer)

(a4) histogram of Mar. 85, Apr. 85, Sep. 85, and Oct. 85 (Equinox)

(a5) histogram of the selected 30 days
(Jul. 5 ... 14, Sep. 1 ... 10, and Dec. 22 ... 31, 1985).

The vertical scale is in % with 100% = 100 mm; the width of one column is 5 mm (columns: 0, −1, +1, −2, +2, ...)

(b1 − 4) for each three seasons, histograms of all results as functions of the time of the day, i.e. 8 different histograms according to the three hour intervals (UT00−03, 03−06, ..., 21−24).

The vertical scale is in % with 100% = 100 mm, the width of one column is 3 mm, and the space between zero-columns of two adjacent histograms is 24 mm.

(c) histograms of all results (Mar. 85 ... Feb. 86) as functions of the hand-scaled K-values, i.e. 9 different histograms.

The scale, width and space information is the same as in case (b).

drawings: abbreviated here

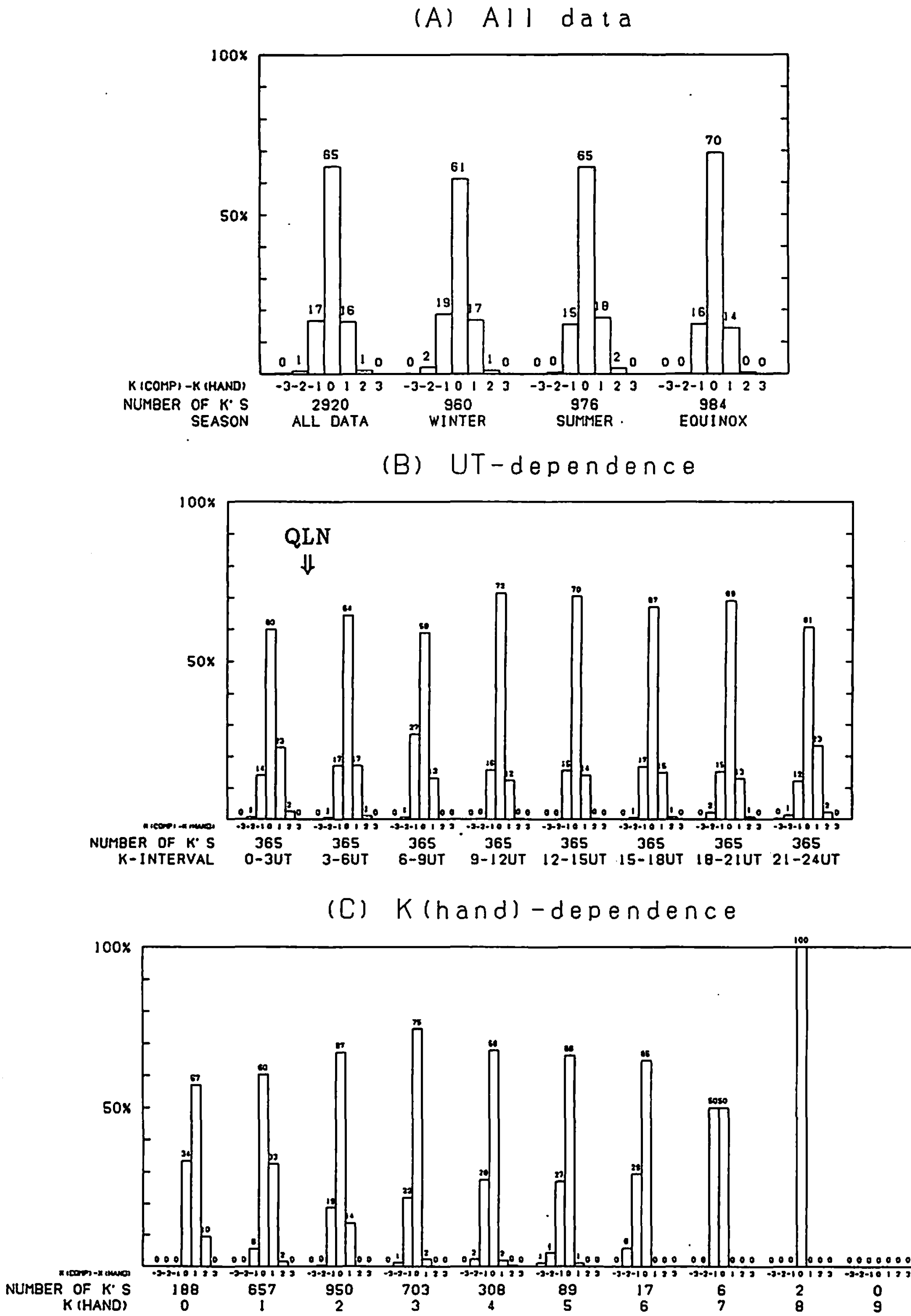

Figure 1 Histograms of the differences between the hand-scaled (K(hand)) and computer-derived (K(comp)) K-values for Memambetsu (MMB); (A) histograms of all results and of the results for three seasons; (B) histograms of all results as function of the time of the day, where 'QLN' indicates noon of the quasi-local-time explained in section 2; (C) histograms of all results as function of the K(hand).

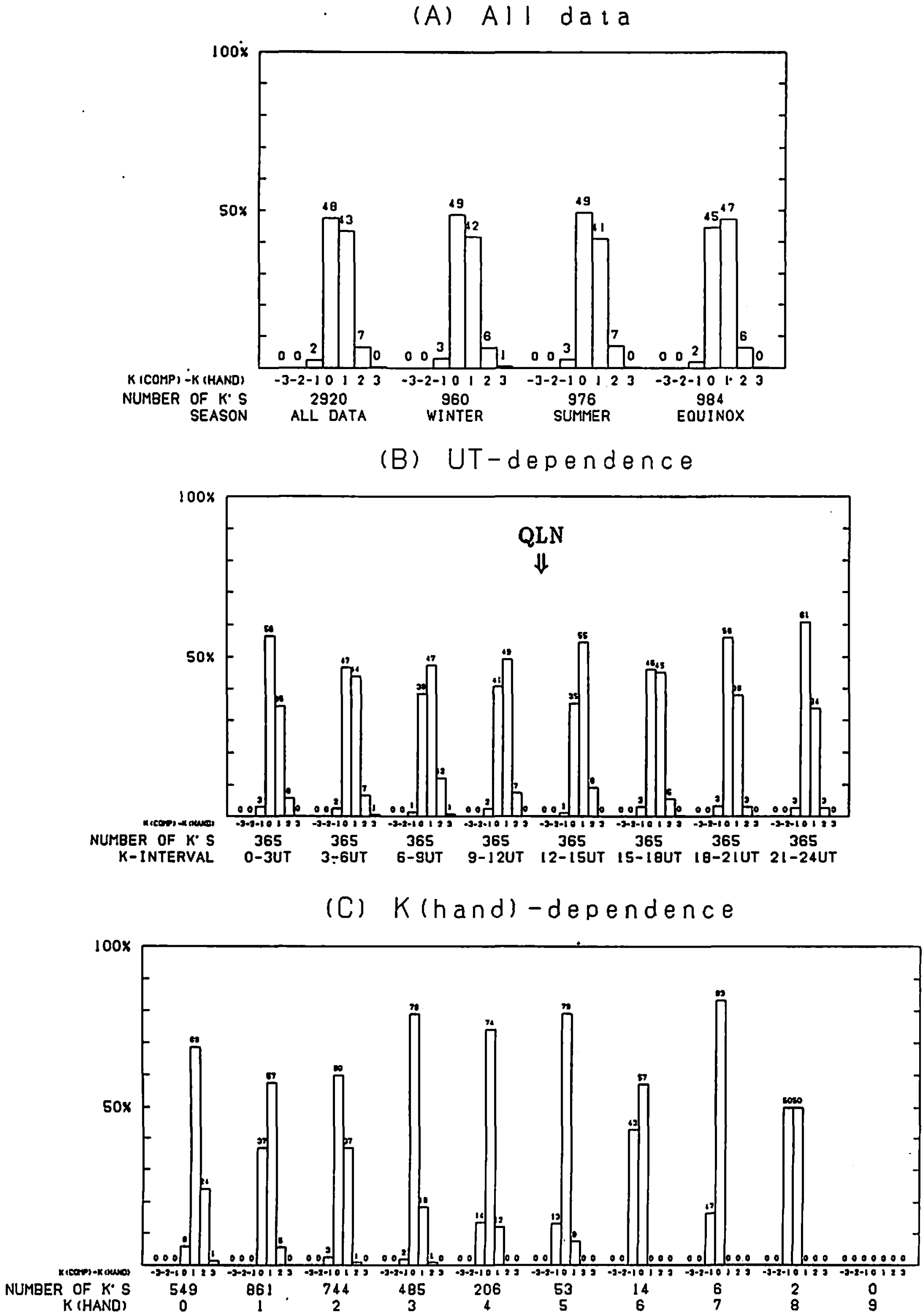

Figure 2 Histograms for Hermanus (HER), as in Figure 1.

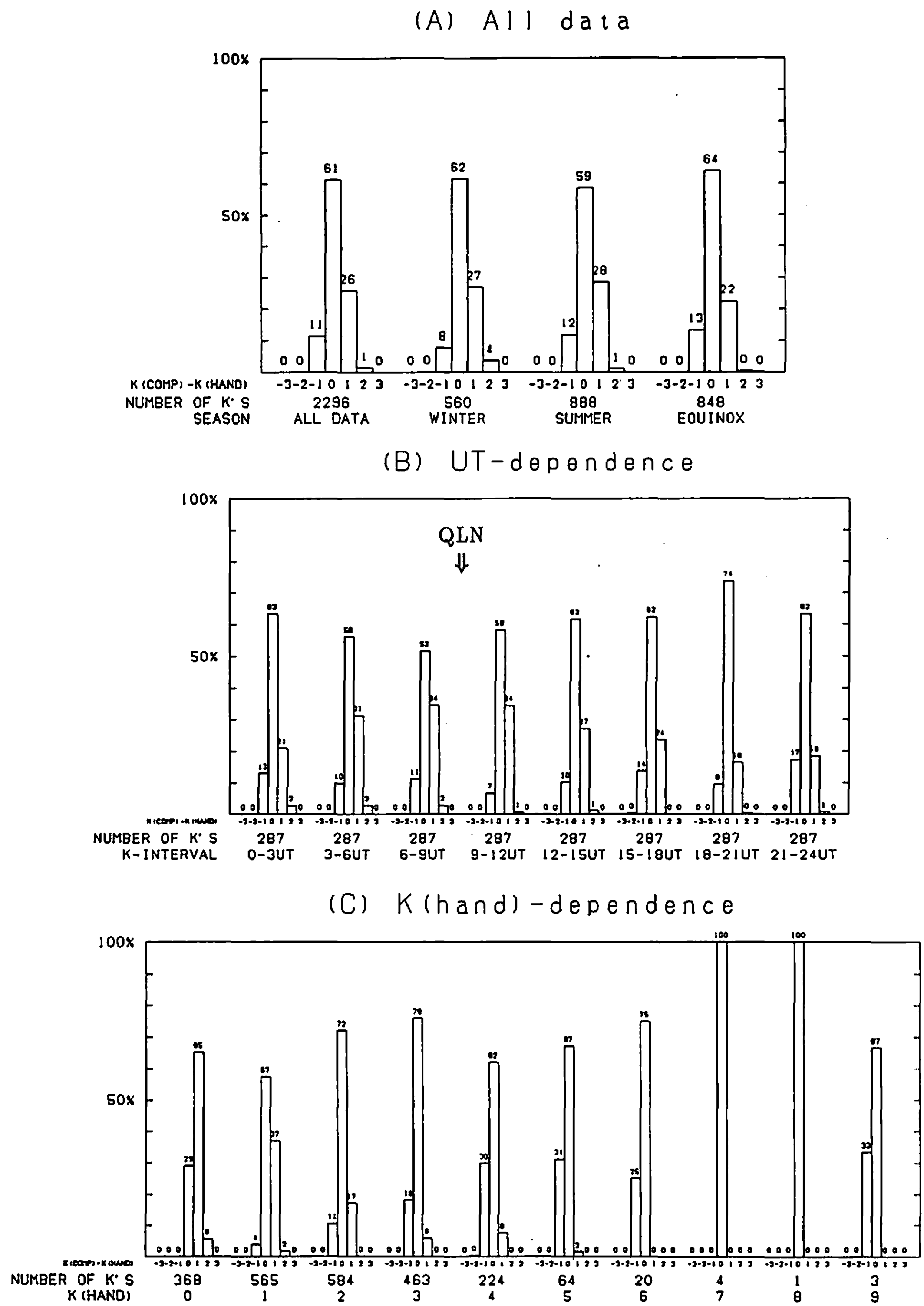

Figure 3 Histograms for Crozet (CZ), as in Figure 1.

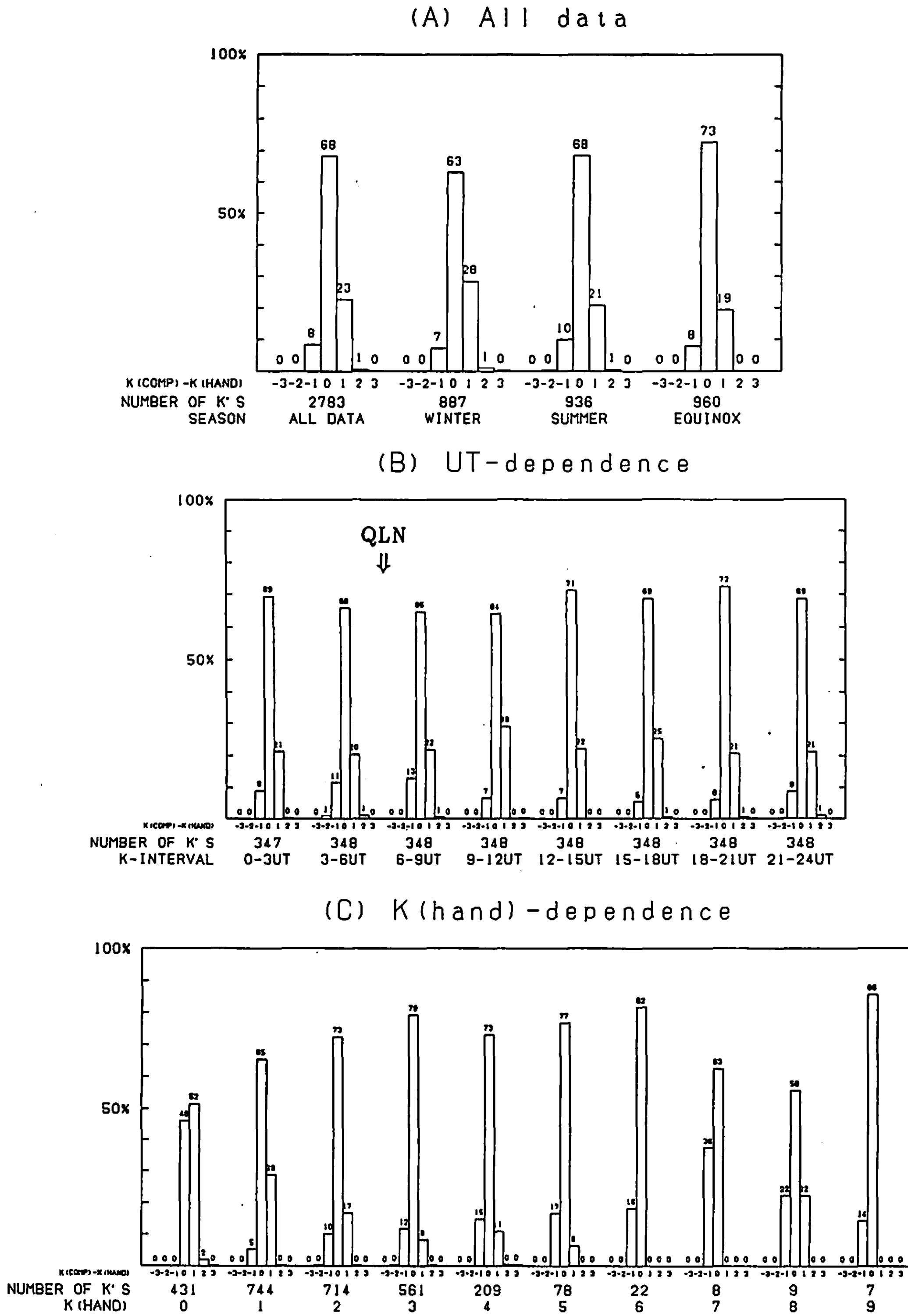

Figure 4 Histograms for Kerguelen (KG), as in Figure 1.

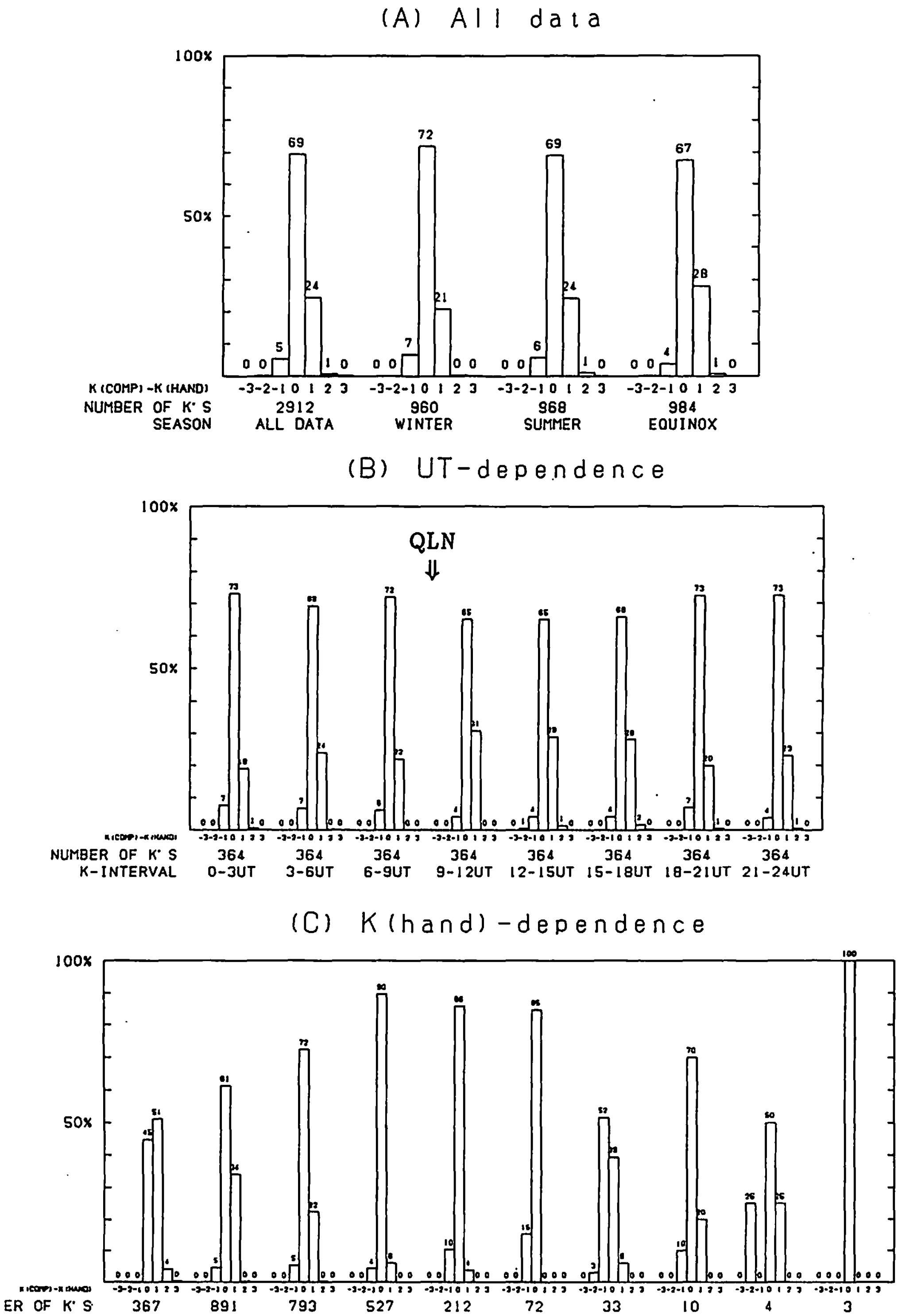

Figure 5 Histograms for Nurmijärvi (NUR), as in Figure 1.

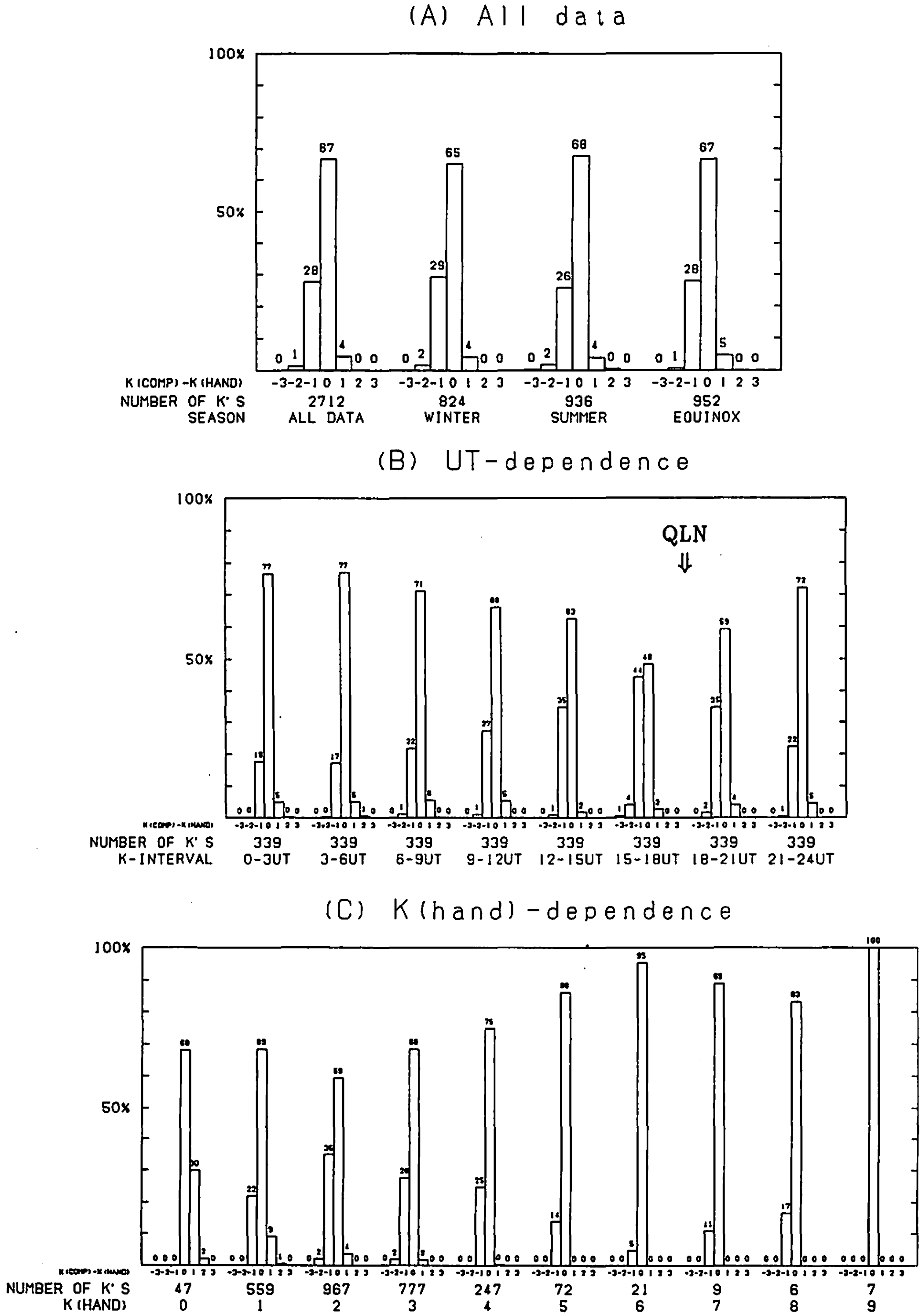

Figure 6 Histograms for Ottawa (OTT), as in Figure 1.

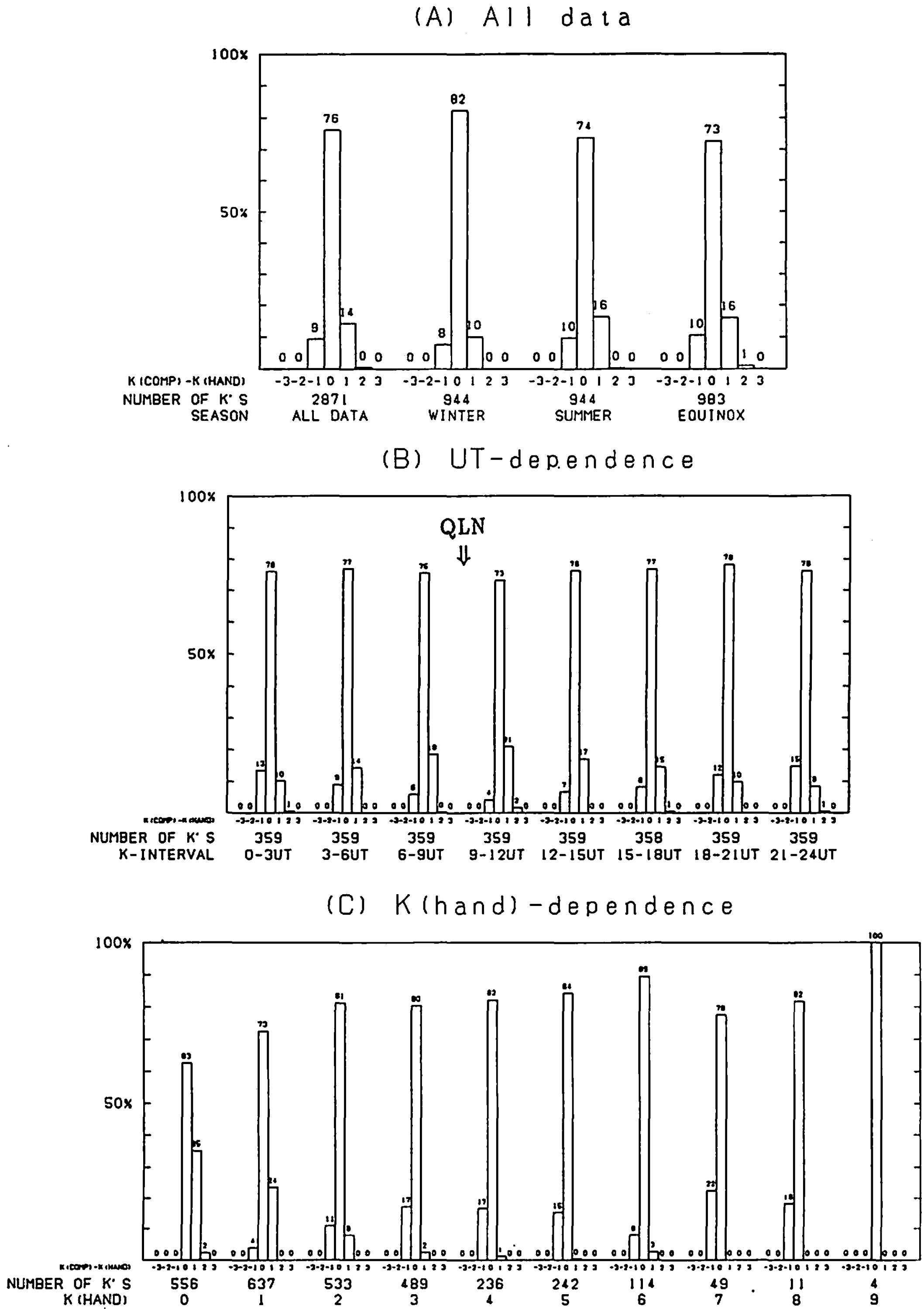

Figure 7 Histograms for Sodankylä (SOD), as in Figure 1.

直交化 S_R 法によって得られたK指数の評価

門倉真二

概　　要

K指数を計算するアルゴリズムの一つ，直交化 S_R 法(OSM)のテストを行なった．テスト方法は，IUGG/IAGAの打ち合わせでの取り決めにしたがい，世界各地7観測点(Memambetsu, Hermanus, Crozet, Kerguelen, Nurmijärvi, Ottawa, Sodankylä)の1985年3月から1986年2月までの期間について，観測者によるK指数(K(hand))との差のヒストグラムによった．結果はHermanus(HER)を除き(1) 60から75%の一致，(2) 1を越える差の起こる頻度は2%以下で十分に少ない，と良好であった．HERの結果がよくないことは，むしろK(hand)が正しくない事を示唆するものらしい．OSMは，統計原理上線形の方法の中では最適化されたものであるが，実際のデータにおいても十分使えることがわかった．