CORRELATION OF WIND DIRECTION OBSERVATIONS AND OTHER SURFACE ELEMENTS

by O. ESSENWANGER (*)

Summary — Many tabulations of climatological data present the distribution of specified elements with wind direction. It is shown that statistical characteristics expressing the relationship between cloud cover or visibility and *wind direction* can be computed with meaningful meteorological interpretation. The U.S. Navy Summaries of Monthly Aerological Records (SMAR) are used. Several distribution characteristics are suggested and the method of computing is discussed with samples given.

The linear correlation coefficient r_{yx} between cloud cover or visibility classes and wind directions seems to be a fairly good characteristic but because of the nonlinearity of the relationship, the correlation ratio η_{yx} is more efficient.

Climatological features may also be studied by the linear correlation coefficient R_{yx}, which expresses the unweighted relation between cloud cover (visibility) classes and wind direction. The mathematical formulation of the relationship by polynomials indicates that higher order terms affect the functional connection.

Zusammenfassung — In vielen klimatologischen Tabellierungen werden Häufigkeitsverteilungen verschiedener meteorologischer Elemente in Abhängigkeit von Windrichtung geben. In diesem Artikel wird gezeigt, dass statistische Parameter berechnet werden können, welche die Beziehung zwischen *Windrichtung* und Himmelsbedeckung beziehungsweise Sichtweite mit sinnvoller meteorologischer Interpretation ausdrücken. Die Studien wurden an Monatszusammenfassungen aerologischer Daten (SMAR) des Marinewetterdienstes der U.S. durchgeführt. Mehrere statistischen Parameter werden vorgeschlagen und die Berechnungsweise mit Beispielen erläutert.

Der lineare Korrelationskoeffizient r_{xy}, der die Windrichtungshäufigkeit in Abhängigkeit von Klassen der Himmelbedeckung bzw. der Sichtweite ausdrückt, erscheint brauchbar, wenn auch nicht zuviel erwartet werden darf, weil die Beziehung vielfach nichtlinear ist. Daher ist das Korrelationsverhältnis η_{xy} mehr zu empfehlen.

Klimatologische Zusammenhänge können auch mit Hilfe des Koeffizienten R_{xy} ausgedrückt werden. Dieser stellt den linearen Anteil einer Polynomreihe zwischen Klassenwerten der Himmelbedeckung (bzw. Sichtweite) dar, ohne den einzelnen Klassen ein Gewicht ihrer Häufigkeit entsprechend zu geben. Bei der vollen Berechnung der Beziehung durch Polynome kann man den nichtlinearen Anteil erkennen.

Das spezielle Ergebnis der vorliegenden Arbeit ist, dass die Darstellung der Sichtweite oder Himmelsbedeckung bei Windrichtungsklassen wenig klimatologischen Effekt offenbart, da die Form der Verteilungsfunktion von Sichtweite und Himmelbedekkung dominiert und klimatologische Effekte in den Hintergrund treten lässt.

(*) Dr. OSKAR ESSENWANGER, 809 Watts Drive, SE, *Huntsville*, *Alab.* (USA).

I. *Introduction.*

The relations between wind direction and other meteorological elements have been a subject of interest in climatological history. Wind roses in the form of wind « star » (11) or thermal wind roses, etc. [see also CONRAD & POLLAK (2)] are well known methods of presentations in climatology. They present much information and can be quite useful to the climatologist.

The wind speed in the surface layer is influenced by various factors (friction, orography, etc.) and represents a disturbed streamflow. Despite this fact, the relation between the surface wind direction as a parameter of the air flow and other meteorological elements is important and many tabulations of detailed frequency distributions have been accumulated at the National Weather Records Center (NWRC) in the form of the Summary of Monthly Aerological Records (SMAR) (Table 1). While it is possible to study individual wind roses and a few wind roses at different stations, a multitude of tabulations as accumulated in the SMAR is hardly convenient for simple review.

The application of statistical methods can remove this deficiency. If it were possible to develop a few characteristics, the resumé of the tabulations could then be given in a handy summary table or a map form. This subsequently would give information where the relation between wind direction and other elements is strong and useful, while other locations may show little relation. In this way it would be possible to summarize detailed tabulations in condensed form for locations of strong relationship and omit those where the weak relationships are indicated.

The application of statistical methods implies another advantage. The interpretation of a wind rose presentation at an individual station will probably cause no difficulty to the climatologist. Review of a multitude of wind roses can hardly be done objectively without employing some kind of mathematics which permits concentration of information. Therefore, the aim is to present pertinent characteristics, supporting the objective interpretation of the meteorological data.

The problem is stated and the solution indicated. Inspection of the material, however, reveals the mathematical difficulty connected with the problem. The wind direction is a typical periodic function and many of the developed statistical techniques cannot be applied. Use of the regular scale from 0^{o} to 360^{o} introduces problems like the computation of an arithmetic mean. Conventional arithmetic computation gives a mean value which may formally be correct but may be reasonably doubtful.

A common refuge is to split the wind into components like the zonal and meridional component, as does a very recent published article by RODEN & GROVES (12). The relation of the components could then be combined to a multiple correlation coefficient, which can express the linear relation only. Including a nonlinear relationship is difficult, but even when accomplished, parts of the functional relation may be lost by the split into components. For there may exist a strong relationship between wind direction and another meteorological element, while the wind speed may offset it. By splitting the wind velocity into components, the result would probably show the total influence of the wind, but the relation between flow pattern and elements may be obscured.

The author has discussed this problem recently (3,4) and by methods developed in said reports (3,4) application of statistical techniques can be made. This article investigates the feasibility of deriving statistical characteristics between wind direction and cloud amount or visibility. The derivation of characteristics between wind direction and other elements is held for future investigation.

Table 1 - *Summary of monthly aerological records* (SMAR).

Station: Washington Natl. Ap. Elevation: 34.1m; Latitude: 38°51'N; Longitude: 77°2'W. — Period of Record: Jan. 1945 thru Dec. 1952

Dir. code	Frequency of wind speed groups (Knots)								Frequency of total cloud amounts in tenths												Wind
	3.7	8.12	13.20	21.30	31.40	> 40	Calm	Total	0	1	2	3	4	5	6	7	8	9	10	Total	Dir.
11	793	784	798	139	8			2522	566	117	85	71	59	63	68	93	97	120	1014	2353	N
12	650	1378	1579	230	6			3843	576	92	108	92	65	69	76	81	147	164	2117	3587	NNE
22	927	1566	1048	157	4	1		3703	413	78	87	93	58	78	77	94	160	168	2102	3408	NE
32	758	954	637	48	6	2		2405	199	62	54	66	44	57	53	73	110	124	1397	2239	ENE
33	594	678	331	24	1			1628	138	47	46	50	51	40	38	53	68	86	859	1476	E
34	529	657	298	27	9	3		1523	159	51	56	53	56	43	41	54	77	106	709	1405	ESE
44	957	1015	403	23				2398	298	92	111	96	59	57	68	96	130	137	1047	2191	SE
54	1453	2162	930	71	2	1		4619	630	198	259	237	168	163	184	199	270	299	1537	4144	SSE
55	2254	2870	1717	216	3			7060	1189	272	325	290	247	190	267	287	445	448	2305	6265	S
56	2548	2581	2012	451	16	1		7609	1591	293	330	317	205	183	278	322	500	510	2449	6978	SSW
66	1789	1075	465	64	5			3398	834	154	168	121	112	108	107	145	203	227	947	3126	SW
76	1032	847	373	44	3			2299	533	120	104	100	97	82	98	126	144	165	543	2112	WSW
77	956	893	707	168	11			2735	701	163	139	117	117	92	100	125	165	174	569	2462	W
78	1009	1667	2139	790	53			5658	1645	345	319	294	216	199	227	260	283	325	977	5090	WNW
88	1799	2586	2211	433	13			7042	2118	409	333	293	218	194	208	236	339	350	1673	6371	NW
18	1181	1865	1572	233	4			4855	1334	223	208	179	172	121	149	155	221	238	1389	4389	NNW
00							6414	6414	2084	251	283	227	153	114	212	217	300	288	1653	5782	Calm
	19229	23578	17220	3118	144	8	6414	69711	15008	2967	3015	2696	2097	1853	2251	2616	3659	3929	23287	63378	Total
	27.6	33.8	24.7	4.5	.2	.0	9.2	100.0	23.7	4.7	4.8	4.3	3.3	2.9	3.6	4.1	5.8	6.2	36.7	100.0	% of Tot

Dir. code	Frequency of weather								Frequency of ceilings (feet)												Wind
	Rain	Rain Showers	Drizzle	Zr. Zl. E.EW. Lc.	Snow Sg.SW. Sp.	Hail AP	Thunder	No Weather	0-150	151-350	351-550	551-1050	1051-2050	2051-3050	3051-10500	10501-20500	20501-30500	> 30500 and unl im	> 6/10 Clouds	Total	Dir.
11	263	68	61	35	59		17	2059	9	25	57	227	150	116	458	102	76	224	1078	2522	N
12	586	107	181	64	93		22	2855	9	60	201	586	385	244	722	90	75	391	1080	3843	NNE
22	558	108	147	39	53	1	29	2825	19	61	179	533	411	270	724	127	107	362	910	3703	NE
32	301	93	92	11	27		10	1901	9	34	86	299	270	226	542	82	98	222	537	2405	ENE
33	165	78	47	5	21		13	1318	10	14	50	170	156	123	341	72	84	164	444	1628	E
34	156	72	37	8	10		11	1242	2	23	37	97	149	110	289	70	69	203	474	1523	ESE
44	178	132	28	14	15		22	2036	3	34	39	113	154	128	494	140	134	347	812	2398	SE
54	249	154	41	7	14	1	27	4154	18	40	65	162	197	140	641	195	280	959	1922	4619	SSE
55	389	216	58	19	16		31	6371	48	92	100	217	267	174	1066	440	380	1240	3036	7060	S
56	344	272	40	7	24		28	6928	39	64	87	274	270	144	1058	538	488	1312	3335	7609	SSW
66	132	148	18	4	14		34	3077	16	33	41	109	101	62	517	234	165	449	1671	3398	SW
76	65	95	9	2	10		17	2121	3	13	23	59	45	32	422	102	95	320	1185	2299	WSW
77	57	89	13	3	18		22	2556	6	15	22	48	52	36	558	125	87	242	1544	2735	W
78	102	128	17	4	75		16	5338	1	12	34	73	78	74	1141	172	168	477	3428	5658	WNW
88	268	207	43	17	84	1	37	6440	9	28	80	178	178	145	1334	234	236	611	4009	7042	NW
18	321	100	34	14	84		15	4319	6	12	60	194	170	155	913	144	138	518	2545	4855	NNW
00	225	92	82	7	23		10	5993	116	100	141	154	134	105	802	255	272	742	3593	6414	Calm
	4359	2159	948	260	640	3	361	61533	323	660	1302	3493	3167	2284	2022	3122	2952	8783	31603	69711	Total
	6.3	3.1	1.4	.4	.9	.0	.5	88.3	.5	.9	1.9	5.0	4.5	3.3	2.9	4.5	4.2	12.6	45.3	100.0	% of Tot

Dir. code	0 - 1/2 Miles Calm	3.12	13.20	> 20	Total	5/8 - 1 Miles Calm	3.12	13.20	> 20	Total	1 1/4 - 2 Miles Calm	3.12	13.20	> 20	Total	2 1/4 - 3 Miles Calm	3.12	13.20	> 20	Total	Visibility wind spd. wind dir.
11		15	1	5	21		23	8	3	34		79	27	7	113		207	55	10	272	N
12		23	4	5	32		40	19	7	66		154	96	13	263		314	181	42	537	NNE
22		31	4	1	36		37	10	1	48		174	50	11	235		370	138	17	525	NE
32		22			22		17	3	1	21		70	24	1	95		208	62	8	278	ENE
33		12			12		9	1		10		37	12		49		122	25	7	154	E
34		4			4		6	1		7		28	2		30		94	7	3	104	ESE
44		10	1		11		18	2		20		40	2		42		92	14	2	108	SE
54		35	1		36		21			21		53	4		57		172	18	1	191	SSE
55		75	7	1	83		40	2		42		152	10	1	163		332	39	3	374	S
56		50	5		55		59			59		176	15		191		432	39		471	SSW
66		29	1	1	31		29			29		96	4		100		205	5	1	211	SW
76		6			6		13	2	1	16		24	2		26		57	2		59	WSW
77		10		1	11		3	1		4		30		3	33		62	5	3	70	W
78		5	3		8		6			6		26	7	2	35		82	16	6	104	WNW
88		13	3		16		20	4	1	25		80	16	1	97		189	34	5	228	NW
18		11	7	3	21		15	6	2	23		67	14	3	84		185	48	12	245	NNW
00	219				219	143				143	537				53	1004				1004	Calm
	219	351	37	17	624	143	356	59	16	574	537	1286	285	42	2150	1004	3123	688	120	4935	Total freq
	.3	.5	.1	.0	.9	.2	.5	.1	.0	.8	.8	1.8	.4	.1	3.1	1.4	4.5	1.0	.2	7.1	% of Tot

DIR. CODE	4 - 6 Miles Calm	3.12	13.20	> 20	Total	7-10 Miles	> 10 Miles	Total Obs. Report Visib.	0 - 1/2 Miles caused by f. gf. if Calm	3.12	13.20	> 20	Total	5/8 - 6 Miles caused by f. gf. if Calm	3.12	13.20	> 20	Total	0-6 Mi. Caused by Bs. Bn. Bd.	0-6 Mi. Caused By H. K. D.	Visibility wind spd. wind dir.
11		308	92	16	416	700	966	2522		15			16		208	111	16	335	4	491	N
12		457	279	37	773	1096	1076	3843		19		1	19		372	328	66	766	13	829	NNE
22		612	231	36	879	1203	777	3703		28	1		29		462	214	35	711	6	976	NE
32		356	120	12	488	831	670	2405		19			19		230	97	11	338		518	ENE
33		210	47	5	262	543	598	1628		9			9		155	51	9	215		240	E
34		165	28	6	199	497	682	1523		4			4		112	16	3	131		193	ESE
44		296	31	7	334	827	1056	2398		9			9		138	12	5	155		314	SE
54		492	66	14	572	1519	2223	4619		35	1		36		171	35	2	208		646	SSE
55		852	116	20	988	2371	3039	7060		75	7		82		331	65	6	402		1246	S
56		1088	142	20	1250	2592	2991	7609		50	5		55		341	71	3	415		1683	SSW
66		558	26	3	587	1147	1293	3398		29	1		30		180	11		191		803	SW
76		229	5	1	235	634	1323	2299		6			6		92			92		253	WSW
77		164	13	2	179	651	1787	2735		8			8		83	3	1	87		204	W
78		220	42	14	276	1122	4107	5658		3			3		104	11	3	118	1	263	WNW
88		502	106	21	629	1860	4187	7042		10			10		284	40	2	326	5	597	NW
18		424	126	14	564	1457	2461	4855		9	1	1	11		223	89	15	327	1	538	NNW
00	1815				1815	1810	886	6414	212				212	777				777		3270	Calm
	1815	6933	1470	228	10446	20860	30122	69711	212	328	16	2	558	777	3486	1154	177	5594	30	13064	Total freq
	2.6	9.9	2.1	.3	15.0	29.9	43.2	100.0	.3	.5	.0	.0	.8	1.1	5.0	1.7	.3	8.0	.0	18.7	% of Tot

II. *Relation Between Cloud Cover or Visibility and Wind Direction.*

a) *Data Selection and Treatment.*

The author realizes that it would have been desirable to establish the statistical technique first for another element like air temperature. Temperature distributions approximate simpler statistical laws such as the normal distribution law (Gauss). Urgent needs for representations of cloud cover and visibility by wind direction, where both elements drastically differ from the normal distribution law, required that the problem be attacked without too much preparatory study on temperature. The interpretation and discussion which follows necessarily involve relations that are somewhat complex because mathematical terms of nonlinear character are dominant.

One other point should be stressed. The SMAR tabulations combine a priori all 24-hourly observations and cannot be used to study diurnal variations. Thus phenomena related to diurnal variation of the wind are obscured by the SMAR. If studies of diurnal effects are needed, the same statistical techniques can be used. The main purpose here was the development of the statistical characteristics and to determine, through use of a few samples, whether those characteristics are suitable for climatological interpretation.

The observation used here-in have been checked previously and so far as missing observations are concerned, no bias could be discovered. Thus, the set of data can be accepted as reasonably uniform samples.

b) Visibility and Cloud Cover by Wind Direction Elements.

Before the author had developed the technique of computing statistical characteristics for periodic functions, tabulations as exhibited in Tables 2 and 3 have been possible to be computed and assembled. They provide the arithmetic mean values (y_i) of cloud cover (Table 2) and visibility (Table 3) by wind direction classes. Two pilot stations, Washington, D. C., National Airport, January 1945 through December 1952 and Seattle, Washington, March 1945 through December 1952, have been chosen. The Tables 2 and 3 represent already a condensed form of information, as the arithmetic mean reflects some condensed information about the frequency for a wind direction class (16 points of the compass). It is also an objective characteristic, although the numerical values themselves must be interpreted by the climatologist with some care. For example, an average of 6.0 gives no indication that cloud cover of 6 tenths occurs very often. The cloud cover frequency curve is a typical so-called « U-shaped » distribution, where the modes are at 0 and 10 tenths and the arithmetic mean value is a compromise between both modes. In other words, the mean value is more related to the ratio between those modes than to the frequency of occurrence. The columns with the mean cloud cover exhibit the variations by wind direction.

Formally one can also compute the standard deviation, (σ_{yi}) which has a unique property for a normal distribution (Gauss). Certain ranges of the standard deviation contain determined percentages of the total available material. It is quite obvious that this property of the standard deviation is not filled for a U-distribution. It may express a certain relative measure of scatter but precise interpretations are difficult. Similar considerations apply to visibility.

On the above basis, Table 2 may be interpreted as follows. The relations of wind direction versus cloud cover for two stations, Washington, D. C. near the east coast and Seattle, Washington on the west coast of the U.S.A. are compared.

TABLE 2 - *Wind Direction versus Cloud*
Washington, D. C. NAP (January 1945 - December 1952)

Compass Points	Wind Direction Degrees	Annual		January		July	
		$\bar{y}_j$	$\sigma_{\bar{y}_j}$	$\bar{y}_j$	$\sigma_{\bar{y}_j}$	$\bar{y}_j$	$\sigma_{\bar{y}_j}$
16	350-11	6.0	4.3	7.8	3.5	6.4	3.6
1	12-34	7.3	4.0	8.3	3.3	7.2	3.7
2	35-56	7.6	3.7	7.8	3.6	7.2	3.6
3	57-79	7.9	3.4	9.3	2.0	7.1	3.5
4	80-102	7.6	3.6	9.3	2.2	7.0	3.6
5	103-124	7.2	3.7	8.6	2.7	5.9	3.8
6	125-147	6.8	3.9	7.9	3.5	6.7	3.6
7	148-169	6.2	3.9	7.3	4.0	5.9	3.6
8	170-192	6.1	4.0	7.1	3.7	5.8	3.8
9	193-214	5.8	4.2	7.1	4.0	5.5	3.8
10	215-237	5.3	4.2	6.8	4.2	5.5	3.8
11	238-259	5.2	4.1	5.6	4.4	6.3	3.7
12	260-282	4.8	4.1	4.9	4.2	6.0	3.3
13	283-304	4.3	4.0	4.7	4.1	5.0	3.8
14	305-327	4.6	4.2	4.9	4.4	5.0	3.8
15	328-349	5.1	4.3	6.2	4.2	5.3	3.9

Seattle, Washington (March 1945 - December 1952)

Compass Points	Wind Direction Degrees	Annual		January		July	
		$\bar{y}_j$	$\sigma_{\bar{y}_j}$	$\bar{y}_j$	$\sigma_{\bar{y}_j}$	$\bar{y}_j$	$\sigma_{\bar{y}_j}$
16	350-11	4.5	4.0	4.9	4.2	2.6	3.6
1	12-24	6.0	4.0	6.7	4.3	4.0	3.8
2	35-56	6.3	4.0	7.2	4.0	4.8	4.2
3	57-79	7.8	3.4	7.6	3.8	6.6	3.8
4	80-102	7.9	3.2	8.2	3.2	6.7	4.1
5	103-124	8.0	3.2	9.4	1.8	5.4	4.5
6	125-147	7.9	3.2	9.4	1.9	5.5	4.0
7	148-169	8.4	2.8	9.6	1.5	4.8	4.3
8	170-192	8.7	2.4	9 4	1.4	7.2	3.6
9	193-214	8.7	2.3	9.1	1.6	7.9	3.3
10	215-237	8.2	2.8	8.1	2.5	7.3	3.6
11	238-259	7.8	3.0	8.6	2.2	5.8	4.0
12	260-282	7.6	3.4	5.7	4.2	5.7	4.0
13	283-304	7.5	3.5	7.4	3.1	7.2	3.1
14	305-327	5.6	4.0	5.8	4.0	3.4	4.0
15	328-349	4.3	4.1	5.4	4.0	2.0	3.1

$\bar{y}_j$ Mean cloud cover in tenths sky cover.
$\sigma_{\bar{y}_j}$ Standard deviation in tenths sky cover.

Table 3 - *Wind Direction versus Visibility.*
Washington NAP (January 1945 - December 1952)

Wind Direction Degrees	Annual $\bar{y}_j$	Annual $\sigma_{\bar{y}_j}$	January $\bar{y}_j$	January $\sigma_{\bar{y}_j}$	July $\bar{y}_j$	July $\sigma_{\bar{y}_j}$
350-11	4.8	1.3	4.1	1.6	4.9	1.2
12-34	4.5	1.4	3.9	1.5	4.9	1.1
35-56	4.4	1.3	4.1	1.3	4.8	1.0
57-79	4.6	1.2	4.0	1.2	5.0	1.1
80-102	4.9	1.2	4.2	1.2	5.2	0.9
103-124	5.1	1.1	4.5	1.3	5.5	0.8
125-147	5.1	1.1	4.6	1.4	5.4	0.8
148-169	5.2	1.1	4.4	1.5	5.5	0.7
170-192	5.3	1.2	4.5	1.5	5.3	0.8
193-214	5.0	1.2	4.7	1.4	5.2	0.8
215-237	4.9	1.2	4.5	1.4	5.2	0.9
238-259	5.4	1.0	4.9	1.3	5.4	0.8
260-282	5.5	0.9	5.3	1.1	5.5	0.8
283-304	5.6	0.8	5.4	1.0	5.6	0.8
305-327	5.4	0.9	5.2	1.1	5.4	0.8
328-349	5.2	1.1	4.9	1.2	5.4	0.8

Seattle, Washington (March 1945 - December 1952)

Wind Direction Degrees	Annual $\bar{y}_j$	Annual $\sigma_{\bar{y}_j}$	January $\bar{y}_j$	January $\sigma_{\bar{y}_j}$	July $\bar{y}_j$	July $\sigma_{\bar{y}_j}$
350-11	5.1	1.2	4.7	1.5	5.7	0.6
12-34	5.0	1.3	4.2	1.8	5.6	0.6
35-56	4.7	1.5	4.2	1.9	5.5	0.8
57-79	4.6	1.5	4.3	1.8	5.3	0.6
80-102	4.6	1.5	4.2	1.8	5.2	0.9
103-124	4.8	1.2	4.3	1.7	5.2	1.0
125-147	4.8	1.1	4.4	1.2	5.2	0.8
148-169	4.8	1.0	4.6	1.0	5.2	0.9
170-192	4.8	0.9	4.7	0.9	5.2	0.8
193-214	4.9	0.9	4.7	0.9	5.2	0.8
215-237	4.9	1.0	4.8	0.9	5.1	0.9
238-259	4.9	1.0	4.9	0.8	5.2	0.9
260-282	4.7	1.2	4.4	0.8	5.0	0.9
283-304	4.8	1.4	4.4	0.9	5.3	0.6
305-327	4.9	1.3	4.5	1.5	5.6	0.7
328-349	5.1	1.2	4.4	1.6	5.7	0.6

$\bar{y}_j$ Mean Visibility in Miles.
$\sigma_{\bar{y}_i}$ Standard Deviation in Miles.

The annual mean value of cloud cover by wind direction shows a greater variation at Seattle. The standard deviation is practically constant at Washington, D. C. The latter fact may be influenced by the maximum mean value, which is 7.9 at Washington and 8.7 at Seattle. Hence, the standard deviation for this 8.7 value should be expected smaller than for 7.9, as 8.7 is closer to the 10.0 boundary. This causes a larger variation of the standard deviation at Seattle.

Table 2 reveals further that the highest value for the mean cloud cover at Washington appears with easterly winds, while for Seattle this extreme occurs with southerly winds.

January and July portray similar patterns at both stations. Differences from the annual value appear in January at Washington and in July at Seattle. January shows some variation for the standard deviation for Washington compared with the year or July. A more complex relation between the mean value of cloud cover and the wind direction, indicated by the 3 peaks at easterly, southerly and westerly directions is obvious at Seattle in July.

Table 3 pictures the relation of visibility versus wind direction. From this table one would conclude that there is little relationship. In the later part of the report it will be proven that this conclusion is deceptive. The frequency distribution of visibility in a linear scale does not follow a normal distribution and this is a typical sample where the mean value is a poor characteristic and hence, the conclusions therefrom may be wrong.

c) *Wind Direction by Classes of Visibility and Cloud Cover.*

The variation of the arithmetic mean value of cloud cover or visibility by wind direction classes have been discussed in the preceding section. The variation of the mean wind direction by cloud cover classes or classes of the visibility shall be investigated now, using the technique established to compute mean direction values without splitting the data into components. Results are given in Tables 4 and 5.

Before starting the interpretation, Figures 1 through 4 should be inspected. They represent the frequencies of the wind direction by cloud cover or visibility classes. Nearly all distributions display two modes. This must be kept in mind in the interpretation of the Tables 4 and 5. However, the mean value is now either a part of one overweighing group (collective), or the compromise between one and the other. Also the standard deviation tends toward normal representation.

Consider now Table 4, the mean wind direction value and cloud cover. The greater cloud cover for easterly winds in the annual summary for Washington, D. C. is very striking by the abrupt change of the mean value from the 9 to 10 tenths cloud cover class.

This discontinuity does not exist for January or July. One would not suspect from Table 2, that in July the southerly winds are connected more closely with cloudy weather than the westerly winds. This can only be seen from Table 4. Figure 1 describes the details of this result. The group with mode at NW gradually is decreasing toward 10 tenths cloud cover. This causes the mean value (the black arrow in the graph) to shift from 232° to 171°. The percentage numbers in the right top of the box give the total of occurrence of the cloud class. A nothwest collective is found for the 0 tenths cloud class, while it is weak in the 10 tenths class. This effect is completely obscured in the form of Table 2, but clearly expressed by the variation of the mean value of wind direction in Table 4.

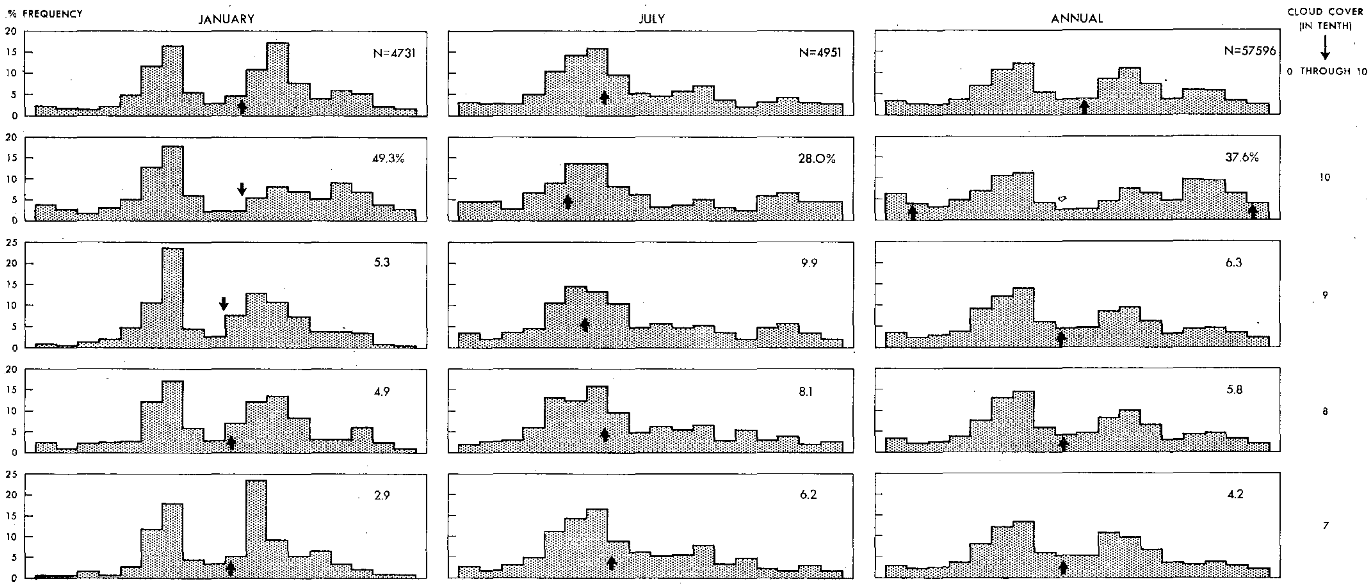

Fig. 1 - *Washington (D.C.) National Airport, Jan.* 1945 - *Dec.* 1952 *wind direction versus clouds.*

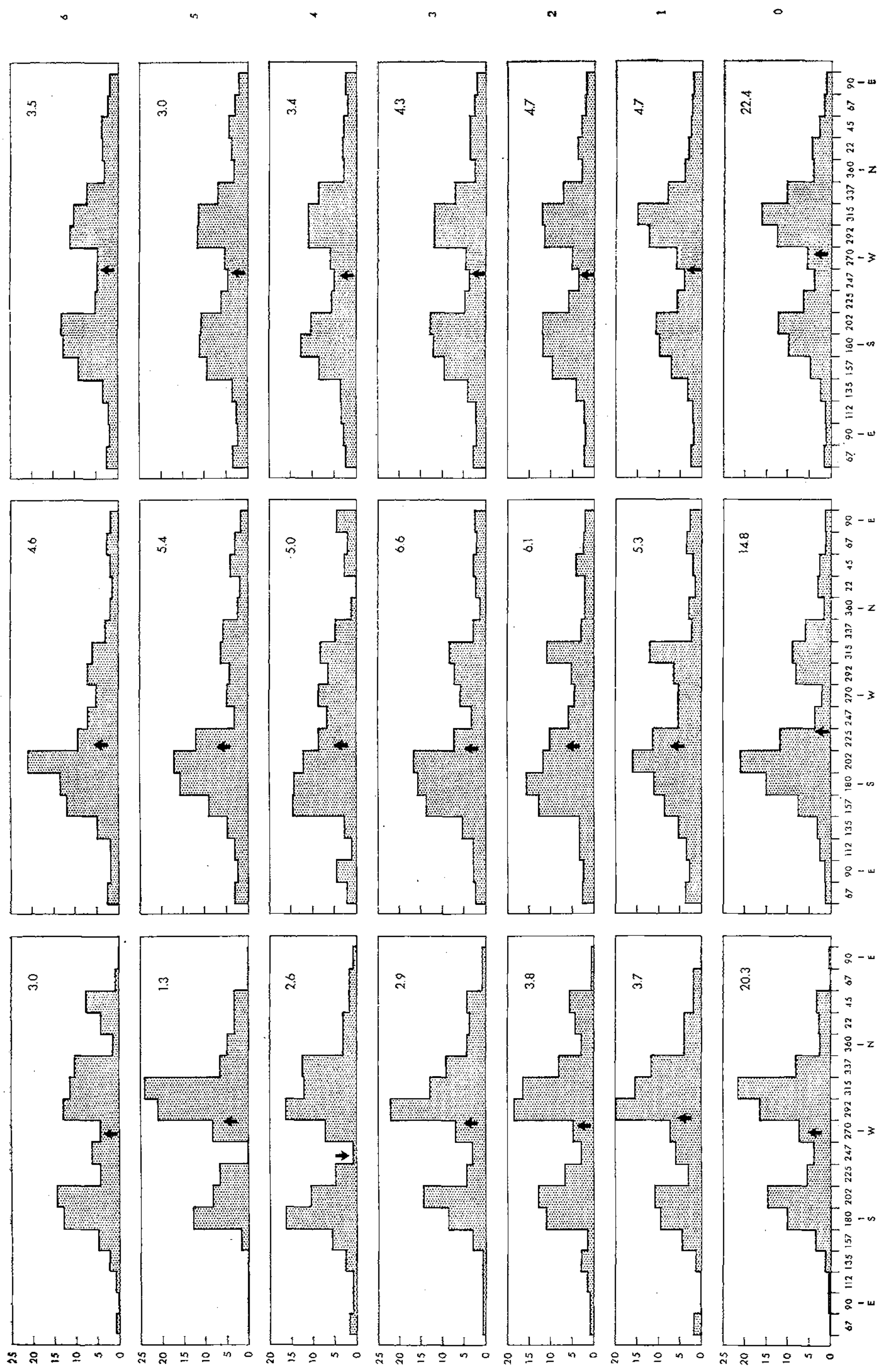
6
5
4
3
2
1
0
3.5
3.0
3.4
4.3
4.7
4.7
22.4
4.6
5.4
5.0
6.6
6.1
5.3
14.8
3.0
1.3
2.6
2.9
3.8
3.7
20.3

Fig. 1 b.

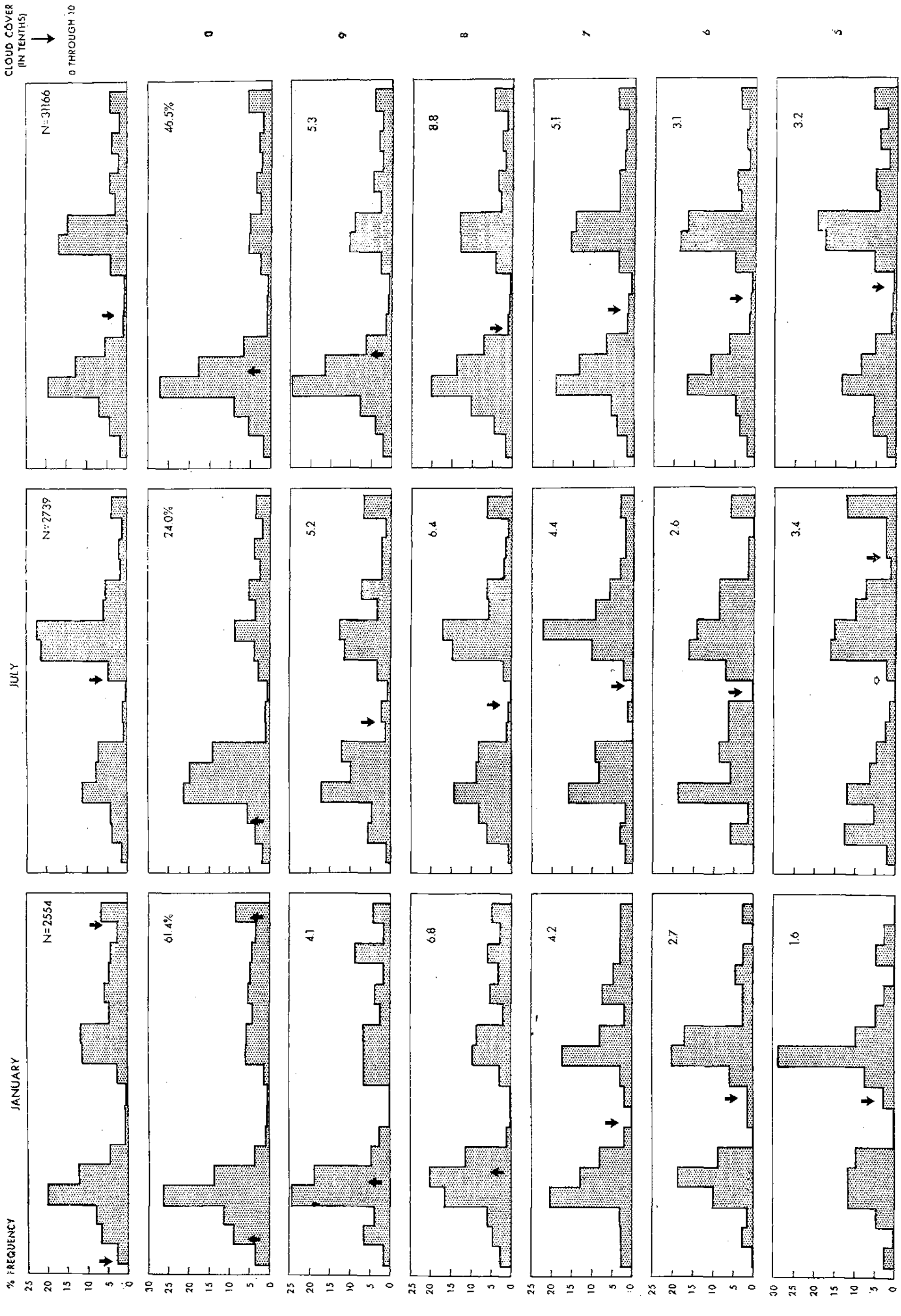

Fig. 2 - *Seattle (Washington), March 1945 - Dec. 1952 wind direction versus clouds.*

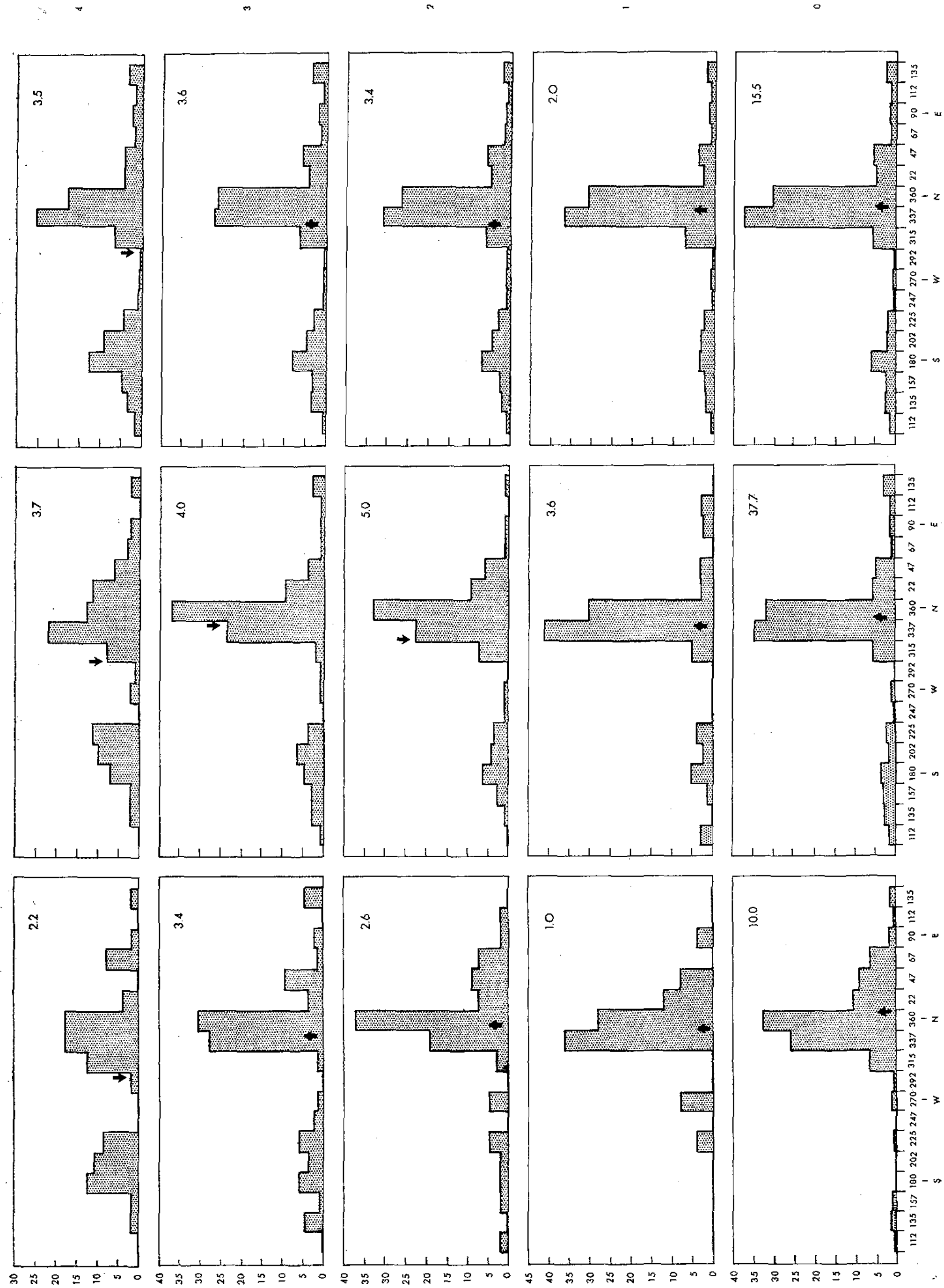

Fig. 2 b.

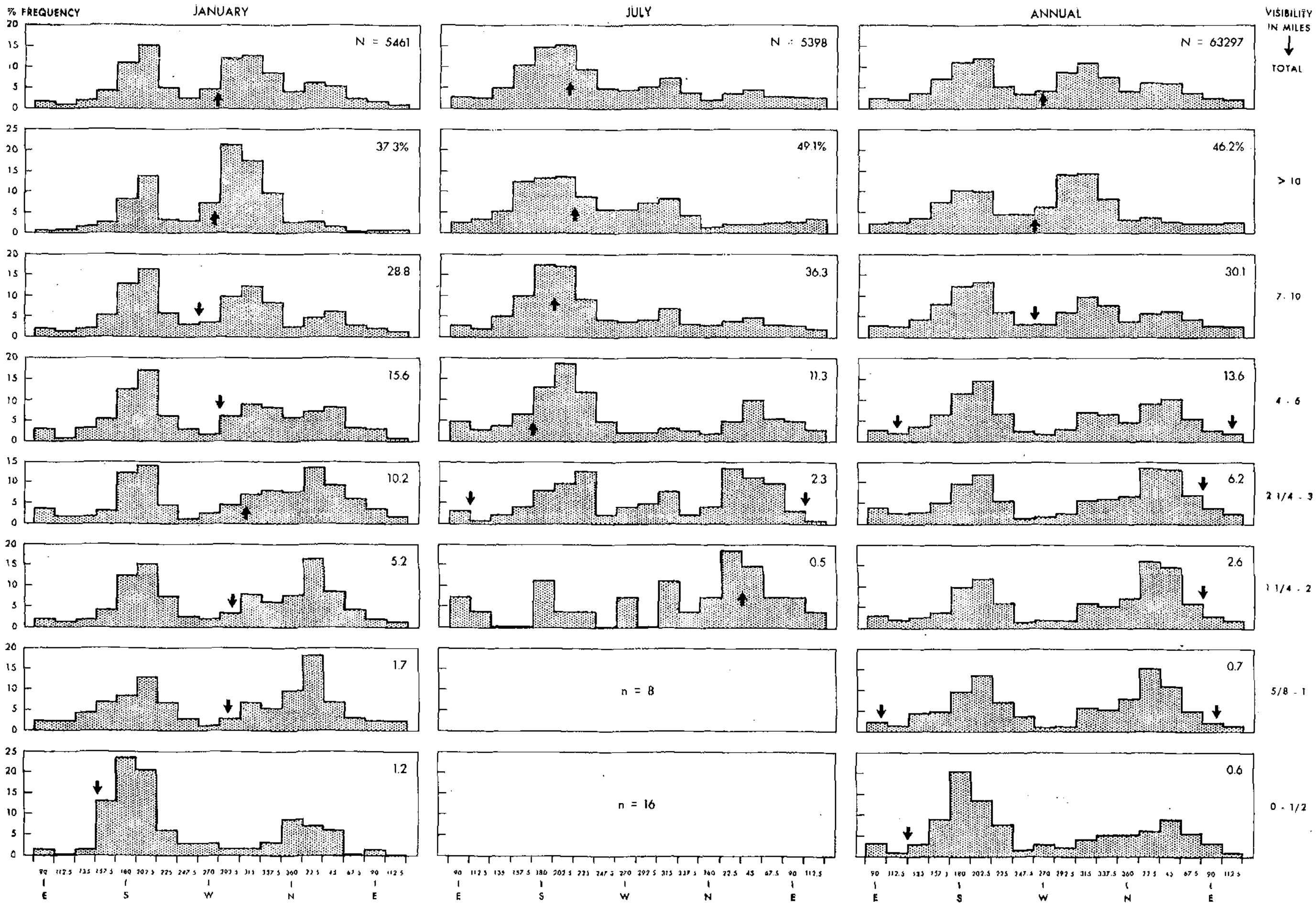

Fig. 3 - *Washington (D.C.) National Airport, Jan. 1945 - Dec. 1952 wind direction versus visibility.*

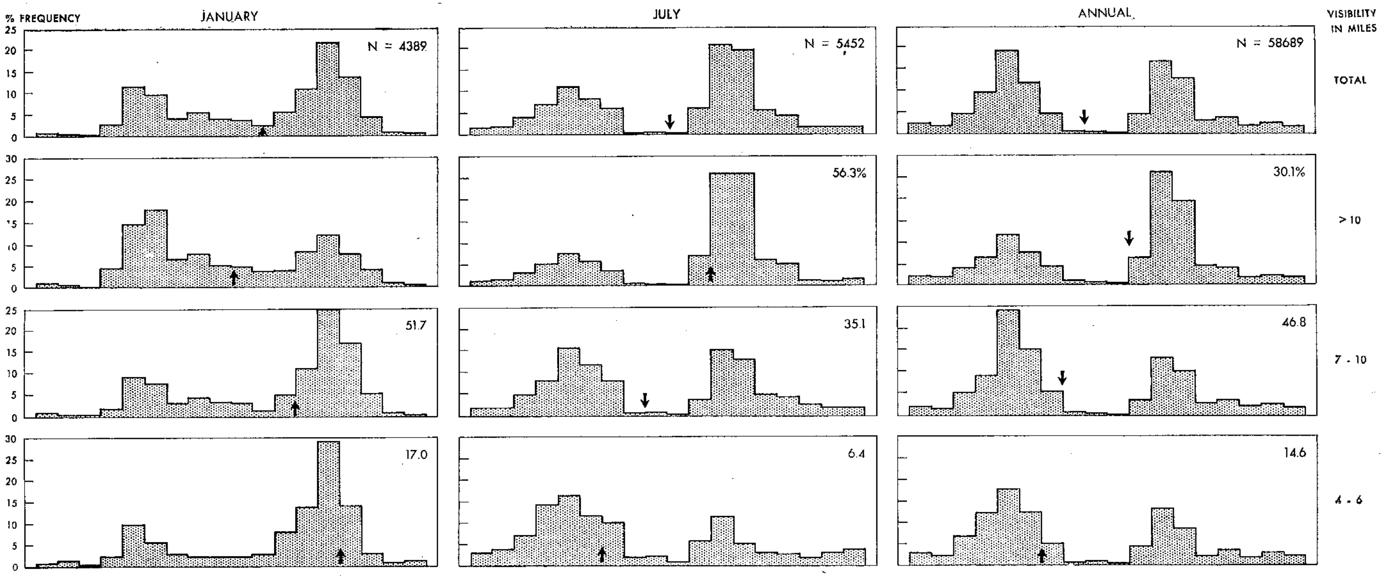

Fig. 4 - *Seattle (Washington), Jan. 1945 - Dec. 1952 wind direction versus visibility.*

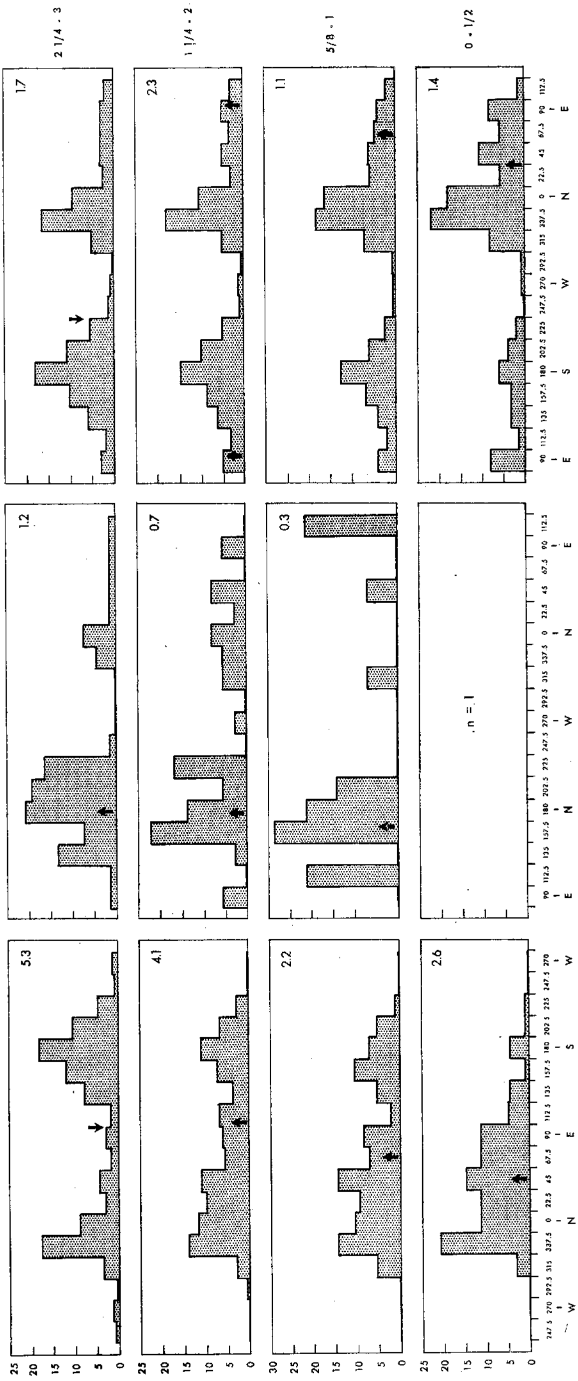

Fig. 4 *b*.

Two groups of prevailing winds can also be observed in January. The mode obviously increases for the southerly group and decreases for the northerly group, this latter group tending to flatten. What the group loses in the mode is gained in broadness. This explains that the mean value for the classification by wind direction shows little variation.

The bimodal character causes a relatively high scatter. This makes understandable that the standard deviation lies generally between 70 and 100 degrees and undercuts the value only in a few cases.

Now turn to Seattle, Washington, in the lower part of Table 4. The annual summary, as well as January and July, show a clear functional relationship between the mean wind direction and the cloud cover classes. Southerly winds bring overcast sky and northerly winds clear sky. This is supported by Figure 2, which shows the decrease of the northerly group towards 10 tenths cloud cover. Quite instructive is the reduction of the scatter expressed by the standard deviation for wind direction data in the cloud cover classes 0 and 1 for January and the year.

One difficulty, which may occur in the bimodal case by application of the introduced technique of Table 4, will be demonstrated now. The complexity of the problem may lead to cases similar to the presented example, however, ways can be found to overcome the difficulty.

There is one weak spot in the tabulation for Seattle, bottom part of Table 4. The mean value for the 5 tenths cloud class in July is listed with 80°. This falls out of line. Consider, however, that in the bimodal case this mean value is a compromise. By the definition explained in reference (3) the value with the minimum variance had to be selected. In the bimodal case two minima exist and the main minimum should be chosen. This presents itself in the 5 tenths class at 80°, while in the adjacent classes, this main minimum variance appears at 289° or 306°. Further analysis of the distribution in class 5 tenths reveals that there would be a second mean value possible at approximately 300° and for the distribution of 4 and 6 tenths cloud cover class one at about 90°. Objectively the selection must lead to the listed values of 289°, 80° and 306° for the respective distributions. If climatological conditions demand it, one can modify this in a subjective way and use the 300° instead on the 80° for conformity.

This discrepancy is caused by the complexity of the frequency distributions in the meteorological field.

Now turn to Table 5. It contains the statistical summary for visibility by wind direction classes. Figures 3 and 4 exhibit again the bimodal character of the wind direction distribution. Hence it can be expected that the relation between the two elements is not of simple nature. However, the prior view of little relationship, obtained by the interpretation of Table 3, must be revised. Good visibility at Washington, D.C. is related to westerly winds in the annual summary and in January, though in July southerly flow seems to predominate. Poor visibility ($< 1/2$ mile) appears with southerly air movement. This is supported by Figure 3, inspecting the classes > 10 miles and 0 to 1/2 mile. The predominance of the northerly group for visibility > 10 miles in the annual combination of data and in January, and the prevailing southerly winds for visibility $\leq 1/1$ mile can be noticed. The transition zone from these two extreme visibility classes shows variation of the mean value between these two groups. In July northeasterly winds bring poor visibility. With improving visibility the wind shifts to South. This contrast between January and July may cause the somewhat parabolic relation for the annual summary when going from poor visibility to excellent.

TABLE 4 - *Cloud versus Wind Direction.*

Washington, NAP (1945-1952)

Cloud Cover (tenth)	Annual $\bar{\alpha}_i$	Annual $\sigma_{\bar{\alpha}_i}$	January $\bar{\alpha}_i$	January $\sigma_{\bar{\alpha}_i}$	July $\bar{\alpha}_i$	July $\sigma_{\bar{\alpha}_i}$
0	274	77	269	65	232	75
1	262	81	277	68	217	79
2	252	83	275	74	212	81
3	253	85	278	70	210	76
4	252	84	243	77	212	75
5	258	87	280	63	210	80
6	253	84	268	78	213	71
7	242	85	264	71	219	78
8	242	87	266	79	210	79
9	240	88	257	73	191	85
10	84	97	276	93	171	88

Seattle, Washington (1945-1952)

Cloud Cover (tenth)	Annual $\bar{\alpha}_i$	Annual $\sigma_{\bar{\alpha}_i}$	January $\bar{\alpha}_i$	January $\sigma_{\bar{\alpha}_i}$	July $\bar{\alpha}_i$	July $\sigma_{\bar{\alpha}_i}$
0	350	60	8	42	351	64
1	344	63	349	46	339	60
2	328	74	355	58	328	70
3	328	81	342	75	343	72
4	298	90	299	85	306	82
5	285	96	287	86	80 *)	91
6	278	92	290	86	289	91
7	261	94	263	96	300	93
8	244	92	207	91	278	95
9	214	90	195	88	259	95
10	197	84	129	77	148	86

*) See remark in text for correctness of this value.

$\bar{\alpha}_i$ Mean Wind Direction in degrees.

$\sigma_{\bar{\alpha}_i}$ Standard Deviation in degrees.

TABLE 5 - *Visibility versus Wind Direction.*
Washington, NAP (1945-1952)

Visibility in Miles	Annual		January		July	
	$\bar{\alpha}_i$	$\sigma_{\bar{\alpha}_i}$	$\bar{\alpha}_i$	$\sigma_{\bar{\alpha}_i}$	$\bar{\alpha}_i$	$\sigma_{\bar{\alpha}_i}$
0-1/2	126	92	149	85	—	—
5/8-1	97	95	288	97	—	—
1 1/4-2	81	93	295	92	37	85
2 1/4-3	81	92	304	96	99	96
4-6	112	98	278	93	165	85
7-10	258	94	263	86	194	82
> 10	260	82	274	66	214	79

Seattle, Washington (1945-1952)

Visibility in Miles	Annual		January		July	
	$\bar{\alpha}_i$	$\sigma_{\bar{\alpha}_i}$	$\bar{\alpha}_i$	$\sigma_{\bar{\alpha}_i}$	$\bar{\alpha}_i$	$\sigma_{\bar{\alpha}_i}$
0-1/2	35	77	43	61	—	—
5/8-1	63	89	67	76	162	59
1 1/4-2	93	94	76	81	196	91
2 1/4-3	238	94	103	94	197	72
4-6	216	91	196	82	215	87
7-10	235	90	127	86	261	95
> 10	298	89	76	88	324	82

$\bar{\alpha}_i$ Mean Wind Direction in degrees.
$\sigma_{\bar{\alpha}_i}$ Standard Deviation in degrees.

For Seattle, Washington, (Figure 4 and lower part of Table 5), a similar change appears between January and July. The reader's attention is called to the different starting point of the scale for January. Year (annual) and July start at E., January at W. This had to be done in order to center the arrow for the mean wind direction. In July, southerly winds bring poor visibility, northerly winds good, > 10 miles. In January northerly winds are mainly connected with fog (< 1/2 mile), south winds have medium visibility and for northerly winds good visibility can be expected. This seasonal change influences also the annual distribution. Northerly winds reveal good as well as poor visibility, the mean values centered respectively at NW and NE.

This sample indicates the necessity of expressing relationship between visibility and wind direction by visibility classes. Presentation of mean values for wind direction other than resultant wind vectors had not been possible in the past. The typical bimodal distribution will cause, however, the resultant wind vectors to tend toward zero for all distributions and no significant relation can be recognized.

The relation to the general circulation pattern shall be discussed in connection with the Tables 6 and 7.

d) *Objective Summary Statistics.*

The presentation of Tables 2 through 5, supported by Figures 1 through 4 permits a detailed interpretation and discussion of the relationship between the two selected stations, Washington, D. C. and Seattle, Washington. For comparison of the two stations or addition of more stations, mapping or further concentration and reduction of parameters is desirable.

By the author's method (3) the mean value can now also be computed for the element of wind direction and hence correlation coefficients or ratios between cloud cover (visibility) and wind direction can be derived. They express the degree of relationship and the statistical and physical significance by objective measures. They also establish a concentration of the information on relationship and enable to evaluate numerous stations, which would be very difficult, if information would be presented merely in forms like Tables 2 - 5. Typical examples of relation parameters are presented in Tables 6 and 7. The detailed computation of the parameters is described in the appendix with example computations.

d-1) *Relation Between Wind Direction and Cloud Cover for 4 Selected Stations.* — Table 6-a and b portray the relation between cloud cover and wind direction. The survey is given for 4 stations, namely Washington (D. C.), Seattle (Washington), Argentia (Newfoundland) and Jacksonville (Florida), for the annual data, the month of January and July. The first line N displays the total number of observations. The second line y represents the mean value of cloud cover, followed by the standard deviation σ_y.

Fourth and fifth lines contain mean wind direction $\bar{\alpha}$ and standard deviation σ_α, respectively, for the total available data. The $\bar{\alpha}$ is the mean wind direction based on summation of all available wind direction data for the reference time period without regard to cloud cover classes.

The sixth line listing is $\bar{x}$, the mean wind direction based on summation of the material for the reference time period under consideration of a contingency table for wind direction data by cloud cover classes. The following line σ_x describes the standard deviation for this contingency table. More details for the computing process are given in the appendix. It should be noticed that $\sigma_x > \sigma_\alpha$ is explainable that for the contingency table the wind direction frequency distribution has been cut and arranged by cloud cover classes into columns, whose mean values fall approximately at the center. This process had been carried for all 0 to 10 cloud cover classes and frequencies within the columns then added across. Therefore, the contingency table increases the scatter range by summation to the marginal distribution (summation along a line through all placed columns). This also explains the divergence of $\bar{x}$ and $\bar{\alpha}$ in some of the cases.

Meteorological interpretation of those parameters at the four stations may be given under consideration, that the period of record is not homogeneous and

TABLE 6a - *Cloud Cover versus Wind Direction. Statistical Summary.*

	Washington, D. C. (NAP) Jan. 1945 - Dec. 1952			Seattle, Washington Mar. 1945 - Dec. 1952		
	Annual	January	July	Annual	January	July
N	57596	4731	4951	31166	2554	2739
y	5.85	6.54	5.88	6.87	7.84	4.41
σ_y	4.17	4.18	3.77	3.83	3.43	4.24
α	264	267	210	260	105	302
σ_α	91	82	82	94	89	88
x	193	272	211	250	192	276
σ_x	123	83	110	102	118	118
Wind Direction Classes						
$r_{\alpha y}$	.01	—.03	—.08	—.40	.44	—.34
$\eta_{\alpha y}$	.26	.32	.18	.49	.51	.52
Cloud Cover Classes						
r_{yx}	—.58	.02	—.17	—.59	—.74	—.64
η_{yx}	.70	.09	.68	.59	.76	.75
R_{yx} (unweighted)	—.64	—.43	—.77	—.99	—.96	—.85
Polynomials, mean wind direction as function of cloud cover classes, unweighted						
z_1^2	.40	.19	.59	.97	.92	.72
z_2^2	.21	.00	.09	.02	.05	.10
z_3^2	.23	.35	.30	.00	.01	.08
z_4^2	.08	.08	.00	.00	.00	.06

y symbol for cloud cover, α symbol for wind direction, x wind direction by cloud cover classes (contingency table), r correlation coefficient, η correlation ratio, z_1^2 percentage reduction for polynomial terms. More details see text.

frequencies are not normally distributed. The latter fact has already been mentioned in the previous sections. Homogeneity of the period would be definitely necessary in a pure climatological study, but in this article more weight has been given to the demonstration of the method than especially climatological aspects and readily available material has been used. Under this restriction one can conclude, that Argentia has the most cloud cover from all stations and Jacksonville has the least amount for the years and January. Seattle shows the least amount

TABLE 6b - *Cloud Cover versus Wind Direction. Statistical Summary.*

	Argentia, Nfld. Mar. 1945 - Nov. 1956			Jacksonville, Fla. Jan. 1945 - Dec. 1958		
	Annual	January	July	Annual	January	July
N	67739	5299	5666	75702	5478	6533
y	7.80	7.99	7.61	5.29	4.84	5.86
σ_y	3.29	3.24	3.33	3.84	4.12	3.27
α	240	282	196	122	265	168
σ_α	88	90	61	102	91	74
x	194	183	198	159	273	170
σ_x	100	117	61	119	96	75
Wind Direction Classes						
$r_{\alpha y}$	—.05	—.08	—.10	—.002	.01	.001
$\eta_{\alpha y}$	.28	.46	.20	.24	.23	.16
Cloud Cover Classes						
r_{xy}	—.40	—.49	—.14	—.55	.07	—.01
η_{yx}	.52	.65	.17	.62	.40	.16
R_{yx} (unweighted)	—.42	—.44	—.59	—.86	—.18	—.18
Polynomials, mean wind direction as function of cloud cover classes, unweighted						
z_1^2	.17	.20	.35	.73	.03	.03
z_2^2	.31	.39	.17	.12	.44	.58
z_3^2	.28	.05	.02	.06	.39	.06
z_4^2	.12	.31	.00	.04	.03	.00

y symbol for cloud cover, α symbol for wind direction, x wind direction by cloud cover classes (contingency table), r correlation coefficient, η correlation ratio, z_1^2 percentage reduction for polynomial terms. More details see text.

for July. Cloud cover is practically equally high in January for Seattle and Argentia.

The mean wind direction $\bar{\alpha}$ exhibits westerly flow in the annual summary except for Jacksonville. A sharp break between January and July is obvious at Seattle. The three other stations exhibit little variation, although a shift exists. The displayed mean direction for the stations agree with the general wind shear pattern as published in mean pressure maps like in HAURWITZ & AUSTIN (6) or

KOEPPE & DELONG (7). The exception proves to be the mean westerly flow in January at Jacksonville, which can be seen in Table 6, while isobars on sea level pressure maps would require easterly winds. However, westerly winds as computed in this tabulation are supported by the mode of the frequency distribution, which had been checked in the basic SMAR tabulations.

The change in cloud cover is remarkable for Seattle, and is connected with the shift of the mean wind direction between January and July. In January, when this area seems to be under low pressure influence, cloud cover is high, while in July anticyclonic influence predominates. This interpretation is confirmed in Figure 2, where high cloud cover is connected with southerly winds and low cloud cover with northerly. Both wind components transport maritima air masses parallel to the mountain range, so the relation of the difference to the pressure system seems reasonable.

The parameters listed in the upper part of Tables 6-a and b represent statistics of single elements. The relation between elements is expressed in the lower part by use of the linear correlation coefficient and the correlation ratio.

The line of σ_x is followed by the linear correlation coefficient $r_{\alpha y}$ and the correlation ratio $\eta_{\alpha y}$ whose computing details are given in the appendix. The parameters have been obtained in sorting the data by wind direction classes (*). Almost no relation and practically no linear relation is seen for Washington, D. C., while the result is essentially better for Seattle. Argentia and Jacksonville are in line with Washington, D. C., although differences in the sign appear.

A definite improvement of the relation is noticeable by inspection of the two lines subsequent to $r_{\alpha y}$ and $Y_{\alpha y}$. They are linear coefficient r_{yx} and the correlation ratio η_{yx} for the contingency table established in placing the data by cloud cover classes. This demonstrates strikingly, that relation of cloud cover and wind direction in principle is expressed best by configuration of the material by cloud cover classes. The relation is still very weak for Washington, D. C. in January, and Argentia and Jacksonville in July. An increase of the correlation ratio over the correlation coefficient is attributed by characteristics involved in the functional relation between the two meteorological elements.

A further coefficient R_{xy} can be seen in Tables 6-a and b. This parameter was introduced, because the correlation coefficient or ratios $r_{\alpha y}$, $\eta_{\alpha y}$, r_{yx} and η_{yx} are weighted by the factor of observation in cloud cover or wind direction classes. For example, the value of the coefficient for the U-shaped frequency distribution of cloud cover is almost determined by the relation between the zero class (clear sky) and the ten class (overcast) data. Thus, an unweighted linear correlation coefficient R_{xy} has been computed. It sets forth the functional relation of Table 4 (unweighted).

It would be meaningless to compute a correlation ratio for this unweighted material.

This value would be one, see PEARSON (9). Further details on the problem whether to use the weighted or unweighted functional relation is treated in more details in the appendix.

The nonlinear part of the relation can be judged by computing the so-called « percentage reduction z_i^2 » for the polynomial coefficient whose computing details

(*) In the appendix it will be proved, that coefficients $\geq$ .3 are significant while smaller coefficients are meaningless. Hence their significance is of no interest.

are given in the appendix or in WADSWORTH [14]. The z_i^2 values can add up to 100%, which is a total explanation of the variance and a perfect fit.

The numbers displayed in the z_i^2 part of Table 6 indicate that the linear term is very strong only for Seattle while the relation at the 3 other stations includes higher degree terms.

Table 6 reveals further that relationship between cloud cover and wind direction is strongest at Seattle. At least for some months of the year no relation exists at the other stations. The relation obtained here does not consider any diurnal variations and it is not surprising that wind and clouds are strongly related at the American West Coast (Seattle), while obviously at the East Coast a more complex connection prevails. The lower value of .59 for r_{yx} at Seattle for the year compared to the higher values in the single months may be attributed to the change of the seasonal circulation pattern, which decreases the annual relation. Thus monthly values may express the relation better than the combined annual summary.

d-2) *Relation Between Wind Direction and Visibility for* 4 *Selected Stations.* — The relation between visibility and wind direction is discussed next. Characteristics have been computed and are exhibited in Tables 7-a and b. Tables 7 list the same parameters as introduced in Tables 6, except the cloud cover element has been replaced by the visibility. The α and σ_α in Tables 6 and 7 should be expected alike. Inspection of the line N in Table 6 and 7 conveys, that both numbers are not equal, although the period of record is the same. This is attained, as obscurations are omitted for cloud cover data besides of some missing data. Thus, there is a slight change in the value of α and σ_α in Tables 7, whose computations include the data eliminated for obscured cloud cover.

The relation is very weak by wind direction classes (expressed by $r_{\alpha y}$ and $y_{\alpha y}$). A definite improvement is noticeable by visibility classes (lines r_{yx} and η_{yx}).

The increase for the unweighted correlation R_{yx} is remarkable, also the high share of the linear term. To be correct, this is not a true linear relationship. The visibility classes are almost of logarithmic progression and these logarithmic classes are renumbered in linear sequence.

The annual correlation ratio η_{yx} is highest at Argentia, pointing toward the connection of visibility and wind. But the single months as January and July show less association. There Jacksonville exceeds the other stations with high « linear relation R_{yx} ». This quasilinear relation is similar for Washington, D. C. and Seattle, where January evidently reveals less connection between wind and visibility. R_{yx} is low at Argentia, because the functional coupling is of second and third order.

IV. *Conclusion and Outlook.*

Statistical characteristics have been developed for objective interpretation and comparison between stations of meteorological data tabulations of the SMAR. Especially the evaluation of the relationship between cloud cover (visibility) and wind direction has been discussed. Three characteristics can be recommended for further investigation of climatological features between wind direction and cloud cover (visibility).

The linear correlation coefficient r_{yx} by cloud cover (visibility) classes is a fairly good measurement, although nonlinear functional relationship is locally very strong. Little information is gained, however, by the linear correlation coefficient $r_{\alpha y}$ by wind direction classes.

TABLE 7a - *Visibility versus Wind Direction. Statistical Summary.*

	Washington, D. C. (NAP) Jan. 1945 - Dec. 1952			Seattle, Washington Mar. 1945 - Dec. 1952		
	Annual	January	July	Annual	January	July
N	63297	5461	5374	58689	4389	5452
$\bar{y}$	5.07	4.73	5.32	4.90	4.55	5.44
σ_y	1.16	1.38	.81	1.15	1.33	.79
$\bar{\alpha}$	263	277	209	256	111	287
σ_α	90	83	83	94	88	92
$\bar{x}$	222	274	198	243	123	292
σ_x	112	84	85	103	94	95
Wind Direction Classes						
$r_{\alpha y}$	—.08	—.04	.05	.07	.07	—.07
$\eta_{\alpha y}$	.29	.33	.23	.12	.13	.10
Visibility Classes						
r_{yx}	.51	—.03	.27	.45	.02	.38
η_{yx}	.60	.22	.29	.49	.43	.41
R_{yx} (unweighted)	.73	.44	.97	.95	.52	.94
Polynomials, mean wind direction as function of visibility classes, unweighted						
z_1^2	.54	.20	.95	.89	.27	.89
z_2^2	.33	.48	.05	.01	.31	.08
z_3^2	.01	.28	.00	.00	.23	.03
z_4^2	.07	.02	.00	.04	.01	.01

Symbols like in table 5, except y for visibility. More details see text.

The correlation ratio η_{yx} proved to be the best characteristic, as it includes nonlinear terms. It loses, however, the coupling with immediate mathematical formulations of the functional connection, which is one very convenient property of the linear correlation coefficient.

The linear unweighted correlation coefficient R_{yx} expresses the relation between cloud cover (visibility) and wind direction without regard to the frequency within the classes. Again, the value may be poor on account of nonlinearity of the function. We may expect linearity only in limited geographical regions, as has been demonstrated by computation of the « percentage reduction », a measure to

TABLE 7b - *Visibility versus Wind Direction. Statistical Summary.*

	Argentia, Nfld. Mar. 1945 - Nov. 1956			Jacksonville, Fla. Jan. 1945 - Dec. 1958		
	Annual	January	July	Annual	January	July
N	99284	8001	8602	107049	8560	9170
$\bar{y}$	4.44	4.52	3.54	5.08	4.75	5.42
σ_y	1.80	1.62	2.32	1.02	1.28	0.66
$\bar{\alpha}$	240	281	195	104	279	168
σ_α	87	86	62	99	91	75
$\bar{x}$	232	266	196	120	278	174
σ_x	165	95	64	108	118	78
Wind Direction Classes						
$r_{\alpha y}$	.13	.08	—.08	.23	.01	—.01
$\eta_{\alpha y}$	.40	.43	.38	.34	.34	.21
Visibility Classes						
r_{xy}	—.19	.14	—.07	—.47	—.35	—.24
η_{yx}	.86	.42	.29	.52	.69	.74
R_{yx} (unweighted)	.42	.36	—.34	.85	—.93	—.85
Polynomials, mean wind direction as function of cloud cover classes, unweighted						
z_1^2	.18	.13	.11	.72	.86	.72
z_2^2	.67	.09	.22	.22	.04	.13
z_3^2	.04	.75	.52	.001	.02	.01
z_4^2	.09	.00	.08	.02	.07	.06

symbols like in table 5, except y for visibility. More details see text express the fraction explained by the pertinent polynomial term in respect to the total variance.

It has been pointed out that the SMAR tabulations, a priori established for combined hours of the day, can show good relationship between cloud cover (visibility) and wind direction for those stations only, where this relation is connected with the general circulation pattern like in Seattle. Relationship of local character with diurnal rhythm such as the land and sea breeze or valley breeze make necessary a division into hours. The same statistical technique as described here can be applied.

Tables 6 and 7 list in concentrated form many of the details disclosed in elaborate work from Tables 2 - 5.

For mapping, the parameters η_{yx} or R_{yx} are highly recommended. It will be admitted that an extension of the study to numerous stations will involve some work of data reduction and computation. A climatological map of the kind recommended would clearly indicate areas with strong relationship and areas of similarity. Supplementary detailed tabulations could then be restricted to a few individual stations. This gives complete information without voluminous listings like the SMAR tabulations for a number of stations.

Appendix

The mathematical program to compute the relation parameters.

a) *The linear correlation coefficient* r.

1) *Classes of wind direction.* — The linear correlation coefficient appears in Tables 6 and 7 in the form of $r_{\alpha y} = r_{xy}$ and r_{yx}. Theoretically, these two values would be identical between the same two elements, if there were no linearizing possible for the case of cloud cover (and visibility) classes. This effect will be discussed later.

It can be shown theoretically that this linear correlation coefficient must be independent of the way of computation. The general formula runs as follows:

$$r_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \tag{1}$$

in which the σ_x and σ_y denote the standard deviations for each element and C_{xy} is the so-called covariance. In our case,

$$C_{xy} = \Sigma \frac{(x_i - \bar{x})(y_i - \bar{y})}{N} . \tag{2}$$

Sometimes the value $N \cdot C_{xy}$ is considered to be the covariance and equation (1) would be modified correspondingly.

The value of the covariance is generally an invariant and the correlation coefficient is thus independent of the way of computation. It will be shown later, that the C_{xy} varies, if one element has a periodic scale like the wind direction.

Application of equations (1) and (2) presumes that $\bar{x}$ and $\bar{y}$ are known. If one of the variables is identified with the wind direction, say x, then the x is the mean value of the wind directions without regard to speed. The solution to this under consideration of the wind direction as a periodic variable is given by the author (3). After that, the correlation coefficient $r\alpha_y$ can be computed by equation (1). Therefore, after definition of $\bar{\alpha}$ the computation may be carried out as described in any statistics text [as reference (1,5,8)]. Cloud cover (and visibility) frequency by wind direction classes involve no further difficulty. It also has been shown by the author (4), that this correlation coefficient furnishes the correlation with the minimum residual error of all possible correlation coefficients.

Computation of the correlation coefficient $r\alpha_y$ is demonstrated in Table 9 for the sample of Table 8. This latter contains data for Washington National Airport annual summary for wind direction versus cloud cover. Let α_j be the wind direction class and $\overline{y_i}$ the mean value for cloud cover in this class, then

$$C_{\alpha y} = \frac{\Sigma n_j (\alpha_j - \bar{\alpha})(\bar{y}_j - \bar{y})}{N}. \tag{3}$$

The

$$\sigma_y^2 = \frac{\Sigma n_i (y_i - \bar{y})^2}{N} \tag{4}$$

and

$$\sigma_\alpha^2 = \frac{\Sigma n_j (\alpha_i - \bar{\alpha})^2}{N}. \tag{5}$$

The actual data extracted from Table 9 are

$$C_{\alpha y} = 3.2559 .$$

Further $\sigma_y{}^2 = 17.3988$ and hence $\sigma_y = 4.17$. The $\sigma_\alpha{}^2 = 8341.75$ and $\sigma_\alpha =$ $= 91.4$. This leads to the correlation coefficient

$$r_{\alpha y} = \frac{C_{\alpha y}}{\sigma_\alpha \sigma_y} \tag{1a}$$

and

$$r_{\alpha y} = \frac{3.2559}{91.4 \times 4.17} = .009 .$$

This correlation coefficient is very poor. From Table 6 can be seen that the reversed process, wind direction versus cloud cover classes, yields a higher value. The correlation ratio will be discussed later.

2) *Cloud cover or visibility classes.* — It has been mentioned that the correlation coefficient r_{xy} would be the same, when none of the elements follows a periodic scale such as the wind direction does. Mean value and standard deviation are independent of the class division.

This can be demonstrated in the following example. Compute the mean value $\bar{y}$ for cloud cover (visibility) by

$$\bar{y}_1 = \frac{\Sigma n_i y_i}{N} \tag{6}$$

where $\overline{y_1}$ denotes value of the cloud cover classes. Also

$$\bar{y}_2 = \frac{\Sigma n_j \bar{y}_j}{N} \tag{7}$$

where the $\overline{y_j}$ is the mean cloud value by wind direction classes. This mean value $\bar{y}_j$ is

$$\bar{y}_j = \frac{\sum_{1}^{n}{}_j\, m_i y_i}{n_j}. \tag{8}$$

TABLE 8 - *Washington, National Airport Annual Summary* (Jan. 1945 - Dec. 1952).

Wind Direction	Central Class Values Degrees	Cloud Cover $y_i = 0$	1	2	3	4	5	6	7	8	9	10	n_j	y_j
$\alpha_j = x_j$														
N	0	566	117	85	71	59	63	68	93	97	120	1014	2353	6.00
NNE	22.5	576	92	108	92	65	69	76	81	147	164	2117	3587	7.26
NE	45	413	78	87	93	58	78	77	94	160	168	2102	3408	7.65
ENE	67.5	199	62	54	66	44	57	53	73	110	124	1397	2239	7.87
E	90	138	47	46	50	51	40	38	53	68	86	859	1476	7.59
ESE	112.5	159	51	56	53	56	43	41	54	77	106	709	1405	7.15
SE	135	298	92	111	96	59	57	68	96	130	137	1047	2191	6.82
SSE	157.5	630	198	259	237	168	163	184	199	270	299	1537	4144	6.19
S	180	1189	272	325	290	247	190	267	287	445	448	2305	6265	6.06
SSW	202.5	1591	293	330	317	205	183	278	322	500	510	2449	6978	5.82
SW	225	834	154	168	121	112	108	107	145	203	227	947	3126	5.27
WSW	247.5	533	120	104	100	97	82	98	126	144	165	543	2112	5.19
W	270	701	163	139	117	117	92	100	125	165	174	569	2462	4.78
WNW	292.5	1645	345	319	294	216	199	227	260	283	325	977	5090	4.30
NW	315	2118	409	333	293	218	194	208	236	339	350	1673	6371	4.60
NNW	337.5	1334	223	208	179	172	121	149	155	221	238	1389	4389	5.07
	n_i	12924	2716	2732	2469	1944	1739	2039	2399	3359	3641	21634	57596	—
	$\bar{x}_i = \bar{\alpha}_i$	274	262	253	253	252	258	253	242	242	240	84	—	—

TABLE 9 - *Computation of the Correlation Coefficient* $r_{\alpha y}$ *for Data of Table* 8.

[$\bar{\alpha} = 264$ from scheme described in reference (4)]

α_j	n_j	$\alpha_i - \bar{\alpha}$	$n_j (\alpha_j - \bar{\alpha})^2$	$\bar{y}_j$	$\bar{y}_j - \bar{y}$	$n_j (\alpha_i - \bar{\alpha}) (\bar{y}_j - \bar{y})$
0	2353	96	21 685 248	6.00	15	33 883
22.5	3587	118	49 945 388	7.26	1.41	596 805
45	3408	141	67 754 448	7.65	1.80	864 950
67.5	2239	163	59 487 991	7.87	2.02	737 213
90	1476	— 174	44 687 376	7.59	1.74	— 446 874
112.5	1405	— 152	32 461 120	7.15	1.30	— 277 628
135	2191	— 129	36 460 431	6.82	.97	— 274 160
157.5	4144	— 107	47 444 656	6.19	.34	— 150 759
180	6265	— 84	44 205 840	6.06	.21	— 110 515
202.5	6978	— 62	26 823 432	5.82	— .03	12 979
225	3126	— 39	4 845 906	5.27	— .58	70 710
247.5	2112	— 17	610 368	5.19	— .66	23 697
270	2462	6	88 632	4.78	— 1.07	— 15 806
292.5	5090	28	3 990 560	4.30	— 1.55	— 220 906
315	6371	51	16 570 971	4.60	— 1.25	— 406 151
337.5	4389	73	23 388 981	5.07	.78	— 249 910
Σ	57596	—	480 451 348	—	—	187 528
$\frac{1}{N} \Sigma$			8341.75			3.2559

y_i	n_i	$n_i y_i$	$y_i - \bar{y}$	$n_i (y_i - \bar{y})^2$
0	12924	0	— 5.85	442 292
1	2716	2716	— 4.85	63 887
2	2732	5464	— 3.85	40 495
3	2469	7407	— 2.85	20 054
4	1944	7776	— 1.85	6 653
5	1734	8695	— .85	1 253
6	2039	12234	.15	46
7	2399	16793	1.15	3 173
8	3359	26872	2.15	15 527
9	3641	32769	3.15	36 128
10	21634	216340	4.15	372 591
Σ	57596	337066	—	1002 099
$\frac{1}{N} \Sigma$		5.852		17.3988

If equation (8) is introduced into (7), it is obtained

(9) $$\bar{y}_2 = \frac{\Sigma n_j \dfrac{\Sigma m_i y_i}{n_j}}{N}$$

which reduces to

(10) $$\bar{y}_2 = \frac{\Sigma\Sigma m_i y_i}{N}.$$

If this is written for a selected class interval y_i, the y_i is constant and equation (10) may be modified to

(11) $$\bar{y}_2 = \frac{\Sigma y_i \Sigma m_i}{N}.$$

The $\Sigma\, m_i = n_i$ and equation (11) takes the form of equation (6). Hence

(12) $$\bar{y}_1 = \bar{y}_2 .$$

This proves that ordinarily the mean value in a contingency table is independent of the class division.

The same relation should be valid for $\bar{\alpha}$. Later computation, however, of the correlation ratio and in this connection the setting up of a contingency table gives an opportunity to compute a linear correlation coefficient r_{yx}, based on this contingency table. As exhibited in Tables 6 and 7, this value is higher than the correlation coefficient $r_{\alpha y}$ and shall be discussed in details now.

First be determined, which mean value is understood. For the presented data of Table 9 the $\bar{\alpha}_1 = 264$ is the solution with the minimum variance, namely

(13) $$\frac{\Sigma\, n_j\, (\alpha_j - \bar{\alpha}_1)}{N} = 0$$

where $|\alpha_j - \bar{\alpha}_1| \leq 180^{\circ}$ [See reference (3)].

By cloud cover (visibility) classes, analogous to (7)

(14) $$\bar{\alpha}_2 = \frac{\Sigma n_i \bar{\alpha}_i}{N}.$$

Each of the single class mean values $\bar{\alpha}_i$ must be computed by the same method as equation (13)

(15) $$\frac{\Sigma\, (\alpha_{ji} - \bar{\alpha}_i)}{n_i} = 0$$

where the $|\alpha_{ji} - \bar{\alpha}| \leq 180^{\circ}$.

This restriction prevents transformation of (14) to (13) and

(16) $$\bar{\alpha}_1 \neq \bar{\alpha}_2 .$$

To avoid confusion about the used mean value, $\bar{\alpha}_1$ may remain $\bar{\alpha}$ and the symbol α_j be used for wind direction classes. The symbol x may denote the wind direction by cloud cover classes, hence $\bar{x} = \bar{\alpha}_2$. Also $x_i = \alpha_i$. The $\bar{\alpha}_2$ in Table 9 amounts to 264, while x from Table 10, computed by formula (14-a), becomes 193°.

(14a) $$\bar{x} = \frac{\Sigma n_i \bar{x}_i}{N}.$$

The $\bar{y}$ remains the same as in formula (6). This has been proved by equation (12). The covariance C_{xy} is computed similar to equation (3), namely

(17) $$C_{xy} = \frac{\Sigma n_i (\bar{x}_i - \bar{x})(y_i - \bar{y})}{N}$$

and the correlation coefficient may be computed by (1). The σ_y is identical to the value computed by equation (4). Covariance and σ_y are known and the σ_x must be computed. It is different from σ_α. It will be shown that the variance $\sigma_x{}^2$ can be computed from

(18) $$\sigma_x^2 = \frac{\Sigma n_i \sigma_{yi}^2}{N} + \frac{\Sigma n_i (\bar{x}_i - \bar{x})^2}{N}.$$

The $\sigma_{yi}{}^2$ is the variance for each cloud cover (visibility) class. It will always be $\sigma_x \geq \sigma_\alpha$.

The computation of the $\sigma_x{}^2$ is demonstrated in Table 10

$$\frac{\Sigma n_i \sigma_{yi}^2}{N} = 7758.2627$$

$$\frac{\Sigma n_i (\bar{x}_i - \bar{x})^2}{N} = 7236.7015\,.$$

This gives $\sigma_x = 122.45$.

Then the correlation coefficient is

$$r_{xy} = \frac{-297.4335}{122.45 \times 4.17} = \frac{-297.4335}{410.616} = -.582\,.$$

The statistical significance of the correlation coefficient is described in any statistics book, for example reference [1,5,8]. It is normally dependent on N. A correlation coefficient $< .3$ expresses a poor relationship, and is of pratically no climatological interest. For a correlation coefficient $\geq .3$, the statistical significance at the 5% level requires $N > 42$, for the 1% level $N > 70$. Tables 6 and 7 display that the correlation coefficients of interest are significat in this report, because $N > 70$ and therefore further detailed discussion on the significance may be omitted.

b) *The Correlation Ratio.*

1) *Computation of the Correlation Ratio.* — The linear correlation coefficient $r_{\alpha y}$ assumes linear relationship and may be poor if the relationship between two elements is nonlinear. In order to establish a characteristic, which expresses whether there is any relation at all independent of functional degree, the correlation ratio η or its squared value η^2 can be used.

This measure is based on a contingency table and utilizes the decrease of the

TABLE 10 - *Computation of the Correlation Coefficient* r_{xy} *from the Data of Table 8.*

y_i	n_i	$\bar{x}_i$	$n_i\bar{x}_i$	$\bar{x}_i - \bar{x}$	$n_i(\bar{x}_i - \bar{x})(y_i - \bar{y})$	σ^2_{yi}	$n_i\sigma^2_{yi}$	$n_i(\bar{x}_i - \bar{x})^2$
0	12 924	274	3 541 176	81	— 6 125 976	5898	76 225 752	84 794 364
1	2 716	262	711 592	69	— 909 860	6617	17 971 772	12 930 876
2	2 732	253	688 464	60	— 631 092	6870	18 768 840	9 835 200
3	2 469	253	624 657	60	— 422 199	7219	17 823 711	8 888 400
4	1 944	252	489 888	59	— 211 896	7032	13 670 208	6 767 064
5	1 734	258	448 662	65	— 95 645	7624	13 220 016	7 347 275
6	2 039	253	515 867	60	18 351	7067	14 409 613	7 340 400
7	2 399	242	580 558	49	134 344	7265	17 428 735	5 759 999
8	3 359	242	812 878	49	352 695	7589	25 491 451	8 064 959
9	3 641	240	873 840	47	538 868	7791	28 367 031	8 042 969
10	21 634	84	1 817 256	— 109	— 9 778 568	9405	203 467 770	257 033 554
Σ	57 596		11 104 838		— 17 130 978		446 844 899	416 805 060
$\frac{1}{N}\Sigma$			192.8		297.434		7758.2627	7236.7015

$\sigma^2_x = 7758.2627 + 7236.7015 = 14994.9642$

$\sigma_x = 122.45$

variance, computing the average of the variance by summation of the variance of the single class. If the functional relation is linear, the η^2 will be identical to r^2.

It is convenient to base the computation of η^2 on the formula

$$\eta_{yx}^2 = 1 - \frac{S_{xi}^2}{\sigma_x^2} \tag{19}$$

where the σ_x^2 is the total variance, computed by equation (4), (5) or (18) depending on the elements involved. The S^2_{xi} may be computed by (*)

$$S_{xi}^2 = \frac{\Sigma n_i \sigma_{xi}^2}{N} . \tag{20}$$

Again, computation by wind direction classes and by cloud cover (visibility) classes is possible. Rewriting equation (19) renders

$$\eta_{\alpha y}^2 = 1 - \frac{S_{\alpha j}^2}{\sigma_y^2} \tag{21}$$

where the $S^2_{\alpha j}$ is the average variance by α_i classes, determined by

$$S_{\alpha j}^2 = \frac{\Sigma n_j \sigma_{\alpha j}^2}{N} . \tag{22}$$

The $\sigma^2_{\alpha j}$ is the variance of the cloud cover (visibility) in wind direction classes. The σ_y^2 may be computed by

$$\sigma_y^2 = S_{\alpha j}^2 + \sigma_{my}^2 \tag{23}$$

where

$$\sigma_{my}^2 = \frac{\Sigma n_j (\bar{y}_j - \bar{y})^2}{N} . \tag{24}$$

The $\bar{y}_j$ follows the definition of equation (8).

The computation for the numerical example of Table 8 is demonstrated in Table 11. The first 4 columns of Table 11 are the same as corresponding values in Table 9. (In practice Table 11 may be enclosed in Table 9).

Table 11 yields $S^2_{\alpha j} = 16.27$, $\sigma_m^2 = 1.16$ and $\sigma_y^2 = 16.27 + 1.16 = 17.43$, This is the same value of Table 9 except of the difference of 0.03 due to rounding. The correlation ratio becomes

$$\eta_{\alpha y}^2 = 1 - \frac{16.27}{17.43} = 0.0666 .$$

$$\eta_{\alpha y} = .26 .$$

The computation of the correlation ratio by cloud cover (visibility) classes is similar.

$$\eta_{xy}^2 = 1 - \frac{S_{yi}^2}{\sigma_x^2} . \tag{25}$$

(*) It is also known as the pooled variance of the subgroups.

The

(26)
$$S_{yi}^2 = \frac{\Sigma n_i \sigma_{yi}^2}{N}$$

and $\sigma_x{}^2$ has been given already in equation (18), which may be modified into the same form of (23).

(18a)
$$\sigma_x^2 = S_{yi}^2 + \sigma_{mx}^2 .$$

The

(27)
$$\sigma_{mx}^2 = \frac{\Sigma n_i \, (x_i - \bar{x})^2}{N} .$$

The necessary values are presented already in Table 10. Further,

$$S_{yi}^2 = 7758.26 \quad \text{and} \quad \sigma_{mx}^2 = 7236.70 .$$

This leads to $\sigma_x{}^2 = 14994.96$. The correlation ratio

$$\eta_{yx}^2 = 1 - \frac{7758.26}{14994.96} = .4826$$

$$\eta_{yx} = .695 \sim .70 .$$

Another coefficient, r_{xy}, based on this contingency table has been computed previously. What is the difference? Obviously the r_{xy} contains the linear part of the η_{yx}. This can be proved easily and detailed derivation need not be given here, It may also be pointed out that r_{xy} of the equation (1) may be transformed to

(28)
$$r_{yx}^2 = \frac{1}{\sigma_x^2} \, [\Sigma n_i \, a_i \, (x_i - \bar{x})]^2$$

where

(29)
$$a_i = \frac{(y_i - \bar{y})}{\sigma_y}$$

the « studentized » variate.

The correlation ratio may be transformed to

(30)
$$\eta_{yx}^2 = \frac{1}{\sigma_x^2} \, \Sigma n_i \, (\bar{x}_i - \bar{x})^2 .$$

It can be shown that

(31)
$$\Sigma n_i \, (\bar{x}_i - \bar{x})^2 = [\Sigma n_i a_i \, (x_i - \bar{x})]^2$$

for the case, that the $\bar{x}_i$ progress linearly with the $\bar{y}_i$. Hence r_{yx} expresses the linear portion of the η_{yx}.

TABLE 11 - *Computation of the Correlation Ratio from the Data of Table 8.*

α_j	n_j	$\bar{y}_j$	$y_i - \bar{y}$	$n_j (y_i - \bar{y})^2$	$\sigma^2_{\alpha j}$	$n_j \sigma^2_{\alpha j}$
0	2353	6.00	.15	52.94	18.44	43 389
22.5	3587	7.26	1.41	7 131.31	15.70	56 316
45	3408	7.65	1.80	11 041.92	13.47	45 906
67.5	2239	7.87	2.02	9 136.02	11.86	26 555
90	1476	7.59	1.74	4 468.74	12.67	18 701
112.5	1405	7.15	1.30	2 374.45	13.84	19 445
135	2191	6.82	.97	2 061.51	15.09	33 062
157.5	4144	6.19	.34	479.05	15.25	63 196
180	6265	6.06	.21	276.29	16.21	101 556
202.5	6978	5.82	— .03	6.28	17.14	119 603
225	3126	5.27	— .58	1 051.59	18.04	56 393
247.5	2112	5.19	— .66	919.99	16.43	34 700
270	2762	4.78	— 1.07	2 818.74	16.63	40 943
292.5	5090	4.30	— 1.55	12 228.73	16.07	81 796
315	6371	4.60	— 1.25	9 954.69	17.97	114 487
337.5	4389	5.07	— .78	2 670.77	18.44	80 933
Σ	57596			66 672.52		936 981
$\frac{1}{\Sigma}\Sigma$				1.16		16.27

2) *The significance of the Correlation Ratio.* — The significance of the correlation ratio may briefly be discussed. There are two evaluations which must be considered. The first reduction eliminates the effect of one value only occurring in each column (line) of a contingency table. Then $\eta^2 = 1$ automatically without significance. The $\sigma^2_{\alpha j}$ and the σ^2_{yi} could then be zero and the corresponding values $S^2_{\alpha j}$ and S^2_{yi}. This gives automatically 1 in equations (21) and (25). The following corrections [see reference (8) and (9)] eliminate this effect.

(32) $$\eta_r^2 = 1 - \left[(1 - \eta^2)\frac{N - 1}{N - m}\right].$$

The η_r^2 denotes the reduced value, the m is the number of columns (lines) and N is the sample size. This equation can be transformed into a more convenient form for better interpretation. Assumed, the new η_r^2 should not differ from the initially computed η for more than an error E. The $\eta_r^2 \leq \eta^2$ would always be valid, hence the E is a negative correction.

(33) $$\eta_r^2 = \eta^2 - E\,.$$

After some length computation,

(34) $$\eta^2 - 1 = E\frac{(N - m)}{(1 - m)}\,.$$

From which the N can be resolved.

$$N = \frac{1 - \eta^2}{-E} (1 - m) + m . \tag{35}$$

This finally may take the form

$$N = m (1 + \varepsilon) - \varepsilon \tag{36}$$

where

$$\varepsilon = \frac{1 - \eta^2}{E} . \tag{37}$$

How large must the sample be in order to stay below a certain error? Table 12 contains for two errors $E = .01$ and $.02$ the necessary information. The m is 7 for the visibility classes, 11 for cloud cover and 16 for wind direction classes.

Comparing the N value of tables 6 and 7 it can be seen that they are potentially enough to guarantee a correction $E \leq .01$. This effect can therefore be neglected.

The second test to be applied is similar to the statistical significance mentioned in connection with the correlation coefficient. From WOO's formula (13) one may conclude that for $N = 2000$ and $m = 16$ the correlation ratio η^2 is significant at the 1% level for all values exceeding .02. This is far beyond the physical level, where $\eta^2 = .10$ is considered to be barely the limit, when η^2 as showing some relation is accepted. Hence, the correlation ratios of interest in this article are also statistically significant.

c) *Functional Relationship.*

Finally one could be interested in expressing the relation between the two elements mathematically. A convenient way among several possible approaches is to use orthogonal polynomials. One may express the cloud cover (visibility) classes versus the mean wind direction. For this functional description polynomial coefficients have been computed.

Computation of orthogonal polynomial coefficients may be found in statistics books and detailed discussion is omitted here. One sample of computations is represented in Table 13. The coefficients c_i of the series f are computed.

$$f = c_0 + c_1\Phi_1 + c_2\Phi_2 + \ldots c_i\Phi_i \tag{38}$$

where

$$c_0 = \frac{\Sigma X_i}{m} \tag{39}$$

and

$$c_i = \frac{\Sigma X_i \Phi_i}{\Sigma \Phi_i^2} \tag{40}$$

further

$$x_i = (\bar{x}_i - \bar{x}) . \tag{40a}$$

Table 12 - *Sample size* N *to eliminate the effect of grouping in contingency tables and to stay below an arbitrarily chosen error* E *of the correlation ratio* η^2. *Evaluation of equation* (36).

a) $E \leq .01$.

η^2	ε	$1+\varepsilon$	$m = 7$ $m(1+\varepsilon)$	$m = 11$ $m(1+\varepsilon)$	$m = 16$ $m(1+\varepsilon)$	$m = 7$ N	$m = 11$ N	$m = 16$ N
.99	1	2	14	22	32	13	21	31
.98	2	3	21	33	48	19	31	46
.70	30	31	217	341	496	187	311	466
.50	50	51	357	561	816	307	511	766
.30	70	71	497	781	1136	427	711	1066
.10	90	91	637	1001	1456	547	911	1366
.00	100	101	707	1111	1616	607	1011	1516

b) $E \leq .02$.

η^2	ε	$1+\varepsilon$	$m = 7$ $m(1+\varepsilon)$	$m = 11$ $m(1+\varepsilon)$	$m = 16$ $m(1+\varepsilon)$	$m = 7$ N	$m = 11$ N	$m = 16$ N
.99	.5	1.5	10.5	16.5	24	10	16	24
.98	1	2	14	22	32	13	21	31
.70	15	16	112	176	256	97	161	241
.50	25	26	182	286	416	157	261	391
.30	35	36	252	396	576	217	361	541
.10	45	46	322	506	736	277	461	691
.00	50	51	357	561	816	307	511	766

The Φ_i and $\Sigma\, \Phi_i^2$ may be found in statistical tables like Pearson (7). To check the efficiency one may compute the percentage reduction z_i^2. It expresses that part of the variance which is explained by the computed relationship.

(41) $$z_i^2 = \frac{c_i^2 \sigma_{\phi i}^2}{\sigma^2}$$

where

(42) $$\sigma_{\phi i}^2 = \frac{\Sigma \Phi_i^2}{n}.$$

The figures are presented in the bottom of Tables 6 and 7. The first polynomial term z_i^2 expresses the linearity in the functional relationship. This term is very high for cloud cover and wind direction at Seattle.

Relationships discussed under points a) and b) of the appendix are taking into account all data N. In other words, they express a weighted relationship

TABLE 13 - *Computation of polynominal coefficient* a_i *and percentage reduction* z_i^2 *for sample of table 8.*

Cloud Cover	Mean Wind Direction $\bar{x}_i$	$X_i = \bar{x}_i - \bar{x}$	$(\bar{x}_i - \bar{x})^2$	$X_i\Phi_1$	$X_i\Phi_2$	$X_i\Phi_3$	$X_i\Phi_4$
0	274	36	1296	— 180	540	— 1080	216
1	262	24	576	— 96	144	144	— 144
2	253	15	225	— 45	— 15	330	— 90
3	253	15	225	— 30	— 90	345	— 15
4	252	14	196	— 14	— 126	196	56
5	258	20	400	0	— 200	0	120
6	253	15	225	15	— 135	— 210	60
7	242	4	16	8	— 24	— 92	— 4
8	242	4	16	12	— 4	— 88	— 24
9	240	2	4	8	— 12	— 12	— 12
10	84	— 154	23716	— 770	— 2310	— 4620	— 924
Σ	2613		26895	— 1092	— 2232	— 5087	— 761
$\frac{1}{N}\Sigma$	$\bar{x} = 237.5$	$\sigma_x^2 = 2445$		— 99.3	—	—	—
$\frac{1}{\varphi^2_i}$		$\sigma_x = 49.4$		— 9.93	— 2.54	— 1.19	— 2.66

i	c_i	c_i^2	$c_i^2\sigma_\phi^2$	z_i^2
1	— 9.93	98.60	986	.403
2	— 2.54	6.45	503	.206
3	— 1.19	1.42	554	.227
4	— 2.66	7.08	184	.075
Σ				.911

between the two elements, where the classes are weighted by its frequency N. The last relation equation (38) however, using polynomial terms, delineates an unweighted relation, giving each class value equal power.

The polynomial terms are closely related to the correlation coefficient and may therefore be computed for the unweighted functions. Its value is

$$R_{xy} = \frac{C'_{xy}}{\sigma_x \sigma_y} \tag{43}$$

where

$$C'_{xy} = \frac{\Sigma\,(x_i - \bar{x})\,(y_i - \bar{y})}{N}\,. \tag{44}$$

The σ_x and σ_y are now standard deviations, computed from the unweighted departures $(x_i - \bar{x})$ and $(y_i - \bar{y})$, respectively. The covariance may be taken from Table 13, as there is the relation

$$C'_{xy} = \frac{1}{N}\,\Sigma X_i \Phi_i\,. \tag{45}$$

The $\sigma_x = 49.4$ and the $\sigma_y = \sqrt{\Sigma \Phi_i^2 / N}$. The $\Sigma\,\Phi_i^2$ is the table value mentioned following equation (40). For Table 13 the first $\Sigma\,\Phi_i^2 = 110$ and hence $\sigma_y = \sqrt{10} = 3.162$. This leads to $R_{xy} = -0.64$ in Table 6. This linear coefficient is presented in the lower part of Tables 6 and 7. Its square should give the percentage reduction z_1^2 for the first polynomial coefficient.

It should be noted that Tables 9, 10, 11 and 13 may be combined and that the published formulae for computations may be modified for special needs of the available computation tool. The author feels that clarity of presentation was served best by breaking these tables into subparts, each containing the information necessary to compute each discussed parameter separately.

Acknowledgement.

I am indebted to Mr. N. Canfield (Chief, Climatic Analysis Section at NWRC, Asheville) for calling my attention to this problem. He also made helpful suggestions in reading the manuscript and supported the technical work. For the somewhat elaborate computational work I want to thank Mrs. G. Whiting, Mrs. E. Pearson and Mrs. N. Nicholes. It is also gratefully acknowledged that the work was sponsored by the Office of Naval Weather Service and its liaison officer at NWRC, Asheville, North Carolina (Lt. Comdr. R. Parker).

REFERENCES

(1) Brooks C. E. P. & N. Carruthers: *Handbook of Statistical Methods in Meteorology*, London. His Majesty's Stationary Office (1953). — (2) Conrad V. & L. W. Pollak: *Methods in Climatology.* Harvard University Press, Cambridge, (1950). — (3) Essenwanger O.: *On the Defining and Computing of Mean Value and Standard Deviation for Wind Directions*, Manuscript. — (4) Essenwanger O.: *On the Computing of the Correlation Coefficient for Wind Direction*, Manuscript. — (5) Ezekiel M.: *Methods of Corre-*

lation Analysis. Wiley and Sons, New York (1956). — (6) HAURWITZ B. & J. M. AUSTIN: *Climatology.* McGraw-Hill Book Co., New York (1944). — (7) KOEPPE C. E. & G. C. DELONG: *Weather and Climate.* McGraw-Hill Book Co., New York (1958). — (8) MILLS F. C.: *Statistical Methods, Applied to Economics and Business.* Henry Holt and Co. (1939). — (9) PEARSON K.: *On the Correction Needful in the Case of the Correlation Ratio.* Biometrica Vol. 8, p. 254-256 (1911). — (10) PEARSON E. J. & H. O. HARTLEY: *Biometric Tables for Statistician,* Vol. I. Cambridge University Press (1958). — (11) SHAW, SIR N.: *The Drama of Weather.* Cambridge (1933) p. 165. — (12) RODEN G. I. & G. W. GROVES: *On the Statistical Prediction of Ocean Temperatures.* Journal Geoph. Research, Vol. 65, p. 249-263 (1960). — (13) WOO T. L.: *Tables for Ascertaining the Significance or Non-significance of Association Measured by the Correlation Ratio.* Biometrica, Vol. 21, p. 1-8 (1929). — (14) WADSWORTH G. P.: *Short Range and Extended Forecasting by Statistical Methods.* Air Weather Service, Technical Report 105-37 (1948).

(*Received* 23 *November* 1961)