AN INVESTIGATION OF THE VALUE OF
SYMMETRY IN FORECASTING

by

EDWARD NORTON LORENZ
A.B., Dartmouth College
1938
A.M., Harvard University
1940

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
1943

Signature of Author.....................
Department of Meteorology, October 6, 1943

Certified by..
Thesis Supervisor

....
Chairman, Dept. Comm. on Graduate Students

28

CONTENTS

The phenomenon of symmetry in atmospheric pressure curves has been the subject of several investigations during the past twenty years, Symmetry may be described as follows:

Let a curve of atmospheric pressure against time be defined by the equation

$$y = f(t) \quad .$$

This curve is symmetric at t_0 if

$$f(t_0 + \tau) = f(t_0 - \tau)$$

for all values of τ. In practice we cannot expect this condition to hold exactly. Instead, we say that the curve is symmetric at t_0 if

$$f(t_0 + \tau) - f(t_0 - \tau)$$

is numerically small for most of the values of τ making $t_0 + \tau$ and $t_0 - \tau$ fall into the time interval under consideration. The exact limitations on the terms "small" and "most", and on the extent of the time interval, are up to the investigator. Different interpretations of these terms lead to the various definitions of symmetry which have been used.

The symmetry just described is called <u>regular</u> or <u>simple</u> symmetry by some investigators. In addition, they consider <u>irregular</u> or <u>double</u> symmetry, which occurs at t_0 if

$$f(t_0 + \tau) + f(t_0 - \tau)$$

is nearly constant for most values of τ under consideration.

It is often not practicable to obtain $f(t)$ for all values of t.

Instead $f(t)$ is determined at equally spaced times $t_1, t_2, \cdots$, and the series $f(t_1), f(t_2), \cdots$ replaces the curve. Regular symmetry then exists at t_k if

$$f(t_{k+i}) - f(t_{k-i})$$

is numerically small for most of the values of i under consideration. An analogous condition defines irregular symmetry.

The first study of symmetry in atmospheric pressure curves was made by L.Weickmann[1] in Leipzig,Germany. Weickmann noted that curves of pressure at European stations often appear to be symmetric. To test for symmetry at t_0, he superposed upon the pressure curve $y = f(t)$ the curve

$$y' = f(t_0 - t)$$

obtained from the original curve by reversing the direction of time along the horizontal axis. If the two curves nearly coincide, there is good symmetry. No quantitative definition of symmetry was made.

Weickmann explained symmetry by treating the atmospheric pressure as a combination of sinusoidal pressure waves. Any sine curve, say

$$y = a \cos k(t - \tau) \quad ,$$

has a regular symmetry when

$$k(t - \tau) = n\pi \quad , \quad n \text{ an integer,}$$

and has an irregular symmetry when

$$k(t - \tau) = (n + \tfrac{1}{2})\pi \quad , \quad n \text{ an integer.}$$

Now if several curves have regular symmetries at $t = t_0$, their sum will have a regular symmetry at $t = t_0$. Hence, if a pressure curve approximates the sum of several sine curves, good symmetries are likely to occur.

The investigation begun by Weickmann was continued at the Geophysical

Institute in Leipzig, Germany. It is described by B. Haurwitz.[2] From the study of pressure curves at numerous European stations, it appears that symmetries usually occur near the solstices rather than near the equinoxes, and that curves tend to show better symmetries in winter than in summer. Harmonic analysis of the curves shows that a 24 day or a 36 day pressure wave tends to predominate over Europe in winter. The presence of such waves supports the concept of pressure curves as sums of sine curves.

Maps of the amplitudes and phases of these waves over the northern hemisphere show that they appear to emanate from the polar regions, where they are most pronounced. They also show that a strong wave over Europe does not necessarily imply such a wave over the United States.

More recently, investigators have used quantitative definitions of symmetry. K. Stumpff[3] has defined a symmetry index S. If the series to be tested for symmetry contains $2n+1$ terms

$$y_{-n},\ y_{-n+1},\ \text{------},\ y_{n-1},\ y_n\ ,$$

then by Stumpff's definition,

$$S = \frac{\sum_{\nu=1}^{n} (y_\nu - y_{-\nu})^2}{\sum_{\nu=1}^{n} (y_\nu - y_{-\nu})^2 + \sum_{\nu=1}^{n} (y_\nu + y_{-\nu})^2 - \frac{1}{n}\left(\sum_{\nu=1}^{n} (y_\nu + y_{-\nu})\right)^2} .$$

For perfect regular symmetry, $y_\nu - y_{-\nu} = 0$ for all ν, and $S=0$, while for perfect irregular symmetry, $y_\nu + y_{-\nu} = 2y_0$, and $S=1$. The index is very useful for observing variations of symmetry with time and location.

Stumpff's index has been modified by E. Wahl.[4] He observed that pressure curves usually show secular variations, which may persist for several months, thereby tending to prevent the occurrence of regular symmetries. An example of such a variation is the normal trend of mean pressure from one month to the next. Wahl defined a new symmetry index which remains unchanged when a linear function is added to the original pressure function, and which reduces to Stumpff's index when the secular variation vanishes. Thus Wahl's index is essentially Stumpff's index corrected for linear secular variations.

An investigation of symmetry is now being made at the Massachusetts Institute of Technology, under the supervision of the Department of Meteorolgy, at the request of the United States Army Air Forces. Its purpose is to determine whether symmetry may be used in long range forecasting. Special attention is being devoted to symmetry in the United States. To date, the most complete phase of the investigation has been the study of regular symmetries in five day mean pressure curves. Such a curve is defined by the relation $y = f(t)$, where $f(t_0)$ is the average pressure during the five day period centered at t_0. These curves have been studied in the portion of the United States east of the Rocky Mountains, during the colder half of the year. In this paper we shall describe the methods used in this study, and present the conclusions which have been drawn.

II

To study all the symmetry points during a winter, we must use a pressure curve extending into both autumn and spring. We have arbitrarily chosen to investigate pressure in the 200 day period beginning on October 3. This period ends on April 20 in common years, and on April 19 in leap years. Such a period will be called a season. The thirty seasons contained in the years 1888-1918 have been studied. They have been numbered in order, the n^{th} season being denoted by S_n. Thus S_1 means the period Oct.3,1888-April 20,1889, while S_{30} means the period Oct.3,1917-April 20,1918.

Each season has been divided into forty consecutive five day periods, which we shall call intervals. They have been numbered in order, the n^{th} interval being denoted by I_n. Thus I_1 means the period Oct.3-7, while I_{40} means the period April 16-20, or April 15-19 in leap years.

The i^{th} interval of the j^{th} season will be denoted by I_{ij}. That is, I_{ij} means I_i of S_j. A comma will be placed between the two subscripts when confusion would exist otherwise.

Eight stations from the eastern two-thirds of the United States have been selected. They have been numbered in order of increasing west longitude, the n^{th} station being denoted by P_n. The stations follow:

P_1:	Boston,Mass.	P_5:	Vicksburg,Miss.
P_2:	Buffalo,N.Y.	P_6:	Duluth,Minn.
P_3:	Charleston,S.C.	P_7:	Kansas City,Mo.
P_4:	Indianapolis,Ind.	P_8:	Denver,Col.

Weather maps of the United States have been issued since the middle of the nineteenth century, first by the War Department and later by the

Weather Bureau. During the years 1888-1918, maps appeared every day at 8 A.M., E.S.T. These morning maps are the exclusive source of our data. Pressures not reported on the maps have been interpolated from the isobars. So the quantity used for the pressure on a given day is the 8 A.M. pressure, reduced to sea level.

The quantity used for the mean pressure in a given interval I_{ij} is the arithmetic mean of the pressures on the five days of I_{ij}. It will be denoted by ψ_{ij}. Thus each ψ_{ij} is a function of the location P. The series of pressures $\psi_{1j}, \ldots, \psi_{40,j}$ during the j^{th} season will be denoted by $[\psi]_j$.

Pressures on these maps appear in inches of mercury, and are always expressed to two decimal places. If the decimal point is omitted, the sum of five pressures as reported may be treated as the arithmetic mean of the pressures, expressed in five hundredths of an inch. For this reason it is convenient to use five hundredth of an inch as the pressure unit in this work.

In some parts of the United States the normal pressure falls rapidly with the approach of spring. Hence we may expect secular variations to interfere with the occurrence of symmetries. We have eliminated the effect of the normal pressure changes through the season by subtracting from each series $[\psi]_j$ the series of normal pressures at the same station.

The normal pressures may be computed from the data for the thirty seasons studied. Thus if

$$a_i = \frac{1}{30}\sum_{j=1}^{30} \psi_{ij} \quad ,$$

we might use a_i as the normal pressure for I_i.

Examination of the curves $a_1, \text{----} a_{40}$ has shown that they are not as smooth as might reasonably be expected. Hence some smoothing process seems desirable. Our process has been chosen for its simplicity. Let

$$b_i = a_{i-1} + a_i + a_{i+1} \qquad , \qquad i = 1, \cdots, 40 \quad ,$$

with the understanding that $a_0 = a_1,\ a_{41} = a_{40}$. Then let

$$c_i = b_{i-1} + b_i + b_{i+1} \qquad , \qquad i = 1, \cdots, 40 \quad ,$$

where $b_0 = b_1,\ b_{41} = b_{40}$. Letting

$$a_i' = \frac{1}{9} c_i \quad ,$$

we find that

$$a_i' = \frac{1}{9}\left(a_{i-2} + 2a_{i-1} + 3a_i + 2a_{i+1} + a_{i+2}\right),$$

provided $a_{-1} = a_2,\ a_0 = a_1,\ a_{41} = a_{40},\ a_{42} = a_{39}$. Examination shows that the values a_i' form reasonably smooth curves. We have chosen them as the normal pressures. It is worth noting that

$$\sum_{i=1}^{40} c_i = 3\sum_{i=1}^{40} b_i = 9\sum_{i=1}^{40} a_i \quad ,$$

so that

$$\sum_{i=1}^{40} (a_i - a_i') = 0 \quad .$$

The values a_i' are now used to eliminate the secular variations. Letting

$$y_{ij} = x_{ij} - a_i' \quad ,$$

we obtain the series $[y]_j$ to be examined for symmetry. Although y_{ij} is actually the departure of x_{ij} from normal, we shall refer to it freely as the pressure at I_{ij}.

III

A quantitative definition of symmetry is now in order. To use a suspected symmetry at $I_{k,j}$ for forecasting, we forecast the pressure $y_{k-i,j}$, which occurred at $I_{k-i,j}$, to occur again at $y_{k+i,j}$, for the values of i being considered. Hence we might choose some constant C as the maximum allowable error for a good forecast, and say that symmetry occurs at $I_{k,j}$ if

$$\Delta_{k,i,j}(P) \equiv \left| y_{k+i,j}(P) - y_{k-i,j}(P) \right| \leqq C \;,$$

for all values of i under consideration.

However, when the expected value of $\Delta_{k,i,j}$ is small, a symmetry showing that $\Delta_{k,i,j} \leqq C$ is less valuable than it would be if the expected value of $\Delta_{k,i,j}$ were large. Thus we ought to modify the condition, requiring that

$$\Delta_{k,i,j}(P) \leqq C_{k,i}(P) \qquad ,$$

where $C_{k,i}(P)$ depends upon the expected difference between two pressures $y_{k+i,m}(P)$ and $y_{k-i,n}(P)$ chosen from arbitrary seasons S_m and S_n.

Such a definition is too cumbersome for our purposes. We have required merely that

$$\Delta_{k,i,j}(P) \leqq c(P) \qquad ,$$

where $C(P)$ depends upon the expected difference of two pressures at P in arbitrarily chosen intervals.

To study this expected difference, consider the 1200 values y_{ij}. Let us temporarily denote them by $\eta_1, \ldots \eta_{1200}$ in any order.

Then consider the $(1200)^2$ differences

$$\delta_{ij} = \eta_i - \eta_j \quad .$$

These are the possible departures of two pressures at P, chosen from arbitrary intervals.

We shall compute the standard deviation S of the values δ_{ij}.

By definition,

$$S^2 = \frac{1}{(1200)^2}\sum_{i,j=1}^{1200} \delta_{ij}^2 - \frac{1}{(1200)^4}\left(\sum_{i,j=1}^{1200} \delta_{ij}\right)^2$$

Now $\delta_{ii} = 0$ and $\delta_{ij} = -\delta_{ji}$. For $i \neq j$, every quantity δ_{ij} may be paired off with its negative, δ_{ji}. Hence

$$\sum_{i,j=1}^{1200} \delta_{ij} = 0$$

So

$$\begin{aligned}
(1200\ S)^2 &= \sum_{i,j=1}^{1200} \delta_{ij}^2 \\
&= \sum_{i=1}^{1200}\sum_{j=1}^{1200}\left(\eta_i^2 - 2\eta_i\eta_j + \eta_j^2\right) \\
&= 1200\sum_{i=1}^{1200}\eta_i^2 - 2\sum_{i=1}^{1200}\eta_i\sum_{j=1}^{1200}\eta_j + 1200\sum_{j=1}^{1200}\eta_j^2 \quad . \\
S^2 &= 2\left[\frac{1}{1200}\sum_{i=1}^{1200}\eta_i^2 - \frac{1}{(1200)^2}\left(\sum_{i=1}^{1200}\eta_i\right)^2\right] \\
&= 2\sigma^2 \quad ,
\end{aligned}$$

where σ is the standard deviation of the 1200 quantities η_i, i.e., of the 1200 quantities y_{ij}.

Thus we should define $c(P)$ in terms of S, or of σ. Arbitrarily, we have chosen

$$c(P) = \frac{1}{2}\ S(P) = \frac{1}{\sqrt{2}}\ \sigma(P) \quad .$$

To complete the definition of symmetry we must specify the values of i which are under consideration. We shall say that a <u>symmetry of duration</u> n exists at P at I_{hj} if

$$(1)\quad \Delta_{h,i,j}(P) \leqq c(P), \qquad i = 1, \cdots, n,$$

and

$$(2)\quad \Delta_{h,n+1,j}(P) \quad \begin{cases} > c(P), & \text{or} \\ \text{does not exist.} \end{cases}$$

Similarly, a symmetry of duration n exists at P between I_{hj} and $I_{h+1,j}$ if

$$(1)\quad \Delta'_{h,i,j}(P) \equiv \left| y_{h+i,j}(P) - y_{h-i,j}(P) \right| \leqq c(P), \quad i=1,\cdots,n,$$

and

$$(2)\quad \Delta'_{h,n+1,j}(P) \quad \begin{cases} > c(P), & \text{or} \\ \text{does not exist.} \end{cases}$$

In general, we shall say that two pressures $y_{ij}(P)$ and $y_{hl}(P)$ <u>agree</u> if

$$\left| y_{ij}(P) - y_{hl}(P) \right| \leqq c(P) \quad .$$

In computing $c(P)$ we note that

$$\sum_{i=1}^{40} \sum_{j=1}^{30} y_{ij} = \sum_{i=1}^{40} \sum_{j=1}^{30} (x_{ij} - a_i')$$

$$= \sum_{i=1}^{40} (30\, a_i - 30\, a_i') = 0 \quad .$$

Hence

$$\sigma = \sqrt{\frac{1}{1200} \sum_{i=1}^{40} \sum_{j=1}^{30} y_{ij}^2} \quad ,$$

and σ and c are readily computed. We give the values c at the eight stations, in five hundredths of an inch.

P	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
$C(P)$	67	57	42	52	43	59	56	55

Symmetries of duration 1 or 2 hardly merit the designation of symmetry. Certainly no such symmetry could be used for an extended forecast. Hence no symmetries occuring from I_1 to I_3 or from I_{38} to I_{40}, where $\Delta_{k,i,j}$ is not defined for $i > 2$, have been enumerated. All symmetries in the remainder of the seasons have been located, and listed according to duration.

We present a table showing for each station the total number T_n of symmetries of duration $\geq n$, for $n = 1, \dots, 8$. The final column gives the arithmetic mean $\overline{T}_n$ of $T_n(P_i)$, $i = 1, \dots, 8$.

n	$T_n(P_1)$	$T_n(P_2)$	$T_n(P_3)$	$T_n(P_4)$	$T_n(P_5)$	$T_n(P_6)$	$T_n(P_7)$	$T_n(P_8)$	$\overline{T}_n$
1	823	802	839	754	804	713	752	832	789.9
2	330	305	348	285	305	242	262	319	299.5
3	117	122	127	98	111	88	92	114	108.6
4	43	43	67	43	36	27	37	31	40.9
5	14	18	23	15	15	11	13	15	15.5
6	4	7	8	8	5	2	7	3	5.5
7	3	2	0	3	3	2	0	1	1.8
8	2	1	0	0	3	0	0	0	0.8

A symmetry of duration 11, occurring at P_5, was the only symmetry encountered lasting more than eight intervals.

Examination shows that T_n varies rather little with P. The most important exceptions occur at P_3, where in general $T_n > \overline{T}_n$, and at P_6, where in general $T_n < \overline{T}_n$.

We shall compare the observed values of T_n with the values T_{n0} of T_n to be expected if symmetry occurred merely by chance. Let q be the probability that two arbitrary pressures $y_{ij}(P)$ and $y_{kl}(P)$ agree. Then

$$T_{n0} = q^n N_n$$

where N_n is the number of possible occurrences of symmetries of duration $\geq n$.

We shall determine q theoretically. As determined previously by the $(1200)^2$ values S_{ij}, the standard deviation of all differences $y_{ij}(P) - y_{kl}(P)$ is $2c(P)$. Now the probability[5] that a quantity, taken from a normal distribution with arithmetic mean 0 and standard deviation σ, will lie between x and $-x$ is

$$\phi\left(\frac{x}{\sigma\sqrt{2}}\right),$$

where

$$\phi(x) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)\, n!}.$$

If the differences $y_{ij}(P) - y_{kl}(P)$ form a normal distribution, we should expect

$$q = \phi\left(\frac{c(P)}{2c(P)\sqrt{2}}\right) = \phi\left(\frac{1}{2\sqrt{2}}\right)$$

$$= 0.382927 .$$

This quantity 0.382927 will be denoted by q_0.

The values of T_{n0}, computed under the assumption that $q = q_0$, appear below, following the values of N_n. For comparison, the values of $\overline{T}_n$ are repeated.

n	N_n	T_{n0}	$\overline{T}_n$
1	2070	792.7	789.9
2	2070	303.5	299.5
3	2070	116.2	108.6
4	1950	41.9	40.9
5	1830	15.1	15.5
6	1710	5.4	5.5
7	1590	1.9	1.8
8	1470	0.7	0.8

The agreement between T_{n0} and $\overline{T}_n$ is truly remarkable. Values of $T_n(P)$ agree fairly well with T_{n0}. Thus it is reasonable to state that symmetry, as defined here, occurs no more often than it would if it were only a result of chance. Immediately one wonders whether symmetry is nothing more than a chance phenomenon.

Closely related to the above discussion is the question of persistence of symmetry, which offers a suggested method of forecasting. At the end of an interval I_{kj} we may observe that a symmetry of duration at least n has occurred at $I_{k-n,j}$. We then assume that the symmetry at $I_{k-n,j}$ will persist, and forecast the pressures $y_{k-2n-1}, y_{k-2n-2}, \ldots$ to occur again at $I_{k+1}, I_{k+2}, \ldots$.

Such a forecast will verify well only if the probability p_n, that a symmetry of duration $\geq n$ will also have duration $\geq n+1$, is considerably greater than q. Now if $p_n \geq q$ for $n = 1, \ldots, k$, then

$$p_1 p_2 \cdots p_k \geq q^k .$$

Moreover, the probability of occurrence of a symmetry of duration $\geq k$ at any given time is $p_1 p_2 \cdots p_k$. We have just seen that this probability is not noticeably greater than q^k, if $q = q_0$. So in general p_n cannot greatly exceed q unless for some $m < n$, $p_m < q$. Or, certain symmetries will tend to persist only if other symmetries tend not to persist.

Symmetries of duration $n-1$ persist if

$$\frac{T_n}{N_n} > q \frac{T_{n-1}}{N_{n-1}} .$$

Values of $r_n = \dfrac{T_n N_{n-1}}{T_{n-1} N_n}$ and of $\bar{r}_n = \dfrac{\bar{T}_n N_{n-1}}{\bar{T}_{n-1} N_n}$ appear below, for $n = 1, \cdots, 6$. They are to be compared with q.

n	$r_n(P_1)$	$r_n(P_2)$	$r_n(P_3)$	$r_n(P_4)$	$r_n(P_5)$	$r_n(P_6)$	$r_n(P_7)$	$r_n(P_8)$	$\bar{r}_n$
1	0.397	.387	.405	.364	.388	.344	.363	.402	0.382
2	.401	.380	.415	.378	.379	.338	.348	.383	.379
3	.355	.400	.365	.344	.364	.365	.351	.357	.363
4	.397	.373	.560	.466	.344	.326	.427	.289	.400
5	.346	.447	.366	.372	.444	.434	.374	.516	.404
6	.306	.416	.372	.570	.356	.195	.576	.214	.380

In most cases these ratios do not differ greatly from q_0, except where T_{n-1} is too small to make the results significant. A notable exception occurs at P_3, where 67 out of 127 symmetries of duration $\geqq 3$ persist. These figures probably indicate that symmetry lasting three intervals at P_3 does tend to persist. But use of this result will yield only a few forecasts a season, none of them extending more than five days ahead, and many of them not verifying. In general, the ratios support the idea that symmetry is a chance phenomenon.

We now observe that even if we could recognize a symmetry upon its arrival, our forecasts would be rather infrequent. The total number of symmetries of duration $\geqq 4$ is 327, or an average of 1.36 per season per station. The most favorable average, at P_3, is 2.23. So we could make pressure forecasts only about twice a season at each station.

Using symmetries of duration $\geqq 3$, we could make forecasts about four times a season. The majority of these would not extend more than fifteen days ahead. Moreover, at every station except P_5 there were some seasons during which no forecasts at all could be made.

As mentioned before, a symmetry of duration 1 or 2 can hardly be called a symmetry. Such symmetries were located entirely for statistical purposes.

The scarcity of symmetries, as defined here, makes it seem hardly worth while to look for a method of recognizing them upon their arrival. The fact that their occurrence resembles that of symmetries in a random series suggests that perhaps such a method does not exist. Instead, it seems more desirable to look for another definition of symmetry which will not have the failures of the first definition.

IV

The examples of symmetry given by Weickmann show good agreement between the curves $y = f(t)$ and $y' = f(t_0 - t)$, but usually the curves diverge widely at a few points. Stumpff's index admits disagreements in pressure, since not all the differences $|y_\nu - y_{-\nu}|$ need be s mall provided their sum is reasonably small. Similarly, in the study of symmetry in the United States, we note intervals which would possess symmetries of long duration were it not for one or two pairs of disagreeing pressures. Such intervals should be classed as possessing symmetries, since their symmetric properties, if recognized, would be of definite forecasting value.

In our second definition, we shall s ay that a symmetry of duration n exists at I_{kj} if

(1) $\Delta_{k,i,j} \leqq c$ for at least $\frac{3}{4} n$ of the values $i = 1, \dots, n$.

(2) There is no $n' > n$ such that $\Delta_{k,i,j} \leqq c$ for at least $\frac{3}{4} n'$ of the values $i = 1, \dots, n'$.

Analogous relations will define symmetry occurring between I_{kj} and $I_{k+1,j}$.

We observe that for any integer m, a symmetry of duration $4m$, $4m+1$, or $4m+2$ can have at most m disagreements. Symmetries of duration $4m+3$ do not exist, since fulfillment of condition (1) above would imply failure of condition (2).

We have applied this definition to all the pressure curves at P_1, P_5, and P_7. We give below the number D'_n of symmetries of duration n, and the number T'_n of symmetries of duration $\geqq n$, for $n = 4, \dots, 17$. The arithmetic means $\overline{D}'_n$ and $\overline{T}'_n$ of D'_n and T'_n also appear.

n	$D'_n(P_1)$	$D'_n(P_5)$	$D'_n(P_7)$	$\overline{D}'_n$	$T'_n(P_1)$	$T'_n(P_5)$	$T'_n(P_7)$	$\overline{T}'_n$
4	203	203	149	188.3	353	346	287	332.0
5	75	62	68	68.3	150	143	138	143.7
6	23	27	30	26.7	75	81	70	75.3
8	26	30	20	25.3	52	54	40	48.7
9	17	8	11	12.0	26	24	20	23.3
10	1	7	5	4.3	9	9	9	11.3
12	1	3	4	2.7	8	6	4	7.0
13	4	2	0	2.0	7	4	0	4.3
14	0	3	0	1.0	3	1	0	2.3
16	1	1	0	0.7	3	0	0	1.3
17	2	0	0	0.7	2	0	0	0.7

No symmetries of duration > 17 were encountered.

Looking at $\overline{T}'_4$, we see that the stations possess an average of 11 symmetries a season, each lasting at least four intervals. By an actual count it was found that if these symmetries had been recognized upon their arrival, pressure could have been forecast for an average of 25.2 of the last 35 intervals of each season, at each station. Moreover, at P_5 there was no season when forecasts could not have been made for at least one half of the intervals, while P_1 and P_7 possessed only three such seasons apiece. We have thus found a definition yielding enough symmetries for fairly regular forecasts.

We shall now compare the observed values of D'_n and T'_n with the values D'_{n0} and T'_{n0} to be expected by chance. Again we assume $q = q_0$. The values follow.

n	D'_{n0}	$\bar{D}'_n$	T'_{n0}	$\bar{T}'_n$
4	168.2	188.3	324.1	332.0
5	69.6	68.3	155.9	143.7
6	24.9	26.7	86.3	75.3
8	31.8	25.3	61.4	48.7
9	13.1	12.0	29.6	23.3
10	5.1	4.3	16.5	11.0
12	6.1	2.7	11.4	7.0
13	2.6	2.0	5.3	4.3
14	1.1	1.0	2.7	2.3

The sets of values agree fairly well. There are, however, some discrepancies worth observing. In general, the number of symmetries of duration 4 is greater than expected, while the number of symmetries of duration ≥ 8 falls below expectation. Also, P_7 shows considerably fewer symmetries than does P_1 or P_5. These discrepancies need not be ascribed to chance. They can be explained.

In computing the expected occurrences, we have assumed that q is independent of time and location. As a check, we consider the probability q_n that two pressures $y_{kj}(P)$ and $y_{k+n,j}(P)$, n intervals apart, will show an agreement. By an actual count of the number of agreements, we have calculated q_n at P_1, P_5, P_7, for $n = 1, \cdots, 20$. We present the values, together with the arithmetic means $\bar{q}_n$ of q_n. The figures in the final row are the observed probabilities of an agreement between two pressures $\leqq$ 20 intervals apart.

n	$q_n(P_1)$	$q_n(P_5)$	$q_n(P_7)$	$\bar{q}_n$
1	0.437	.401	.394	.411
2	.392	.396	.376	.388
3	.394	.387	.359	.380
4	.387	.383	.372	.381
5	.399	.390	.384	.391
6	.369	.400	.384	.384

n	$q_n(P_1)$	$q_n(P_5)$	$q_n(P_7)$	$\bar{q}_n$
7	.380	.405	.393	.393
8	.391	.386	.403	.393
9	.410	.377	.371	.386
10	.359	.387	.380	.375
11	.383	.393	.386	.387
12	.377	.382	.385	.381
13	.370	.388	.384	.381
14	.383	.382	.358	.374
15	.356	.377	.352	.362
16	.366	.371	.346	.361
17	.378	.390	.372	.380
18	.373	.389	.386	.383
19	.359	.397	.360	.372
20	.377	.422	.382	.394
	.3846	.3919	.3773	.3846

The final average of 0.3846 agrees very well with q_0, and indicates that the assumption $q = q_0$ is justified. But the variation of $q_n(P)$ with n and P is significant. For small values of n, q_n tends to exceed q_0, while for most of the larger values of n, $q_n < q_0$. We might anticipate the observation $q_1 > q_0$ from the continuity of the pressure function. So although the pressures y_{ij} may form a normal distribution, their order within a season is not entirely random.

Since the expected occurrence of symmetries of duration 4 depends mostly upon $q_1, \cdots, q_8$, while symmetries of longer duration depend also upon the remaining probabilities, the discrepancies between D'_{n0} and $\bar{D}'_n$ agree qualitatively with the arrangement of the probabilities q_n. Also since q_n tends to be smaller at P_7 than at P_1 and P_5, it is not surprising that P_7 shows fewer symmetries.

The results make one wonder again whether there is any method of recognizing a symmetry upon its arrival.

V

We conclude with a few statements which we believe have been definitely established, and with some conjectures which seem to follow.

Concerning five day mean pressure curves in the portion of the United States east of the Rocky Mountains, the following may be stated:

1. The phenomenon of symmetry occurs often enough so that it would be of great use in forecasting if a symmetry could be recognized upon its arrival.

2. The phenomenon of symmetry occurs no more often than it would in a series of terms with a normal frequency distribution and a random order.

3. Symmetry shows little tendency to persist.

No research has been done here to verify the following conjectures concerning symmetry. They are merely presented as conclusions suggested by this investigation.

Conjecture 1: Symmetry in curves of five day mean pressure, in the portion of the United States east of the Rocky Mountains, is but a chance occurrence. It is therefore not worth while to look for a method of forecasting the occurrence of symmetry, or of recognizing it.

Conjecture 2: Symmetry in daily pressure curves in the United States is also a phenomenon most of whose occurrences are due to chance.

Conjecture 3: Symmetry cannot be used as a method of forecasting pressure in the United States. At best it may become an aid to some other method, which will be workable without the use of symmetry.

REFERENCES

1. L.Weickmann. Wellen im Luftmeer. Sächsischen Akad. der Wiss. Abhandlungen der Math.-Phys. Klasse Vol.39,no.2,1924.

2. B.Haurwitz. Investigations of Atmospheric Periodicities at the Geophysical Institute,Leipzig,Germany. Monthly Weather Rev., Vol.61,Aug.1933

3. K.Stumpff. Untersuchungen über die Symmetrieeigenschaften von Luftdruckkurven. Veröff des Meteor.Inst.der Univ.Berlin,Vol.3,no.1,1938

4. E.Wahl Eine Erweiterung des Begriffes der Symmetrie von Luftdruckkurven. Meteor.Zeit.Vol.56,Dec.1939

5. E.Whittaker and G.Robinson The Calculus of Observations, p.179-183. D.Van Nostrand Co., New York,1924.

ACKNOWLEDGEMENT

The writer wishes to express his most sincere thanks to all the employees of Project 6136, Division of Industrial Cooperation, whose work in assembling data and performing numerous computations has made this research possible.