THE PRESENT STATUS OF CORRELATION INVESTIGATION IN METEOROLOGY [1]

By FRANZ BAUR

(Translated by W. W. Reed)

Since merely first attempts at an application of correlation reckoning to the phenomena of daily weather and related problems are made here, this report was limited to those investigations that are, as a rule, in mind when there is discussion of meteorological correlation investigation, that is, to the investigation of weather phenomena and their relations.

These investigations are divided into two groups which are, on the whole, not to be sharply differentiated:

1. Investigations of the relations of weather anomalies to coexistent anomalies in other regions, and
2. Investigations of the relations of weather anomalies to antecedent or subsequent anomalies at the same place or even at another place.

The most comprehensive investigations of the first class are those made by G. T. Walker. According to his view

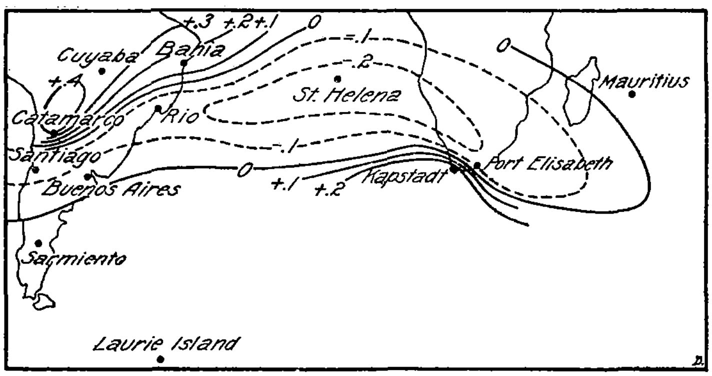

FIGURE 1.—Lines of equal correlation coefficients between January pressure at Laurie Island (South Orkneys) and the coexisting pressure in the high pressure area of the South Atlantic. 1903-1923.

we have on the earth's surface three pressure oscillation systems:

(*a*) The North Atlantic oscillation characterized by negative correlation between the pressure over the Azores and the coexistent pressure over Iceland;

(*b*) The North Pacific oscillation, which appears in a corresponding negative correlation between the air pressure at Honolulu and that over the Aleutian Islands; and

(*c*) The so-called see-saw of the Southern Hemisphere, which consists in a negative correlation between the pressure over the Indian Ocean and that in the South Pacific Ocean, in which, in general, the regions of high pressure in the Southern Hemisphere have a pressure oscillation following the course of that of the South Pacific, while the regions with low pressure in summer have a pressure oscillation paralleling that of the Indian Ocean.

The first two oscillation systems are very evidently connected with that total of current systems that we are accustomed to call the "general atmospheric circulation." This connection comes about in that:

1. The velocity of the west-east current in the temperate zone depends on the amount of the pressure gradient from the subtropical high pressure region to the low pressure region in about latitude 65° N. and;
2. Low pressure near Iceland and low pressure over the Aleutian Islands brings about zonal pressure differences, which make possible the exchange of air masses between the subtropical and the polar regions in adjacent meridional currents.

On the other hand, despite all endeavors devoted to the purpose, there has been thus far no success in fitting the air pressure see-saw of the Southern Hemisphere into the system of the general circulation of that region in the same clear and natural manner. It must be pointed out here that the correlation coefficients characterizing the see-saw are considerably smaller and more changeable from year to year than is the case with the coefficients in the northern oscillation systems, especially that in the North Atlantic.

On theoretical grounds, however, it is probable that meridional pressure compensations exist in the Southern Hemisphere also. At present they can be demonstrated only with very great difficulty, since there are no suitably located observation stations in the pressure trough between 50° and 70° S. The only station here is Laurie Island (South Orkney Islands), from which there is available at present only a short series of observations. The pressure correlations, calculated by myself, for this station with the different stations in the Southern Hemisphere in latitudes 15° to 45° permit the conclusion—as will appear from Figures 1 and 2—that there is a negative correlation of pressure between lower and higher latitudes; and it appears that the central region of the South Atlantic High has the most marked negative pressure correlation with Laurie Island, thus, that in the South Atlantic there exists a pressure compensation system such as we have long recognized in the North Atlantic.

As we delve deeper and deeper into weather phenomena it appears that the so-called "general atmospheric circulation" is a rather complicated system, made up of a whole series of subsystems and that, in the temperate zones especially, the essential part of this general circulation is not the west current or—more correctly stated—the excess of westerly wind components, but consists in band-like meridional circulations that bring about exchange between subtropical and polar air. In view of this there arises the question as to how far it is correct to speak of a "general atmospheric circulation" in the sense of a homogeneous system with homogeneous oscillations. The question here proposed is most closely connected with another which I have designated as the basic question of macro-meteorology, which can be formulated thus: In the oscillations of intensity in the different current systems, whose sum total we designate as the gen-

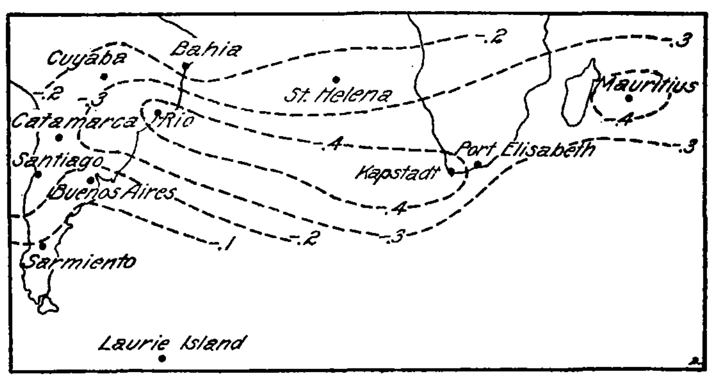

FIGURE 2.—The same except for July instead of January.

[1] *Zeitschrift für Geophysik, Jahrg. 5, Heft 8. 1929.* The paper from which this one is abstracted appears in *Meteorologische Zeitschrift, Heft 2. 1930.*
See also Gilbert T. Walker's discussion of this paper in *Meteorologische Zeitschrift Heft 6, s. 229-231* (in German).

eral atmospheric circulation, are there involved processes which in the main have their basis in themselves, that is, in which a state B is conditioned in greatest part by a preceding state A (atmospheric or hydrospheric), or is the cause of the changes to be sought in greatest part outside the atmospheric mechanism, thus in cosmic influences or, for example, in dust turbidity due to volcanic eruption?

In his *Untersuchung der Schwankungen der allgemeinen Zirkulation* A. Wagner has shown that in the consideration of 10-year average values there appear between different decades systematic differences in air pressure, temperature, yearly temperature amplitude, and precipitation which permit us the inference of changes in the general circulation as a whole and make it appear probable that these changes have their cause in changes in solar radiation. When, however, we make a numerical comparison between the changes found by Wagner in the 10-year mean values and the anomalies that we encounter in the monthly means there is found such a vast difference that there appears to be full warrant for doubt that these anomalies, which determine the actual character of the weather, are caused chiefly by exterior influences and not by inherent conformity to law in the individual current systems.

These inherent laws bring it about that the oscillations in the individual current systems naturally take place not independently and yet not in homogeneous manner. We have confirmation of this in the negative correlation of air pressure over Iceland with coexistent pressure over the Aleutian Islands in winter and spring, also in extremely low correlation (in spring even negative correlation) of pressure over the Azores with that over Honolulu, and again in very low correlation between air pressure over the Azores and over Iceland, too, with the coexistent pressure in the centers of the subtropical high pressure belt in the Southern Hemisphere; in addition even the correlations for annual means likewise show to some extent not even the sign that would be expected with homogeneous oscillation in all parts of the atmospheric circulation. Then, too, the synchronous correlations within the southern high-pressure belt are only small and not systematic.

In good agreement with these facts there stands the additional one that up to the present time no worthwhile relation could be pointed out between cosmic processes, especially oscillations in sunspots, and terrestrial weather phenomena, since if the oscillations in the individual parts of the atmospheric circulation were to be referred, in the main, directly to oscillations in solar radiation then they must appear uniformly in all current systems. G. T. Walker has calculated for a series of individual stations and for entire regions the correlation coefficients of three-month means of air pressure, temperature, and precipitation with the synchronous relative sunspot numbers. With the exception of those for temperature at Batavia and on Samoa, which are negative in all seasons, the correlation values are generally very small. In agreement with this is the fact that in periodogram analyses of the temperature course at points in the temperate zone the amplitude of the 11-year period always proves to be about the same as, or but little greater than, that to be expected. There could be shown no connection between amount of pressure gradient from the Azores to Iceland and the relative sunspot numbers and none between gradient and change in number from month to month. It may be noted, in conclusion, that Clayton's correlations with Abbot's solar constant are either rather low or—in so far as they show worthwhile values—are derived too unsystematically and from too scant material for conclusions of any certainty to be drawn therefrom.

All these facts lead to the conclusion that the oscillations actually occurring in the several circulation systems and the weather anomalies connected with such oscillations can not be explained exclusively by oscillations in the general atmospheric circulation or by outside influences causing the same, and further that it is rather the peculiar laws inherent in the individual current systems and the relations existing between successive weather anomalies that are of deciding significance.

Correlation reckoning is an important aid in the investigation of the relation between non-synchronous weather anomalies in different regions of the earth. But since there is involved in this the determination of statistical relations between antecedent and subsequent phenomena it can be understood that from the beginning there was connected with these investigations the thought of probable attainment, along this way and at some future time, of long-range weather predictions. It is probably due in part to this combination with the idea of a subsequent long-range forecast that the value of these investigations was not always judged objectively; on the one hand it was over-estimated relative to significance in long-range weather forecasting and on the other hand it was underestimated relative to general scientific value.

This pointed criticism has its basis in large part in the fact that the import of the values and the method of correlation calculation were not always correctly recognized and in the further fact that correlation calculation is, on the whole, not yet a fully developed branch of knowledge, but must be elaborated for the problems propounded to it.

The first error relative to correct discernment of the significance of the values was that of being of the opinion that the correlation coefficient shows directly by what fraction of its amount the one variable is contained in the other. As the result of this idea the value of the first relation equations was overestimated since this fraction, or percentage, is given, not by the correlation coefficient, but by the square of that value.

Consequently the requirement is now made that a relation equation intended for use in forecast purposes shall show at the very least a total correlation coefficient or correlation index of 0.71, since only then has there been comprehended in the elements contained in the relation equation at least the half of all influences.

If we consider in this light some relation equations that have become well known hitherto, we recognize that only a small percentage of these satisfy the condition. Out of the twelve relation equations derived by Walker for different regions and seasons with a view to the prediction of precipitation in India only three fulfill the imposed condition. Among these three equations we find, to be sure, one with a total correlation coefficient of 0.94. However, if this equation is considered more closely it is found that in it are contained a series of questionable correlations.

A high total correlation coefficient is still not enough sufficient for the assumption that a relation equation can serve as basis for a forecast. In addition to this it is necessary that the individual relations be stable and real. The investigation of stability and reality of the relations is unconditionally necessary in order that the correlations may be rightly interpreted.

A good example of the fact that even relations showing noteworthy correlations in a period of 50 years may be unstable is given by the relation between air pressure in Argentina from April to June and the following winter temperatures in northwestern Germany, which gives for the period 1874–1923 a coefficient of—0.48. If we go back to 1858 and calculate the coefficients for 23-year periods, we have:

1858 to 1880	+0.03
1881 to 1903	−.51
1904 to 1923	−.10

The correlation is, therefore, obviously unstable. This result opens the way to suspicion that the numerous other relations connected with the air pressure over South America in the autumn of the Southern Hemisphere, and hitherto considered as dependable, are unstable.

By real correlations we are to understand those in which there comes to light a coherence in the oscillations of two phenomena in contrast to the apparent, "symptomatic" correlations, in which a correlation is simulated by a secular variation either direct or inverse. The remarkable correlations by Groissmayr between the annual precipitation at Charleston and the Nile flood two years later is ovbiously an outstanding symptomatic one, as can be easily proven by the calculation of the Tschuprow divergence coefficient.

In meteorological correlation investigations we must now make the requirement that stability and reality of correlations be investigated in detail. The most important requisite for a wise application of correlation reckoning in meteorology (and in geophysics) is the formulation of a clear and studied statement of the problem in a theoretical manner. The meagerness of the results of previous correlation investigations arises from the fact that these requirements were, for the greater part, not fulfilled.

There are given, in conclusion, some results of the recent investigations of the writer relative to cases of conformity to law found in the air pressure oscillations in the belts of circulation over the North Atlantic Ocean. In strictest adherence to theoretical considerations and calculations, which showed that the air pressure on the middle meridian of belt of circulation oscillates in dependence on the preceding zonal differences in pressure and temperature, there were calculated for the following table the correlation coefficients of the monthly means of zonal pressure and temperature differences with (*a*) the air pressure gradient Ponta Delgada–Iceland from the month under consideration to the next, and (*b*) with the change in air pressure over Iceland in the same period.

Correlation coefficients of the zonal pressure and temperature differences

	I–I/II	II–II/III	III–III/IV	IV–IV/V	V–V/VI	VI–VI/VII	VII–VII/VIII	VIII–VIII/IX	IX–IX/X	X–X/XI	XI–XI/XII	XII–XII/I	Mean
		(*a*) With change in pressure gradient Ponta Delgada-Iceland to the next month (1874–1923)											
Pressure difference:													
Rome-Ponta Delgada	+0.36	+0.48	+0.37	+0.53	+0.59	+0.46	+0.35	+0.32	+0.16	+0.48	+0.21	+0.27	+0.38
Indianapolis-Ponta Delgada	+.42	+.52	+.66	+.56	+.41	+.44	+.48	+.29	+.23	+.46	+.41	+.34	+.44
Haparanda-Stykkisholm	−.39	−.58	−.55	−.58	−.26	−.59	−.58	−.19	−.30	−.43	−.35	−.16	−.41
Jacobshavn-Stykkisholm	−.32	−.45	−.47	−.47	−.06	−.36	−.23	−.31	−.50	−.39	−.44	−.36	−.36
Temperature difference:													
Tromso-West Greenland	−.16	−.42	−.44	−.41	−.21	−.59	−.16	−.43	−.53	−.51	−.39	−.34	−.38
		(*b*) With change in pressure to the next month over Iceland (1874–1923)											
Pressure difference:													
Haparanda-Stykkisholm	+.41	+.61	+.58	+.66	+.41	+.61	+.67	+.19	+.38	+.43	+.42	+.27	+.47
Jacobshavn-Stykkisholm	+.41	+.57	+.55	+.59	+.22	+.34	+.34	+.30	+.54	+.42	+.57	+.47	+.44
Temperature difference:													
Tromso-West Greenland	+.18	+.42	+.49	+.48	+.21	+.60	+.16	+.45	+.54	+.55	+.47	+.38	+.41

NOTE.—Explanation of the preceding table: The first correlation coefficient, +0.36, is the correlation coefficient of the air pressure difference Rome-Ponta Delgada in January with the change in pressure gradient Ponta Delgada-Iceland from January to February; that is, with the difference February mean—January mean of this pressure gradient.

In agreement with the theoretical requirements it appeared that after a supernormal pressure difference Rome-Ponta Delgada there follows pressure rise at Ponta Delgada and therefore increase in the pressure gradient between Ponta Delgada and Iceland, and that after a greater than normal pressure difference Haparanda-Stykkisholm or Jacobshavn-Stykkisholm there follows pressure rise at Stykkisholm and therefore decrease in pressure gradient from Ponta Delgada to Stykkisholm. Knowledge that the relation to the meridional pressure gradient is effected indirectly at first is gained from the fact that the correlations of the two northern zonal pressure differences with the air pressure change at Stykkisholm is still greater than the correlations with the change in pressure gradient. The same is also the case relative to the correlation of the temperature difference Tromso-West Greenland, which is likewise in harmony with theoretical considerations. All of these correlations show the same sign in all months of the year. The order of the correlation coefficient is not the same, however, in all months of the year since the period of action is not always the same in the course of the year. The correlations are of greatest importance in the months of February, March, and April and later in October. Detailed investigation shows that in these months the relations are linear, stable, and real. From the relations to the four zonal pressure differences, the relation to the temperature difference Tromso-West Greenland, and the correlation with the preceding meridional temperature gradient there result relation equations with total correlation coefficients of 0.80 to 0.81. The value of R^2 is therefore 0.64 to 0.67. The change in pressure gradient is determined thus at some two-thirds of the zonal pressure differences, the temperature difference between the eastern and the western parts of the circulation belt, and the meridional temperature gradient—a new proof of the determinant significance of previous weather history.

For those months that gave rather small correlation values with the use of monthly means the extent of the period of action must first be obtained. Such investigations are now under way. There is in progress, too, a study as to whether the length of the period of action is dependent on sunspots or other influences.

Although these investigations relative to conformity to law in the oscillations of the North Atlantic air circulation have for the time being no direct value in long-range forecasting, since the weather depends not alone on the air pressure gradient from the Azores to Iceland, they mark a point of progress on the long, difficult road to long-range weather forecasting in that there are involved here the first relation equations that were obtained in strict adherence to theoretical considerations and calculations and which in the demonstrated stability and reality of each individual relation give a close connection between a meteorological value and a preceding complex of phenomena.