Popular Long-Range Weather Forecasts.

THE 50-day weather forecasts published in the *Daily Mail* have attained such popularity that an authoritative pronouncement from science as to whether they have or have not any real value has become desirable, and Capt. Cave is to be congratulated on opening an investigation on the subject.

The principles on which the curves are prepared are largely empirical, so that it is only by long-continued comparison of expectations with facts that their dependability can be estimated or improvement can be effected. Their preparation entails a large amount of labour and involves the maintenance of an office with technical assistance. They have never received the support of official meteorology, so that, without the public-spirited enterprise first of the *Field* and now of the *Daily Mail*, the work could not have been carried on. From the point of view of science, two consequences of this are unfortunate :

(1) The form in which they appear is popular rather than scientific.

(2) It has not been possible to publish full details of the methods used without disloyalty to those who have backed them.

Before trying to compare the expectation diagrams with facts, it is necessary to be quite clear as to what is aimed at in them. Since they are published for popular use and not for scientific study, the explanation given with them is not so complete as a scientist might wish. It is stated that—

(1) Nothing approaching infallibility is claimed.

(2) The diagrams indicate the expectation of rain; the higher the curve the greater the expectation.

(3) The curves are not intended as day-to-day forecasts, and their author is quite satisfied if the timing of his expectations is correct within 24 hours either way.

This was explained in some detail, but in popular form, in the *Daily Mail* of April 30, 1926.

For the benefit of more scientific readers their author would like to add :

(4) The datum line and the shaded and blacked areas are intended only as a guide to the eye and have no precise significance.

(5) While the curves primarily indicate degree of expectation, it is reasonable that they should also show some relation to rain amounts. At one time the investigations were based on rain or no-rain, 50 mm. being counted as exactly the same as 0·2 mm., but now (as from June 17, 1926) the factor of amounts has been introduced into the expectations, and their author aims at checking his results for each station by plotting the actual rain amounts to the vertical scale shown in the following table, which is roughly proportional to the logarithms of the amounts. It will be realised that this is an extremely searching test, and one which even 12-hour forecasts would not be able to stand.

Rain Amount.	Vertical Scale.
Nil	− 2 units
Up to 0·1 mm.	− 1 ,,
,, 1·0 ,,	0 ,,
,, 2·0 ,,	+ 1 ,,
,, 5·0 ,,	+ 2 ,,
,, 11·0 ,,	+ 3 ,,
,, 25·0 ,,	+ 4 ,,
,, 56·0 ,,	+ 5 ,,

(6) The diagrams for the British Isles as a whole show the extent to which rain is expected to be general over the British Isles.

It is to be regretted that Capt. Cave did not study the *Daily Mail* of April 30, 1926, or communicate with the author of the diagrams before making his tests, for it will be seen that neither his " weather numbers " nor his " forecast numbers " are appropriate. Both his criticism and his diagrams, therefore, being based on those numbers, fail to apply.

Further, since the complete set of curves is made for the 50 days independently every week, each one has a character of its own, and should be considered as a whole. To select the fourth to the tenth days out of each 50 for criticism is like criticising a 12-hour forecast by examining the weather for an interval of 103 minutes out of each day. For this reason, too, the comparison made in Capt. Cave's diagrams is not a fair one.

The *Daily Mail* curves are based on the combination of fifteen cycles which are selected for each of the four stations every week. Some of these cycles are well known, and no originality is claimed for them ; others have been developed for this special purpose, and have stood the test of time ; others again are still in the experimental stage. The method of selecting the cycles depends partly on permanent factors, which can be anticipated for any length of time ahead, and partly on temporary factors, which may change from time to time. While, therefore, some of the cycles usually remain the same from week to week, often changing conditions introduce new ones, causing more or less variations in the expectation curves. Such variations do not seem unscientific or in any way unreasonable. They must be admissible even in a 12-hour forecast—if a forecaster at 8 A.M. expects a heavy shower at 2 P.M. and at noon modifies his forecast and expects a moderate shower at 3 P.M. instead, we do not regard him as a fraud or his forecasts as worthless. Whether or not the diagrams under discussion are invalidated by the modification is a question of degree which can only be decided by a careful comparison of a series of complete diagrams over a long period.

For this purpose two sets of diagrams are here shown. Fig. 1 is a series of ten expectation curves as published in the *Daily Mail* for the British Isles as a whole, covering the period June 18–Oct. 7, 1926. This series is a fair average sample, and it shows the extent of the weekly variation in the

curves for the British Isles; at individual rain gauges obviously somewhat greater variation must be expected. As a convenient check on the results, a curve of facts has been added in which the ordinates are made proportional to the number of rain gauges, out of the 43 British reporting stations, which have recorded rain (0·2 mm. or more) during the 24 hours.

Fig. 2 is a series of five expectation curves for the Scilly Isles, covering the period July 2 – Sept. 16, 1926, together with a curve of facts plotted according to the logarithmic scale explained above. This series is selected from among those which Capt. Cave criticised, and is perhaps above the present average of achievement.

Such a juxtaposition of a series of expectations with facts seems to be the only fair and conclusive method of testing them; but it will be seen that the method is a severe one, and it will perhaps be agreed that, if any series of 12-hour forecasts were graphed in the same way and checked for every quarter hour, they would probably not compare with facts nearly so well as these do.

The diagrams for the Scilly Isles were also tested on the basis of day-to-day expectations checked with rain or no-rain facts, and the degree of success attained was compared mathematic-

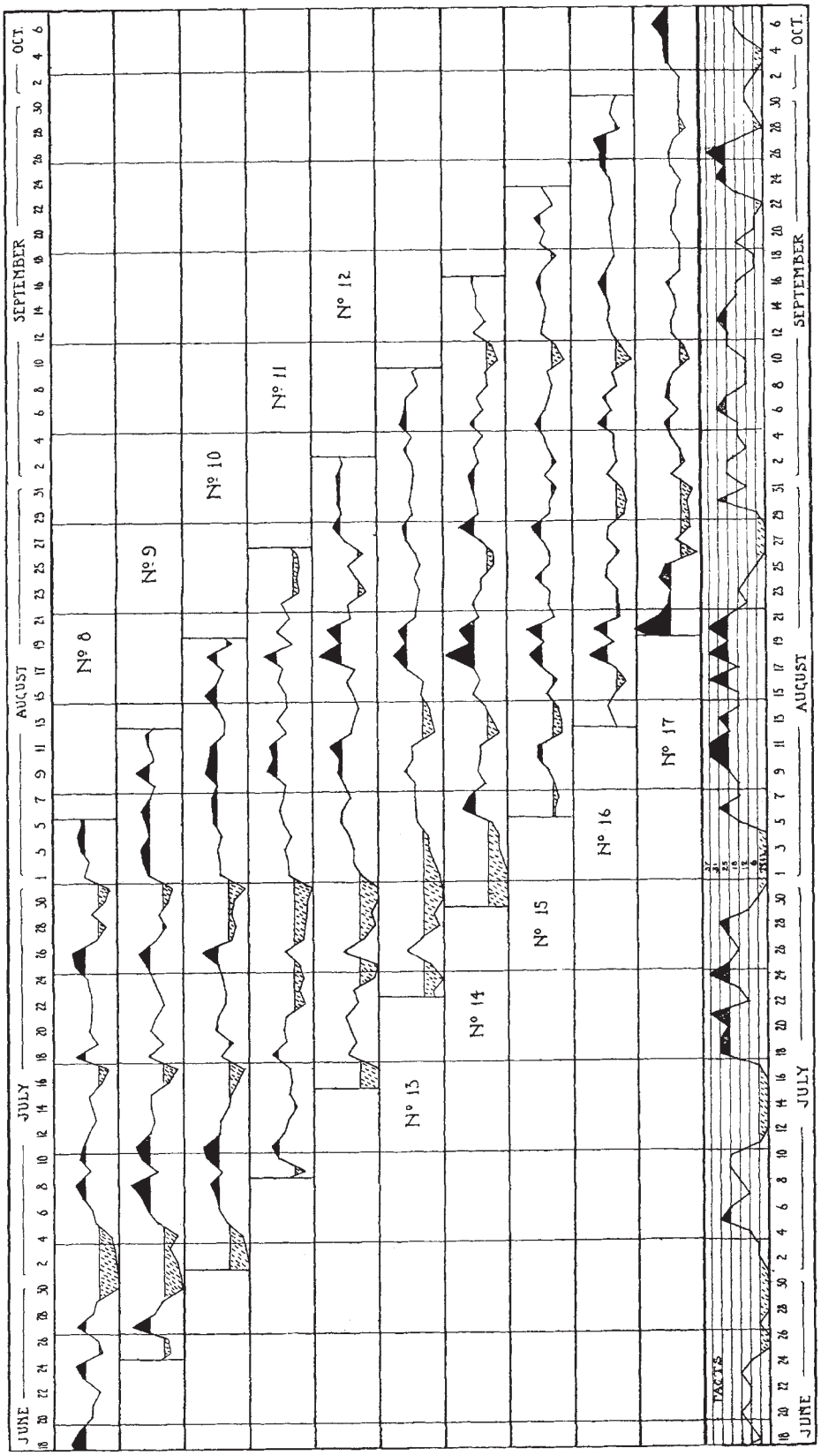

FIG. 1.—Great Britain and Ireland. Comparison of *Daily Mail* charts with observations of ten diagrams for 1926.

ally with what might be expected of a random guesser.

The data were obtained as follows: In each published curve, each day which was shown blacked was regarded as a rain expectation; if 0·2 mm. or more of rain was officially recorded that day, it was considered a success; if less than 0·2 mm. of rain, a failure. Similarly, each day which was shown shaded was regarded as a no-rain expectation; if less than 0·2 mm. of rain was officially recorded that day, it was considered a success; if 0·2 mm. or more, a failure.

In the diagrams for the Scilly Isles shown in Fig. 2, the rain expectations have been marked with an arrow pointed upwards, the no-rain expectations with an arrow pointed downwards. It should be mentioned that the published diagrams always begin on a Friday, but, since they are completed and leave their author's hands on the Wednesday, for the purpose of his own checking he always includes the Thursday's expectation, which is not published, and it has been done for the purpose of this check too. This is slightly to his advantage, for 12 definite expectations for the first Thursday are included in the series here considered, and of these 8 were successes and 4 were failures.

Over the period June 17–Dec. 3, 1926, for the Scilly Isles 640 definite expectations were made. Of these, 412 were successes and 228 were failures on the day-to-day basis explained above. At first sight this may not seem a very striking measure of success, and for a small number (say 12 expectations) such a proportion could reasonably be attributed to chance; but for 640 expectations, the odds against a random guesser making 412 successes are incredibly large. If any one doubts it, let him try spinning a coin 640 times, and he will find that he never gets anything approaching 412 of either heads or tails.

The odds against the random guesser were tested according to the theory of probability. Given that the facts for the days considered were 348 rain and 292 no-rain, also that the guesser makes 640 guesses, knowing that the normals for the period are 95 rain and 74 no-rain, so that he would guess rain 360 times and no-rain 280 times.

Assuming the above data only:

The chance that the guesser would get exactly 412 successes is

$$\frac{\lfloor 348 \times \lfloor 292 \times \lfloor 360 \times \lfloor 280}{\lfloor 240 \times \lfloor 108 \times \lfloor 120 \times \lfloor 172 \times \lfloor 640} = 7{\cdot}00 \times 10^{-13}.$$

By adding the chances of more than 412 successes we get $1{\cdot}01 \times 10^{-12}$.

The odds are therefore $9{\cdot}90 \times 10^{11}$ to 1 against a random guesser scoring 412 or more successes under the conditions assumed.

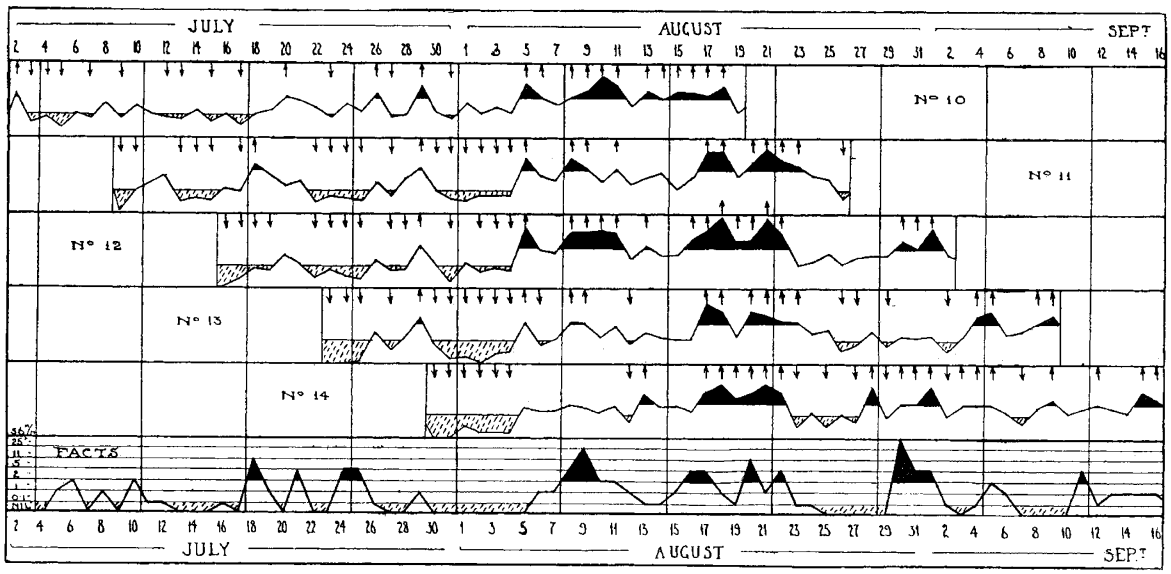

FIG. 2.—Scilly Isles. Comparison of expectations with observations, 1926.

These calculations were made by Mr. W. Hope-Jones, of Eton College, and confirmed by Dr. R. A. Fisher, of Rothamsted Experimental Station.

Admittedly there may be factors which perhaps ought to be included, and might, if included, reduce this figure, but it would appear that, in whatever way the problem may be tackled, the odds against a random guesser doing so well are practically infinite.

The expectations and facts may be tabulated as follows:

	Expectations.		
	Rain.	No Rain.	Total.
Rain Facts . .	243	105	348
No-Rain Facts .	123	169	292
Total	366	274	640

It has been the purpose of this article to show that—

A. Owing to the popular form in which these expectations are produced, Capt. Cave formed an erroneous conception of their aims, so that his arguments were unsound and his conclusions misleading.

B. A careful examination of a series of diagrams indicates that

(1) Variations from week to week are not great, and generally tend towards improvement as time advances.

(2) Although it is not claimed that rain amounts are correctly anticipated, the peaks and dips of the expectation curves correspond with those of the facts curve in many cases to the actual day, and in most cases to within 24 hours.

(3) There is a degree of correspondence between expectations and facts which cannot be wholly due to chance.

C. The above is confirmed definitely and conclusively by the calculations based on the theory of probability.

If these propositions are accepted, and it would seem hard to dispute them, then it must be agreed that the claims made in the *Daily Mail* are more than substantiated, and that what Capt. Cave condemns as no better than fortuitous prediction has been proved to be something of the order of a billion times better, and represents in fact a notable advance in meteorological science.

R. P. BUTLER.

CAPT. BUTLER complains that my weather and forecast numbers are unfair, and appeals to the *Daily Mail* for April 30. The only relevant sentence seems to be that the greater the height of the curve above the datum line the greater the probability of rain, but not necessarily the amount of rain; I do not think that this fact invalidates anything I have said. I have already explained the method of comparing the forecast diagrams with the weather, and I am quite content to leave it to the readers of NATURE to say whether the method is fair or unfair.

It is also said that the forecasts are not meant to be day-to-day forecasts; they are, however, given in a day-to-day form, and are, I think, generally so taken by the public; Capt. Butler, however, emphasises the point and says that the author is quite satisfied if the timing of his expectations is correct within twenty-four hours either way. The number of wet days with 0·04 inch of rain or more is about 120 in the year for the south-east of England, or one day in every three; any one forecasting rain for to-day and claiming a success if rain comes yesterday, to-day, or to-morrow, is putting his forecast in a very favourable position. If one were to forecast by drawing counters out of a bag, the successes under the above conditions would be very marked.

I do not understand Capt. Butler's complaint that I only took one forecast for each week for the purpose of comparing them with the weather in my diagrams. I took the complete week nearest to the actual happenings as being the most fair to the author. Capt. Butler says that each 50-day forecast has a character of its own, which is exactly what I maintained. I have selected one of the seven varying forecasts for each week and compared it with the actual weather, and I have done this for the whole period from April 15 to Oct. 27. To say, as Capt. Butler does, that doing this is like "criticising a 12-hour forecast by examining the weather of 103 minutes out of each day," seems to me to be a statement devoid of meaning. If the forecasts are in the main similar, there is nothing unfair in taking any one in preference to the six other forecasts for the same week; if they differ, it is reasonable to suppose that the one nearest to the period for which the forecast is made would be the most correct; I therefore took the first complete week of the latest forecast. Capt. Butler's diagrams show exactly what I maintained, that the forecasts for each week vary so materially as they are issued week by week that they cannot all be guides to the coming weather. Nor are his diagrams of facts compared with forecasts very striking. I pointed out that August was a particularly favourable month for the forecasts, especially in south-west England; I do not see anything very remarkable in the diagrams as extended to July and September.

The method of the forecasting is still wrapt in obscurity on the plea that to disclose the method would be an act of disloyalty to those who have backed the forecasts. Such an attitude to scientific truth has probably never before made its appearance in the pages of NATURE. It precludes one from examining the worth of the forecasts except in so far as the results declare it, and nothing that Capt. Butler has brought forward changes my opinion that chance operates largely, if not entirely, in the relation of forecasts to facts. His probability figures do not impress me very much. He has evidently treated each of the seven weekly forecasts as entirely independent, which he himself claims not to be the case. If they were not treated as independent variables, I fancy that the impressive figures he brings forward would dwindle to very modest proportions. In any case he has attempted to prove too much; if the forecasts are a billion times better than would be expected on pure chance, failures should practically never occur, whereas it is obvious, even from his own selected diagrams, that the method cannot be relied on by the farmer for his agricultural operations, or by the man in the street who wants to know whether or not to take out his umbrella.

C. J. P. CAVE.