Estimates of Seasonal Mean Temperature, Using Persistence between Seasons

Harry van Loon and Roy L. Jenne

National Center for Atmospheric Research,[1] *Boulder, Colo. 80303*

(Manuscript received 15 May 1975; in revised form 11 July 1975)

ABSTRACT

We have tested three methods of estimating the level of a coming season's mean temperature at a station where the statistical association between two selected seasons is as high as one can expect in extratropical regions. The methods are contingency tables, regression equations, and the use of the last few decades if there is a trend at the station which will separate the mean of these decades a fair distance from the long-term mean. A moderate amount of skill was achieved, but the degree of seasonal association in our test case was exceptionally high, and generally these methods will provide only a small improvement over a probability based on knowing only the observed frequency distribution.

1. Introduction

Contingency tables which connect the temperature anomalies of two adjacent seasons seemed, as one of a number of methods discussed by Namias (1964), to provide some skill in estimating the mean temperature level of the second season from that in the first. Using the monthly mean temperatures collected for *World Weather Records,* we have tried this method in a study of worldwide, seasonal temperature persistence and found that there are indeed large areas where the tables apparently give good results, and that often a considerable part of these results is owing to association between seasons on long time scales (trends). Regression equations and the scatter about the regression lines, although they do not raise the skill, are nonetheless preferable to contingency tables since they not only supply the general information contained in the tables, but they provide a continuum of expected means with a clear statement of probability. Apart from the tropical regions where the correlation between seasons is often high, the expected mean which is obtained by regression is frequently not far enough removed from the long-term mean for an estimate of the next season's mean temperature level to be made which is significantly different from the long-term mean.

2. Contingency tables, correlation coefficients, and regression equations

a. Contingency tables

We defined three categories to be used in the contingency tables: *normal*[2] is within half a standard deviation on either side of the long-term mean; this category thus covers one standard deviation and in a Gaussian distribution it contains 38.3% of the cases. Nearly 31% then falls in each of the other two categories, *below* and *above normal.* The tables were tested by the chi square test: the expected distribution of cases in the nine combinations of the categories was computed and then compared with the observed distribution on the assumption that the two seasons are unrelated. If the disparity between the two distributions is too large to be ascribed to chance, that is, if the χ^2 is larger than some selected fiducial limit (see, e.g., Panofsky and Brier, 1963, pp. 53–58), the chance that the two seasons are unrelated is less than the selected limit and the assumption may be rejected. Table 1A is such a contingency table for fall and the following winter at New Haven (41.3°N, 72.9°W). In this instance $\chi^2 = 31.9$, and since the 0.1% limit is at $\chi^2 = 18.5$ with the given number of degrees of freedom, statistically there is considerably less than one chance in a thousand that the seasons are unrelated.

Such good results are not uncommon. Figure 1 is an analysis of χ^2 over a large part of the Northern Hemisphere for the sequence winter to spring. Only the areas with stations having records longer than 50 years have been analyzed. The χ^2 isopleths are labelled according to limiting values, and where $\chi^2 > 9.5$ (i.e., where the chance that the seasons are not related is less than 5%) the region is shaded. It is clear that there are wide areas where the persistence of seasonal temperature anomalies is high as defined by this method. Apparently, there is also something systematic in the distribution of high and low values, since location near or over the ocean frequently seems to influence the size of the χ^2; but one should not generalize from Fig. 1 because the pattern changes from one season to another (see, e.g., Namias, 1964, who shows the results of contingency analyses

---

[1] The National Center for Atmospheric Research is sponsored by the National Science Foundation.

[2] Not to be confused with "climatological normal," usually taken (by international agreement) to be a mean value over a specified 30-year period.

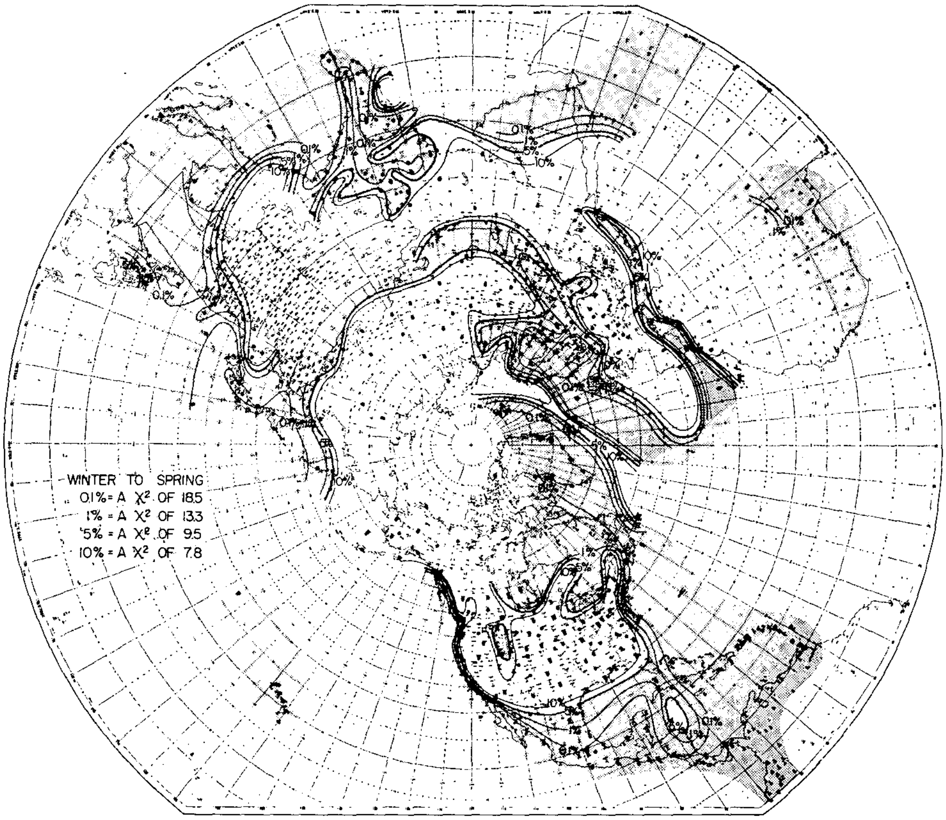

FIG. 1. Lines of equal fiducial limits referring to chi square tests of contingency tables which link winter and spring. Areas below the 5% level (less than one chance in twenty that the seasons are unrelated) are shaded. The dots are the stations used; *all are ≥50 years.*

over the United States for the various seasons vs. the following season).

Let us examine a station where two seasons are significantly related according to the contingency method. New Haven with a $\chi^2=31.9$ for the sequence fall to winter is a good example. In the time series of its fall and winter temperatures, Fig. 2, the curves are placed so that an autumn and the following winter are at the same point on the time axis. The parallel lines in the drawing are plus and minus half a standard deviation from each season's mean, and the interval between the lines in each set thus equals the category *normal*

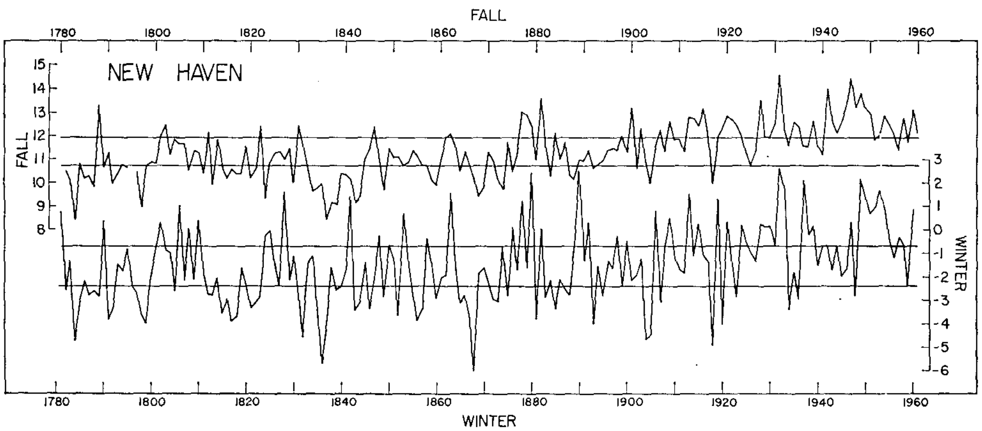

FIG. 2. Autumn and winter mean temperatures (°C) at New Haven. The area between the parallel lines in each season outlines the category *normal,* which is one standard deviation wide and centered on the mean.

in Table 1a. It is obvious that most of the *below normals* in winter following *below normals* in fall are from the earlier half of the period, whereas the *above normals* following *above normals* are mainly from the latter half, which is also seen in Tables 1b and 1c. *The persistence in Table 1a is therefore to a large extent a reflection of the upward trend which is present during a substantial part of the period.*[3]

We can assess how the contingency table changes probabilities by the following considerations. Figure 3 shows how the winter mean temperatures are distributed as a function of the three categories of fall mean temperatures, as each histogram is the frequency distribution of mean temperature in winter following one of the three categories in fall. Those winters which follow autumns when the mean temperature was in the *above normal* category have a mean (Fig. 3a) that is 0.9°C above the 180-year mean and a standard deviation of 1.67°C nearly equal to that of the long-term distribution. The 0.9°C difference places the mean of these winters 0.53 σ above the 180-year mean. Knowing also the standard deviation of the distribution in Fig. 3a, one can calculate that 51% of the cases in a normal distribution based on the mean and standard deviation of Fig. 3a falls in the category *above normal*, defined as the upper 31% of the cases in the Gaussian winter distribution. Using the contingency method, we have thus raised from 31 to 51% the probability that a statement about the coming winter's being placed in the *above normal* category will be true—considering normal distributions instead of the numbers in the boxes in Table 1a. In the same instance we are able to say that the probability is 12% that a winter at New Haven will be *below normal* after a fall which was *above normal*, down from the 31% which has been defined as *below normal* in the Gaussian winter distribution.

The mean of those winters that came after *below normal* autumns is 0.85°C lower than the 180-year mean and has a standard deviation of 1.46°C. Its position is 0.5 σ below the 180-year mean, which is where the category *below normal* begins, and the probability therefore is that 50% of the winters following a *below normal* fall will themselves be *below normal*, whereas 12% will be *above normal*.

TABLE 1. (A) Contingency table, winter after fall at New Haven. Normal ranges from half a standard deviation below to half a standard deviation above the 178-year mean. (B) and (C) show the distribution of the categories in the two halves of the period.

(A)	Winter Below normal	Winter Normal	Winter Above normal	
FALL Below normal	33	20	7	
FALL Normal	18	29	16	1780–1960
FALL Above normal	10	16	29	

(B)	Winter Below normal	Winter Normal	Winter Above normal	
FALL Below normal	28	15	6	
FALL Normal	11	14	6	1780–1870
FALL Above normal	4	3	2	

(C)	Winter Below normal	Winter Normal	Winter Above normal	
FALL Below normal	5	5	1	
FALL Normal	7	15	10	1870–1960
FALL Above normal	6	13	27	

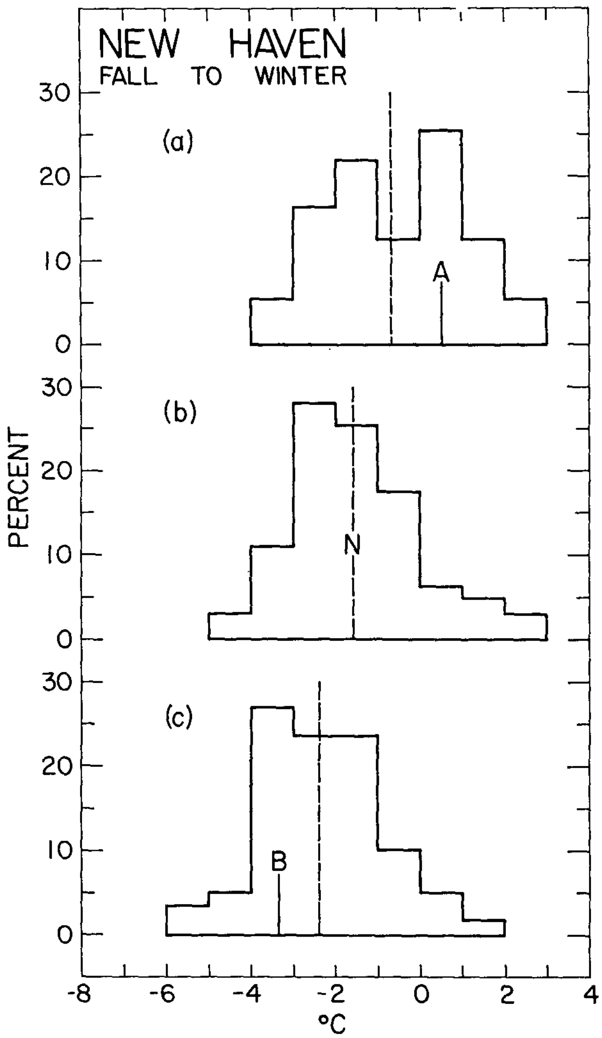

FIG. 3. Frequency distributions of the winter mean temperatures which follow autumn mean temperatures in the category (a) *above normal*; (b) *normal*; (c) *below normal* at New Haven. The vertical dashed lines are the means of these frequency distributions. A, N, and B are the positions of the means of those winter temperatures which belong in each category.

[3] For our purpose, it does not matter whether the trend is owing to natural causes (solar, oceanic, etc.), or if it appears because of urbanization or change of location. We want to stress only the effect of *a* trend on the correlation between two seasons.

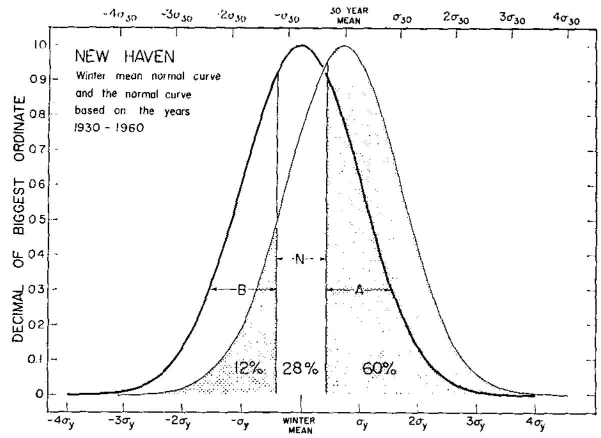

FIG. 4. The heavy curve is the Gaussian distribution of winter mean temperature based on the 180-year mean and its standard deviation at New Haven, Conn. The two vertical lines divide the area under the curve into three equal parts: *below normal, normal, above normal*. The light curve is the normal distribution obtained from the mean of the last 30 years in Fig. 1 and its standard deviation. The percentages show how much of the 30-year normal distribution falls into the categories defined with reference to the long-term mean and standard deviation.

If the relationship between the seasons in our sample were perfect so that a winter always fell into the same category as the preceding fall, the means of the winters following *above normal* autumns would be at A in Fig. 3a, and those following *below normal* would be at B in Fig. 3c. These positions were determined by averaging the winter temperatures by category. Both A and B are a little more than one standard deviation removed from the 180-year winter mean.

New Haven in this seasonal sequence has one of the highest χ^2 values in extratropical latitudes of any seasonal sequence, and the use of the contingency method would improve only modestly our ability to estimate a following season's probable temperature level. *Over large parts of the Northern Hemisphere, however, less significance than at New Haven is achieved by the contingency tables (cf. Fig. 1) which can then be used only locally and under favorable circumstances.*

Until this point we have used a division of the normal distribution into three parts, of which the two extreme categories each contains 31% of the cases; this was how we had defined the categories in the study of the contingency technique. We later decided that it might be more useful to include more cases in the outer categories if the scheme should ever be used for purposes of planning, and therefore changed to a system where the area under the normal curve is divided into three *equal* parts: *below normal, normal, above normal*. This system will be used below.

Because of the trend in Fig. 2, one might disregard the autumn and consider only a recent period instead of the entire record when looking for a probable value of the coming winter, assuming that the trend will not come to an abrupt end. The last 30 years of the record shown in Fig. 2 have been used to calculate a mean and its standard deviation for winter, which may be compared with the 180-year mean and standard deviation in Fig. 4. The long-term mean is $-1.56°C$, $\sigma = 1.69°C$, and the 30-year mean is -0.33 with $\sigma_{30} = 1.61°C$. Owing to the trend, the 30-year record would be somewhat better than the total record for estimating the probability that a given temperature would occur: Since the 30-year mean lies 0.73 σ above the long-term mean, the probability that a coming winter will fall in the *above normal* category (based on the long-term frequency distribution's division into three equal parts) rises from 0.33 to 0.60. The probability that the mean will be in the *normal* category is 28%, and 12% that it will be *below normal*. In the 31–38–31% division of the normal distribution used earlier, the probability would rise from 0.31 to 0.58.

This method is most advantageous if the trend is so steep that the mean of the last few decades differs much from the long-term mean, or if the scatter about the mean of recent years is much smaller than that about the long-term mean, or if both of these circumstances combine. Sydney (33.9°S, 151.2°E) is an example of this: its 112-year mean for winter is 12.6°C, $\sigma = 0.65°C$, whereas the mean of the last 21 years is 13.3°C with $\sigma = 0.35$. The time series of Sydney is in Fig. 5a, and the normal curve based on the 112-year mean and its standard deviation is shown in Fig. 5b, together with

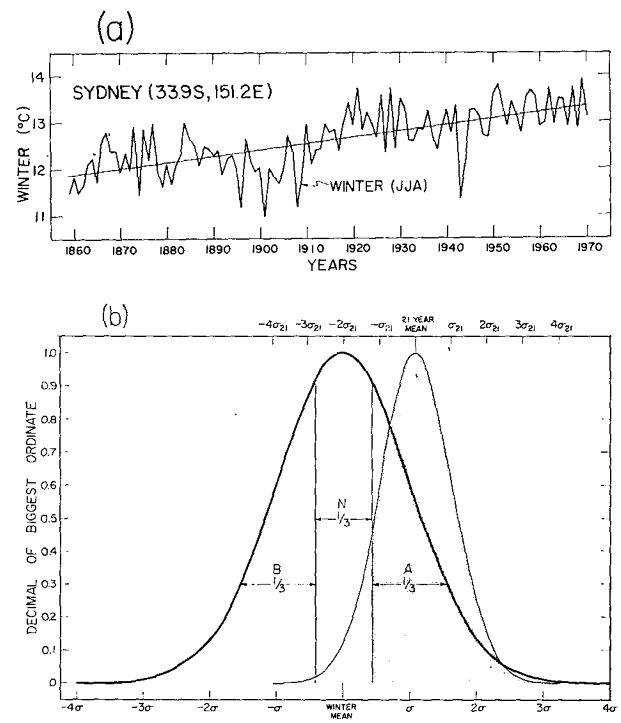

FIG. 5. (a) Winter mean temperatures (°C) at Sydney, Australia. The straight line is the regression line for temperature regressed on time. (b) Same as Fig. 4, but for the 112 and last 21 years in (a).

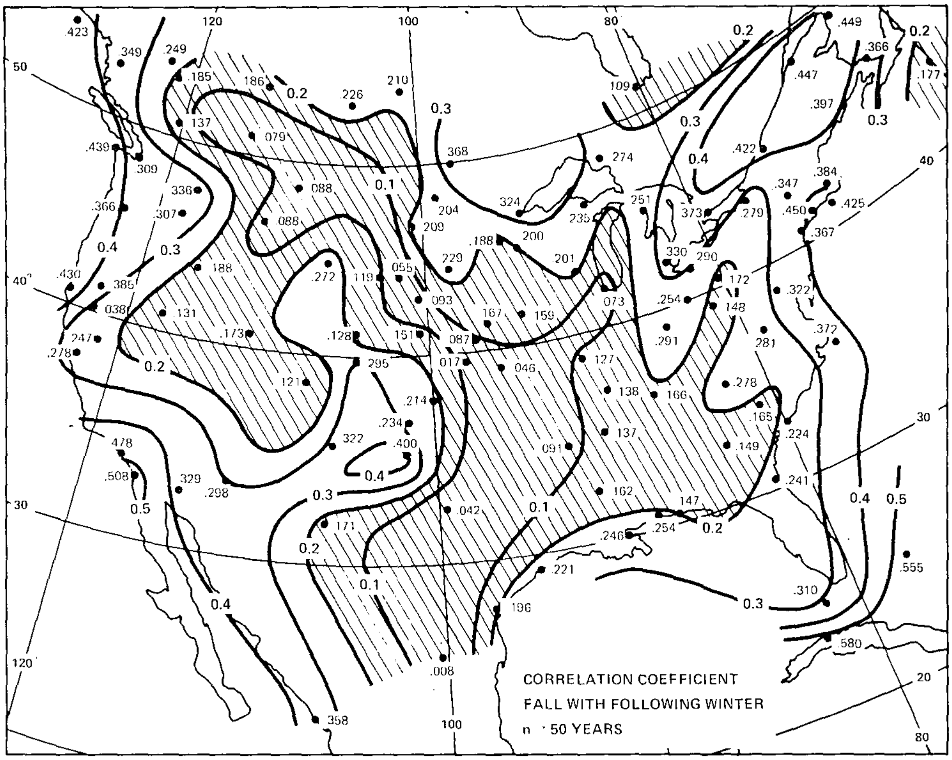

FIG. 6. Correlation coefficients between the mean temperatures (°C) of fall and the following winter. All records are ≥ 50 years.

the normal curve derived from the mean of the last 21 years and its standard deviation. In this favorable instance 89% of the cases in the distribution around the 21-year mean fall into the category *above normal* of the Gaussian distribution around the long-period mean. The definition in this method (and in the contingency method) is not sharp, however, as *above normal*, for example, covers all temperatures which are higher than a certain level above the long-term mean without outlining their probable distribution in the category, but the method does offer an improvement over the contingency tables when a clear trend is present. In addition, it is independent of knowledge of the previous season.

b. *Correlation coefficients*

The correlation coefficient between the two seasonal time series will, of course, also be affected by the existence of a trend in the series. If, for example, the trend is upward in both, the higher temperatures at the end of the first series will be associated with the higher temperatures in the second, and the correlation will be enhanced because of the trend. The correlation between fall and winter at New Haven is 0.450, and the map in Fig. 6 indicates that this is a high correlation in comparison with others in the United States. Yet it means that only 20% of the variance in winter is connected with the variations in fall, and as will be shown in Section 4, a substantial part of that connection is in the longer time scales.

c. *Regression equations*

The statistical relationship between two seasons may also be described by the regression of the second season on the one before. In Fig. 7, 178 winter mean temperatures at New Haven are plotted as a function of the preceding fall; the regression equation connecting the two seasons is

$$y = -9.128 + 0.668x,$$

and the scatter about the regression line $= 1.51°C$. When the mean temperature of the previous season is known, the regression equation will provide an expected mean for the following season. The scatter about the expected mean is smaller than the scatter about the long-term mean by an amount related to the size of the correlation coefficient (see Panofsky and Brier, 1963, p. 96). The correlation between fall and winter at New Haven is 0.450, which means that the scatter about an expected mean, forecast by regression, is 11% less than the scatter about the observed mean.

A displacement of the expected mean by $0.5\,\sigma_y$ from the long-term mean will raise the probability from 0.33 to 0.54, that is, a little more than half of the cases in the distribution around the expected mean will fall into one of two extreme categories. In the instance of fall

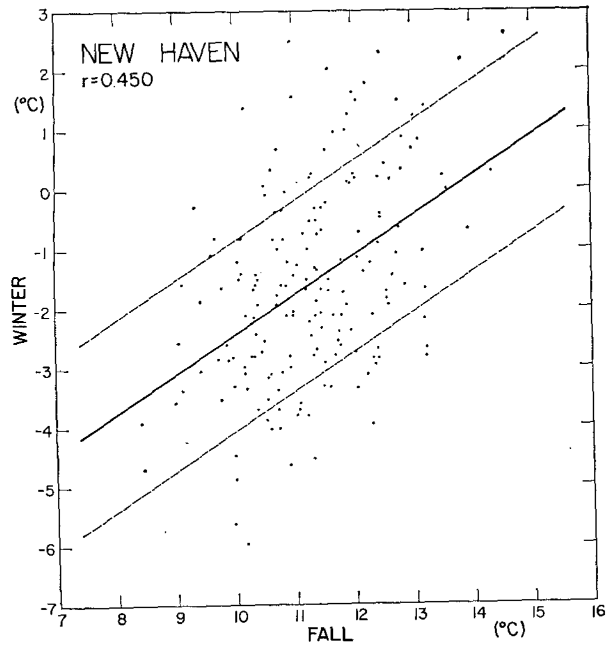

FIG. 7. Scatter diagram, regression line, and scatter about the regression line for winter as a function of the preceding fall at New Haven.

to winter at New Haven, a displacement of the autumn mean by 1.1 σ_x is required to obtain by regression a winter mean which is 0.5 σ_y beyond the long-term mean. To get an expected winter mean that is one standard deviation above the regular mean (at this level 71% fall into *above normal* and 5% into *below normal* categories), the fall mean temperature must be 2.22 σ_x above normal. At this point less than 3% of the fall cases are left in the normal distribution. To arrive at high probabilities by regression, the mean temperature of the predictor season must then be so far removed from its long-term mean that the chances that this will occur are small. Regression does not necessarily improve our skill, but it has the advantage over the other methods of (a) taking a trend into account if the mean temperature of the predictor season is at the level of the trend; and (b) providing a continuum of expected means and the scatter about each instead of just a category (the contingency method) or a fixed mean and distribution (the recent-years method). The expected mean and the scatter can be used to assess the probability for each of the categories; and the scatter is smaller than the scatter about the original mean—the more so the higher the correlation coefficient—which allows a narrower definition of the probability.

3. Seasonal association as a function of frequency

The association between two seasons, as measured in the contingency tables or by correlation coefficients, is composed of contributions from periods of widely different length—from century-long trends to single years. In the following we shall examine the covariance (cov_{xy}) as a function of frequency (f) to discern the contribution of different time scales to the total correlation.

The covariance of the two time series $x(t)$ and $y(t)$ is defined as

$$\mathrm{cov}_{xy} = \overline{(x-\bar{x})(y-\bar{y})},$$

where $\bar{x}$ is the mean of $x(t)$ and $\bar{y}$ is the mean of $y(t)$ (in the instance of New Haven, $\bar{x}$ is the long-term mean of fall, and $\bar{y}$ the long-term winter mean). The relation of the covariance to the correlation coefficient, r, is seen from the definition of $r = \overline{(x-\bar{x})(y-\bar{y})}/\sigma_x\sigma_y$. For practical purposes we shall be using cospectra, $C_{xy}(f)$, which have been normalized so that the co-spectrum averaged over all frequencies is approximately equal to 2 cov_{xy} (see below). The cospectra have been estimated by first expanding $x(t)$ and $y(t)$ into Fourier series and then multiplying the in-phase parts of corresponding harmonics of each time series. Ten percent at each end of the time series was tapered with a cosine bell before the computations (see e.g., Julian, 1971, Appendix). All the cospectra have been smoothed by running averages of five adjacent harmonics and therefore the longer the series the shorter will be the band width (cf. in Fig. 8).

We have chosen three stations in Europe from Fig. 1 for a closer look at the seasonal persistence. They are Copenhagen (55.7°N, 12.6°E) with $r=0.491$, Palma (39.6°N, 2.7°E) with $r=0.501$, and Geneva (46.2°N, 6.2°E) with a correlation of 0.123, the latter representing the belt of low persistence across central Europe. The three cospectra are shown in Fig. 8. The cospectra are normalized so that

$$\int_0^{0.5} C_{xy}(f)df \approx \mathrm{cov}_{xy},$$

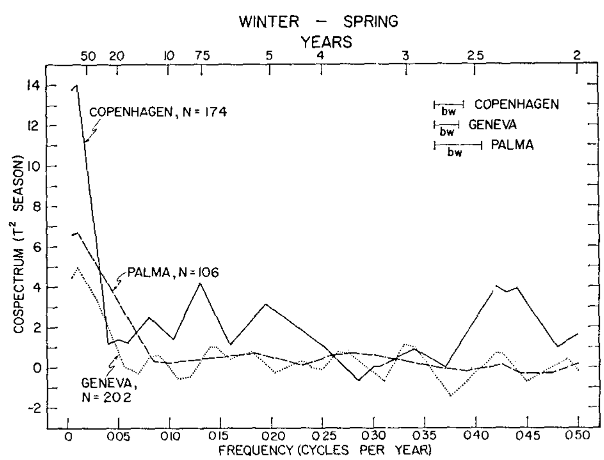

FIG. 8. Cospectra as a function of frequency for three European stations; "bw" is the band width.

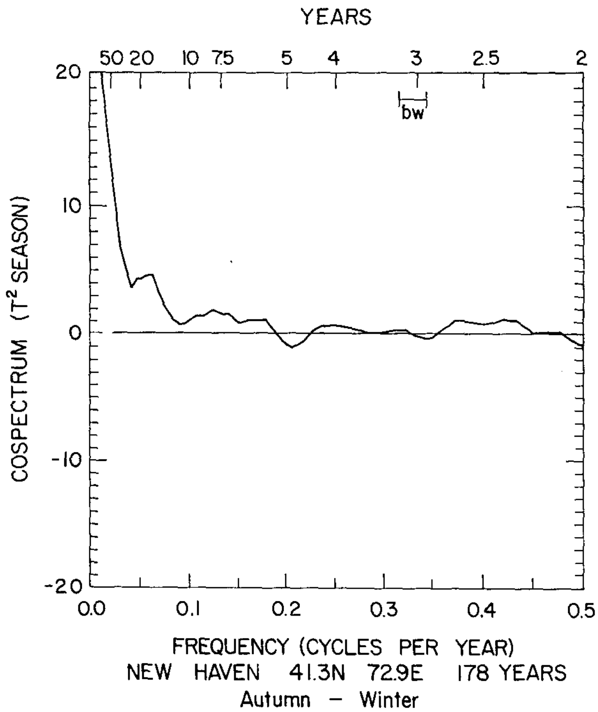

FIG. 9. Cospectrum as a function of frequency for New Haven; "bw" is the band width.

0.5 cycles per year (2 years) being the highest frequency resolvable when the two time series consist of one value (seasonal mean) per year. There will be slight differences between the integral of $C_{xy}(f)$ and cov_{xy} introduced mainly by the tapering procedure. In each instance the lowest frequency components (long-period trends) contribute a large portion to the total variance, and at Geneva and Palma most of the association between the seasons expressed by the correlation coefficient stems from the longest time scales.

The cospectral values at Copenhagen are higher than those at Palma at nearly all frequencies; yet Palma's correlation coefficient is slightly bigger than Copenhagen's. As r is defined as $\mathrm{cov}_{xy}/\sigma_x\sigma_y$, and since

$$\int_0^{0.5} C_{xy}(f)df \approx \mathrm{cov}_{xy},$$

then

$$r \approx \int_0^{0.5} C_{xy}(f)df \bigg/ \sigma_x\sigma_y,$$

and the fact that Copenhagen's correlation coefficient is no larger than Palma's, although the cospectrum averaged over all frequencies is smaller at Palma, must then be owing to a larger interannual variation ($\sigma_x\sigma_y$) at Copenhagen. The standard deviations are: at Copenhagen, $\sigma_x = 1.72$ and $\sigma_y = 1.29$; at Palma, only $\sigma_x = 1.02$ and $\sigma_y = 0.95$.

As regards New Haven, the cospectrum of autumn and winter (Fig. 9) demonstrates that most of the covariance measured by the correlation coefficient of 0.45 occurs at frequencies well below 0.1 cycles year^{-1}, that is, through parallel trends in the two seasons which are appreciably longer than 10 years. Little or no association is found at middle and high frequencies.

4. Conclusion

We do not know a more objective way of obtaining the probable temperature level of the next season than the techniques discussed above. We tested them on a seasonal sequence at New Haven when the connection between the two seasons was as high as one can expect in extratropical latitudes. Under these very favorable circumstances the techniques raised the probability by a moderate amount, but unfortunately the degree of association between two seasons which made this modest improvement possible is comparatively rare at extratropical latitudes. More often the association is slight and the separation between the original and the forecast means is too small to raise the probability by more than a few percent. Even the method that uses only the recent decades of the season about which a statement must be made, and which takes advantage of longer trends, fails when the association is small since high correlation is nearly always associated with a clear trend and low correlation with a weak or with no trend.

We therefore agree with Walker (1936) that

> ". . . predictions can only be issued with restraint if public confidence is to be won. The natural consequence is silence, except when the indications are markedly favorable or unfavorable. In a race with 30 starters a conspicuous good horse may, without undue risk, be backed to come in within the foremost 6, and we may feel confident that a thoroughly bad animal will be in the last 6; but it would be unwise to hazard much on the likelihood that a commonplace individual will finish among the central 6. It may at first sight seem a confession of weakness to issue no forecast when conditions appear roughly normal; but it is better to admit your limitations, and only speak when you can do so with some safety than to issue predictions when they are little more than guesses."

That is, if the association between two seasons is reasonably high, one may venture to give a probability for the second season if the first not only fell into one of the extreme categories, but was also well removed from the mean.

Finally, it should be noted that although the standard deviation has been used to estimate a statistically meaningful displacement from the mean, each time series must be considered on its own when it comes to estimating *socially* meaningful changes. Half a standard deviation may in one place spell crop failure or greatly increased energy consumption, whereas in another place it may be trivial.

Acknowledgment. We are grateful to have had the opportunity often to discuss this paper with R. A. Madden. His suggestions improved the contents in many ways, e.g., the idea of using cospectra to resolve the contribution of different time scales to the correlation coefficient.

REFERENCES

Julian, P. R., 1971: Some aspects of variance spectra of synoptic scale tropospheric wind components in midlatitudes and in the tropics. *Mon. Wea. Rev.*, **99**, 954–965 (see Appendix).

Namias, J., 1964: A 5-year experiment in the preparation of seasonal outlooks. *Mon. Wea. Rev.*, **92**, 449–464.

Panofsky, H. A., and G. W. Brier, 1963: *Some Applications of Statistics to Meteorology*. Pa. State Univ., Univ. Park, Pa., 224 pp., (see pp. 53–58).

Walker, G. T., 1936: Seasonal weather and its prediction. *Smithson. Inst. Rpt. for 1935*, 117–183.