APPLICATION OF CLIMATOLOGICAL ANALYSIS TO ENGINEERING DESIGN DATA (*)

by H.C.S. THOM

Silver Spring, Environmental Data Service - ESSA, U.S.A.

ABSTRACT

The general climatological prediction is defined, and its proper role in engineering design is discussed. The fundamental principles involved in the selection of the best meteorological variables for determining design values are discussed together with their analytical representation.

There are two general types of design values: quantiles on extreme values and quantiles on subextreme values. The extreme value quantile applies where the engineering system must withstand some recurring extreme. Examples are all forms of structures exposed to meteorological forces which could cause failure or destruction. The subextreme value quantile applies where the overloading of the engineering system does not cause destruction but only discomfort or inconvenience with perhaps some tolerable economic loss. Examples are heating and air conditioning systems and storm sewer and small drainage systems.

The proper analysis for extreme value design data comes from the statistical theory of extreme values. Examples of application to wind and snow loads are discussed. For subextreme value data either the recurrence of critical events within an individual time period may be limited or their spacing in time controlled. Both of these devices depend upon the fact that frequent repetition of the critical event is more undesirable than less frequent repetition. The statistical theory of the m^{th} value or the theory of the time interval distribution apply here. Examples of application to heating and air conditioning and to intense rainfall are discussed.

(*) Presented at the Symposium on Urban climates and Building climatology, Brussels, October 1968.

Introduction.

Weather enters into many of man's activities sometimes facilitating sometimes limiting his activity. The weather forecast with its severe limitation of time range makes it useful only for operational decisions being made a few days in advance at most. For the large area of planning decisions for years, even many years in the future, the statistical analysis of historical data resulting in the climatological prediction is required. This use in planning covers a wide variety of problems one of the most important being the application to engineering design discussed here.

From time immemorial builders knew something of the weather factor at least from experience and used this in their designs. It was not, however, until early in this century that the statistical analysis necessary to the development of climatological analysis began to be developed. The coming of the airplane, at about the same time which is so subject to weather factors, added greatly to the needs for climatological analysis. At first the methods were rather crude and heavily empirical or nonparametric. Later with the sophistication of engineering systems brought on by the second world war, climatological analysis began a rapid development with greater and greater use of parametric methods. Unfortunately, there are still relatively few who understand modern climatological analysis; so some work is still crude and even faulty. Notwithstanding there is today a good body of valid climatological analysis available which can give unbiased climatological prediction for the development of engineering design data almost anywhere in the world.

A weather design value may be defined as the magnitude of a meteorological variable which, when used in the design of an engineering system, nonmeteorological factors having been accounted for jointly or independently, will assure with a given probability that the system will adequately meet a set of prescribed design requirements. The choice of the meteorological value naturally depends on the assigned design conditions and on the prescribed value of the probability of being surpassed. The assignment of the design requirements and the probability level is clearly an engineering problem while the analysis to determine a set of probabilities containing the one chosen by the engineer or from which he can make a choice is the problem of the climatologist. The choice of the meteorological design variable and its relationship to the engineering system variables is the joint problem of the climatologist and engineer, for often the choice of meteorological element is a compromise and must be adapted to the problem. Thus, the weather factor design problem consists of three parts: (1) choice of the design variable, (2) the solution of the relationship problem, and (3) the climatological analysis. It is the objective here to outline the general development of design values and to give the actual climatological analysis in certain special areas of application.

The design variable and its prediction.

The weather design variable may be defined as a variable quantity which is rationally related to the engineering variable of the design problem. Since it is also a climatological variable, it is a random variable with an associated distribution function which defines probability over a range of design requirements. The formal analysis may be illustrated by using functional notation applied to a hypothetical design variable [1]. The prescribed design requirements are completely specified by a load t_E which the engineer determines by engineering analysis. This results in the relationship

$$t_E = h(t_W) \tag{1}$$

where t_W is the weather design variable, t_E is the load variable, and h is the relationship function. t_W and h may be given by the engineer or obtained in collaboration with the climatologist, and t_W may be multidimensional. Conversely weather will also be associated with design values. This may be expressed by the inverse of (1) which is

$$t_W = h^{-1}(t_E) \tag{2}$$

By formal climatological analysis a distribution function on t_W is given by $F(t_W)$ [2]. Associated with t_E one must then admit a probability with which failure to meet the load can be tolerated. If $1 - p$ is the probability of failure then the relationship

$$p = F(t_W) \tag{3}$$

defines the weather design value t_W. Substitution from equation (2) defines the design load t_E

$$p = F[h^{-1}(t_E)] \tag{4}$$

t_E may also be thought of as a transformed variable and will then have a transformed distribution which may be expressed by a transformed distribution function

$$p = G(t_W). \tag{5}$$

Occasionally h will be a tractable function in which case G can be found explicitly. Taking the inverse of equation (4) yields

$$h^{-1}(t_E) = t_W = F^{-1}(p) \quad (6)$$

Solving for t_E gives

$$t_E = h\,[F^{-1}(p)] = G^{-1}(p) \quad (7)$$

which shows that the probability of being less than t_W is equal to the probability of being less than t_E which is prescribed by the engineer; hence, (6) defines an equiprobability transformation from t_W to t_E. The weather design value which is a p-quantile on t_W is determined by equation (6). This equation naturally offers the possibility of a selection of design values over any range of design probabilities p.

It is now the climatologist's task to choose a series of t_W in such a way as to form a random variable capable of giving rational probabilities. This in general will be a climatological series which is a stationary series of independent meteorological observations. In many instances the choice of weather design variable must be a compromise, for the observed meteorological elements do not always fit engineering relationships and, therefore, must be transformed or modified.

The relationship problem

Meteorological variables are purely physical variables which really apply to nothing but weather. In order, therefore, to use them in non-meteorological applications a means of conversion to the applied variable must be provided so that the climatological prediction can be transformed to the applied variable. This is done as illustrated above by establishing a functional relationship between an appropriate meteorological variable and the applied variable. The problem of finding such a relationship is called the relationship problem. It is often solved jointly by the engineer and the climatologist.

The solution to the relationship problem is always some form of functional relationship between the meteorological variable and the application variable. In some instances the function is known from fundamental principles or can be derived from them; in others varying amounts of fundamental principles and empiricism must be employed. In all situations the problem has the formal structure set out below:

Assuming that the variable of the application is given, the first step is the choice of the proper meteorological variable. This must be selected from the elements being observed or combinations of them. Not only must these elements be a part of the observation program, but they must have been observed for a sufficient period of record so that a satisfactory climatological analysis may be made. Let the application variable be y and the meteorological variable x. Then the required relationship may be expressed by the equation

$$y = h(x), \tag{8}$$

where x may be a set of variables. This relates the application variable y to the meteorological variable x, and serves to transform the prediction on the climatological variable to one on the applied variable.

The solution of the relationship problem must have the close collaboration of the climatologist and the engineer. The choice of the variable y is, of course, the engineer's task, but because of the limited availability of meteorological elements, and the varied possibilities of developing a climatological prediction equation, the climatologist must have a hand in choosing the variable x. In addition, the two together must provide a suitable series of data on y and x concurrently for estimating the function $h(x)$ as a regression unless this relationship is already available from purely theoretical analysis. This series need not be a climatological series, i.e., one value per year, but for complete validity in carrying out the regression analysis, variable x must form a series in which the successive values in time are independent of each other. This may usually be accomplished by spacing the values of x sufficiently in time so that successive values are independent. Advantage may be taken of all data by averaging the several regressions based on the lagged data. The climatological prediction, however, must be made only on the basis of a climatological series, i.e., a stationary series of independent values.

Extreme value design data

A set of extreme values will constitute the weather design variable wherever the engineering system cannot otherwise be protected against the occurrence of an extreme [3]. This set will always consist of a series of extremes for successive periods of one year or more because weather has an annual fluctuation. The reasoning back of this is as follows: If the period were less than one year, it would be impossible to form a proper climatological series of extremes for the series for one part of the year must always be different from the series for the other part of the year so the composite series would not form a proper random variable. On the other hand series of extremes for periods longer than a year would form a proper series but the distributions of such series can be shown to be identical with the annual series; hence, the result would be an undesirable reduction in sample size where lengthy homogeneous series are already seldom available. Therefore, the series of annual extremes is the best climatological series for establishing extreme value design data. The following discussions of climatological analysis for wind and snow load are examples of extreme value design data.

Wind load design data

Since dangerous loads on structures are often the result of the shorter bursts of wind such as those of the same scale as the structure, and these are of only a few seconds duration, it has always appeared desirable to use the regularly observed wind speed with the shortest duration. This is one mile or one minute in the U.S. and sometimes less and sometimes more elsewhere. To go from the observation to the shorter duration speed a gust conversion factor is applied. While there are some weaknesses to this procedure, it has appeared to be the best procedure available at present.

In the U.S. it has been found that fitting the Fréchet extreme value distribution to the annual extreme winds provides a satisfactory basis for obtaining design loads. This distribution function may be expressed as

$$F(v) = \exp\left[-(v/\beta)^{-\gamma}\right] \qquad (9)$$

where F gives the probability of the speed being less than v and β and γ are statistical parameters. Maps of several quantiles from the distribution have been provided for the use of engineers [4].

Since the wind speed is readily converted to pressure, the relationship problem involves mainly the application of a gust factor and the distribution of the forces over the structure. These have been largely handled by engineers with some assistance from meteorologists.

Since the earth is very inadequately covered with observations of extreme wind speeds, the writer recently developed a procedure for obtaining approximate design wind speeds from much more widely available or readily estimated data [5]. The method is based on two principles: (1) that an extreme wind speed is composed of a general level of wind on which a wind speed due to a violent meteorological disturbance is superimposed and, (2) that based on a large number of observations the extreme wind distributions may be divided into two classes — one for extratropical storms and thunderstorms, the other for tropical cyclones.

For extratropical cyclones and thunderstorms a study of a large number of sample extreme wind records showed that the shape parameter γ tended to 9.0 as the number of years of record became large. The scale parameter β was determined by the level of the wind which was assumed to be the highest mean monthly speed $\bar{u}$ for the year. The estimate β^* was related to $\bar{u}$ by the relationship

$$\beta^* = (347.5\ \bar{u} + 364.5)^{1/2} - 19.1\ . \qquad (10)$$

Thus, the extreme wind distribution for extratropical storms and thunderstorms becomes

$$F_E(v) = \exp\left[-(v/\beta^*)^{-9.0}\right]. \qquad (11)$$

Tests of this relationship over a wide range of climatic conditions of small sample sizes gave results within a few miles per hour of the empirical estimate.

For tropical storms it was found for the longest and best records that γ averaged near 4.5. The scale parameter, as was to be expected, followed the same relationship as for extratropical storms, i.e., equation (10). The extreme value distribution may then be expressed by

$$F_T(v) = \exp\left[-(v/\beta^*)^{-4.5}\right]. \qquad (12)$$

Equations (11) and (12) may be readily inverted by taking logarithms to obtain design values of v.

In the tropical storm area of the Atlantic Ocean and in part of the typhoon area of the Pacific extratropical storms also occur. This results in a mixture of the two distributions (11) and (12) giving

$$F_{ET}(v) = (1 - p_T)\ F_E(v) + p_T\ F_T(v)\ . \qquad (13)$$

Here p_T is the probability of an annual extreme wind speed from a tropical cyclone. The mixture parameter p_T was estimated by a relationship with the mean annual frequency of tropical cyclone passages per five degree square. Since the frequency of tropical storm passages is about three times as large in the maximum square east of the Philippine Islands as it is in the maximum square east of Miami, the relationship for p_T had to be different in the two tropical storm areas. For the Atlantic Ocean the relationship was found to be

$$p_T = 1/[1 + 99 \exp(-3.0 f)] \quad (14)$$

and for the Pacific

$$p_T = 1/[1 + 99 \exp(-4.4 f)] \quad (15)$$

where f is the mean annual frequency of tropical storm passages for the five degree square under consideration. Equations (10), (11), (12), (14) or (15) provide the required substitution in equation (13) to obtain approximations to the extreme wind distribution for all areas of the earth. The result cannot be inverted directly to obtain design speed but must be inverted empirically by plotting several values of F_{ET}.

Snow load design data

The variable which appeared to have the best prospects of measuring the snow load on structures is the maximum annual weight of the snow pack on the ground expressed in depth of water. In the U.S. this observation is called the water equivalent of the snow pack. It is superior to the depth of the snow pack with an assumed density since the density has been found to vary widely and in fact, have its own distribution [6]. Except for very large accumulations the density of the snow pack even appears to be uncorrelated with the depth.

The annual extremes of water equivalent of the snow pack greater than zero were found to follow a logarithmic normal distribution given by

$$G(\ln x) = N[\ln x;\ \alpha_1(\ln x),\ \sigma(\ln x)] \quad (16)$$

where x is the water equivalent, $\alpha_1(\ln x)$ is the mean of $\ln x$, and $\sigma(\ln x)$ is the standard deviation of $\ln x$. This distribution, of course, holds only where a snow pack exists every year. In most southern areas a snow pack does not exist every year although when it does occur dangerous snow loads may be produced. This makes the universal water equivalent distribution a mixed distribution given by

$$G(\ln x) = (1 - p) + p\, N[\ln x;\ \alpha_1(\ln x),\ \sigma(\ln x)] \quad (17)$$

Here p is the probability of a snow pack greater than zero and $1 - p$ is, of course, the probability of no snow pack.

The design water equivalent on the ground is then found by inverting (17) and substituting $z = \ln x$ and the statistical estimates $\bar{z}$, $s(z)$ and p yielding the G^{th} quantile

$$z(G) = s(z)\, N^{-1}\, [G(z) - q)/p] + \bar{z} \qquad (18)$$

Design snow loads on the ground are given by multiplying $z(G)$ by a suitable constant k to convert depth of water to pressure.

The relationship problem is complex, but the best solutions to it are given in the National Building Code of Canada [7]. Here coefficients are given to convert snow load on the ground to snow load on roofs for a wide variety of structures.

Subextreme value design data

There are a number of areas where it is not economically feasible or desirable to design for the extreme value. Here the objective is either to space a critical event so that its recurrence in time is not intolerable or to limit the risk in a way which prevents excessive recurrence in a single season. The first may be accomplished by use of a time interval distribution and the second by use of an m^{th} value distribution. Occasionally a situation arises where the m^{th} values cannot be readily made available. Then roughly similar results can be attained by treating the problem by an extreme value distribution on a subextreme statistic such as the daily average.

Intense rainfall design data

In the design of storm drainage systems for buildings and groups of buildings it is often not economical to design against overflows which cause only moderate inconvenience or insignificant damage. In the past the so-called frequency or mean recurrence interval of depth for a given storm duration has been employed to design such drainage systems. Since the time interval between events is averaged, such design values give little information on the spacing of critical rainfall amounts. The spacing appears to be the important factor since the system would be considered a failure if events were spaced too close together. Thus, it appears desirable to fix the rainfall amount for the given duration and give the probability that a rainfall greater than this amount will recur in a certain time interval.

The discrete events of high intensity rainfall follow a Poisson time interval distribution [8] which may be expressed by

$$dp = \alpha \exp(-\alpha t)\, dt \qquad (19)$$

where p is probability, t is time, and α is the mean frequency of occurrence per year. The distribution function is found by integration of (19) giving

$$F(T) = 1 - \exp(-\alpha T) \qquad (20)$$

where T is the time interval in years and F (T) is the probability of a spacing less than T years. The mean recurrence interval is the expected value of t which is $1/\alpha$. Design T's are easily obtained by inverting (20) by taking logarithms.

In some areas estimates of the mean recurrence interval $1/\alpha$ are available for various durations from which the time distribution function is immediately available from equation (20).

The relationship problems in this instance need not be of concern to the climatologist since they have already been fully analyzed by engineers.

Air conditioning design data

In this area mean frequency design data have occasionally been employed which fail completely to reflect the risk of failure of the system. As in other subextreme value problems failure of the system is temporary and an inconvenience but not critical. What is desired is a design value which balances the risk with the potential discomfort and the cost of installation.

The deficiency of the mean frequency design data lay in its failure to reflect within seasonal risk. Thus, the mean frequency data gives only the risk of failure on the average and ignores the possibility of a run of high meteorological values causing system failure in a single season. Clearly the user of the system will be most concerned about successive failures to carry the load; hence, it is necessary to limit the within season risk as well as the between season risk. This can best be accomplished by using the m^{th} value distribution [9].

In the U.S. the air conditioning season is often assumed to be the four months June through September. This period contains 122 days; thus, the highest value in this period has a probability of about 0.0125 of occurring, the second highest 0.025, and the m^{th} highest 0.0125 m. Hence, if the m^{th} value distribution, of which the Type I extreme value distribution is a special case for $m = 1$, is fitted to an annual series of m^{th} values for the air conditioning season, the within year risk is $m/122$ and the between year risk is given by the distribution function. The complement is expressed by the integral

$$H(t) = \frac{m^m}{\beta_m \Gamma(m)} \int_t^\infty \exp\left[m\left(\frac{\alpha - t}{\beta}\right) - m \exp\left(\frac{\alpha - t}{\beta}\right)\right] dt \qquad (21)$$

where t is the design value, m is the rank of the variable from the highest value, α is a location parameter and β is a scale parameter. Tables for obtaining $H(t; m)$ are available. The distribution is applied to both the wet bulb and dry bulb temperatures although the dry bulb load is small compared to the wet bulb load. Design values are obtained from (21) by empirical inversion as for equation (13).

The conversion to the applied variables is made through well known equations for heat transmission and sensible and latent heat loads. Radiation is always an appreciable part of the sensible heat load and is usually taken as that produced by the clear day value for highest sun.

Heating design data

The heating design variable problem is somewhat different from the air conditioning problem because the highest loads often occur during hours when the heating system is operating at comparatively low output or there is little concern for the system meeting the load. Thus, the extreme value is not an appropriate weather design variable nor is an m^{th} highest value useful since it does not reflect diurnal variation which is the problem here. The most desirable variable would be that for a chosen hour but then the choice of hour would be difficult. Finally the daily average was chosen as the variable since it also occurs at some hour, and more important, it was highly correlated with heating load [1]. The climatological series chosen was that of annual minimum of the daily average temperature.

The appropriate distribution is the Type I extreme value distribution for minimum values. This is expressed by the distribution function

$$F(t) = \exp\left[-\exp\left(\frac{t - \alpha}{\beta}\right)\right] \qquad (22)$$

where t is the weather design variable, α is the location parameter, β is the scale parameter, and $F(t)$ is the probability of meeting heat load associated with the minimum daily average t. Design values may be obtained by simple inversion of (22).

The relationship problem has been solved by engineers using heat transmission coefficients and quantities of infiltration and ventilation by outside air.

REFERENCES

[1] Thom, H.C.S., « Revised Winter Outside Design Temperature », *Trans. ASHAE,* Vol. 63, 1957, pp. 111-112.

[2] Thom, H.C.S., « Some Methods of Climatological Analysis », *WMO Tech. Note No.* 81, 1966, Geneva.

[3] Thom, H.C.S., « Frequency of Maximum Wind Speeds », *Proceedings, ASCE,* Vol. 80, 1954, p. 2.

[4] Thom, H.C.S., « New Distribution of Extreme Winds in the United States », *Journal Structural Div. Proceedings ASCE,* July 1968, pp. 1787-1801.

[5] Thom, H.C.S., « Toward a Universal Climatological Extreme Wind Distribution », Paper No. 24, *Proceedings Seminar on Wind Forces on Buildings and Structures,* Sept. 1968, pp. 669-683.

[6] Thom, H.C.S., « Distribution of Maximum Annual Water Equivalent of Snow on the Ground », *Monthly Weather Review,* Vol. 94, No. 4, April, 1966, pp. 265-271.

[7] National Building Code of Canada, NRC No. 3188, National Research Council, Ottawa, 1953.

[8] Thom, H.C.S., « A Time Interval Distribution for Rainfall », *Journal Hydraulics Div. Proceedings, ASCE,* July 1959, pp. 83-91.

[9] Thom, H.C.S., « A Rational Method of Determining Summer Weather Design Data », *Trans. ASHRAE,* Vol. 66, 1960, pp. 248-263.