A STATISTICAL MODEL FOR THE INCIDENCE OF LARGE HAILSTONES ON SOLAR COLLECTORS

MICHAEL COX and PETER R. ARMSTRONG
Altas Corp., 500 Chestnut Street, Santa Cruz, CA 95060, U.S.A.

(Received 4 January 1980; revision accepted 7 October 1980)

Abstract—A statistical model is described for estimating the risk of impacts by large hail on any ground installation (such as a solar collector array). The model is based on data for three frequency distributions: hailstone size, hailfall count (number of hailstones per m² per storm), and number of haildays per year. Other than parameters derived from meteorological data, the parameters of the model required to describe a particular installation are the number of years of surface exposure and the area of the exposed surface. The independent variable is the critical hailstone diameter, D. The result given by the model is the probability of a hailstone of diameter D or greater striking a given surface area in a given number of years. Thus it is possible to determine the “probable maximum hailstone size”, a convenient index of hail risk. Alternatively, the “mean time between hits” may be computed for a given size of hailstone. However, the meteorological data for estimating hail risk are sparse at this time, covering few geographic locations; much of the information available is deficient in sampling consistency and/or sample size. For general application of the model, more detailed data on hailfall in many geographic locations is needed. This model improves on previous work in including all three of the distributions needed to characterize the variability of large hail incidence: hailstone size, count and storm frequency, and in identifying better analytical expressions for these distributions: a special B function for hailstone size, a y function for hailfall count, and a Poisson function or a negative binomial function for annual hailday frequency. The independence of these three random variables is also discussed.

1. INTRODUCTION

One factor in the cost of collecting solar energy is the cost of maintaining solar collectors. A potentially significant component of this maintenance cost is the cost of repairing hail damaged collectors. It is important to be able to estimate the incidence of damaging hail so that the cost of solar power can be compared with other power sources on an equitable basis, and so that hail resistant collector designs can be optimized for different hail prone regions.

Sufficiently detailed hailfall data have been collected in recent years with which to develop the structure of a viable hail risk model. With this model, reasonably accurate hail risk estimates can be made for regions where detailed hailfall data exist. However, until detailed hailfall data are collected in many geographic locations, only crude risk estimates can be made in other regions.

In this article we describe the hail phenomenon, discuss the data available for estimating hail risk, derive a statistical model for estimating hail risk, and show how the model may be applied in assessing the risk of hail damage to a solar collector.

2. THE HAIL PHENOMENON

Though hail is formed in storms across the U.S., as shown in Fig. 1, large hail is formed almost exclusively in
Thunderstorms lee of the Rockies, on the Great Plains, and in the midwest. This "hail belt" is where property damage (e.g. solar flat plate collector glazing breakage) is most likely to occur. Although hail the size of golfballs sometimes occurs outside the hail belt (e.g. Los Angeles[1]), such reports are extremely rare, qualifying for front page newspaper headlines. Crop damage can result from storms with many small hailstones, which are often observed even in California, Washington, Florida and Maine. But this type of hail is not likely to puncture roofing or break collector glazing. Therefore, this study focuses on the occurrence of large hailstones, which have reportedly pierced automobile bodies and killed large farm animals.

Thunderstorms form from atmospheric instability, when warm, moist air is forced to rise. Since temperature is colder with altitude, the rising air is even warmer with respect to the environment. It is, as a result, more buoyant, and accelerates as it rises. At the earth's surface, when a low pressure zone is created and air is sucked into the developing thunderhead (explaining why winds generally blow in the direction of an approaching storm).

In the beginning stages of thunderstorm development, virtually all local air movement is upward. The humidity of the air mass sustains its ascent because condensation liberates heat which counters the adiabatic cooling of the rising air. The ascent slows and eventually stops at some altitude where most of the humidity has been depleted and a "cap" forms on the thunderhead.

At this point the developing storm can be observed from the ground, but precipitation has not yet begun. Water droplets formed in the cloud are supported by the updrafts and coalesce to larger sizes. Hail is formed when particulate matter is carried through a supercooled region of the thunderhead. In this region, supercooled water droplets (as cold as -20°C) readily adhere to ice nuclei and freeze[3]. Eventually, the weight of precipitation is too great to be supported by the updrafts, and torrential hail or rains fail.

Since evaporation or sublimation of this downpour cools the air, it becomes heavier with respect to the surrounding air and strong downdrafts develop. These downdrafts become outrushing "squall winds" at the ground level. At this stage of thunderstorm evolution, up and downdrafts are strongest. The strongest updrafts generally develop on the leading side of a thundercloud with downdrafts immediately behind, in the core of the system. If a growing hailstone is blown from an updraft (say, by strong horizontal gusts) and is then recaptured by the updraft, it can be carried through the supercooled region of the cloud again, accreting more and more layers of ice. In this way, hailstones can be "bounced" up and down in an thunderhead many times until they are heavier than can be supported by the updraft. As long as the fallspeed of a given stone is less than the updraft velocity, it will be supported and can continue to accrete layers of ice and grow. It is interesting to note that, twice in this century, hailstones over 12 cm in diameter have been discovered, indicating updrafts of over 75 m/sec (corresponding to a Mach number of greater than 0.2).

A number of factors affect the size of hail that will be formed. For example, the steeper the "lapse" rate (the vertical temperature gradient in the atmosphere) the faster will be the updrafts generated. More hailstones will therefore be supported longer and will reach larger sizes before falling. Another factor affecting hail size is the vertical extent of the thunderhead. A taller storm potentially allows a longer traverse for a growing hailstone. Thunderstorms in the "hail belt" often extend to altitudes above 12 km.

Three types of thunderstorms produce hail. Possibly the most severe of these is the "frontal" thunderstorm, which comprises a line of storm cells, sometimes hundreds of miles long. At one time, these storms were quite hazardous to transcontinental flights because of the strong up and downdrafts, powerful horizontal gusts and hail encountered. Frontal storms result when either a warm, moist air mass meets colder air and rises (warm front), or when cold wind undercuts a warm air mass, forcing it aloft (cold front). A weather front need not appear at ground level. Air masses can meet at higher altitudes, and are likely to produce taller thunderheads.

Orographic (mountain effect) storms result when warm, moist air blows into a mountain range and is forced to rise. This effect can be seen on the slopes and up to 300 km east of the Rockies.

The "convectional" or "thermal" thunderstorm is a result of local surface heating. On warm, (generally summer) days, moist air close to the earth is often heated by thermal loss from the ground. The air naturally rises, encountering colder air above.

With this brief overview of the hail phenomenon and the factors influencing hail formation, possible sources of data for evaluating the risk of hail damage may be considered.

3. DATA ON HAILFALL AND HAIL DAMAGE

The risk of hail damage may be estimated in either of two ways. The two approaches are based on quite different sets of data.

The direct approach is to consult historical data on hail damage. Insurance companies collect data on crop damage and property damage in the form of damage claims. Crop damage claims are of little use because even the sturdiest plants can be obliterated by pea size hail if it falls with sufficient intensity[4]: larger hailstones do not necessarily cause greater crop damage. Structural damage, on the other hand, results almost exclusively from the impact of golf ball (or larger) sized hail. Insurance claims for hail damage to real property are therefore a much more likely source of data for predicting hail risk. However, most claims do not contain sufficient information to estimate the impact thresholds of damaged structures. This, and the fact that insurance companies do not generally file or tabulate hail damage claims separate from other causes of real property damage (e.g. fires, hurricanes and other phenomenon generally protected by "extended coverage") disqualifies this source of data from analysis at the present time.

The indirect approach requires a determination of the size of hailstone needed to damage a given structure and an estimate of the probability that a hailstone of this size
or larger will strike the structure in a given period of time. The determination of damage thresholds is fairly straightforward. A structural surface is generally tested by propelling artificial hailstones against it at the velocities with which natural hailstones of like masses would fall. (The terminal velocities of hailstones are discussed in Section 8.) The risk of large hail incidence, on the other hand, has eluded analysis. There are, nevertheless, several sources of meteorological data which can form the basis for a hail incidence model[5].

The distribution of hailstorm durations is potentially useful. However, a large number of storms in which both hailstorm duration and hailstone size frequencies are observed is needed to quantify the hail threat. At present, such joint observations have not been recorded.

Another set of joint observations, which is sufficiently complete to be useful in quantifying hail risk, consists of hailstone size observations and observations of hailfall count (hailstones per unit area per storm). From the data on hailstone sizes it is possible to identify an analytical function representing the size distribution for a local parent population of hailstones. From the data on hailfall count it is possible to identify an analytical function representing the hailfall count distribution for a local hailstorm population. These characterizations, together with the familiar and accessible National Weather Service data on annual hailday frequencies, form a complete basis for hail risk prediction. A model based on such characterizations is derived in the next section followed by a discussion of the three distributions used in the model.

4. HAIL RISK MODEL

The hail risk model is based on the distributions of three random variables associated with hailfall: hailstone size, hailfall count (per unit area per storm), and annual hailday frequency. The risk of encountering a hailstone above some threshold diameter, D, is also a function of the surface area exposed to hail and of the duration of exposure. The parameters of the three hailfall distributions are location specific; the area and duration of exposure are installation specific; the hailstone threshold diameter, D, is the independent variable.

A rather confusing but unavoidable feature of the proposed model should be noted at the outset. Hail risk is first derived in terms of the probability that all stones encountered are smaller than diameter D. The alternate probability, that of encountering one hailstone of size D or larger, is "one minus the probability of encountering only stones smaller than D".

The term "trial" may also cause some confusion since it can refer to the random selection of a single hailstone, a single storm, or a year of storms. Forewarned of this, the reader can deduce the proper meaning of "trial" from the context in which it is used.

Now, the cumulative probability that a randomly selected stone will be smaller than D can be expressed as

\[ P_s(<D) \]

and the probability that the hailfall count in one trial storm will be \( n \) hailstones/m² can be expressed as

\[ P_f(n). \]

If \( P_s \) is independent of \( n \), the probability that, in the course of a hailstorm, the hailfall count is \( n \) and all of the \( n \) stones are smaller than \( D \) is

\[ P_s(n) \cdot P_f(<D)^n. \]

For an arbitrary exposure area, \( A \), the probability that \( n \cdot A \) stones will be intercepted during a storm, and that all stones will be smaller than \( D \) is

\[ P_s(n) \cdot P_f(<D)^n \cdot A. \]

The composite probability, \( F \), of encountering only stones smaller than \( D \) on area \( A \) in a trial storm (regardless of the number of hailstones that fall) is

\[ F(<D) = P_s(1) \cdot P_f(<D)^{1 \cdot A} + P_s(2) \cdot P_f(<D)^{2 \cdot A} \\
+ P_s(3) \cdot P_f(<D)^{3 \cdot A} \\
+ \cdots = \sum_{n=1}^{\infty} P_s(n) \cdot P_f(<D)^{n \cdot A}. \quad (4.1) \]

In other words, \( F(<D) \) is the probability of an exposed area, \( A \), not encountering a damaging hailstone in one hailstorm. The probability of experiencing \( m \) storms in a year and of encountering only stones smaller than \( D \) is simply the probability of encountering \( m \) hailstorms in 1 yr, \( P_s(m) \), multiplied by the probability of encountering only stones \( < D \) in \( m \) trials (each storm is a trial). That is

\[ P_s(m) \cdot F(<D)^m. \]

The probability of encountering only stones \( < D \) in a trial year, regardless of the number of storms (any \( m \)) is thus

\[ G(<D) = \sum_{m=0}^{\infty} P_s(m) \cdot F(<D)^m. \quad (4.2) \]

In other words, \( G(<D) \) is the probability of an exposed area, \( A \), not encountering a damaging hailstone in \( 1 \) yr. The probability, therefore, of encountering only hailstones \( < D \) on area \( A \) in \( K \) years is

\[ G(<D)^K. \]

The alternate probability, that of encountering at least one stone of diameter \( D \) or greater in \( K \) years on area \( A \) is simply

\[ P(\geq D, A, K) = 1 - G(<D)^K. \quad (4.3) \]

For convenience, we will refer to this result as the "ultimate risk". Expressed in terms of the three hailfall distributions, \( P_s(<D) \) (distribution of hailstone size), \( P_f(n) \) (distribution of hailfall count) and \( P_s(m) \) (distribution of annual number of haildays) and the in-
stallation parameters, A (exposure area) and K (exposure duration), the ultimate risk is

\[ P(D,A,K) = 1 - \left( \sum_m P_3(m) \cdot \left( \sum_n P_2(n) \times P_1(<D)^m \right) \right)^K. \]  

To evaluate (4.4), \( P_1, P_2 \) and \( P_3 \) must be defined.

5. HAILFALL DISTRIBUTIONS

Hailstone size distribution, \( P_1 \)

To determine the incidence of large hail across the country, one would ideally have a distribution of hailstone size for each geographic location under consideration. Since such detailed records do not exist, it is necessary to assume that the hailstone size distribution does not vary excessively between hailprone locations. Large samples of hailfall data are available for only three locations in North America: Illinois[6-9], north-eastern Colorado[10] and Alberta[11-13], Canada.

Alberta data

Alberta farmers with wire-mesh baskets collected a total of 67 useful stone samples from 1957 to 1963; these samples have been tabulated in histograms and fit to an exponential of the form:

\[ N_i = N_0 \exp \left( -mD_i \right) \]  

where \( N_i \) is the number of stones in diameter category \( D_i \), and \( N_0 \) and \( m \) are the estimated distribution parameters. This is the functional form most commonly fit to hailstone size distributions[10-13]. The exponential appears as a straight line on logarithmic plots such as Figs. 3 and 4.

Note that the exponential does not correspond well with observed frequencies of hailstones greater than about 1.5 cm in diameter. The National Hail Research Experiment (NHRE, conducted in Colorado) investigators, who have not been particularly concerned with the incidence of large stones, acknowledge this. They chose the simplest analytical function which

![Fig. 2. 1972 National Hail Research Experiment hailstone size distribution showing best fitting exponential distribution.](image)

![Fig. 3. Several experimental and fitted hailstone size distributions.](image)

![Fig. 4. Extrapolation of hailstone size distributions. Here the fitted curves of Fig. 3 are extended to larger hailstones sizes where their divergence is significant.](image)
The incidence of large hailstones on solar collectors

adequately describes the bulk of the data, i.e. the smaller diameters. In this study, we are interested in the incidence of unusually large hailstones. That (in the four samples) an exponential always overestimates the occurrence of large stones casts doubt on its use to characterize the high tail of the size distribution.

The inadequacy of the exponential distribution can be attributed to the fact that it has no upper bound. Using this function, there is a finite, albeit small, probability of encountering hailstones the size of basketballs. (Needless to say, hail this size has neither observed nor theoretically postulated.) The knowledge of a maximum possible stone size is potentially valuable information in fitting an analytical distribution to the data, adding, in essence, an additional constraint to the curve. This constraint causes the cumulative probability of hailstone diameter to intersect the point \( (D_{\text{max}}, 1) \). The maximum theoretical size\(^{[14]} \) fortunately agrees quite well with literature reports of the largest observed hailstones\(^{[15]} \).

If these two pieces of evidence are borne out by statistical analysis of the data, then an upper bounded function should be used to represent the hailstone size distribution.

The \( \beta \) distribution is often used to represent naturally occurring distributions which have finite upper and lower bounds. The general \( \beta \) function has two more parameters than the exponential (three rather than one) and can resemble an exponential quite closely except at its upper extreme. The \( \beta \) function was therefore considered a likely candidate distribution to represent the hailstone size distribution.

We have selected the NHRE data for statistical analysis because it is the only large sampling to have been published in usable form.

**NHRE data**

From 1972 to 1974, a total of 1250 hailpads in a total of 37 storms were struck by hail, registering about 150,000 hailstone indentations. Because a new sampling technique was instituted in 1974 (and subsequently deemed unreliable), 1974 data have been disregarded (losing only about 10,000 of 150,000 total hailstones from the sample).

We have fit a special case of the \( \beta \) distribution to the remaining NHRE data (1972 and 1973). This distribution has a probability density of the form

\[
p(x) = (a + 1)(a + 2)(1 - x)^{a+1}x
\]  

(5.2)

where

\[
0 < x < 1.
\]

In fitting a continuous probability function to discrete data (particularly data which have been classified into histogram cells), the integral of the chosen function must be fitted to the cumulative frequency of the data. Integrating (5.2), we have

\[
P(< D) = \frac{x}{D_{\text{max}}} = \frac{D}{D_{\text{max}}}
\]

(5.4)

where \( D \) = stone diameter and \( D_{\text{max}} \) = maximum hailstone diameter (estimated statistically or otherwise). \( P(x) \) was fitted to the classified cumulative data by varying the parameters \( a \) and \( D_{\text{max}} \) in (5.3) and minimizing the root mean square (RMS) error. Ideally, the best fit would be found when \( D_{\text{max}} = 12.5 \text{ cm} \). That the minimum RMS error was obtained with \( D_{\text{max}} = 10 \text{ cm} \) is rather reassuring. With \( D_{\text{max}} = 10 \text{ cm} \) and \( a = 41.1 \), the RMS error was, for 1972 data, < 0.2 per cent and for 1973 data, < 1.0 per cent.

The best fitting cumulative special \( \beta \) function, then, is

\[
P(< D) = 42.1(1 - D/D_{\text{max}})^{3.1} - 43.1(1 - D/D_{\text{max}})^{42.1} + 1
\]

(5.5)

where \( D_{\text{max}} = 10 \text{ cm} \). Figure 4 shows how the \( \beta \) distribution diverges from the exponential at large hailstone diameters.

**Hailfall count distribution, \( P_2 \)**

Of the three large hailfall samples gathered in North America, only the Colorado data (NHR Experiment) were used to quantify the distribution of hailfall count (stones per \( m^2 \) per storm). The data from Alberta and Illinois are not published in a useable form.

The NHRE hailfall count data are plotted in Fig. 5. Although a large number of isolated hailpads (1250) were struck by hail, the number of storms involved was much smaller—only 37. This small sample of storms is probably responsible for bumpiness of the observed hailfall count distribution.

In fitting an analytical function to these data, two constraints are immediately apparent: the function should be lower-bounded (in this case, zero), and the function should have no upper bound, since no evidence...
parameters, p and k. In this case, the method of maximum likelihood must be employed. Calculations for both not always give a satisfactory estimate of the moments. However, the sample mean and variance do vary, s^2, can be estimated by the method of the distribution and, given a sample mean, \( \bar{m} \), and sample variance, \( s^2 \), is given by:

\[
\bar{m} = \frac{1}{N} \sum_{i=1}^{N} m_i 
\]

which N = the number of years of hailday observations and \( m_i \) = the number of haildays observed in the ith year.

Thom [17] has found that when \( \bar{m} < 2 \) the Poisson distribution generally represents the observed hailday distribution better than the negative binomial distribution.

The negative binomial allows another degree of freedom, not requiring \( \mu = \sigma^2 \). The probability function for the negative binomial distribution is given by

\[
f(x) = \frac{(k+x) p^x (1+p)^{-x-k}}{(x+1)k(1-p)^{k-1}}. \tag{5.8}
\]

The two dimensionless parameters, p and k, of the negative binomial are related to the mean and variance of the distribution and, given a sample mean, \( \bar{m} \), and sample variance, \( s^2 \), can be estimated by the method of moments. However, the sample mean and variance do not always give a satisfactory estimate of the parameters, p and k. In this case, the method of maximum likelihood must be employed. Calculations for both moments and maximum likelihood are detailed in Refs. [17, 35].

The calculations required to fit a Poisson distribution to the hailday data are much simpler than the calculations for fitting a negative binomial distribution. The Poisson is defined by just one parameter, \( \lambda \), the mean of the distribution. Thus, one needs only a map or tables of the mean annual number of haildays (available, for example, in Refs. [18, 19]) to calculate this probability function for any location.

Because the Poisson fit is so simple, it has been used almost exclusively to characterize hailday frequencies. However, the Poisson does not represent the stochastic processes of hail in all cases. As we shall see, when \( m \neq s^2 \), misuse of the Poisson can lead to large errors in ultimate risk estimates.

**Independence of the three hailfall distributions**

The derivation of eqn (4.4), giving the probability of large hail incidence, was based on the assumption that the distribution of hailstone size is independent of hailfall count and that hailstone size and hailfall count are both independent of hailstorm frequency at a given location. Hailfall data analyzed to date have been inadequate either in sample size or quality to unequivocally establish the independence of these distributions. The prospects for obtaining adequate data, and the conclusions suggested by existing evidence are therefore quite important at present.

It is conceivable that severity of hailstorms (in terms of either hailfall count or relative frequency of large stones) is correlated with hailstorm frequency. For example, in years when more than the average number of storms occur, the storms may tend to be more severe. Because the magnitude of a hail monitoring program to adequately test such a hypothesis is huge, no such program has ever been completed.

One might also hypothesize—and perhaps more plausibly—a correlation between hailfall count and the distribution of hail size. This hypothesis has been examined by several researchers with various degrees of rigor.

One approach has been to seek a correlation between hailstone size distribution and the duration of hailfall. Gokhale [15] describes the results of two such efforts:

"No correlation was found between the size pattern and the duration of hailfall" by Beckwith [20] in a study of 450 hail reports from Denver, Colorado and vicinity. Likewise, "no relationship was apparent between the duration and size of hailstones" according to Sulakvelidze [21].

Only Carte and Basson [22] have reported any correlation between duration of continuous hailfall and maximum reported hail size. Their investigations were confined to Transvaal, and they observed only a weak dependence.

A second approach is to seek a correlation between size distribution and precipitation rate. Size distributions from several data sources have been compiled by Atlas and Ludlam [23]. Varying degrees of bias are present in the relative frequencies of small diameter stones due to
The incidence of large hailstones on solar collectors

103 melting on the ground and other biases peculiar to the various sampling techniques used. However, the distributions are all quite similar for diameters greater than about 1.5 cm in spite of a hundred-fold variation in precipitation rate among the storms sampled.

The most direct approach, of course, is to seek a correlation between size distribution and hailfall count. Douglas [11] found the correlation between hailfall count and size distribution to be insignificant in his analysis of 25,589 hailstones collected in 57 point occurrences of hailfall in Alberta. He therefore postulated "the existence of a simple and unique parent population".

The independence of hailfall count and hailfall size distribution in storms sampled by the Colorado and Illinois hailpad networks has apparently never been tested. Such an analysis is currently underway at Altas. At the present time, however, it can only be assumed that hail size distribution is uncorrelated with hailfall count since the evidence indicates that they are, at best, poorly correlated.

Likewise, there is no evidence that either hailfall count or size distribution is correlated with annual number of haildays. Again, analysis of the Colorado and Illinois hailpad data should provide a more definite answer to this question. Until such data are analyzed it must be assumed that the distributions are independent.

Equation (4.4) can, nevertheless, be made valid for a climate where \( P_i \) is dependent on \( P_2 \) by replacing \( P_i(<D) \) with \( P_i(<D, n) \). However, the reduction of eqn (4.4) to probability charts such as Figs. 6-10 is not possible if \( P_i \) is correlated with \( n \). The variation of size distribution with location alone, on the other hand, would not reduce the applicability of probability charts since these are location dependent anyway.

6. PROBABILITY CHARTS FOR ESTIMATING HAIL INCIDENCE

Having identified the probability functions for the three hailfall variables, the expression for ultimate risk may be evaluated for various values of \( A, K, \) and \( D \). Figures 6–10 show the results of many evaluations of the risk model. The ultimate risk—i.e., the probability that one stone of diameter \( D \) or larger will strike the exposed surface, \( A, \) in \( K \) years—is plotted on the abscissa in Figs. 6–9.

It will be noted that the ordinate axis in these figures does not represent hailstone size, \( D \). This means that the figures cannot be read directly to determine the ultimate risk of encountering larger than a \( D \)-cm hailstone. The ordinate value, \( P_i(<D)^A \) must first be evaluated using one of the hail size distributions of Figs. 3 and 4.

In Figs. 6–10, either coordinate may be viewed as the independent variable. If a value of \( D \) is specified, \( P_i(<D)^A \) is the implied independent variable and the ultimate risk of encountering a hailstone \( \geq D \) is read from the graph. Alternatively, a level of risk can be specified; reading the corresponding value of \( P_i(<D)^A \) from the graph, the largest size hailstone, \( D \), likely to be encountered at that level of risk may be calculated. These two views are illustrated below.

**Ultimate risk as the dependent variable**

Suppose we wish to know the ultimate risk of encountering a hailstone \( \geq D \) on an exposed surface area, \( A, \) in \( K \) years at a particular location.

Recall from eqn (4.1) that \( A \) enters the risk model as follows:

\[
\sum_n P_3(n) \cdot P_i(<D)^A \cdot A.
\]

![Fig. 6. Relation between hailstone size distribution, \( P_i(<D) \), and ultimate risk of large hail impact, \( P(\geq D, K, A) \), for Albuquerque, NM. \( K \) is in years and \( A \) is in m².](image-url)
Fig. 7. Relation between hailstone size distribution, $P_i(<D)$, and ultimate risk of large hail impact, $P(\geq D, K, A)$, for Abilene, TX. $K$ is in years and $A$ is in $m^2$.

Fig. 8. Relation between hailstone size distribution, $P_i(<D)$, and ultimate risk of large hail impact, $P(\geq D, K, A)$, for Kansas City, Mo. $K$ is in years and $A$ is in $m^2$.

Note, however, that this is equivalent to
$$\sum_n P_i(n) \cdot [P_i(<D)^A]^n.$$ So the entire term, $P_i(<D)^A$, may be taken as the independent variable. The results in Figs. 6–10 were obtained, not for specific values of $D$ and of $A$, but for a plausible range of the term $P_i(<D)^A$. Thus Figs. 6–10 are independent of the distribution used to represent hail size.

For example, suppose we are interested in the risk of encountering a 3 cm hailstone on either of two flat plate collector installations, one 10 $m^2$ and one 100 $m^2$. From Fig. 4, $P_i(<3 \text{ cm}) = 0.999994$ (using the Colorado size distribution).

For the smaller of the two installations, $A = 10 \text{ m}^2$, $P_i(<3)^{10} = 0.99994$. For the larger installation, $A = 100 \text{ m}^2$, $P_i(<3)^{100} = 0.99994$. The ultimate risk of encountering a 3-cm stone in both cases can now be read from one of Figs. 6–9. Suppose we are interested in the risk of
The incidence of large hailstones on solar collectors

Fig. 9. Relation between hailstone size distribution, \( P(\leq D) \), and ultimate risk of large hail impact, \( P(\geq D, K, A) \), for Cheyenne, WY. \( K \) is in years and \( A \) is in \( m^2 \).

Fig. 10. Relation between hailstone size distribution, \( P(\leq D) \), and exposure period, \( K \), at an ultimate risk of 0.50. This chart is used to estimate the largest hailstone likely to be encountered in \( K \) years, or to estimate the mean time between impacts of hailstones with diameter \( D \) or larger.

encountering such a hailstone in Kansas City, MO; suppose also that both collector installations will be operating for 5 years (\( K = 5 \)).

For smaller installation, \( P(\leq 3)^{10} = 0.99994 \), so ultimate risk = 0.82. For the larger installation, \( P(\leq 3)^{100} = 0.9994 \), so the ultimate risk is greater than 0.99 (very nearly 1.00).

Following this procedure, one can determine the risk of encountering a hailstone \( \geq D \) on area \( A \) in \( K \) years for many combinations of parameters, with one exception. The parameter \( K \) in Figs. 6–9 has been calculated only to \( K = 10 \). For utilizing the model for longer time periods, Fig. 10 has been plotted to show the hailstone size predicted at an ultimate risk = 0.50 for values of \( K \) as large as 210 yr.

Ultimate risk as the independent variable

Suppose that we do not know the hailstone size that will cause damage to our ground installation, or we are interested in determining "hail-proof" design specifications. We might be interested in the size we can expect to encounter at a given level of risk.
In this case, we would go from the abscissa value in Figs. 6-9 to the corresponding ordinate value of \( P_i(< D)^A \) (for a given value of \( K \)). Before this can be correlated with hailstone size via Fig. 4, we must take the \( A \)-th root. That is,

\[
[P_i(< D)^A]^{1/A} = P_i(< D).
\]

For example, suppose we want to know the hailstone size that is likely to be encountered with an ultimate risk of 0.30 in Kansas City in 10 yr. Using curve \( "K = 10" \) from Fig. 10, we find that \( P_i(< D)^A = 0.999992 \). Suppose, again, that we have two collector installations, one 10 m\(^2\) and one 100 m\(^2\). Where \( A = 10 \) m\(^2\),

\[
P_i(< D)^{10} = 0.999992
\]

\[
P_i(< D) = 0.999992.
\]

From Fig. 4, \( D = 3.3 \) cm. Where \( A = 100 \) m\(^2\),

\[
P_i(< D)^{10} = 0.9999992
\]

\[
P_i(< D) = 0.9999992.
\]

From Fig. 4, \( D = 3.75 \) cm. Following this procedure, values of \( P_i(< D)^A \) at an ultimate risk of 0.50 have been plotted against the number of years of exposure, \( K \), in Fig. 10. The hailstone size that is predicted by the risk model at an ultimate risk = 0.50 may be interpreted as the “probable maximum hailstone size” that will be encountered at a particular location on area \( A \) in \( K \) years, and as such, is perhaps of most general interest.

Alternatively, the value of \( K \) corresponding to ultimate risk = 0.50 for a given \( D \) may be read from Fig. 10. This value is the mean time between impacts by hailstones \( \geq D \).

Table 1 shows the probable maximum hailstone size that is predicted in the four locations analyzed (see Table 2) for various values of \( A \) and \( K \).

<table>
<thead>
<tr>
<th>Number of years of exposure (K)</th>
<th>Surface area ( A ), m(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>COLORADO</td>
<td>2.75</td>
</tr>
<tr>
<td>ALBERTA</td>
<td>4.05</td>
</tr>
<tr>
<td>ILLINOIS</td>
<td>4.45</td>
</tr>
<tr>
<td>Abilene, Texas</td>
<td>4.38</td>
</tr>
<tr>
<td>Kansas City, Missouri</td>
<td>4.35</td>
</tr>
<tr>
<td>Cheyenne, Wyoming</td>
<td>4.35</td>
</tr>
<tr>
<td><strong>Albuquerque, New Mexico</strong></td>
<td><strong>5.15</strong></td>
</tr>
<tr>
<td><strong>COLORADO</strong></td>
<td><strong>5.15</strong></td>
</tr>
<tr>
<td><strong>ALBERTA</strong></td>
<td><strong>5.05</strong></td>
</tr>
<tr>
<td><strong>ILLINOIS</strong></td>
<td><strong>5.00</strong></td>
</tr>
<tr>
<td><strong>COLORADO</strong></td>
<td><strong>5.15</strong></td>
</tr>
<tr>
<td><strong>ALBERTA</strong></td>
<td><strong>5.05</strong></td>
</tr>
<tr>
<td><strong>ILLINOIS</strong></td>
<td><strong>5.00</strong></td>
</tr>
</tbody>
</table>
7. SENSITIVITY ANALYSIS

It is of interest to learn how estimated hail risk varies with parameters of the three hailfall distributions. This exercise indicates which deficiencies in hailfall data are most crucial.

Sensitivity to hailday distribution parameters

The principal hailday statistics of four locations chosen to test the model's sensitivity are listed in Table 2. These locations represent the range of hailday statistics encountered in the U.S. The number of hailstorms in Cheyenne, WY, can be considered a "worst case", with the highest value of mean number of haildays found in the literature, and probably one of the most dispersed distributions of annual hailday frequency, with a sample variance of \( s^2 = 2.5 \bar{m} \). The dispersion of hailday distributions has not been given adequate attention and, as will be seen, is a significant factor in the incidence of large hail.

Only one of the four data sets appearing in Table 2 was better fit by a Poisson probability density function than by a negative binomial. This is typical of the 25 hailday data samples.

Among the 25 locations analyzed, there were no distributions with a mean annual number of haildays of between 5 and 8. From Fig. 1, it can be seen that only a few, very localized regions have a mean annual number of haildays of greater than 5.

In Figs. 6-10, ultimate risk is plotted for four values of exposure period, \( K \). The solid line represents the results when the negative binomial was used to describe hailday frequencies, the dashed line when the Poisson was used. It can be seen in Figs. 6-8 (not in Fig. 9) that in the "mid-range" of ultimate risk, (with values between 0.2 and 0.8) there is not much difference between the results when either distribution is used. However, as the number of years of exposure increases, the negative binomial becomes increasingly divergent from the Poisson distribution. (Nevertheless, the simple criterion suggested by Thom[17] for deciding when a Poisson distribution is acceptable—the \( \bar{m} < 2 \) criterion—is not unequivocally borne out by these examples.) It is interesting to note the great difference between the Poisson and negative binomial distributions found for Cheyenne, WY (Fig. 9). For \( K = 10 \), the two distributions give results that hardly seem to apply to the same geographic location. This is best explained by pointing out that the \( s^2 = \bar{m} \) property of the Poisson distribution is violated most seriously by the Cheyenne hailday distribution where \( s^2 = 2.5 \bar{m} \).

It is clear that proper choice of the analytical function for \( P_3(m) \) is crucial in certain locations. This is especially important when \( K \) is large, as it must be for most solar collector installations to be economically viable. The foregoing analysis also demonstrates the importance of hailday dispersion. In the case of Cheyenne, use of the Poisson distribution is tantamount to assuming a hailday frequency variance of only 40 per cent of the true variance, causing order-of-magnitude errors in ultimate risk estimates.

Sensitivity to exposure period

Next, the model was tested for sensitivity to increases in the number of years of collector exposure, \( K \). It has already been noted that for each location shown in Table 2, the Poisson and negative binomial distributions led to increasingly different values of ultimate risk as \( K \) increased from 1 to 10. However, for a common base of comparison between locations, only one distribution was chosen to describe \( P_3(m) \) as the model was tested for sensitivity to increasing \( K \); the negative binomial was used because it provided an adequate fit of the raw hailday data in all four locations whereas the Poisson provided a marginally better fit in only one case. In Fig. 10, \( P_3(< D)^k \) is plotted against \( K \) for four locations (four different hailday distributions). For each curve, the ultimate risk is 0.50. That is, each curve represents the largest size hailstone (via \( P_3(< D)^k \)) likely to be encountered in \( K \) years.

One might expect the probable maximum hailstone size to be roughly proportional to the mean annual number of haildays. But note in Fig. 10 that there is little difference between values of \( P_3(< D)^k \) predicted for Albuquerque, Abilene and Kansas City. This means that if the same \( P_3(< D) \) distribution is used, the hailstone size predicted by the model is nearly the same in these locations. Table 1 shows actual sizes predicted.

In contrast, the results for Cheyenne, WY seem to depart radically from those of the other three locations. This can be seen in Figs. 9 and 10. Recall that the most significant departure of the Poisson from the negative binomial is found for this hailday distribution. And clearly, the results for large values of \( K \) are quite deviant.
from those of the other locations tested. In Fig. 10, the results for Cheyenne extend beyond the range of the graph at a value of $K = 35$ yr.

These observations indicate that hailday dispersion is more important than has been generally appreciated. The sensitivity of large hail incidence to $s^2$ is comparable to its sensitivity to $\bar{m}$.

Sensitivity to size distribution parameters

The semi-log plots of Figs. 3 and 4 and Figs. 6-10 emphasize the importance of how the size distribution is extrapolated. The extrapolated tails of these distributions are a major uncertainty in the model because of the lack of data on large hailstone frequencies. We are now in a position to quantitatively assess the importance of assuming a particular extrapolation.

It will be seen in Table 1 that hailstone sizes predicted using three size extrapolations of Figs. 3 and 4 differ in a systematic manner. As $A$ and $K$ increase, each distribution for $P_r(<D)$ leads to increasingly divergent results. The Alberta and Colorado size distributions are both analytically fitted functions (an exponential for Alberta, a $\beta$ distribution for Colorado). As was discussed previously, use of an exponential for extrapolating a size distribution is questionable, since it has no upper bound.

The use of an unbounded function will, at some large value of $A$ and $K$, result in the prediction of a supernatural hailstone. At what value of exposure an exponentially fitted size distribution becomes unacceptable is unknown. However, over the range of $A$ and $K$ represented in Table 3, neither the exponential fit nor the $\beta$ fit predicts implausible hailstone sizes. Perhaps the best value for hailstone size can be taken as somewhere between the $\beta$ and exponential predictions. (Without further analysis of the Illinois raw data, we recommend against the use of the extrapolation of Illinois size data appearing in Figs. 3 and 4.) The maximum probable hailstone diameter estimates resulting from $\beta$ and Alberta exponential size distributions don't seem to disagree excessively in many cases. For example, the hailstone predicted in Kansas City in 20 yr on 100 m$^2$ is between 3.3 and 4.6 cm. However, the terminal kinetic energies of falling hailstones of these two diameters differ by a factor of more than 3 as is shown in the next section.

8. HAILSTONE TERMINAL VELOCITY AND KINETIC ENERGY

The kinetic energy of a falling hailstone is a good measure of its destructive potential. The stone's kinetic energy is related to its mass and velocity.

A fully formed hailstone falling from a thunderhead will accelerate until its weight is just balanced by the aerodynamic drag acting on its surface. At this point, the hailstone's velocity is, by definition, its "terminal velocity".

A more complete discussion, and derivation of an expression for hailstone terminal velocity can be formed in Refs. [14, 15, 24, 28, 29, 35]. However, equating drag with weight, we obtain an expression for the terminal velocity of a hailstone, $V_t$:

$$ V_t = \left( \frac{4\rho_D Dg}{3\rho_s} \right)^{1/2}. \tag{8.1} $$

The drag coefficient, $C_d$, for a sphere is a weak function of Reynolds number only. The dimensionless Reynolds number is defined as

$$ Re = \frac{VD\nu}{\nu} $$

where $\nu$ = the kinematic viscosity of the fluid, air, cm$^2$/sec.

A plot (from Ref. [26]) of drag coefficient vs Reynolds number for smooth spheres is shown in Fig. 11. Because hailstone shape and surface roughness are extremely difficult to characterize [20, 27-29], we shall assume that hailstone drag coefficients are the same as the smooth sphere drag coefficient. Since smooth spheres generally have somewhat lower drag coefficients than rough or irregular objects that are "almost spherical" this represents a slightly conservative approach. Note that there is a sudden drop in $C_d$ at $Re \approx 2 \times 10^4$. This is due to a transition in the boundary layer from laminar to turbulent flow. (The "wake" begins to separate from the hailstone further toward the rear end because of the
The incidence of large hailstones on solar collectors

The incidence of large hailstones on solar collectors turbulence, reducing the drag.) The drop in $C_d$ corresponds to the "critical Reynolds number".†

It will be noted in Fig. 12 that there are two values of terminal velocity for hailstones between 6 and 8.5 cm dia., the "double valued" region of terminal velocity. For hailstones this size, the terminal Reynolds number is near the critical value of about $2 \times 10^5$. For a slight increase in Reynolds number, there is about a factor of two decrease in $C_d$. At just under $Re = 2 \times 10^5$, a stone falling through still air would normally have a laminar boundary layer, hence the lower of the two terminal velocities. However, some investigators have proposed that a real hailstone might actually attain the faster of these fall speeds by exceeding the critical Reynolds number ($Re > 2 \times 10^5$) as it falls through atmospheric turbulence[14]. Various other mechanisms have also been suggested by which a hailstone in the double valued region could attain the higher velocity[28]. There is a general consensus that a stone of a diameter such that its terminal velocity is in the double valued region will tend to remain at the faster velocity once a disturbance has caused the initial transition to a turbulent boundary layer.

The "terminal kinetic energy", $E_k$, of a falling hailstone is given by the simple relation

$$E_k = \frac{1}{2} m \cdot V_t^2$$

where $m = \rho D^3$ is the mass of the hailstone.

The terminal kinetic energies of hailstones have been calculated and plotted in Fig. 13. It can be seen by combining (8.1) (8.2) that the kinetic energy of a falling hailstone varies approximately as the $4$th power of the diameter. That is,

$$E_k \propto D^4.$$  

(8.3)

With small increases in stone size, one finds large increases in the terminal kinetic energy. The accuracy with which a hail risk model predicts the maximum probable hailstone size, $D$, is thus quite important.

The kinetic energies of impact sufficient to damage

†Willis et al.[28] have noted that large dry hailstones behave as if rough, resulting in a boundary layer transition to turbulence at lower Reynolds numbers than for smooth stones. (That is, critical Reynolds number is observed for smaller hailstones.) However, it is difficult to envision natural hailstones with surfaces sufficiently dry (and therefore rough) to observe this phenomenon in nature.
various collector glazings and reflector materials have been determined under a number of testing programs [30–33]. At present there are also some testing programs in progress including one at Atlas Corporation. In some cases materials are tested by dropping steel balls on the surface. Bags of steel shot are also dropped, producing impacts with the same kinetic energy but much lower maximum local stress. Natural hailstone impacts of the same kinetic energy can be expected to cause an intermediate level of damage. To determine damage thresholds more precisely, investigators have propelled ice balls at their test surfaces to simulate hailstone impacts. In these tests, 4.8 mm (3/16 in.) tempered lights have broken with impact energies ranging from 10 to 40 Nm. For comparison, 6.4 mm (1/4 in.) tempered lights are broken by the impact of a steel ball with about 10 Nm kinetic energy and by the impact of a shot bag with over 100 Nm kinetic energy. Even the validity of ice ball tests is debatable since natural hailstones, having some air and liquid water content, are softer. The properties of dried clay balls or refrozen crushed ice balls may better duplicate the properties of large natural hailstones. With further progress in ongoing investigations it should be possible to answer some of these questions and to compile and present comprehensive data on damage thresholds for collector materials.

9. CONCLUSION

We have shown that the probability of a solar collector being struck by a given or larger hailstone is related to the distributions of three random variables: hailstone size, hailfall count per storm, and hailfall frequency. Evidence for the independence of these distributions has been presented. The independence assumption leads to the expression for ultimate risk (eqn 4.4).

To evaluate (4.4), analytical functions for the three hailfall distributions must be identified. Of the three distributions, hailday frequency, \( P_3 \), is probably the most geographically variable. Fortunately, the data needed to identify this distribution are readily available for many locations [34] and the distribution parameters can be evaluated using moments and maximum likelihood estimates presented in Ref. [35].

In contrast to hailday observations, the sampling programs needed to collect sufficient data for characterizing hail size and hailfall count distributions are vast and complex. Useful data have been collected in only three areas of North America. Hailfall count per storm is probably less geographically variable than hailstone size. Therefore, the \( \gamma \) distribution, fitted to Colorado hailfall count data in Section 5, can probably be used with confidence for all of North America. The curves of Fig. 10 use this hailfall count distribution, and the hailday frequency distribution parameters of the four locations represented in Fig. 10 span the range of local conditions found in hail prone areas. Furthermore, these curves are independent of which hailstone size distribution is used. Figure 10 may therefore be used in assessing hail risk for many applications to avoid detailed analysis of hailfall count and hailday frequency data.

Hailstone size distributions probably vary somewhat more with location than hailfall count. Of the three distributions, moreover, hailstone size requires the largest sample size because variation of mean size between storms is great. The problem is compounded by the fact that the large stones, whose relative frequency is of greatest concern, are so rare that it is necessary to extrapolate the upper tail of the size distribution. Thus, the form of analytical function chosen to represent the distribution is quite crucial. The special \( \beta \) distribution, fitted to Colorado size data in Section 5, is the simplest analytical form that is reasonable for this extrapolation. The evidence that size distributions in all hailprone areas are fairly similar is sufficient to justify the use of this \( \beta \) distribution for all North American locations for the present. Analysis of the Illinois and Alberta hailsize and count observations may alter this view. However, what is really needed are extensive hailstone size and count monitoring programs in many more locations. With better characterization of hailstone size distribution the hail risk model presented herein can be used to accurately optimize the design of solar collectors susceptible to hail damage.

Acknowledgement—We are indebted to Robert H. Bushnell, meteorologist and solar engineer, Boulder, CO, and to Mike Edesess, SERI engineer and statistician, for reviewing the manuscript and offering many helpful suggestions. This work has been supported by the Conservation and Solar Applications Branch of the U.S. Dept. of Energy under Contract EM-78-C-04-4291.

REFERENCES

The incidence of large hailstones on solar collectors


Resumen—Se presenta un modelo estadístico para la estimación del riesgo de impacto de piedras de granizo de tamaño apreciable sobre cualquier instalación terrestre (como una instalación de colectores solares). El modelo está basado en datos sobre tres distribuciones de frecuencia: tamaño de piedra de granizo, número de piedras (número de piedras por m² por tormenta), y número de días con granizo por año. Aparte de los parámetros derivados de datos meteorológicos, los parámetros requeridos para describir una instalación específica con el modelo son los años de operación de la instalación y el área de las superficies expuestas. La variable independiente es el diámetro de la piedra crítica, D. El resultado del modelo es la probabilidad de impacto de la superficie por una piedra de diámetro D o más grande durante el período determinado. Por otra parte, se puede computar el "tiempo medio" entre impactos para un dado tamaño de piedra. Sin embargo, los datos meteorológicos necesarios para estimar los riesgos son muy escasos en la actualidad, y son para muy pocos lugares. Mucho de los datos son deficientes en cuanto se refiere a consistencia de muestras y tamaño de muestras. Para aplicación general del modelo, se necesitan más datos sobre caídas de piedras en muchas áreas geográficas. Este modelo ha mejorado los trabajos previos por el hecho de incluir las tres distribuciones necesarias para describir la incidencia de piedras grandes: el tamaño de piedra, el número de piedras, y el número de tormentas, y por el hecho de identificar mejores expresiones analíticas para estas distribuciones: una función especial de tipo "beta" para el tamaño de la piedra, una función gamma para el número de piedras, y una función Poisson o una función de binomio negativa para el número de días de tormenta. Hay una discusión de la independencia de estas tres variables estadísticas.