26. Statistical Characterization of Hazards and Risk in Coastal Areas

Donald T. Resio, Mark A. Tumeo, Jennifer L. Irish

We examine the foundation for hazard/risk assessment and its application to coastal problems. Historically, emphasis was on specifying expected values of wind waves and storm surges; however, as shown by the recent tsunamis in Southeast Asia in 2004 and in Japan in 2011, there are critical parts of the world where tsunamis represent the dominant threat to coastal communities. Recently, there has been an increased awareness of the combined effects of heavy rainfall and/or river discharge with surge levels and strong winds. This forcing combination played an important role in the flooding in southern Louisiana during Hurricane Isaac in 2012, where water levels exceeded the 500-year return interval levels. Such forcing combinations complicate both the modeling systems required for their simulation and the treatment of the multivariate probabilities that define the relative importance of their impacts.

We begin with a set of consistent hazards and risk definitions, along with comparative definitions from other fields. This should help readers who have focused primarily on traditional coastal hazards and risks understand the broader context of risk assessment and also allow readers with a broader perspective gain insight into the specific nature of coastal hazards and risk. Following this, we introduce the basic concepts used in estimating coastal hazards and risks. We then examine the historical perspective for the evolution of coastal risk assessment, beginning with early deterministic methods and culminating in a recent transition to probabilistic methods. The steady increase in the ability of probabilistic methods to deal with persistent problems such as the lack of data and uncertainty is documented as a part of this transition.

26.1 Overview of Risk and Uncertainty

Technically, *risk* is defined as the mathematical probability that a specific outcome will occur multiplied by the magnitude of the impact of that occurrence. However, there are many other definitions in use, and while a technical working definition such as the one presented is useful, it is more important that there is a contextual understanding of the concept.

While *risk* is now used in some manner across almost every discipline, the entire cultural framework upon which the concept is based is only a few hundred years old (and relatively young as a basis compared to other engineering and mathematical concepts). Most ancient cultures took the stance that the future was either fated (deterministic) or subject to the whims of powers beyond our control or understanding. There is no concept of *risk* in such cultures beyond the fear of certain punishment should you anger the powers that be. The roots of our conception of risk arose in the Renaissance, as cultures in the west broke away from and challenged long-held beliefs. One of the critical turning points for the concept of risk as we know it was the development of the concept of *probability* – the idea that the future was not determined or unpredictable, but that there was the possibility to make predictions based on numbers, not oracles or soothsayers. The origin of this truly revolutionary concept is recounted by *Bernstein* as follows [26.1]:

> *In 1654, a time when the Renaissance was in full flower, the Chevailier de Mere, a French nobleman with a taste for both gambling and mathematics, challenged the famed French mathematician Blaise Pascal to solve a puzzle. The question was how to divide the stakes of an unfinished game of chance between two players when one of them is ahead. The puzzle had confounded mathematicians since it was posed some 200 years earlier by the monk Luca Paccioli … Pascal turned for help to Pierre de Fermat, a lawyer who was also a brilliant mathematician. The outcome of their collaboration was intellectual dynamite. What might appear to have been a seventeenth-century version of the game of Trivial Pursuit led to the discovery of the theory of probability, the mathematical heart of the concept of risk.*

Hence, imbedded in the concept of risk is the inherent belief in predictability, at least in a probabilistic sense. This concept, however imbedded in our current culture, stands in stark contradiction to the well-known adage that one cannot accurately predict the future. In the modern world, this contradiction is resolved through the belief that if we knew enough, we could accurately predict the future. This belief in scientific determinism (the same concept from ancient times that was *overturned* by the concept of risk) was articulately stated by *Pierre-Simon Laplace* in his landmark 1812 treatise *Essai philosophique sur les probabilités* (*A Philosophical Essay on Probabilities* as translated by Truscott and Emory in 1902) [26.2]:

> *We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.*

We have therefore, in many ways, come full circle as a culture. Through the concept of risk we have moved away from the underpinning belief in determinism to accept that there is an unknown and unknowable future. Yet, through the application of mathematical probabilities, which gave rise to the concept of risk, we also have developed the idea of a scientific determinism, wherein if all lack of knowledge could be eliminated through scientific inquiry, we would once again have a certain and determined future.

The cultural concept of scientific determinism through the elimination of uncertainty and evaluation of undetermined but manageable future options are the heart of risk assessment. The purpose of the study of risk is to manage the future by using our best knowledge and current understanding to reduce uncertainty and quantify the remaining uncertainty through probabilistic prediction.

26.1.1 Definitions of Basic Terms

Because the concepts of risk and uncertainty are now so ingrained into our society, the use of terms associated with *risk and uncertainty*, including *hazard*, *probability*, *risk assessment*, *risk analysis*, and in fact the words *risk* and *uncertainty* themselves, are used frequently in differing and often confusing ways. When one adds to this situation the fact that risk and uncertainty inherently rely on other ill-defined and often misused concepts such as *error* and *variability*, it becomes apparent that before entering into a discussion of the application of hazard probability and risk assessment to coastal systems, it is necessary to ensure that there is a consistent and uniform understanding of terms. A full list of definitions for probability and risk related terms used in describing coastal hazards is given in the Sect. 26.A; however, several key definitions are retained here that are most pertinent to discussions within this chapter:

- *Probability*: In its most basic sense, probability is the chance or likelihood that something will happen. Qualitatively, the more likely an event is to happen, the more *probable* the event. Quantitatively, probability is a value between 0 and 1, with 1 representing absolute certainty of the event occurring. The probability of an event is typically measured as the ratio of the number of times an event occurred over the total number of times the event could have occurred. For example, if we consider the event to be the occurrence of precipitation on any given day, we would collect information on whether it rained on any given day for a period of time (say 1 year). The probability of rain would then be

$$\text{probability of precipitation} = \frac{\text{number of days with precipitation}}{\text{Total number of days}} .$$

- *Probability distribution*: The basic definition of probability works for discrete events, but when the event or parameter being evaluated is continuous in nature (i. e., the maximum temperature on any given day, the stage of a river on any given day), then a single probability is not sufficient. Instead, possible values are grouped into discrete ranges and the number of occurrences in that range are counted, then divided by the total number of measurements.
- *Risk*: The potential for realization of unwanted, adverse consequences to human life, health, property, or the environment. The estimation of risk is usually quantified using the expected value of the conditional probability of the event occurring times the consequence of the event given that it has occurred. This definition is currently used by the Society of Risk Analysis. Mathematically, this is expressed as

  $$R = P(A) * P(B|A) ,$$

  where R – Risk (a probability from 0 to 1), $P(A)$ – probability of event (A) occurring, $P(B|A)$ – probability of a consequence (B) occurring given that event A occurred.
  This technical definition is what will be used through this handbook. However, there are other definitions in other fields of which the knowledgeable practitioner should be aware (Sect. 26.A). Besides having multiple potential definitions, risk can also be differentiated by types. Different types of risk, even if they are of the same quantitative value, are often managed differently or even ignored. Some of the common risk types that greatly influence both how these risks are assessed, managed and communicated are defined below (Sect. 26.A for full list).
  – *Actual risk*: A scientifically verifiable risk. For example, it is well researched and documented that smoking places you at-risk for cancer [26.3–6]. Actual risk is sometimes referred to as *objective risk*, but whether risk is subjective or objective is related more to its ability to be measured than it is to its actual verifiability.
  – *Perceived risk*: Risk that is thought to exist by an individual or group that either is understated or is exaggerated. This often occurs in situations where the public is misinformed or in which media reports instill unnecessary panic. Food safety concerns often top the list of such events and lead to significant public policy debates, see, for example [26.7].
  – *Assumed risk*: Risk that is taken by choice. Assumed risk can be quantifiably large or small, and actual or perceived. For example, individuals who choose to partake in risky activities (skydiving, mountain climbing) choose to assume the relatively large risks associated with these activities, but choosing to drive a car, take medicines, or be involved in day-to-day activities all involve some assumed risk.
- *Imposed risk*: Risk that is forced upon an individual, either without the knowledge of the individual or if known, without consent. For example, second-hand smoke exposure is seen as an imposed risk [26.8–10]. Natural events such as earthquakes, tropical cyclones (e.g., hurricanes, typhoons), and extreme weather events are to a large extent *imposed* risks, but to some extent individuals assume that risk based on where they choose to live. For an interesting approach to this, see the work by [26.11].
- *Uncertainty*: Uncertainty arises from the fact that any model (including risk assessment models) cannot provide 100% *certainty* in its results. This uncertainty arises from two major sources.
  – *Epistemic uncertainty*: Epistemic uncertainty arises from the fact that there are things about the natural system we are analyzing that we do not know. This *lack of knowledge* can be because the knowledge is not yet scientifically available – and as such, it can be reduced (along with the resulting epistemic uncertainty) going forward through additional data, experimentation, theoretical development, and scientific inquiry. However, lack of knowledge also includes things that we do not even know that we do not know. This area of *lack of knowledge* is more difficult, because it is not easily identifiable and as such, becomes included in either variability, or is called *error* when comparing model results to reality.
  – *Alleatory uncertainty*: Variability is a range of potential values for a given parameter. Variability is an inherent characteristic of natural processes (also called natural variation). It is describable using probability distributions. The result of natural variation is sometimes called alleatory uncertainty. Variability in natural processes (and the resulting alleatory uncertainty) cannot be reduced, as it is an inherent property of the process itself.

26.2 Quantifying Coastal Hazards/Risks

A detailed treatment of all coastal hazards and risks is beyond the scope of the work presented here. Environmental/health hazards and risks such as those related to water quality, chemical spills, and salinity intrusions will not be treated here; however, it is hoped that the methods discussed in this section will be useful, at least in a general sense, to those working in these areas. Additionally, due to the fact that much of the recent research on coastal hazards/risks has been undertaken in response to Hurricane Katrina, much of the focus relating to recent developments will be on storm surges produced by tropical cyclones. Although extratropical storms play a very large role in many coastal areas, the lack of a recent catastrophic event, such as flooding the subway system in New York City before Hurricane Sandy (which was a hybrid mix of tropical cyclone and extratropical characteristics) has meant that methods used for evaluation of these hazards/risks have received far less attention than they probably deserve. A single section (Sect. 26.3.6) will be devoted here to specific issues related to extratropical storms, following our general treatment of extreme and related hazards/risks.

Before launching into our treatment of hazard probabilities, it should be recognized that both epistemic and alleatory uncertainties need to be considered within any thorough analysis of hazards and associated risks. Limitations in theoretical/numerical predictions come from many sources, including simplifications in our models, the lack to make nature fit our models, the lack of details in the information utilized by models, and shortcomings in our theoretical knowledge on which model are based. Similarly, the lack of sample size continues to be a serious problem in our estimations of extreme values. In some studies, many years of record are simulated and statements are made which might lead the reader to believe that a commensurate gain in information over the existing historical record has somehow been achieved. However, a detailed review of these studies reveals the obvious fact that alleatory uncertainty cannot be overcome by simply executing numerical models over many randomly posed simulations, since the information upon which the parent distribution is based remains the same.

Many available textbooks provide excellent formal reviews of statistics, derivations of distributions and applications specific to hazard/risk assessment [26.12–16]. Some recent examples of similar books which focus more on risk issues include [26.17, 18]. Here we will present only a summary of information which should allow a reader to gain sufficient familiarity with these basic statistical methods. As shown in Sect. 26.1, risk is associated with the probability of consequences. Since design thresholds are selected in a manner intended to limit consequences to an *acceptable* level, the study of natural hazards and associated risks tends to focus to either on extremely large or extremely small values in a distribution. Here we will focus on the upper end of the distribution, since these represent the major design consideration for most natural disasters in coastal areas.

We start here with the conventional definition of the probability density function (PDF) of a continuous variable x with the following properties

$$p(x) \geq 0 \,, \tag{26.1a}$$

$$\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1 \,, \tag{26.1b}$$

$$\int_{a}^{b} p(x)\,\mathrm{d}x = P(a < x < b) \,, \tag{26.1c}$$

where $p(x)$ is the probability density function for the variable x and $P(a < x < b)$ is the probability of x falling between the values of a and b. It is clear from the form of (26.1c) that the PDF is not a dimensionless function. Since $\mathrm{d}x$ has the same units as x, the dimensions of $p(x)$ must be x^{-1}. As an example if we allow the PDF to be given by

$$p(x) = k\mathrm{e}^{-qx} \,, \tag{26.2}$$

k and q must both have dimensions of x^{-1}; otherwise the integral would not produce results with the correct dimensions. A simple example with $q = 1$, $a = 1$ and $b = 3$ yields a value for (26.1c) equal to $\mathrm{e}^{-1} - \mathrm{e}^{-3} = 0.318$.

The cumulative distribution function (CDF) of x is equal to the probability that a value is less than or equal to a particular value of x,

$$F(x) = \int_{-\infty}^{x} p(x)\,\mathrm{d}x \,, \tag{26.3}$$

where $F(x)$ is the CDF of x. $F(x)$ is a dimensionless function since it is the product of two functions with dimensions x and x^{-1}; however, in engineering applications, it is very important to recognize that it becomes implicitly related to time when we convert to expected mean recurrence intervals. This is because, even though the time between samples does not appear explicitly in $F(x)$, it enters into the CDF implicitly via the time interval between samples. When we convert (26.3) into a form for estimating the expected interval between exceedances of a particular value of x, we obtain a relationship of the form

$$T(x) = \frac{1}{\lambda(1-F(x))} , \tag{26.4}$$

where $T(x)$ is the mean return period (or interval), and λ is the sampling frequency. Even though (26.4) is often written without λ in the denominator, its omission can lead to a dimensional inconsistency. The units used to define λ (decade^{-1}, year^{-1}, month^{-1}, day^{-1}, etc.) provide the information which characterizes $T(x)$. In engineering, it is common practice to use years for $T(x)$, in which case the expected value annual probability of exceeding a particular value of x is given by $1/T(x)$. Thus, a 100-year mean return interval is associated with an event that has an annual exceedance probability (AEP) of 0.01.

Here we will use an example of some results taken from wave hindcasts (the reconstruction of wave conditions from simulations of past events) for a location in the southern Gulf of Mexico area to show the importance of considering the frequency term in our estimation of return periods. In this region, two different meteorological phenomena can generate large waves, tropical cyclones, and *Northers*. The latter type of wave generation occurs when a cold front crosses the Gulf of Mexico and produces very strong surface wind speeds behind it. For both hurricanes and Northers, the results are given for sites located in deep water and in nominal depths of 20 m. In Fig. 26.1, we plot $-\ln\{-\ln[F(x)]\}$ against x, using the 20 data points available in each specific sample. Points that plot along a straight line in such a graph are consistent with a Gumbel distribution.

From Fig. 26.1, it appears that Norther and hurricane hazards are roughly comparable in this region of the Gulf of Mexico; however, if we include the frequency parameter and estimate the mean return interval, we obtain Fig. 26.2. In this case, the hazards related to hurricanes are shown to be considerably larger than those related to the Northers.

Since the concept of a mean interval between occurrences (return period) is simpler for some people to grasp, it is often used instead of the AEP concept in applications; and for that reason we will use it in some of our figures and discussions. However, it is very important to recognize that a *100-year event* is not constrained to occur only once every 100 years. For example, we can define the probability of exceeding a given value of x_c as 1 minus the probability of that level not being exceeded in any year, in which case the probability of having no exceedances in n years is given by

$$\begin{aligned} &P(x < x_c | n \text{ years}) \\ &= 1 - P(x < x_c|\text{year } 1) \times P(x < x_c|\text{year } 2) \ldots \\ &\quad \times P(x < x_c|\text{year } n) \\ &= 1 - \prod_{i=1}^{n} P_i(x < x_c) = 1 - [1 - F(x_c)]^n , \end{aligned} \tag{26.5}$$

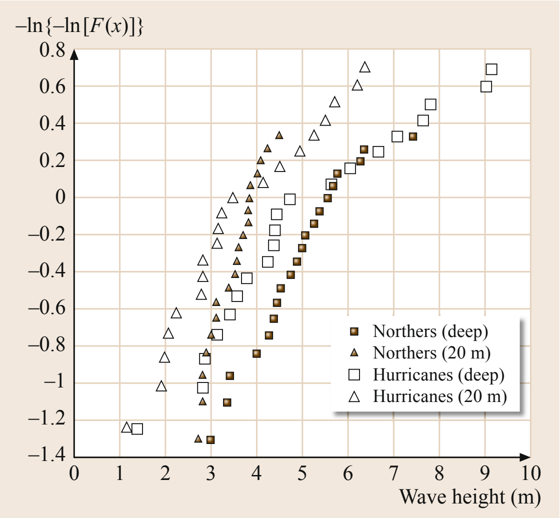

Fig. 26.1 Plot of the $-\ln\{-\ln[F(x)]\}$ versus x for waves generated at points in the southern Gulf of Mexico by hurricanes and *Northers* at sites in deep water and a depth of 20 m

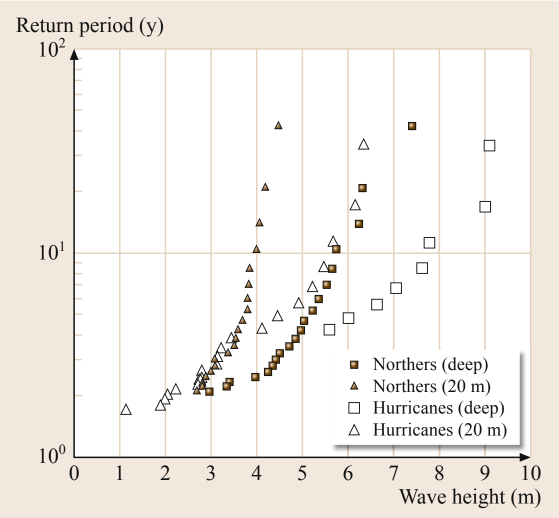

Fig. 26.2 Plot of the mean return period versus wave height for waves generated at points in the southern Gulf of Mexico by hurricanes and *Northers* at sites in deep water and a depth of 20 m

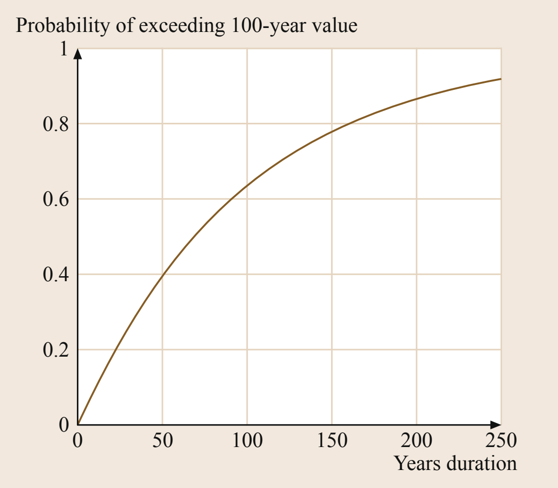

Fig. 26.3 Graph showing the probability of exceeding a value expected on the average only once every 100 years as a function of the number of years of duration

where n is the number of years and i is the year counter. Figure 26.3 gives the estimated encounter probability for an AEP of 0.01 as a function of the expected *design life*, the number of years being considered for the design. It can be seen that the actual exposure to hazards/risks is a clear function of the number of years considered for the design life. Over 30 years the probability of encountering a 100-year exceedance would be 0.2603; and over 50 years the probability of exceedance would be 0.3950.

Equation (26.5) is a bit tedious to use but is instructive regarding the reason why the probabilities become relatively large over a number of years. A second, perhaps simpler form for estimating the probability that a value, x_c, will not be exceeded in n years is to use a Poisson distribution with the integer value for occurrences equal to zero. In this case, if we define $\lambda =$ number of years over which the non-occurrence is being estimated divided by the mean recurrence interval for x_c, we can estimate the equivalent probability to (26.5) as $1-e^{-\lambda}$. For example, the equivalent value for the 30-year example above would be $\lambda = 0.3(= 30/100)$, or $1-e^{-0.3}$, which would equal 0.0259 and the equivalent value for the 50-year example would be $1-e^{-0.5}$, which would equal 0.3935. It should be noted that this would yield the same result for any combination of values of number of years over which the non-occurrence is being estimated and return value for x_c, such as 100 years of non-occurrence for a value with a 200-year recurrence interval for the latter example.

As an interesting side note to the above discussion, we can examine the common practice today of construction deterministic *flood zones* on maps for insurance purposes. If a property had an AEP for flooding of 0.009, it would lie outside of the primary flooding area covered under the program of the Federal Emergency Management Agency (FEMA) in the United States; yet the probability of flooding over 30 and 50 years would still be 0.2375 and 0.3637, respectively, which is only 8 to 9% different than for the property with the AEP of 0.01. Thus a definitive line on a flood map can be very misleading, with people on one side believing that their likelihood of flooding is much different than on the other side, which is certainly not the case.

Equation (26.5) shows that the accurate estimation of the CDF is fundamental to the estimation of hazards and risks. *Fuller* [26.19] credits George W. Rafter with initially pointing this out in 1896; however, the first published study directly linking hazards to a probability distribution in the United States (using the normal distribution) was that of [26.20]. Shortly afterward, *Hazen* [26.21] noted that logarithms of flood stages tended to provide more consistent estimates of CDF, essentially hypothesizing that floods followed a log-normal distribution. Since those early beginnings about a century ago, many studies have improved our understanding of the statistical behavior of extremes. Early studies typically attempted to fit specific distributions directly to empirically selected distributions. An excellent review of these early studies and the fitting methods used in them is contained in [26.22]. As shown there, many different distributions (for example, normal, lognormal, Poisson, generalized extreme value, and Pearson distributions) were developed, tested, and applied to a number of applications through the 1960s. The gist of the fitting methods was based on the concept that the CDF for each distribution could be related to the ranked data. For example, Table 26.1 shows the relationship between the estimated return period and the ranked order of the data, termed the plotting position, for a number of different distributions.

A detailed treatment of all of distributions used for estimating extremes is beyond the purpose of our discussion and Table 26.1 is only included to show how many different empirical approaches had been developed by the middle of the last century. For our purpose, it is sufficient to recognize that distributions can be classified into two general classes, parametric and non-parametric. Parametric distributions are characterized by assumed theoretical forms which include a small number of constants which must be determined by a statistical fitting technique; thus, such distributions constrain the number of degrees of freedom in the data. All of the distributions included in Table 26.1 are parametric. Nonparametric distributions are allowed to vary

Table 26.1 Plotting position formulae recommended by different investigators for the estimation of the return period

Name	Date	Formula[a] for T or $1/P(X \geq x)$
California	1923	$\frac{N}{m}$
Hazen	1930	$\frac{2N}{2m-1}$
Weibull/Gumbel	1939/1959	$\frac{N+1}{m}$
Beard[b]	1943	$\frac{1}{1-0.5^{1/N}}$
Chegodayev	1955	$\frac{N+0.4}{m-0.3}$
Blom	1958	$\frac{N+1/4}{m-3/8}$
Tukey	1962	$\frac{3N+1}{3m-1}$
Gringorten	1963	$\frac{N+0.12}{m-0.44}$

[a] $N =$ total number of items; $m =$ order number of the items arranged in descending magnitude, thus $m = 1$ for the largest item.
[b] This formula applies only to $m = 1$; other plotting positions are interpolated linearly between this and the value of 0.5 for the median event.

in a form consistent with the data elements; hence, they are not necessarily constrained to a small number of degrees of freedom.

Two different types of parametric distributions have been widely used in studies of natural phenomena, distributions which encompass the full range of data and distributions which focus only on the portion of the data with the largest observed values. Examples of coastal applications of the first type can be found in many studies of coastal waves in the 1960s through the 1980s, as discussed in [26.23, 24]. *Hogben* and *Lumb* [26.25] provide one of the earlier comprehensive collections of wave data around the world. Examples of subsequent studies using information covering the entire range of data measured can be found in [26.26–29]. However, here we will focus on the latter class of distributions, since extremes represent the primary variable range in most designs. This helps us avoid many of the issues with samples that have significant serial correlations, due to the shortness of the sampling interval. Most methods of the second type originally used a constant sampling interval (typically 1 year which implied a value of λ equal to one and the time units of years). However, as will be discussed in more detail in the historical perspective section, there are many cases where it is advantageous to allow the time interval between samples to vary. For example, if we are only interested in winds, waves and/or surges from tropical cyclones, the sampling time interval will vary and the λ in (26.4) would be equal to the mean frequency of tropical cyclone occurrence (or one over the mean interval between tropical cyclones). This parameter behaves analogously to the frequency term in the Poisson distribution and hence is often called the Poisson frequency, the average number of storms per year included in the analysis.

Before proceeding, it is important to note that wave extremes are usually posed in two different but related contexts: extremes related to the overall energy content in the local wave field (sometimes termed the sea state) and extremes related to individual waves. In the 1940s, it began to be recognized that the water surface was not well-described by simple unidirectional, monochromatic mathematical formulations. Instead, the wave field was better described as a random surface produced by the superposition of a continuous distribution of wave components in frequency and direction. The total energy within a unit area of water surface can be related to the variance of that surface around its mean position. The so-called *significant wave height* is approximately equal to the average of the one-third highest wave heights over an interval of time. More recently, it has been defined as $H_s = 4\sqrt{E_0}$, where H_s is the significant wave height and E_0 is the total variance of the water surface. The distribution of individual wave heights depends on the total energy and some additional factors which influence the degree of nonlinearity within the wave field, but these topics are beyond the main scope of what we will treat here. Significant wave heights tend to be used in applications dependent on time-averaged properties of waves (such as wave set-up, wave overtopping rates, beach erosion rates, etc.), while individual wave heights tend to be used in applications dependent on instantaneous forces (such as structural components in offshore facilities and some elements of ship design).

In nonparametric estimation of extremes, the probability of future events is obtained via methods that resample the available historical sample, under the assumption that future distributions will follow the same general form [26.30–34] This method is very powerful for application to extremes when the number of samples is large and the value range of interest falls within the range contained in the historical (initial) sample. In the context of coastal applications to extremes of waves, winds, and surges, probably the most widely used method of this kind is that of *Scheffner* et al. [26.35], which is termed the empirical simulation technique (EST). In this method, the CDF is estimated in terms of the rank of the ordered values (m) and the total number of observations N,

$$F(x) = \frac{m}{N+1} , \quad (26.6)$$

which can be seen as an equivalent form to the Gumbel plotting position given in Table 26.1; the return period there is defined as the inverse of the nonexceedance probability. However, the manner in which such information is used in nonparametric methods is very different to its use in parametric methods.

To help demonstrate the basic premise of resampling, let us assume that our data includes 30 samples taken over 50 years. Thus, the range of the CDF covered would be $1/31(0.0323)$ to $30/31(0.9677)$, or more generally $1/(N+1)$ to $N/(N+1)$, where N is the number of samples. Based on (26.4), we would estimate that the return periods explicitly covered by this data cover the range 1.72–51.67 years, recalling that the value of λ would be $30/5$ for this example. If a specific design requirement fell into this range, the results from the EST could be used directly, without any extrapolation. However, two key points should be understood in applications of this methodology. First, outside of the CDF range covered by the data, it is necessary to extrapolate parametrically to obtain estimates of the CDF as a function of x; so in this example, for return periods larger than 51.67 years or less than 1.72 years, some fitting algorithm similar to that used in parametric methods must be utilized. Second, any sample drawn from nature cannot be regarded as the *parent* population constructed from of all possible samples.

Let us assume for simplicity that we have a fitting methodology, which extends the range of the CDF from 0 to 1. In this case, we can use a random number generator to construct a random sequence of values for both the number of storms in a particular year (based on a Poisson distribution) and the value of the variable of interest in all of the storms that occur. The latter of these is obtained by assuming that each random number is a CDF value and inverting this value (using the defined CDF-x relationship) to obtain an estimate of x. If we simulated 100 sets of 50 years using this method, we could obtain an estimate of the magnitude of the variability we might expect in a particular 50-year sample; however, it would still be assuming that the initial sample represented the population characteristics. In fact, the actual variability in a subsequent 50-year interval will include contributions from the potential deviation of the initial sample from the overall population sample plus the contribution of deviations due to sampling given an initial distribution.

26.3 Historical Perspective

It is important to recognize the role of thresholds in coastal hazards and their relevance to risk. For example, a flood over a particular value might overtop a levee, or a wave height greater than some limit might impact critical components of an offshore oil platform, and a toxin level over some threshold might endanger the environment. Any of these occurrences might have serious consequences on lives and property. Thus, the success or failure of engineering designs and decision making in general can be very sensitive to whether or not a particular threshold is exceeded. However, is it really this simple? Are we certain that any value lower than an exactly specified threshold (even by a minuscule amount) will not have serious consequences, while any value equal to or greater than the threshold will definitely have serious consequences? Given that we never have complete information about the processes and the probabilities of different combinations of environment conditions, it is intuitively obvious that such thresholds should be considered only as an approximation to a level at which consequences are initiated and our estimates of the probability of these thresholds being surpassed cannot be exact. Here we see that both alleatory and epistemic uncertainty affect our estimate of this threshold probability.

Traditionally, the role of conservatism is at least in part to compensate for uncertainties in our knowledge base, both alleatory and epistemic uncertainty, via some factor used to shift the design level in a manner that compensates for the lack of knowledge (i. e., an attempt to account for inaccuracies in our theories and models of nature and/or in our quantification of the probabilities of factors which influence these levels). In many engineering fields, it has been common practice to include the effects of such uncertainty through the application of a codified *safety factor* or some similar device; however, newer methods are beginning to evolve toward the consideration of risk in a more formal probabilistic sense.

Historically, two fundamentally different approaches have been used to estimate threshold values for the design of critical infrastructure: approaches based on the establishment of a deterministic estimate of an upper limit for a hazard level and approaches based on the probabilistic estimation of a design level associated with a selected value for an acceptable hazard/risk level. Our review here will show that the field of engineering has been shifting through the years from the first approach to the second and that this trend is continuing.

Examples of the deterministic upper-limit approach can be found in early concepts used for designs of harbor protection, primarily based on somewhat subjective extrapolations of visual observations [26.36]. They have continued to play a dominant role in the development of design specifications throughout much of the twentieth century, including the estimation of environmental conditions for siting nuclear power plants in coastal areas, the formulation of design parameters for large dams, and the design of large coastal structures for protection from tsunamis. These methods tacitly assumed that there was no epistemic uncertainty in their estimated maximum values; hence, the estimated deterministic levels could be interpreted as absolute upper limits which could not be exceeded (i. e., the annual exceedance probability was zero or, conversely, the average return period was infinite). As will be discussed subsequently, applications of ultimate limits assuming no uncertainty can be under-conservative compared to probabilistic design levels; however, functioning structures based on this earlier approach still exist in many areas of the world.

Examples of the second approach include design specifications for most site-specific coastal structures around the world since the middle of the last century, when the combination of measurements and predictive tools attained sufficient skill to permit such estimates. Initially, these estimates were based predominantly on historical storms; but with time, there has been a slow transition from designs based only on events occurring in the historical record at a given site toward designs that consider all events that could occur at that site. Early during the development of the probabilistic approach, the accepted hazard level was taken to be associated with an annual exceedance probability of 0.01 (a mean interval between exceedances of 100 years). As shown earlier in (26.5) and the discussion following that equation, such a probability level does not represent a very low risk for many critical designs or applications, since there is a 1 in 3 chance that such a value will be exceeded in about 40 years.

26.3.1 The Development of Deterministic Methods

Before modern risk concepts, analytical tools, and computer codes existed, it was still necessary to estimate design parameters for various engineering applications. For example, in the 1950s, nuclear power began to be harnessed for commercial applications; and the issue of plant location quickly became extremely important to the overall scientific and engineering communities, as well as the public at large. Initially, a major consideration was given to issues related to national defense; and although the importance of this issue faded with time, it influenced the siting of many of the earliest nuclear power plants. In addition to economic and national defense issues, these plants were expected to be located in areas where natural hazards would not adversely affect performance and safe operation over their expected lifetime. The list of potential hazards considered was quite comprehensive and included the effects of earthquakes (both direct such as ground motions and indirect such as the loss of cooling water); atmospheric phenomena (lightning, tornadoes, tropical cyclones, etc.); flooding due to coastal storms, tsunamis and precipitation; and wave action superimposed on top flooding events.

A well-documented example of a deterministic design approach can be found in the methodology for siting of nuclear power plants in US coastal areas affected by hurricane storm surges. In 1959, the US Army Corps of Engineers (USACE) contracted the National Weather Service (NWS) to develop a hypothetical hurricane that could be used to design hurricane protection projects along the Gulf and Atlantic coasts of the United States. At that time the NWS, as part of its National Hurricane Research Project, set out to define *the most severe storm that is considered reasonably characteristic of a region*, defining a storm with such characteristics as the *Standard Project Hurricane* (SPH) [26.37]. Since their storm definition was related to nuclear power plants, the estimated maximum value possible was consistent with the concept of a very low allowable risk. *Graham* and *Nunn*'s report [26.37] contains a description of the derivation of the storm parameters for SPH along US Gulf and East coasts.

It is clear that all project designs do not require a level of acceptable risk as low as that used for nuclear power plants; so in 1979, the NWS Technical Report 23 [26.38] redefined the SPH in a fashion that would permit it to be more broadly applied. In this report, the SPH was defined as:

> *a steady state hurricane having a severe combination of values of meteorological parameters that will give high sustained wind speeds reasonably characteristic of a given region.*

This removed the idea that the SPH was related to the *most severe storm* expected in a particular area. Since many coastal projects are examined on an economic basis over a fixed interval of time, this re-definition of the SPH offered more general guidance on the relation cost/benefits for a project. In fact, NWS goes on to say that this revised

> *concept of the SPH has been developed for Gulf and Atlantic coasts as a benchmark against which to judge the hazards for a particular community.*

Also in NWS 23 [26.37], the concept of a *probable maximum hurricane* (PMH) was introduced as:

> *a hypothetical steady-state hurricane having a combination of values of meteorological parameters that will give the highest sustained wind speed that can probably occur at a specified coastal location.*

It is clear from this definition that the PMH was intended to be an event that was much rarer than the SPH; but it is not clear that an objective definition can be gleaned from what is written in NWS 23. In its executive summary, NWS 23 states that the central pressure of the PMH *is simply the lowest sea-level pressure at the hurricane center*. Estimated central pressures for the PMH within the Gulf of Mexico had values ranging from about 887 mb at Port Isabel, Texas to about 891 in the vicinity of Apalachicola, Florida and then diminishing to about 885 at Fort Myers, Florida (Fig. 26.4). These are all values which are substantially lower than the lowest central pressure available in any observations within the Gulf of Mexico at that time and no methodology for their derivation is given. Thus, it is likely that a considerable amount of *expert judgment* was utilized in estimating these values. The treatments of the other parameters used to define the PMH (peripheral storm pressure p_0, radius to maximum winds R_{max}, forward speed of the storm V_f, storm heading θ_f) and the latitude of the eye of the storm φ_c) are all relatively straightforward; however, correlations among the parameters were not addressed.

Besides the apparent subjectivity in the SPH and PMH definitions, other problems with this deterministic approach have arisen. For example, the use of *sustained wind speed* as the fundamental parameter for specifying the combination of PMH characteristics implies that the highest sustained wind speed will always produce the highest storm surge and that other storm parameters are less significant to storm hazard levels and associated risk. However, *Irish* et al. [26.39, 40] have clearly shown that storm size can be equally important to the magnitude of the storm surges. Figure 26.5 from [26.39] shows that Hurricane Katrina, although a much weaker storm than Hurricane Camille at the time of landfall in Mississippi, produced much

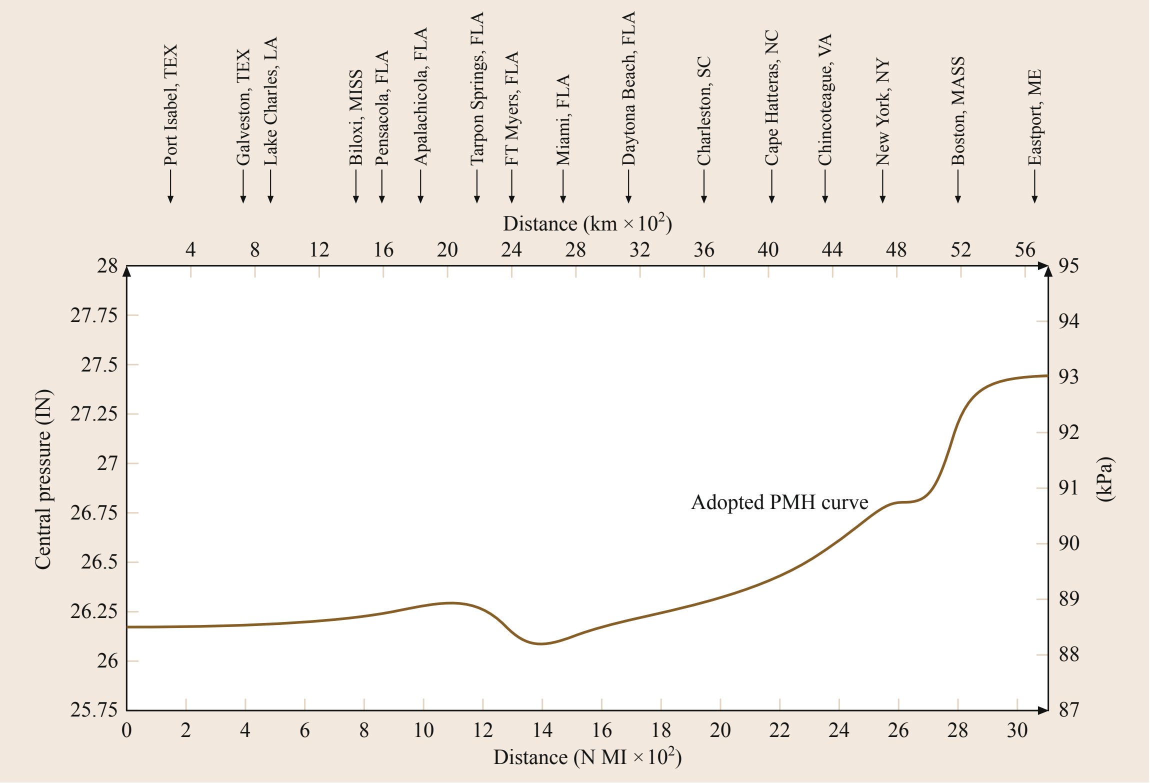

Fig. 26.4 Central pressure for PMH along US Gulf and East coasts (after [26.37])

higher surges there, consistent with results produced by idealized simulations. Additionally, storm approach direction can be important even along straight coastlines but can become exceptionally important on coasts with irregular shapes.

This particular example also affords a good opportunity to examine problems related to overly simplistic representations of the physics of processes in hazard/risk analyses. In the 1950s, the state of the art in surge prediction was quite crude and assumed that inland surges could be estimated via a two-step procedure, with the first step producing an estimate of the wind-driven surge levels at the coastline (neglecting the contributions of wave setup to these water levels) and the second step using hydrologic flow models to convey the water inland (neglecting both wind and wave contributions to inland surge levels). Subsequently, many studies demonstrated the need to estimate inundation in inland areas by including the effects of winds, topography, precipitation, and river discharge within an area [26.41–44]. Wind–wave contributions to coastal flooding have also been shown to be very significant in areas where the nearshore slope is high [26.41]. Based on all these studies, two primary results should be recognized. First, it is not possible to utilize a single factor to characterize waves and flooding in coastal and near-coast areas; and second, the older paradigm in which water levels were first estimated at the coast, using equations appropriate for surge prediction, and then allowed to propagate inland, using equations appropriate for hydrologic flow modeling, is inconsistent with today's state of the art.

There is an obvious analog between good computational models (Chap. 26) and good measurements in the estimation of extremes. No one would trust a design based on measurements known to have deficiencies leading to potentially large biases and random errors. In a similar vein, it is essential to obtain reliable, unbiased estimates from models of winds, waves, and surges for applications to hazard/risk assessment. Only after a modeling system is shown to be capable of generating accurate estimates is it worth expending the huge resources necessary to execute the number of computational simulations in such a system that would be required to provide reliable estimates of extremes.

26.3.2 The Development of Probabilistic Methods

During the latter half of the twentieth century two factors began to emerge which combined to allow significant progress in our capabilities for estimating coastal extremes: an improved theoretical foundation for the estimation of extremes and significantly increased quantities of data for analysis. The latter of these was associated with the development of automated measurement systems for waves and water levels in the 1960s and 1970s. For some purposes, such as daily operations at ports and coastal shipping routes, these direct measurements afforded an excellent source of quantitative information, since requirements for these purposes did not focus on extremes. For other applications, such as the design of breakwaters or the estimated recurrence of coastal flooding levels, the limited data collected from direct measurements proved to be insufficient, due to the short duration of available measurements.

An improved theoretical framework for extremes began to emerge when researchers started to question the application of empirical formulae to extremes based only on its *goodness-of-fit* to a given sample. Pioneers in this field such as Gumbel, Jenkinson, and Gringorten built upon earlier mathematical foundations of *Fisher* and *Tippett* [26.45], and *Gnedenko* [26.46] and began to develop improved methods for treating the behavior of the upper *tail* of the distribution of extremes. Following the work of these individuals, the concept of a generalized extreme value (GEV) distribution was formed. *Borgman* and *Resio* [26.47] offer an example of the application of GEV method to coastal hazards. More recent general reviews of the estimation of extremes from the sequences of random variables can be found in [26.16, 48].

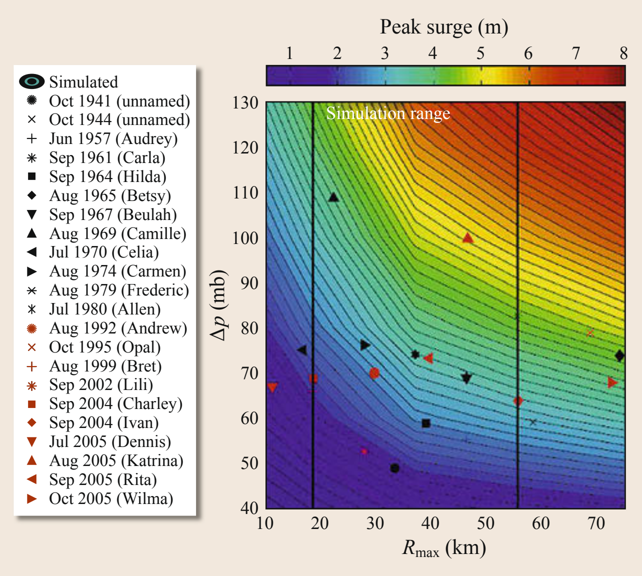

Fig. 26.5 Simulated peak surge as a function of hurricane size (R_{max}) and intensity (Δp) for a 1 : 10 000 bottom slope case. Historical R_{max} and Δp observations are superimposed on the numerical results to indicate peak surge potential of the historical storms (after [26.39], courtesy of American Meteorological Society)

The fundamental concept for the GEV is based on the analysis of asymptotic characteristics of probability functions. In this context, two CDFs are said to be of the same type if there exists constants $a > 0$ and b such that $F_1(ax + b) = F_2(x)$, where the subscripts denote the different CDFs. It was shown by *Fisher* and *Tippett* [26.45] that, excluding certain improper distributions characterized by discontinuities in the CDF, there are only three possible limiting distribution types which meet this constraint. *Jenkinson* [26.49] showed that all three of these asymptotic distributions could be written in a common form

$$F(x) = \exp - \left[1 + \varepsilon \left(\frac{(x-c)}{d} \right)^{\frac{-1}{\varepsilon}} \right] , \tag{26.7}$$

where c, d, and ε are parameters of the distribution. For the case of $\varepsilon = 0$, this equation is defined over a range of x from $-\infty$ to $+\infty$ and reduces to the two-parameter Gumbel distribution, sometimes termed the Fisher–Tippett Type I distribution.

$$F(x) = \exp[-\exp(-y)], \quad \text{where } y = \frac{x-c}{d} . \tag{26.8}$$

For $\varepsilon > 0$, the resulting distribution is a Frechet (or Fisher–Tippett Type II) distribution, which is bounded at the low end of the distribution but is unbounded for large x; and for $\varepsilon < 0$, the resulting distribution is a Weibull (or Fisher–Tippett Type III) distribution. The Weibull distribution has an explicit upper limit, but is unbounded in the negative direction.

Returning to Figs. 26.1 and 26.2, we see why a simple plot of x versus $-\ln\{-\ln[F(x)]\}$ will form a straight line if the sample follows a Gumbel distribution. If the points tend to curve upward on such a plot, the distribution is of the Fisher–Tippett Type III form; and if the points tend to curve downward, the distribution is of the Fisher–Tippet Type II form, which predicts larger values of x than the Gumbel distribution for a given value of the CDF.

Shortly after applications with GEV distributions began to become widespread, a second distribution with a somewhat different theoretical basis was introduced. This distribution, the generalized Pareto distribution (GPD), was first published by *Pickands* [26.50] and has been further developed by many subsequent researchers, with some of the more topical for our purposes being *Resnick* [26.14] and *Davison* and *Smith* [26.51]. Methods for fitting the GPD can be found in articles such as [26.52] and are now available in many statistical packages. The GPD can be written in the form

$$F(x|x_c) = 1 - \left(1 + \varepsilon' \frac{x_c}{\mu'} \right)^{\frac{-1}{\varepsilon'}} , \tag{26.9}$$

which has three parameters (ε', μ', and x_c), similar to the GEV, and also similar to GEV; the GPD has three basic forms depending on ε. With $\varepsilon' = 0$, the GPD approaches an exponential form given by

$$F(x|x_c) = 1 - \exp\left(-\frac{x}{\mu'} \right) . \tag{26.10}$$

Since the exponential and double exponential (Gumbel) distributions have been shown to converge for mean recurrence intervals greater than 7 years (based on annual data), significant deviations between the GPD and GEV for the case of $\varepsilon = \varepsilon' = 0$ tend toward zero for estimates of values with mean recurrence intervals greater than 7. When ε' is positive, the GPD will be similar to a Fisher–Tippett Type II and when ε' is negative, the GPD will be similar to a Fisher–Tippett Type III.

It is sometimes argued that the GPD is superior to the GEV, due to its formal incorporation of a threshold into its derivation; however, the essence of this distinction lies in the assumption that a sample of values from nature represents a *homogeneous population* in a statistical sense. As will be shown later in this chapter in Sect. 26.3.4, samples in nature often contain a mixture of populations. For example, large surges and waves can be generated along coasts by both extra-tropical and tropical storms. In smaller basins, squall lines and thunderstorms can become dominant wave and surge producers. Wherever possible it is advisable to try to avoid mixing populations from clearly different sources, but sometimes the ability to recognize more subtle variations can be obscured. A good example of this can be found in the work of *Resio* [26.53], who showed that what appeared to be a Fisher–Tippett II distribution for hindcast wave heights in the Cape Hatteras area was actually the sum of different types of storms definable in terms of how far off the coast they were when they passed Cape Hatteras. Estimated values within the record length are usually not affected by this mixed population problem however, it can become very significant when extrapolating to larger values.

Given the previous discussion, and hopefully without unduly disparaging the excellent theoretical work which has been accomplished in the field of extreme statistics, it should be recognized that given the nuances in the different theories upon which the GEV and GPD are based, it is difficult to argue definitively that one class of distribution is superior in applications to the other. Both distributions represent either a two-parameter distribution (for the limiting case $\varepsilon = \varepsilon' = 0$)

or a three-parameter distribution (for the case that either $\varepsilon \neq 0$ or $\varepsilon' \neq 0$). As noted here, samples of coastal winds, waves, and surges often contain very significant effects related to mixed populations, such as the obvious differences in statistical populations in surges generated by tropical cyclones making landfall some distance away from a site versus surges generated by a direct tropical-cyclone hits. These effects typically far outweigh any differences in theoretical justifications for using a GEV versus a GPD analysis. Because of this, it is probably better to develop a good level of familiarity with a tool of preference and use it appropriately (either the GEV or the GPD) rather than to be able to apply both distributions only marginally.

Returning to the problem of record lengths that are short relative to the mean recurrence interval of the design event, it is necessary to quantify the impact of randomness in the samples on the estimated values before we can assess the importance of this issue. As mentioned at the end of the last section, the quality of individual samples is an important factor in the estimation of extremes. Once routine measurements began to become available, there was an implicit expectation on the part of some that measurements would replace model estimates as the primary source of information used in the estimation of extremes. However, two sources of uncertainty can affect estimated extreme values: epistemic uncertainty in the sampled data (for example, theoretical unknowns and errors in model results) and alleatory uncertainty in the sampled data (natural randomness in the sample). Thus, there is a trade-off between using a short interval of very high quality measurements, removing most of the epistemic uncertainty, and using a long interval of lesser quality model results, which would reduce the alleatory uncertainty but would increase the epistemic uncertainty component.

Many studies in the 1960s and 1970s utilized a combination of theoretical formulations and computer simulations to establish guidelines for the standard deviations of estimates (sometimes referred to as confidence bands) around the value of x at a specific AEP. In most cases, the range of interest for design applications is beyond the range of observations. For example, the typical mean return interval was 100 years in most early design studies. However, in mapping flood zones for FEMA studies today, values up to the 500-year mean recurrence interval are required. For other applications, such as large offshore structures and coastal defenses in many areas of the world, the mean recurrence intervals of interest tend to be in the 1000s or even 10 000s of years, and probabilistic estimates for highly critical infrastructure such as nuclear power plants have been established to use mean recurrence intervals in the range of 1 000 000 to 10 000 000 years. Such estimates must consider the inherent uncertainties in extrapolations that are many times the length of the observational record. Estimates of standard deviations can be obtained by a number of methods based on the GEV or GPD forms [26.49, 54–56].

A useful form for the sampling uncertainty was given by *Gringorten* [26.57, 58]. Using a combination of theoretical characteristics and numerical testing, he showed that the expected root-mean-square (RMS) error of an estimated value as a function of mean recurrence interval in a Gumbel distribution could be estimated from the relationship

$$\sigma_T = \sigma\sqrt{\frac{1.1000y^2 + 1.1396y + 1}{N}}\,, \tag{26.11}$$

where σ is the distribution standard deviation; σ_T is the rms error at return period, T; N is the number of samples used to estimate the distribution parameters; and y is the reduced Gumbel variant given by $y = (x-\alpha)/\beta$; x is the value of the variable; and α and β are the parameters of the Gumbel distribution.

For a Gumbel distribution, the reduced variate and return period are related by

$$y = -\ln\left[\ln\left(\frac{T}{T-1}\right)\right]\,, \tag{26.12}$$

which for $T > 7$ years approaches an exponential form given by

$$\left(T - \frac{1}{2}\right) \to \mathrm{e}^{y}\,. \tag{26.13}$$

Equation (26.11) shows that the rms error at a fixed return period is related to the distribution standard deviation and the square root of a dimensionless factor involving the ratio of different powers of y (y^2, y^1, and y^0) to the number of samples used to define the parameters. In our previous description of the Gumbel distribution, we saw that the CDF for such a distribution has a double exponential form.

By the method of moments, the Gumbel parameters can be shown to be given by

$$\mu = c - \gamma d, \quad \sigma = \frac{\pi}{\sqrt{6}}d\,, \tag{26.14}$$

where γ is Euler's constant ($= 0.57721\ldots$), μ is the distribution mean, and σ is the distribution standard deviation.

Thus, the distribution standard deviation is related to the coefficient of variability for the distribution and can be used for estimating the expected width of the confidence bands for a specified return period. The dependence of the confidence limits as a function of the mean recurrence interval and the presence of the square root of the sample number in the denominator show that these limits will become very large when one is extrapolating to several times the record length. Figure 26.6 gives an example for return periods greater than 100 years given a distribution standard deviation of only 0.25 m, which is somewhat smaller than found in many natural data sets. This figure shows the importance of record length to alleatory uncertainty (due to natural variation, see Sect. 26.A). The estimation of a 10% exceedance or 1% exceedance above the Gumbel value at a specific return period can be obtained by multiplying the estimated value of σ_T by 1.08 and 2.33, respectively, since the assumption of normality is inherent in this estimation method.

Although (26.11) through (26.14) were initially derived for applications to annual maxima, they can be adapted to any time interval for data sampling in a straightforward manner. For the case of tropical cyclones, the average interval between storms (the inverse of the Poisson frequency used in the compound Gumbel–Poisson distribution) can be used to transform (26.11) into the form

$$\sigma_T' = \sigma\sqrt{\frac{1.1000y'^2 + 1.1396y' + 1}{N'}}\,, \tag{26.15}$$

where σ is the distribution standard deviation; σ_T' is the rms error at return period, $T' = T/\hat{T}$, where $\hat{T}$ is the average years between hurricanes, and N' is the number of samples used to estimate the distribution parameters $(N/\hat{T})$.

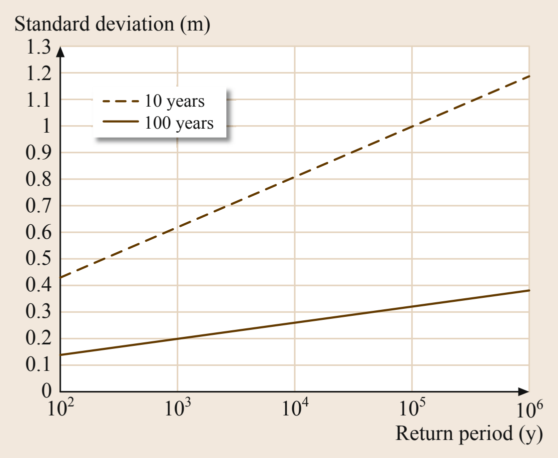

Fig. 26.6 Example standard deviation for confidence limits for a Gumbel distribution with a standard deviation of 0.25 m, shown here for sample lengths of 10 years and 100 years

Many decision-makers believe that the *best-fit* line for a distribution of extremes provides an unbiased estimate of the encounter probability of a value equal to or greater than a specific threshold; and this would be true if there were no alleatory or epistemic uncertainty in our sample. *Resio* et al. [26.44] have shown that uncertainty of both types can affect the expected encounter probabilities. In particular, their study shows that the estimated standard deviation provides a good means of quantifying the spread of the probabilities around the deterministic estimate; in this context, the probability of encountering a given surge value can be written in terms of an integral in two dimensions with a delta function to reduce it back to a single dimension

$$p(\eta) = \int_0^\infty \int_{-\infty}^\infty p[\hat{\eta}(T) + \varepsilon_\eta|\hat{x}]p(\hat{\eta})p(\varepsilon_\eta) \times \delta(\hat{\eta} + \varepsilon_\eta - \eta)\,\mathrm{d}\varepsilon_\eta\,\mathrm{d}\hat{\eta}(T)\,, \tag{26.16}$$

where $\hat{\eta}(T_r)$ denotes the deterministic estimate of η for a given return period and ε_η denotes the deviation from the deterministic surge estimate.

In (26.6), the estimate for $p(\varepsilon_\eta)$ is taken as a Gaussian distribution with the mean value at $\hat{\eta}$ and the standard deviation taken from the equation for estimating confidence bands, which will have the form

$$p(\varepsilon_\eta|\hat{\eta}) = \frac{1}{\sigma_T\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{\varepsilon_\eta}{\sigma_T}\right)}\,, \tag{26.17}$$

where σ_T is the standard deviation of the estimate at a given value of $\hat{\eta}$.

Equation (26.16) can be integrated directly to convert to a return period representation using estimates of population parameters and standard deviation as shown by *Resio* et al. [26.44]

$$T(\eta) = \frac{1}{1 - F(\eta)}\,, \tag{26.18}$$

with

$$F(\eta) = \int_0^\infty \int_{-\infty}^\infty p[\hat{\eta}(T) + \varepsilon_\eta|\hat{\eta}]p(\hat{\eta})p(\varepsilon_\eta) \times H(\hat{\eta} + \varepsilon_\eta - \eta)\,\mathrm{d}\varepsilon_\eta\,\mathrm{d}\hat{\eta}(T)\,, \tag{26.19}$$

where $H(g)$ is the heaviside function, equal to 1 if $(g) \geq 0$ and equal to 0 if $(g) < 0$.

To give an example of the impact of extrapolating many times the record length of the sample on the expected encounter probabilities, the *Resio* et al. [26.44] study analyzed central pressures from a 70-year interval for the west coast of Florida. Since surges are related more to tropical cyclones which make a landfall than to exiting tropical cyclones, they stratified their storm sample to include only storms with a general west to east motion in a latitude-longitude box with boundaries at 81°W and 85°W longitude and 25°N and 30°N latitude. Additionally, to eliminate very weak storms which might behave differently to well-organized storms, they further limited the storms included within their sample to those with central pressures less than 990 mb within the selected geographic area. Using the most recent reanalysis data available from NCDC (National Climate Data Center), it was determined that most storms before about 1940 did not report central pressures; consequently, the analysis was limited to the interval 1940–2009.

Figure 26.7 shows the results of this integration for three Florida cases compared to the original deterministic estimate from [26.44]. Case 1 is the deterministic solution for the return period based on the best-fit Gumbel distribution. Case 2 is a test of the numerical algorithm generated by representing the probability of the deviations as a delta function (i. e., all of the probability exactly on the deterministic line), which was computed as a test case for the integration algorithm. As expected, this case exhibited no significant deviations from case 1, so the case 1 and case 2 lines are identical. Case 3 shows the results using standard deviations equal to the estimated distribution standard deviations divided by 2, simply to provide an indication of how nonlinear the dependence is on the standard deviation in the confidence band when compared to case 4; case 4 shows the results for the estimated standard deviation. These results indicate that in this region of the Gulf of Mexico, the encounter probabilities of central pressure (a surrogate for tropical cyclone intensity) are markedly affected by sample uncertainty.

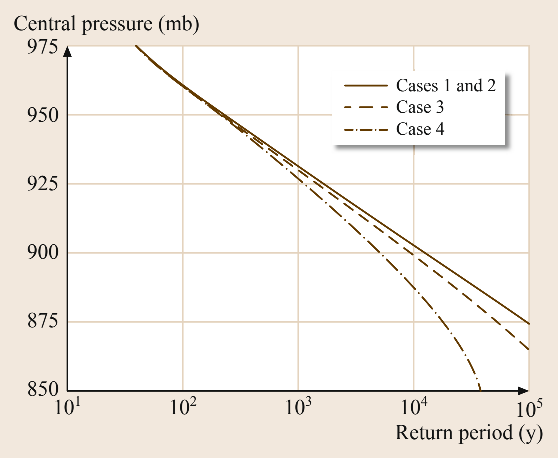

Fig. 26.7 Comparison of expected return periods for central pressures without uncertainty considered (cases 1 and 2) and with uncertainty considered (cases 3 and 4) for the West Florida region using data from 1940–2009

A curious aspect of these results is the increasing deviation between the lines as the return period increases. Whereas the 10 000-year value for the distribution which includes the full uncertainty (case 4) is slightly higher than the maximum possible intensity (MPI) value of 880 mb assumed here, the 40 000-year value has a much lower value (850 mb). This suggests that the role of uncertainty in very low probabilities is to act as a sort of filter for how long of a time period one can extrapolate before the uncertainty begins to become a dominant contributor to the distribution. Even at the 10 000-year value for the central pressure, the pressure differential has already been increased by about 20% over its value with uncertainty neglected. Since values for maximum possible surges in many areas along the west Florida coast are around 10 m, the impact of including this uncertainty would add about 2 m to the design surge levels over the deterministically estimated values.

26.3.3 The Development of the Historical Storm Method for Estimating Coastal Extremes

By the 1980s, it was widely recognized that record length was a significant factor in the confidence that could be placed on estimates of extremes. This led to a quandary of sorts. How could estimates of 100-year values be meaningfully estimated when observational data from measurements covered only relatively short time spans? In addition to the lack of long record lengths for waves and surges, careful review of the available measurements showed that data from large storms were often missing, further compromising the ability of measurements alone to provide a suitable base for the estimation of design values. The answer came via the development of improved physics-based models for predicting waves and surges from available meteorological information. Since meteorological information extended much farther back in time, it was hypothesized that reconstructions of winds and pressures for past events could be used to drive wave and surge models to obtain a surrogate for measurements during these storms. As noted previously, however, it is critical that the computational models used for this purpose are sufficiently accurate to provide accurate unbiased estimates of the extreme values.

This focus on storm hindcasts (the estimation waves and/or surges during past intervals of time) soon became the primary method used in many offshore designs [26.60] as well as in many coastal designs [26.61–64]. Over time, it began to be evident that the selection of storms for hindcasts of waves and surges often exhibited different CDF characteristics in different magnitude ranges. This led to the development of the points over threshold (POT) method for estimating extremes. In this method, only the sample values above some threshold would be included within the sample to be fit by the selected parametric form.

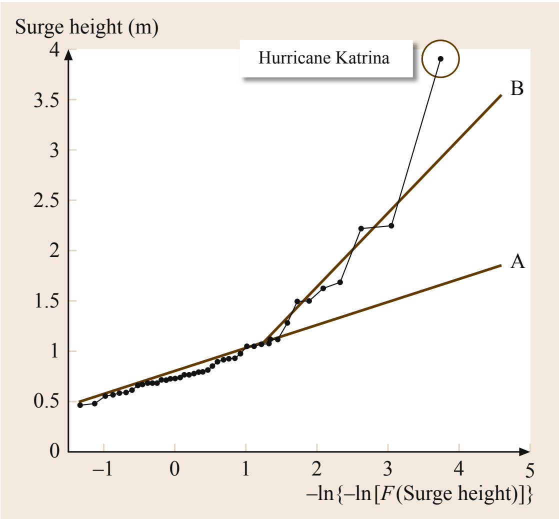

Fig. 26.8 Graph of a sample of hindcast surge heights as a function the double-negative logarithm of the CDF. The *line labeled A* represents storms which made landfall a significant distance away from the site, while the *line labeled B* represents hurricanes which made a landfall in the immediate vicinity of the site. The *single circled point* represents the surge at this site from Hurricane Katrina

Figure 26.8 is a good example of the situation faced by many planners and engineers when dealing with real-world data and helps to demonstrate the need for the POT approach in many situations. In this example, the data is actually more recent than the data that was used initially to justify the POT in earlier studies; however, the point is the same. The data shown here are simulated hindcast storm surges for a site in Lake Pontchartrain, Louisiana. A simplistic concept would be to use a three-parameter GEV (or GPD) distribution to fit the overall distribution. This would imply that Hurricane Katrina was only about a 65-year event in terms of the surges that it generated at this site; however, based on several independent analyses, it appeared that such a surge level would be expected considerably less frequently than this.

In Fig. 26.8, we have added two lines labeled A and B, representing a stratification of the surges at this site based on the proximity of landfall to the site. As can be readily seen and as expected from the physics of surge generation, there is a substantial difference between these two groups of events. This finding is consistent with the findings of researchers who had for years been advocating for the necessity of stratifying the sample before fitting the sample data to some parametric distribution. An interesting additional issue arises in Fig. 26.9. This relates to the issue of *outliers*. An outlier is an event that is believed to represent something

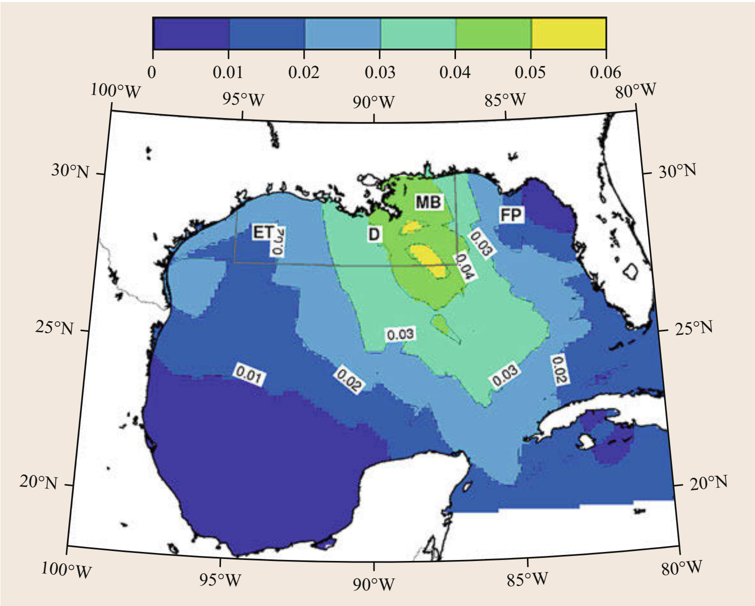

Fig. 26.9 Analysis of hurricane frequency within the Gulf of Mexico based on an analysis using an optimized spatial kernel (after [26.59]

that would only happen much less often than the time span of the sample record. It is intuitive that even the 1 000 000-year event must happen in some 10-year sample, so the historical treatment for these points was to consider them as not representative of the record length within which they were sampled. Such points were typically simply omitted from the distribution fitting process; however, it seems obvious that this omission for all points in a region neglects a potentially valuable piece of information from all of the analyses in an area.

26.3.4 The Development of Alternative Methods

The tragedy of Hurricane Katrina emphasized the point that reliable estimates of coastal hazards and risks should include the effects of all storms that can occur in an area and not just the events that occurred within a particular historical sample. However, as early as the 1970s, this had been recognized, and alternative methods started to be developed which afforded improved estimates of coastal hazards and risks. Probably the two most widely used of these methods have been the joint probability method (JPM), along with a derivative of the JPM termed the JPM-OS (joint probability method optimal sampling), and a second approach which, in this handbook, we will term the empirical track method (ETM). Material from a US Army Corps of Engineers White Paper [26.65] and a reduced version of the White Paper [26.66] will form a substantial portion of our discussions of these developments, and a considerable amount of details on the specifics can be found there.

The Development of the Joint Probability Method (JPM)

The JPM was developed for application to coastal surges in the 1970s [26.67, 68] and subsequently extended by a number of investigators [26.38, 69] in an attempt to circumvent problems related to limited historical records. In this approach, information characterizing a small set of storm parameters was analyzed from a relatively broad geographic area. In applications of this method in the 1970s and 1980s, the JPM assumed that storm characteristics were constant along the entire section of coast from which the sample was drawn. This assumption is inconsistent with more recent studies (Fig. 26.9) which show that storm frequencies vary substantially in this area.

The JPM used a set of parameters, including central pressure, radius of maximum wind speed, storm forward speed, storm landfall location, and the angle of the storm track relative to the coast, to generate parametric wind fields. Furthermore, initial applications of the JPM assumed that the values of these five parameters varied only slowly in storms approaching the coast; therefore, the values of these parameters at landfall could be used to estimate the surge at the coast, and these values could be treated as a constant during the approach of a storm to the coast. Recent data shows that this is not a good assumption [26.70, 71]; specifically, these studies suggest that tropical cyclones decay as they approach land. *Kimball* [26.72] has shown that such decay is consistent with the intrusion of dry air into a tropical cyclone during its approach to land, although other hypotheses for this consistent decay in intensity, such as the lack of energy production from parts of the tropical cyclone over land and increased drag in these areas have also been advanced. In any event, the evidence appears rather convincing that major tropical cyclones begin to decay substantially before they make landfall, rather than only after landfall as previously assumed.

The initial formulation of the JPM used computer simulations of straight-line tracks with constant parametric wind fields to define the maximum surge value for selected combinations of the basic five storm parameters. Each of these maximum values was associated with a probability

$$p(c_\mathrm{p}, R_\mathrm{max}, v_\mathrm{f}, \theta_\mathrm{l}, x)\ , \tag{26.20}$$

where c_p is the central pressure, R_max is the radius of maximum wind speed, v_f is the forward velocity of the storm, θ_l is the angle of the track relative to the coast at landfall, and x is the distance between the point of interest and the landfall location.

These probabilities were treated as discrete increments and the CDF was defined as

$$F(x) = \sum p_{ijklm} | x_{ijklm} < x\ , \tag{26.21}$$

where the subscripts denote the indices of the five parameters used to characterize the tropical cyclones.

The JPM represents a straightforward application of the *response surface* method for estimating extremes. In this method, the response of a particular variable (such as waves and surges) is derived as a function of several parameters, such as the set $(c_\mathrm{p}, R_\mathrm{max}, v_\mathrm{f}, \theta_\mathrm{l}, x)$ used in the above example. Given that the multivariate probability of the five-parameter space is known and a method for defining the response from this set of variables is known, the equation for the CDF for the response can be written for this example as

$$\begin{aligned} F(\eta) = \int\int\int\int\int & p(c_\mathrm{p}, R_\mathrm{max}, v_\mathrm{f}, \theta_\mathrm{l}, x) \\ & \times H[\eta - \Lambda(c_\mathrm{p}, R_\mathrm{max}, v_\mathrm{f}, \theta_\mathrm{l}, x)] \\ & \times \mathrm{d}c_\mathrm{p} \mathrm{d}R_\mathrm{max} \mathrm{d}v_\mathrm{f} \mathrm{d}\theta_\mathrm{l} \mathrm{d}x\ , \end{aligned} \tag{26.22}$$

where Λ represents the model used to convert the parameters to an estimate of maximum surge for each event as a function of $(c_p, R_{max}, v_f, \theta_l, x)$; however, it is easy to see that this approach can be generalized to any set of n parameters for which a response surface can be defined, along with the multivariate probability distribution. Variations on the methodology for evaluating the parameter probabilities and the storm set to be simulated have been presented by *Toro* et al. [26.59], *Niedoroda* et al. [26.73], *Resio* et al. [26.70], and *Irish* et al. [26.74]. Each has its advantages and disadvantages, and as long as the multivariate probability function is sufficiently resolved in the integration process, the results should be in reasonable agreement with each other.

Similar to the EST, this method is nonparametric with respect to the form of the CDF; however, the type of extrapolation used to extend the JPM result beyond the record length is very different to that used by the historical storm method. As shown by *Irish* et al. [26.75], the former of these is based on probability estimates which are much more smoothly varying than the local surge response inherent in the historical data, which means that when applied the JPM provides a significant decrease in sampling uncertainty, when compared to the historical storm method.

A potential advantage of the JPM over methods which depend heavily on historical storms is that the JPM attempts to consider all storms that might happen in an area; whereas, the EST considered only storms that did happen in that area. For example, the 100-year return period storm set to contain a full suite of $(c_p, R_{max}, v_f, \theta_l, x)$; whereas, the historical storm method will have just one or two storms available in this parameter space. Assuming that, for the purpose of surge generation, storm characteristics can be represented adequately by the set of parameters used, it is possible to construct a Katrina-like storm (high intensity combined with large size) even if one has not happened previously. Likewise, it is possible to interpolate between re-curved storms such as Hurricanes Opal and Wilma to understand probabilities of possible hurricane impacts in the Tampa, Florida area, even though neither of these storms produced significant surges in the Tampa area.

Perhaps the biggest controversy in JPM applications during the 1970s and 1980s centered on the sufficiency of this five-dimension set of parameters used in the joint-probability function to produce accurate wind fields. In addition to this concern, the lack of data on historical storms prior to 1950 made it very difficult to derive representative distributions, even for extended sections of coast. For example, information on storm size (R_{max}) was lacking for most historical storms; consequently, a statistical estimate of R_{max} (as a function of latitude and central pressure) was frequently substituted for actual values in the probability distribution. One wind field factor not considered in early JPM applications was the variable peakedness of tropical cyclone wind fields. This term is represented in terms of the Holland B parameter [26.76] in recent tropical cyclone wind models.

It is clear that the number of primary dimensions within the JPM must be capable of representing wind fields to sufficient accuracy so that they provide reasonable, relatively unbiased skill when used to drive coastal wave and surge models. For the case of extratropical storms, there is no known simple set of parameters that meets this criterion, and some extension of the EST or POT method may be the suitable choice for such applications, at least for recurrence intervals which are not too long (less than 100 years or so). For the case of tropical cyclones, dynamic models of tropical cyclone wind fields [26.77, 78] have been shown to capture a substantial portion of the wind field structure, when driven with the parameters listed above plus the Holland B parameter.

An argument against using the parametric wind fields in the JPM is that each historical tropical cyclone will tend to exhibit some degree of deviation from the theoretical parametric (planetary boundary layer) estimates. At any fixed time, such deviations could be produced by strong storm asymmetries, variations in R_{max} around the storm, enhanced spiral bands, etc. Hence, a *best-estimate* wind field crafted by experts to assimilate all the observations in a given tropical cyclone will typically represent the details of that particular storm much more faithfully than is possible via a parameterized theoretical model. Such wind fields today are produced primarily by groups like the National Oceanic and Atmospheric Administration (NOAA) Hurricane Research Division [26.79]. These wind fields are absolutely essential for advancing our understanding of tropical cyclone winds relative to wave and surge forcing in offshore and coastal areas. An excellent discussion of the reconstruction of historical wind fields and the impact of different approaches on model accuracies can be found in [26.80].

It is obvious that *best-estimate* wind fields contain an extremely large number of degrees of freedom in their formulation. Given the relatively small number of historical tropical cyclones, it is unlikely that we can understand/quantify the probabilistic nature of all the interrelated detailed factors that create these deviations. If these details were absolutely critical to coastal wave and surge estimation, we would be able to represent a past tropical cyclone very accurately but would know little about the probability of future tropical cyclones unless we retained the same number of degrees of free-

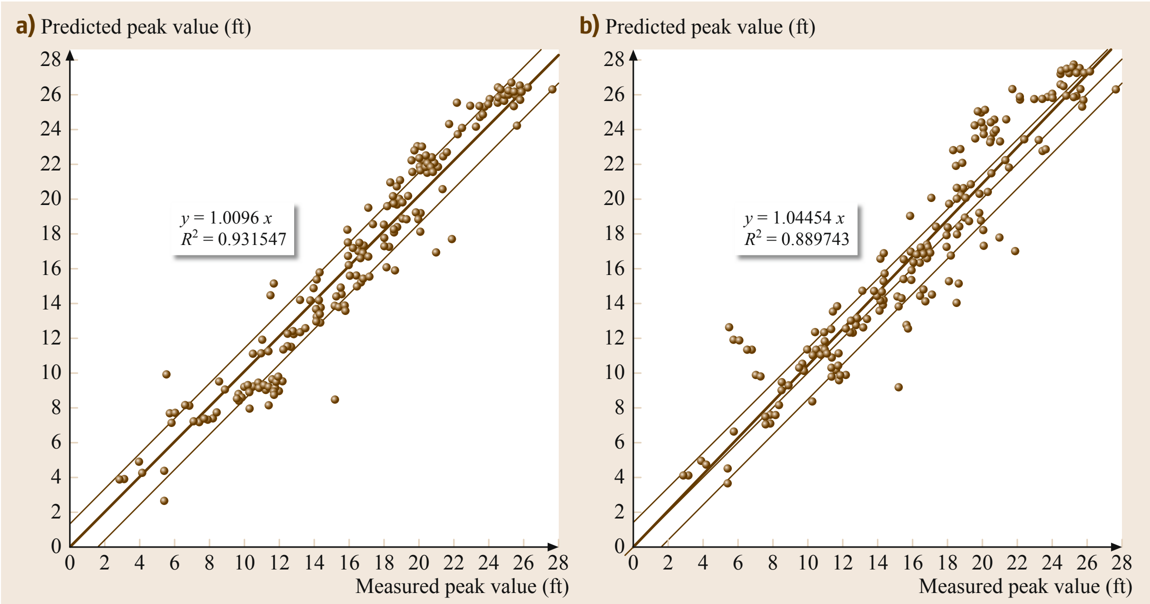

Fig. 26.10 **(a)** Comparison of observed high water marks (HWM, [26.81] for Hurricane Katrina and the computational simulation using *best-wind* wind fields. Points are the values at the recorded HWMs. Lines display a 1 : 1 correlation as well as 1.5 ft standard deviation on each side. **(b)** Comparison of observed HWM for Hurricane Katrina and the computational simulation using the best hand-analyzed post-event wind fields. *Points* are the values at the recorded HWMs. Lines display a 1 : 1 correlation as well as 1.5 ft standard deviation on each side

dom, including their expected variability in estimates of future storm surge and wave estimates. However, Figs. 26.10a and 26.10b show comparisons of hindcast storm surges using *best-estimate* wind and parametric wind fields in a hindcast of Hurricane Katrina. As can be seen in these figures, the differences in the surge model results are relatively small for this case.

Following Hurricane Katrina, improvements to the JPM were made to include additional physical factors affecting the storm and to improve the efficiency of the method [26.40, 66, 70, 74]. These newer methods have played a large role in the estimates of design levels for the new Hurricane Risk Reduction System for New Orleans. A more detailed discussion of these improvements can be found in [26.65].

Empirical Track Method (ETM)

Vickery et al. [26.78] presents a method for modeling hurricane risk in the United States. This method has been adopted for the development of design wind speed maps within the US (American National Standards Institute (ANSI), [26.82, 83]). The ETM uses a Monte Carlo approach to sample from empirically derived probability and joint probability distributions. The central pressure is modeled stochastically as a function of sea surface temperature along with storm heading, storm size, storm speed, and the Holland *B* parameter. This method has been validated for several regions along US coastlines and provides a rational means for examining hurricane wind risks associated with geographically distributed systems such as transmission lines and insurance portfolios.

A key requirement for the application of the ETM within its Monte Carlo framework is the ability to execute storms over many, many years (20 000 years in the application by *Vickery* et al. [26.78]). As such, a basin-scale study area with an average of three tropical cyclones occurring per year occur would require the simulation of 60 000 storms. Whereas this is not too demanding for the estimation of wind probabilities using an efficiently written planetary boundary layer (PBL) wind model, it is well beyond the range of current computer capacity for existing large, high-resolution ocean and coastal response models (wave models and surge models). For this reason, the ETM is typically not used for applications in which the computational burden for individual events is very high.

In its present form, the ETM is based on an autoregressive approach in which storm parameters at a subsequent time step are determined numerically from the same set of parameters at the current time step, utilizing a set of relationships derived from statistical and theoretical considerations. Although it is recognized that such predictions would contain very large statistical errors for a specific storm, particularly in terms of predictions of storm intensity and size, the predictions appear to approximate certain statistical characteristics of the storm population reasonably well, such as the mean number of storms in an area. What is not so clear is whether they capture the complex multivariate structure that is inherent in interrelationships among storm characteristics, such as the interaction of rates of change of these characteristics and the effect of land proximity. The effect of the omission of these higher order interactions on the statistics of predicted extremes remains a topic for future research.

26.3.5 Probabilistic Analyses of Extratropical Storms

As noted earlier, extratropical storms tend to be more frequent and affect larger coast regions than tropical cyclones due to the relatively large size of these storms. Based on these two considerations, it has been assumed that, for most coastal areas, the use of the historical storm method is adequate for quantifying hazards/risks. However, there are four reasons why this may not be a justifiable assumption given the geometric complexity of many coastal areas, particularly inland areas such as major bays and lakes:

- In some areas, the use of in situ observations to characterize the extremes may contain mixed populations, with a few observations from storms in which strong winds blow along the optimal surge/wave generating direction mixed with many other observations with winds from other directions. In such situations, the former events often have the appearance of outliers. Since a record length of 30–40 years may only contain a single interval of intense storms in which winds line up with the optimal generation direction, this creates a difficult situation for the analysis of extremes and a difficult situation for planners and engineers to assess.
- Although the GPD is specifically derived to fit a distribution above a particular threshold, it is still subject to the statistical constraint that all the samples are drawn from a homogeneous population. Like the GEV, the GPD is still only a three-parameter distribution, and if too many small events are mixed into the analysis, the upper tail of the distribution is very likely to be misestimated. Thus, in applications where researchers choose the lower limit to be given by an arbitrary (pre-selected) number of events, this can lead to a very poor estimation of the actual hazards/risks.
- Earlier, we noted that storm characteristics and frequencies seem to be significantly influenced by multidecadal variations in atmospheric circulation patterns. In these situations, it is exceptionally difficult to quantify the effects of the climatic variability and relatively short-term sample durations may not be adequate for accurately quantifying expected extremes. This is an especially important point relating to the difference between storm events that can occur and storm events that have occurred.
- Due to all of the issues raised in items 1–3 and inherent effects of uncertainty on encounter probabilities, the use of historical storm methods for very low probabilities, as is needed for certain critical infrastructure vulnerability assessment and design, is very difficult to justify without the inclusion of some means to add substantial conservatism into the analysis. Additional research developing a variation of the JPM to fit extratropical storms would be very valuable in meeting this need.

26.3.6 Future Directions and Final Comments

Irish and *Resio* [26.40] show that the effects of storm size, storm forward speed, landfall location, and storm track angle at the coast on storm surges all have asymptotic upper limits in tropical cyclones. Although these upper limits are site dependent, the forms of these limiters can be written in terms of some relatively simple functions. This information can help simplify the estimation of very low probability extremes; since in this very extreme range of surges, it reduces the surge probabilities back to a univariate distribution, in this case a function of a single parameter, storm central pressure. However, as emphasized previously, sampling uncertainty must still be considered in the estimation process. It is likely that, in many areas, the inclusion of sampling uncertainty within relatively small sample sizes will potentially make the estimates of central pressure for AEP values of 10^{-6} unrealistically low. In such situations, it will likely be necessary to combine probabilistic estimates with theoretically derived upper limits to be able to estimate coastal surges. Presently, this theoretical upper limit, the MPI, is still developed via relatively simplistic theoretical formulations combined with empirical envelopes of lowest observed central pressures around the world [26.84–86]. In their present form, this limit is estimated primarily as a function of sea surface temperature; however, more theoretical work will likely show the importance of additional terms to the estimation of the MPI. It should certainly be recognized that any estimate of an MPI will directly contain epistemic errors and indirectly contain alleatory errors due to the sampling distribution used to derive and calibrate theories.

26.4 Summary

On the one hand, we can see that considerable progress has been made in the estimation of hazards and associated risks in coastal areas over the last 100 years. Like all decision-making processes which become codified, it is difficult to change a methodology used for such estimates very quickly; hence engineering practice has often significantly lagged the recognized state of the art of our understanding of coastal hazards and risks, tending to wait until the occurrence of a natural disaster, such as Hurricane Katrina and the tsunami affecting Fukushima, Japan, to foster a willingness to change. Prudent design concepts in areas with critical infrastructure and communities that can be affected by hazards/risks must continue to consider the effects of both alleatory and epistemic uncertainty.

While the discussion herein focuses primarily on the coastal storm problem, the key concepts translate to the wide array of hazards impacting the coast. While each of these hazards will have its own unique set of physical characteristics and statistical challenges, the general concepts discussed here are appropriate for evaluating their probability and risk, as well as for understanding the factors contributing to uncertainty in these extreme estimates.

26.5 Nomenclature

a	specified limit or coefficient
b	specified limit or coefficient
c	coefficient
c_p	storm central pressure
d	coefficient
k	coefficient
n, m	indices
$p(x)$	probability density function
q	coefficient
v_f	storm forward speed
x	variable
x_c	coefficient
AEP	annual exceedance probability
CDF	cumulative distribution function
EST	empirical simulation technique
ETM	empirical track method
$F(x)$	cumulative distribution function
GEV	generalized extreme value distribution
GPD	generalized Pareto distribution
JPM	joint probability method
MPI	maximum possible intensity
N	total number of items
P	probability of exceedance
PDF	probability density function
PMH	probable maximum hurricane
POT	peaks over threshold
$Pr(A)$	probability of event A occurring
$Pr(B\|A)$	probability of a consequence (B) occurring given that event A occurred
R	risk
R_max	storm radius to maximum winds
SPH	standard project hurricane
T	return period (or return interval)
T'	dimensionless return period (or return interval)
ε	coefficient
ε'	coefficient
ε_η	deviation from deterministic surge estimate
η	surge value
θ_l	storm track angle relative to coast at landfall
λ	sampling frequency
μ	coefficient
y	reduced Gumbel variate
σ	distribution standard deviation
σ_T	rms error at return period T
σ'_T	dimensionless root-mean-square error at dimensionless return period T'

26.A Appendix: Glossary of Probability and Risk Terms

- *Confidence intervals*: A confidence interval is a statistical representation of how certain one is that a given variable will lie within a given range. A confidence interval includes two components – the *interval* (the value will be between x and y) and the level of *certainty* (a 90% confidence interval indicates one can say that there is a 90% probability the value *will* fall between the cited interval. As the interval gets smaller, the certainty that the values will fall within the interval goes down. Confidence intervals are related to uncertainty in that they are a statistical method of measuring uncertainty of a predicted value based on past data that provides the probability distribution and variance of the specific parameter of interest.

- *Error*: Error itself has multiple meanings. In the modeling/predictive sense, the concept of error requires that there be a *correct answer* against which to compare a predicted answer (modeling context) or sample result (statistical context). Error is then defined as the difference between the actual and predicted answer or measured sample. Therefore, in the sense of predictive risk assessment (looking forward in time), error is not a useful concept, as there is not yet any *actual* or *true* value against which to measure or compare. Error should not be confused with *mistake*, which is the result of an incorrect assumption, calculation or model formulation. The hallmark of a mistake is that it is avoidable.
- *Error* (*Types I and II*): Type I error and Type II error are technical terms used in statistics to describe particular types of erroneous results in a testing process. The terms relate to the acceptance or rejection of the hypothesis being tested (the null hypothesis). If the null hypothesis was rejected (found to be false) when it actually is true, then a Type I error occurred. Conversely, if the null hypothesis is not rejected (found to be true) when it actually is not true, a Type II error has occurred. These definitions correlate to the concepts of false positives (Type 1 error) and false negatives (Type II error) in testing. For a good technical tutorial on Type I and Type II errors [26.87].
- *Exposure*: The process of the receptor coming into contact with the hazard.
- *Exposure pathway*: The route by which the receptor is exposed. For animals, this is either by ingestion, inhalation, or dermal absorption.
- *Exposure assessment*: Quantitative analysis of how much (concentration, duration) of the hazard reaches the receptor.
- *Hazard/stressor*: An event (storm, accident), agent (chemical, radiation), situation or action with a potential for an undesirable consequence, such as harm to property, the environment, and human health or life. This term is often used synonymously with threat.
- *Probability*: In its most basic sense, probability is the chance or likelihood that something will happen. Qualitatively, the more likely an event is to happen, the more *probable* the event is. Quantitatively, probability is a value between 0 and 1, with 1 representing absolute certainty of the event occurring. The probability of an event is typically measured as the ratio of the number of times an event occurred over the total number of times the event could have occurred. For example, if we consider the event to be the occurrence of precipitation on any given day, we would collect information on whether it rained on a given day for a period of time (say 1 year). The probability of rain would then be

$$\text{probability of precipitation} = \frac{\text{number of days with precipitation}}{\text{total number of days}} .$$

- *Probability distribution*: The basic definition of probability works for discrete events, but when the event or parameter being evaluated is continuous in nature (the maximum temperature on any given day, the stage of a river on any given day), then a single probability is not sufficient. Instead, possible values are grouped into discrete ranges, and the number of occurrences in that range are counted then divided by the total number of measurements.
- *Receptor*: A receptor is the specific thing or entity being affected by the hazard/stressor. In a human health risk assessment – the receptor is a person.
- *Risk*: The potential for realization of unwanted, adverse consequences to human life, health, property, or the environment. The estimation of risk is usually quantified using the expected value of the conditional probability of the event occurring times the consequence of the event given that it has occurred. This definition is currently used by the Society of Risk Analysis. Mathematically, this is expressed as

$$R = P(A) * P(B|A) , \tag{26.23}$$

where $R =$ risk (a probability from 0 to 1), $P(A) =$ probability of event (A) occurring, and $P(B|A) =$ probability of a consequence (B) occurring given that event A occurred.

This technical definition is what will be used through this handbook. However, there are other definitions in other fields of which the knowledgeable practitioner should be aware. Table 26.2 presents a summary of other definitions used in various fields and a list of citations to which one may refer for more information. In each case, the same mathematical representation can be used, but there are assumptions made that result in the definition being slightly different. This is also presented in Table 26.2.

Besides having multiple potential definitions, risk can also be differentiated by *types*. Different types of risk, even if they are of the same quantitative value, are often managed differently or even ig-

Table 26.2 Risk definitions in different applications

Source	Risk definition	Mathematical representation	References
Occupational Safety and Health Administration (OSHA)	The probability that an adverse effect will occur	In this approach, the event is assumed to happen as the hazard is the item being regulated. For example, the absence of the railing in a factory catwalk would be the hazard. For OSHA purposes, if the hazard exists, then the probability of the accident becomes the risk. Mathematically, the probability of the even $[P(A)]$ is assumed to be 1.0. Therefore, the risk equation becomes: $R = 1 * P(B\|A) = P(B)$	[26.88]
Ecological and Human Health Risk – US Environmental Protection Agency (EPA)	Risk is the overall the probability of injury, disease, or death to humans or damage to the environment resulting from exposure to a chemical, stressor, or occurrence of a hazard. Risk has three components: a hazard, a receptor, and a method for the hazard to reach the receptor (involves both transport and exposure pathway)	The modification in this formulation is that the risk of the effect is not only dependent on the occurrence of A, but on the ability of the hazard to reach the receptor to cause the consequence. Hence, if the occurrence is a release of a toxic chemical, and the outcome of the release on a person is cancer (B), there will be no risk if there is no pathway (P) to the receptor. Mathematically, this means that the probability of the event becomes a conditional probability on the occurrence of the pathway. Therefore, the risk equation becomes: $R = P(A\|P) * P(B\|A)$	[26.89–95]
Medicine, psychology, health and social services	The potential for realization of unwanted adverse consequences or events	This approach, much like the OSHA approach assumes the hazard to be a reality (in this instance, an illness or negative condition such as poverty or exposure to domestic violence). Therefore, there is no probability of a conditional event – the illness/condition *is* the adverse outcome (i. e., $P(A) = 1$). This is also referred to as being *at risk* for some illness or negative outcome. Therefore, the risk equation becomes: $R = 1 * P(B\|A) = P(B)$	[26.96–103]
Nuclear Regulatory Commission [26.104]	Risk is the combined answer to three questions that consider (1) what can go wrong, (2) how likely it is, and (3) what its consequences might be	While posed as three questions, the first two are the definition of the probability of a hazard (e.g., what could go wrong is a pump failure (A), and the probability of that failure is by definition $P(A)$. The third question is the conditional probability of the negative consequence given the occurrence of A. Therefore, mathematically: $R = P(A) * P(B\|A)$	NRC Glossary (Web) [26.105]

Table 26.2 (continued)

Source	Risk definition	Mathematical representation	References
Financial Professions	Risk is the chance of loss with respect to person, liability, or the property of the insured or the possibility that the actual return on an investment will be different from its expected return. This includes such things as the danger or probability of loss to an insurer, the amount that an insurance company stands to lose, the variability of returns from an investment, and the chance of non-payment of a debt. Also often defined as the standard deviation of the return on total investment or the degree of uncertainty of return on an asset	In this formulation, the B (outcome) is defined as the deviation from an expectation. Note – this can be either positive or negative (return can exceed expectation), but the scenarios examined typically are only those which will produce a negative impact. If we assume the expected return of an investment to be x, the actual return realized to be x_a, and the difference to be defined as $D = (x_a, -x)$, then the risk financially if some event A occurs is: $R = P(A) * P(D\|A)$	[26.106–113]
United Nations	Risk is defined as the uncertainty that surrounds future events and outcomes. It is an expression of the likelihood and impact of an event with the potential to influence the achievement of the organization's objectives and goals. Risks in the United Nations context are normally referred to as programmatic and operational areas that have the greatest exposure to inefficiencies, ineffectiveness, fraud, waste, abuse, and mismanagement	In this formulation, the only difference is that the specific *hazard* is pre-defined as inefficiencies, ineffectiveness, fraud, waste, abuse, and mismanagement. Therefore, the mathematical formulation remains the same: $R = P(A) * P(B\|A)$	[26.114]
The ISO 31000 (2009)/ ISO Guide 73:2002	Risk is the effect of uncertainty on objectives. In this definition, uncertainties include events (which may or may not happen) and uncertainties caused by ambiguity or a lack of information. It also includes both negative and positive impacts on objectives. This definition was developed by an international committee representing over 30 countries and is based on the input of several thousand subject matter experts	This approach is closest to the financial professions in that risk is measured with respect to a desired objective (an expected outcome in financial terms). Using the same approach as financial risk, define the expected outcome as $\bar{o}$; the actual as o_a and the difference $D = (o_a - \bar{o})$, then the risk if some event A occurs is: $R = P(A) * P(D\|A)$ The only difference between this and financial risk is that the event A being analyzed explicitly may cause o_a to increase (a positive result)	ISO31000:2009 Risk Management Standard [26.115]

nored. Some of the common risk types that greatly influence both how these risks are assessed, managed and communicated are defined below.

- *Actual risk*: A scientifically verifiable risk. For example, it is well researched and documented that smoking places you at-risk for cancer [26.3–6]. Actual risk is sometimes referred to as *objective risk*, but whether risk is subjective or objective is related more to its ability to be measured than it is to its actual verifiability.
- *Perceived risk*: Risk that is thought to exist by an individual or group that is non-existent or exaggerated. This often occurs in situations where the public is misinformed or in which media reports instill unnecessary panic. Food safety concerns often top the list of such events and lead to the significant public policy debates, see, for example, [26.7].
- *Assumed risk*: Risk that is taken by choice. Assumed risk can be quantifiably large or small, and actual or perceived. For example, individuals who choose to partake in risky activities (skydiving, mountain climbing) choose to assume the relatively large risks associated with these activities, but choosing to drive a car, take medicines or be involved in day-to-day activities all involve some assumed risk.
- *Comparative risk*: Risk placed in context through comparison with another, perhaps better known risk. For example, stating that one is more likely to be hit by a meteor than to be injured in a plane crash compares a risk that people perceive as low (being hit by a meteor) with one they *perceive* as high but which actually is not. This helps place the risk in a conceptual framework.
- *Imposed risk*: Risk that is forced upon an individual, either without the knowledge of the individual or if known, without consent. For example, second-hand smoke exposure is seen as an imposed risk [26.8–10]. Natural events such as earthquakes, hurricanes. and extreme weather events are, to a large extent, *imposed* risks, but to some extent individuals assume that risk based on where they choose to live. For an interesting approach to this, the reader is referred to the work of *Parsad* [26.11].
- *Relative risk*: Risk of a particular outcome compared between two different groups or conditions. The relative risk is calculated as

$$\text{relative risk} = \frac{\text{risk under condition 1}}{\text{risk under condition 2}} .$$

  Thus relative risk proved a value of the risk for Condition 1 as a multiple of the risk for Condition 2. For example: suppose property bordering a coastline (Condition 1) has a 30% probability of being inundated during a storm surge, while properties over 100 m away from the coastline (Condition 2) have a 25% probability of inundation. The relative risk of inundation is, therefore, 1.2 times higher along the coastline (Condition 1).

  It is imperative that the underlying information about the actual baseline risk be given (i. e., 30% and 25% probability of inundation). Relative risk statistics where no baseline information is given can be very misleading. For example, if there is a 1 in 1 000 000 (10^{-6}) chance of an event at location A, and a 1 in 10 000 000 chance of the same event at location B, saying that location A has a 10 times greater risk than B belies the fact that the risk is still exceedingly small at location A.
- *Percent increased risk*: Risk of a particular outcome compared between two different groups or conditions measured as a relative difference of the two risks related to a base risk. Percent increases risk is calculated as

$$\begin{aligned}&\text{percent increased risk}\\&\quad = \frac{\begin{matrix}\text{risk under}\\ \text{Condition 1}\end{matrix} - \begin{matrix}\text{risk under}\\ \text{Condition 2}\end{matrix}}{\text{risk under Condition 2}} \times 100 .\end{aligned}$$

  Using our example from conditional risk, the percent risk increase for inundation by living on the coast would be

$$\begin{aligned}\text{percent increased risk} &= \frac{30-25}{25} \times 100\\ &= 20\%\ \text{increase} .\end{aligned}$$

  Just as with relative risk, it is imperative that the underlying information about the actual baseline risk be given (i. e., 30% and 25% probability of inundation). Percentage increased risk can be even more misleading than relative risk, especially where small risks are involved. For example, if there is a 1 in 1 000 000 (10^{-6}) chance of an event at location A, and a 1 in 10 000 000 chance of the same event at location B, the percent increased risk at location A would be 900%, which clearly presents a very different picture than saying that location A has a *one in a million* chance of the event.

- *Risk analysis*: The overall name given to the application of risk concepts to decision making. It

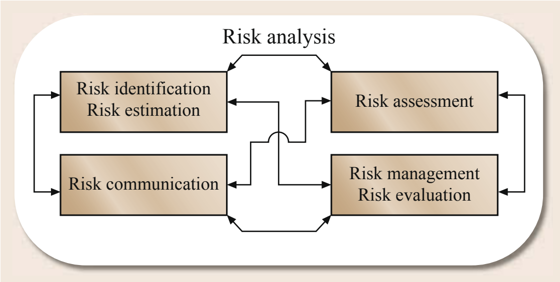

Fig. 26.11 Graphical representation of relationship of the elements of risk analysis

involves detailed examinations performed to understand the nature of unwanted, negative consequences to human life, health, property, or the environment. The process includes identification of potential events (scenarios), quantitative and/or qualitative assessment of risk, analysis of risk management alternatives, and communication of that risk to the necessary stakeholders; an analytical process to provide information regarding undesirable events; and the process of quantification of the probabilities and expected consequences for identified risks. Figure 26.11 gives a graphical representation of the iterative and highly interactive relationship of the various aspects of risk analysis.

- *Risk assessment*: The use of scientifically supported relationships to evaluate the magnitude and probability of adverse impacts on selected endpoints of specific actions, events or hazards/stressors. The assessment may be either qualitative or quantitative. An example of a qualitative risk assessment is the prediction of cancer based on decreased ozone in the atmosphere conducted by the World Meteorological Organization. The risk assessment predicted that if conditions did not change, there would be 50 million additional skin cancer cases due to sunburn by the year 2000 [26.116]. An example of a qualitative risk assessment concerning the same topic (cancer from sunburn) is that a person's risk for melanoma – the most serious form of skin cancer–doubles if he or she has had five or more sunburns [26.117].
- *Risk estimation*: The scientific determination of the characteristics of hazards/threats, usually in as quantitative a way as possible. This includes the magnitude, spatial scale, duration, and intensity of adverse consequences and their associated probabilities, as well as a description of the cause and effect links.
- *Risk evaluation*: A component of risk assessment in which judgments are made about the significance and acceptability of risk.
- *Risk identification*: Recognizing that a hazard exists and trying to define its characteristics. Often risks exist and are even measured for some time before their adverse consequences are recognized. In other cases, risk identification is a deliberate procedure to review, and it is hoped, to anticipate possible hazards.
- *Stochastic*: The property of having inherent random variation. The variation in a stochastic process, while random, is describable through probability theory (see *uncertainty* and *variability* below).
- *Threat*: See *hazard/stressor* above.
- *Uncertainty*: Uncertainty is a widely used term that is unfortunately often misused as a *catchall* term. In some instances, it is erroneously applied to the concept of confidence interval, which is actually a method of quantifying uncertainty. For the purpose of risk assessment, modeling, and/or prediction, uncertainty arises from three main components: error, variation, and lack of knowledge (see definitions herein and in the chapter text).
- *Variability* (*alleatory uncertainty*): Variability is a range of potential values for a given parameter. Variability is a natural characteristic of natural processes (also called *natural variation*). It is describable using probability distributions. The result of natural variation is sometimes called *alleatory uncertainty*. Variability in natural processes (and the resulting alleatory uncertainty) cannot be reduced, as it is an inherent property of the process itself.
- *Lack of knowledge* (*epistemic uncertainty*): The result of a lack of knowledge in risk assessments is also sometimes called epistemic uncertainty. A lack of knowledge can arise because the knowledge is not yet scientifically available – and as such, it can be reduced (along with the resulting epistemic uncertainty) going forward through additional data, experimentation, theoretical development, and scientific inquiry. However, lack of knowledge also includes things that we do not even know we do not know. This area of *lack of knowledge* is more difficult, because it is not easily identifiable and as such, becomes included in variability, or is called *error* when comparing model results to reality.

References

26.1 P.L. Bernstein: *Against the Gods: The Remarkable Story of Risk*, 3rd edn. (Wiley, New York 1996)

26.2 P.S. Laplace: *A Philosophical Essay on Probabilities* (Wiley, London 1902), translated by F.W. Truscott, F.L. Emory

26.3 S. Bišanović: The length of cigarette smoking is the principal risk factor for developing COPD, Int. J. Collab. Res. Intern. Med, Public Health **4**(1), 45–54 (2012)

26.4 H.G. Coleman, S. Bhat, B.T. Johnston, D. McManus, A.T. Gavin, L.J. Murray: Tobacco smoking increases the risk of high-grade dysplasia and cancer among patients with Barrett's Esophagus, Gastroenterology **142**(2), 233–240 (2011)

26.5 G.C. Kabat, M.Y. Kim, J. Wactawski-Wende, T.E. Rohan: Smoking and alcohol consumption in relation to risk of thyroid cancer in postmenopausal women, Cancer Epidemiol. **36**(4), 335–340 (2012)

26.6 G.C. Kabat, N. Shivappa, J.R. Hébert: Mentholated cigarettes and smoking-related cancers revisited: An ecologic examination, Regul. Toxicol. Pharmacol. **63**(1), 132–139 (2012)

26.7 O. Johansson-Stenman: Mad cows, terrorism and junk food: Should public policy reflect perceived or objective risks?, J. Health Econ. **27**(2), 234–248 (2007)

26.8 A. Besaratinia, G.P. Pfeifer: Second-hand smoke and human lung cancer, Lancet Oncol. **9**(7), 657–666 (2008)

26.9 L. Lazuras, A. Rodafinos, J.R. Eiser: Adolescents' support for smoke-free public settings: The roles of social norms and beliefs about exposure to secondhand smoke, J. Adolesc. Health **49**(1), 70–75 (2010)

26.10 J. Meadowcroft: Economic and political solutions to social problems: The case of second-hand smoke in enclosed public places, Rev. Political Econ. **23**(2), 233–248 (2011)

26.11 S. Parsad: *Planning for Human Settlements in Disaster Prone Areas* (Manglam Publ., Delhi 2004)

26.12 E.J. Gumbel: *Statistics of Extremes* (Columbia Univ. Press, New York 1958)

26.13 M.R. Leadbetter, S. Lindgren, H. Rootzen: Extremes and related properties of random sequences and processes, Z. Wahrsch. Geb. **65**, 291–306 (1983)

26.14 S. Resnick: *Extreme Values, Point Processes and Regular Variation* (Springer, New York 1987)

26.15 N. Cressie: *Statistics for Spatial Data* (Wiley, New York 1993)

26.16 S. Coles: *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics (Springer, London 2001)

26.17 T. Aven: *Foundations of Risk Analysis: A Knowledge and Decision-Oriented Perspective* (Wiley, Chichester 2003)

26.18 M. Rausand: *Risk Assessment: Theory, Methods, and Applications* (Wiley, Hoboken 2011)

26.19 W.E. Fuller: Flood flows, Trans. Am. Soc. Civ. Eng. **77**, 564–617 (1914)

26.20 R.E. Horton: Frequency of recurrence of Hudson River Floods, US Weather Bur. Bull. **2**, 109–112 (1913)

26.21 A. Hazen: Discussion on flood Flows, Trans. Am. Soc. Civil Eng. **77**, 626–632 (1914)

26.22 V.T. Chow: *Handbook of Applied Hydrology* (McGraw-Hill, New York 1964)

26.23 Y. Goda: Random Waves and Spectra. In: *Handbook of Coastal and Ocean Engineering*, Vol. 1, ed. by J.B. Herbich (Gulf Publ. Company, Houston 1990) pp. 175–212

26.24 Y. Goda: Distribution of Sea State Parameters and Data Fitting. In: *Handbook of Coastal and Ocean Engineering*, Vol. 1, ed. by J.B. Herbich (Gulf Publ. Company, Houston 1990) pp. 371–404

26.25 N. Hogben, F.E. Lumb: *Ocean Wave Statistics* (HMSO, London 1967)

26.26 J.A. Battjes: Long-term wave height distribution at seven stations around the British Isles, Dtsch. Hydr. Z. **25**(4), 179–189 (1972)

26.27 M. Isaacson, N.G. Mackenzie: Long-term distributions of ocean waves, J. Waterw. Port Coast. Ocean Div. **107**(2), 93–109 (1981)

26.28 C. Graham: The parameterization and prediction of wave height and wind speed persistence statistics for oil industry operational planning purposes, Coast. Eng. **6**(4), 303–329 (1982)

26.29 S. Kuwashima, N. Hogben: The estimation of wave height and wind speed persistence statistics from cumulative probability distributions, Coast. Eng. **9**(6), 563–590 (1986)

26.30 B. Efron: Bootstrap methods: Another look at the jackknife, Ann. Stat. **7**(1), 1–26 (1979)

26.31 B. Efron: Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods, Biometrika **68**(3), 589–599 (1981)

26.32 B. Efron: *The Jackknife, the Bootstrap, and Other Resampling Plans*, CBMS-NSF Monographs, Vol. 38 (Society of Industrial and Applied Mathematics, Philadelphia 1982)

26.33 B. Efron: Better bootstrap confidence intervals, J. Am. Stat. Assoc. **82**(397), 171–185 (1987)

26.34 C.F.J. Wu: Jackknife, bootstrap and other resampling methods in regression analysis (with discussions). Ann, Stat. **14**, 1261–1350 (1986)

26.35 N. Scheffner, L. Borgman, D. Mark: Empirical simulation technique applications to a tropical storm surge frequency analysis of the coast of Delaware, Proc. 3rd Int. Conf. Estuar. Coast. Model. (1993)

26.36 V. Harcourt: *Harbours and Docks*, Vol. 1 (Clarendon Press, Oxford 1895)

26.37 H.E. Graham, D.E. Nunn: *Meteorological considerations pertinent to the Standard Project Hurricane, Atlantic and Gulf Coasts of the United States*, National Hurricane Res. Proj. Rep. No. 33 (Weather Bureau, Washington 1959)

26.38 R.W. Schwerdt, F.P. Ho, R.R. Watkins: *Meteorological criteria for Standard Project Hurricane and*

Probable Maximum Hurricane Windfields, Gulf and East Coasts of the United States, Tech. Rep. NOAA-TR-NWS-23 (National Oceanic and Atmospheric Administration, Washington 1979)
26.39 J.L. Irish, D.T. Resio, J.J. Ratcliff: The influence of storm size on hurricane surge, J. Phys. Oceanogr. **38**(9), 2003–2013 (2008)
26.40 J.L. Irish, D.T. Resio: A Hydrodynamics-based surge scale for hurricanes, Ocean Eng. **37**(1), 69–81 (2010)
26.41 D.T. Resio, J.J. Westerink: Modeling the physics of hurricane storm surges, Phys. Today **61**, 33–38 (2008)
26.42 S. Bunya, J.C. Dietrich, J.J. Westerink, B.A. Ebersole, J.M. Smith, J.H. Atkinson, R. Jensen, D.T. Resio, R.A. Luettich, C. Dawson, V.J. Cardone, A.T. Cox, M.D. Powell, H.J. Westerink, H.J. Roberts: A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for Southern Louisiana and Mississippi. Part I: Model Development and Validation, Mon. Weather Rev. **138**, 345–377 (2010)
26.43 J.C. Dietrich, M. Zijlema, J.J. Westerink, L.H. Holthuijsen, C. Dawson, R.A. Luettich Jr., R.E. Jensen, J.M. Smith, G.S. Stelling, G.W. Stone: Modeling hurricane waves and storm surge using integrally-coupled, scalable computations, Coast. Eng. **58**(1), 45–65 (2010)
26.44 D.T. Resio, J.L. Irish, J.J. Westerink, N. Powell: The effect of uncertainty on estimates of hurricane surge hazards, Nat. Hazards **66**(3), 1443–1459 (2013)
26.45 R.A. Fisher, L.H.C. Tippett: Limiting forms of the frequency distribution of the largest or smallest member of a sample, Math. Proc. Camb. Philos. Soc. **24**(02), 180–190 (1928)
26.46 B.V. Gnedenko: Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math. **44**, 423–453 (1943)
26.47 L.E. Borgman, D.T. Resio: Extremal statistics in wave climatology. In: *Topics in Ocean Physics*, ed. by A. Osborne, P.M. Rizzoli (North-Holland, Amsterdam 1982) pp. 387–417
26.48 P. Embrechts, C. Klüppelberg, T. Mikosch: *Modelling Extremal Events for Insurance and Finance* (Springer, Berlin, Heidelberg 1997)
26.49 A.F. Jenkinson: The frequency distribution of the annual maximum (or minimum) values of meteorological elements, Quart. J. R. Meteorol. Soc. **81**, 158–171 (1955)
26.50 J. Pickands: Statistical inference using extreme order statistics, Ann. Stat. **3**(1), 119–131 (1975)
26.51 A.C. Davison, R.L. Smith: Models for exceedances over high thresholds (with discussion), J. R. Stat. Soc. B **52**(3), 393–442 (1990)
26.52 E. Castillo, A.S. Hadi: Fitting the generalized Pareto distribution to data, J. Am. Stat. Assoc. **92**, 1609–1620 (1997)
26.53 D.T. Resio: Some aspects of extreme wave prediction related to climatic variations, Proc. 10th Annu. Offshore Technol. Conf. (1978), OTC-3278-MS
26.54 N.R. Mann, R.E. Schafer, N.D. Singpurwalla: *Methods for Statistical Analysis of Reliability and Life Data* (Wiley, New York 1974)
26.55 M. Engelhardt, L.J. Bain: On prediction limits for samples from a Weibull or extreme-value distribution, Technometrics **24**, 147–150 (1982)
26.56 J.K. Patel: Prediction intervals – A Review, Commun. Stat. Theory Methods **18**(7), 2393–2465 (1989)
26.57 I.I. Gringorten: A simplified method of estimating extreme values from data samples, J. Appl. Meteorol. **2**, 82–89 (1962)
26.58 I.I. Gringorten: *Extreme Value Statistics in Meteorology – A Method of Application*, Air Force Surveys in Geophys. N. 125 (Air Force Cambridge Research Center, Bedford 1963)
26.59 G.R. Toro, D.T. Resio, D. Divoky, A.W. Niedoroda, C. Reed: Efficient joint probability methods for hurricane surge frequency analysis, Ocean Eng. **37**(1), 125–134 (2010)
26.60 V.J. Cardone, W.J. Pierson, E.G. Ward: Hindcasting the directional spectra of hurricane generated waves, J. Petrol. Technol. **28**, 385–394 (1976)
26.61 S.G. Coles, J.A. Tawn: Statistics of coastal flood prevention, Philos. Trans. R. Soc. Lond. A **332**, 457–476 (1990)
26.62 N.W. Scheffner, L.E. Borgman, D.J. Mark: Empirical simulation technique based storm surge frequency analysis, J. Waterw. Port Coast. **122**, 93–101 (1996)
26.63 E.F. Thompson, N.W. Scheffner: *Typhoon-Induced Stage-Frequency and Overtopping Relationships for the Commercial Port Road, Territory of Guam* (Coastal and Hydraulics Laboratory, US Army Corps of Engineers Engineer Research and Development Center, Vicksburg 2002), ERD/CHL TR-02-1
26.64 H.W. van den Brink, G.O. Konnen, J.D. Opsteegh, G.J. van Oldenborgh, G. Burgers: Improving 104-year surge level estimates using data of the ECMWF seasonal prediction system, Geophys, Res. Lett. **31**, L17210 (2004)
26.65 US Army Corps of Engineers: *Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System, Final Report of the Interagency Performance Evaluation Task Force. Vol. VII – The consequences* (US Army Corps of Engineers, Washington 2006)
26.66 US Army Corps of Engineers: *Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System, Final Report of the Interagency Performance Evaluation Task Force. Vol. VIII – Engineering and Operational Risk and Reliability Analysis* (National Oceanic and Atmospheric Administration, Washington 2009)
26.67 V.A. Myers: *Storm Tide Frequencies on the South Carolina Coast*, NOAA Tech. Rep. NWS-16 (National Oceanic and Atmospheric Administration, Washington 1975)
26.68 F.P. Ho, V.A. Myers: *Joint Probability Method of Tide Frequency Analysis applied to Apalachicola Bay and St. George Sound, Florida*, NOAA Tech.

Rep. NWS, Vol. 18 (National Oceanic and Atmospheric Administration, Washington 1975)
26.69 F.P. Ho, J.C. Su, K.L. Hanevich, R.J. Smith, F.P. Richards: *Hurricane Climatology for the Atlantic and Gulf Coasts of the United States*, NOAA Tech. Rep. NWS, Vol. 38 (National Oceanic and Atmospheric Administration, Washington 1987), completed under agreement EMW-84-E-1589 for FEMA
26.70 D.T. Resio, J.L. Irish, M.C. Cialone: A surge response function approach to coastal hazard assessment: Part 1: Basic Concepts, Nat. Hazards **51**(1), 163–182 (2009)
26.71 D.H. Levinson, P.J. Vickery, D.T. Resio: A review of the climatological characteristics of landfalling Gulf hurricanes for wind, wave, and surge hazard estimation, Ocean Eng. **37**(1), 13–25 (2010)
26.72 S.K. Kimball: A modeling study of hurricane landfalls in a dry environment, Mon. Weather Rev. **134**, 1901–1918 (2006)
26.73 A.W. Niedoroda, D.T. Resio, G.R. Toro, D. Divoky, H.S. Das, C.W. Reed: Analysis of the coastal Mississippi storm surge hazard, Ocean Eng. **37**, 82–90 (2010)
26.74 J.L. Irish, D.T. Resio, M.C. Cialone: A surge response function approach to coastal hazard assessment: Part 2: Quantification of spatial attributes of response functions, J. Nat. Hazards **51**(1), 183–205 (2009)
26.75 J.L. Irish, D.T. Resio, D. Divoky: Statistical properties of hurricane surge along a coast, J. Geophys. Res. **116**, C10007 (2011)
26.76 G.J. Holland: An analytic model of the wind and pressure profiles in hurricanes, Mon. Weather Rev. **108**, 1212–1218 (1980)
26.77 E.F. Thompson, V.J. Cardone: Practical modeling of hurricane surface wind fields, J. Waterw. Port Coast. Ocean Eng. **122**(4), 195–205 (1996)
26.78 P.J. Vickery, P.F. Skerjl, L.A. Twisdale: Simulation of hurricane risk in the U.S. using empirical track model, J. Struct. Eng. **126**(10), 1222–1237 (2000)
26.79 M.D. Powell, P.J. Vickery, T.A. Reinhold: Reduced drag coefficient for high wind speeds in tropical cyclones, Nature **422**, 279–283 (2003)
26.80 V.J. Cardone, A.T. Cox: Tropical cyclone wind field forcing for surge models: Critical issues and sensitivities, Nat. Hazards **51**, 29–47 (2009)
26.81 J.J. Westerink, Personal communication with Dr. D.T. Resio. (2007)
26.82 Minimum Design Loads for Buildings and Other Structures, ANSI A58.1 (American National Standards Institute, New York 1982)
26.83 Standard Minimum Design Loads for Buildings and Other Structures, Report No. ANSI/ASCE 7-95 (American Society of Civil Engineers, New York 1996)
26.84 K.A. Emanuel: An air-sea interaction theory for tropical cyclones. Part I: Steady-state maintenance, J. Atmos. Sci. **43**, 585–604 (1986)
26.85 G.J. Holland: The maximum potential intensity of of tropical cyclones, J. Atmos. Sci. **54**, 2519–2541 (1997)
26.86 H. Tonkin, G.J. Holland, N. Holbrook, A. Henderson-Sellers: An evaluation of thermodynamic estimates of climatological maximum potential tropical cyclone intensity, Mon. Weather Rev. **128**, 746–762 (2000)
26.87 M.C. Ortiz, L.A. Sarabia, M.S. Sánchez: Tutorial on evaluation of type I and type II errors in chemical analyses: From the analytical detection to authentication of products and process control, Anal. Chim. Acta **674**, 123–142 (2010)
26.88 OSHA: Guidance For Hazard Determination For Compliance With The OSHA Hazard Communication Standard. (29 CFR 1910.1200) http://www.osha.gov/dsg/hazcom/ghd053107.html (2014)
26.89 USEPA: Community-based air pollution projects glossary, http://www.epa.gov/airquality/communitybase/glossary.html (2015)
26.90 USEPA: Waste and Cleanup Risk Assessment Glossary, http://www.epa.gov/oswer/riskassessment/glossary.htm (2015)
26.91 USEPA: *Exposure Factors Handbook: 2011 Edition*, EPA/600/R-09/052F (National Center for Environmental Assessment, USEPA Office of Research and Development, Washington 2011)
26.92 USEPA: Environmental Management System Glossary, http://www.epa.gov/region4/ems/glossary.htm (2015)
26.93 USEPA: *Terms of Environment: Glossary, Abbreviations and Acronyms*, EPA 175-8-92-001. (Office of Communications, Education, and Public Affairs, Washington 1992)
26.94 USEPA: Regional vulnerability assessment (ReVA) program glossary, http://www.epa.gov/glossary.html (2015)
26.95 USEPA: Integrated risk information system (IRIS) glossary, http://www.epa.gov/iris/help_gloss.htm (2015)
26.96 USEPA: Radiation protection radiation glossary, http://www.epa.gov/radiation/glossary/index.html (2015)
26.97 USEPA: RadNet Glossary (last updated on February 13, 2012), http://www.epa.gov/radnet/radnet-glossary.html
26.98 A. Barker, J. Kamar, M. Graco, V. Lawlor, K. Hill: Adding value to the stratify falls risk assessment in acute hospitals, J. Adv. Nursing **67**(2), 450–457 (2011)
26.99 J.C. Fowler: Suicide risk assessment in clinical practice: Pragmatic guidelines for imperfect assessments, Psychotherapy **49**(1), 81–90 (2012)
26.100 A.S. Prentiss: Early recognition of pediatric venous thromboembolism: A risk-assessment tool, Am. J. Crit. Care **21**(3), 178–184 (2012)
26.101 N. Scurich, R.S. John: Prescriptive approaches to communicating the risk of violence in actuarial risk assessment, Psychol. Public Policy Law **18**(1), 50–78 (2012)
26.102 G. Côté, A.G. Crocker, T.L. Nicholls, M.C. Seto: Risk assessment instruments in clinical practice, Can. J. Psychiatry **57**(4), 238–244 (2012)
26.103 US Department of Health and Human Services: *Guidance for Industry Development and Use of*

Risk Minimization Action Plans (FDA Center for Drug Evaluation and Research (CDER), FDA Center for Biologics Evaluation and Research (CBER), Washington 2005), online at http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126830.pdf
26.104 Nuclear Regulatory Commission: Glossary, http://www.nrc.gov/reading-rm/basic-ref/glossary.html#R (2015)
26.105 Nuclear Regulatory Commission: Glossary, www.nrc.gov/readng-rm/basic-ref/glossary/html
26.106 USEPA: Environmental insurance and risk management tools glossary of terms, http://www.epa.gov/brownfields/insurance/ei_glossary_06.pdf (2004)
26.107 Farlex, Inc: The Farlex financial dictionary, http://financial-dictionary.thefreedictionary.com/ (2012)
26.108 J. Guinan: *Investopedia: The (I)Investopedia Guide To Wall Speak* (McGraw-Hill, New York 2009)
26.109 A. Chernobai, P. Jorion, F. Yu: The determinants of operational risk in US financial institutions, J. Financ. Quant. Anal. **46**(6), 1683–1725 (2011)
26.110 G. van de Venter, D. Michayluk, G. Davey: A longitudinal study of financial risk tolerance, J. Econ. Psychol. **33**(4), 794–800 (2012)
26.111 K. Watson Hankins: How do financial firms manage risk? Unraveling the interaction of financial and operational hedging, Manag. Sci. **57**(12), 2197–2212 (2012)
26.112 M. Drehmann, K. Nikolaou: Funding liquidity risk: Definition and measurement, J. Bank. Finance **37**(7), 2173–2183 (2013)
26.113 E. Vasile, I. Croitoru, D. Mitran: Risk management in the financial and accounting activity, Anul VII **3**(27), 13–24 (2012)
26.114 United Nations: Risk management, Office of Internal Oversight Services http://www.un.org/Depts/oios/pages/risk_management.html (2010)
26.115 ISO 31000: Risk Management Principles and Guidelines (International Organization for Standardization, Geneva 2009)
26.116 W.M. Organization: Scientific assessment of ozone depletion. In: *WMO Global Ozone Research and Monitoring Project*, Report No. 44, ed. by D.L. Albritton, P.J. Aucamp, G. Megie, R.T. Watson (World Meteorological Organization, Geneva 1998)
26.117 Skin Cancer Foundation: Facts about sunburn and cancer, http://www.skincancer.org/prevention/sunburn/factsabout-sunburn-and-skin-cancer (2013)