Journal of Math-for-Industry, Vol. 4 (2012A-6), pp. 41–48

Math-for-industry
Education & Research Hub

Towards statistical modeling of tsunami occurrence with regional frequency analysis

Jonathan R. M. Hosking

Received on January 17, 2012

Abstract. Regional frequency analysis is a statistical method for frequency estimation of extreme environmental events. Data for several sites are combined to improve the estimates of event frequencies at any one site. The computations are typically based on L-moments, which are summary statistics that have good properties of efficiency and robustness for describing data from heavy-tailed probability distributions. We summarize this work and apply it to a worldwide data set of historical records of tsunami magnitudes, obtaining estimates of the frequency distribution of tsunami runup height for essentially any location in the Pacific basin with exposure to tsunami events. The results have potential application to risk estimation and design of structures in tsunami-prone locations.

Keywords. extreme values, frequency estimation, L-moments, rare events, runup height

1. Introduction

Large tsunamis are enormously destructive events, and knowledge of how often they occur at any given site is of considerable importance for decisions about where construction should be permitted, the design of structures, and preparedness issues such as the identification of evacuation routes. This is one example of frequency estimation for extreme events, a problem that is common in natural hazard estimation. Often there is a need to estimate the return period of rare geophysical or meteorological events for a site or a group of sites. Typically the return periods of interest far exceed the available record length at any site, and statistical methods based on analysing historical event data from a single site are inaccurate and unreliable.

Regional frequency analysis seeks to overcome this problem by using data from several sites to estimate the frequency distribution of the observed data at each site. One method of regional frequency analysis seeks to identify a scaling relationship between frequency distributions at different sites, and estimates the rescaled frequency distribution by averaging summary statistics of the data from different sites. L-moments are summary statistics that are particularly suitable for this purpose. A full description of the approach is in [10]. This approach has been used in large-scale data analyses for precipitation [19, 2], streamflow [16, 11], and earthquake magnitudes [18].

This paper applies the regional frequency analysis methodology of [10] to historical data on tsunami observations from the NGDC/WDC Historical Tsunami Database [13]. The analysis in principle provides estimates of the frequency distribution of tsunami runup height for essentially any location in the Pacific basin with exposure to tsunami events. It should be emphasized that the results are not definitive: there are issues with data coverage and there is scope for alternative decisions at several stages of the analysis. Overall, however, the approach appears promising and worthy of further investigation and improvement.

The structure of the paper is as follows. Section 2 contains a brief introduction to L-moments; Section 3 contains a summary of the regional frequency analysis method described in [10]. Section 4 contains the data analysis, and Section 5 contains some concluding remarks.

2. L-moments

L-moments are summary statistics for probability distributions and data samples. They are based on linear combinations of order statistics. Denote by $X_{k:n}$ the kth smallest observation from a sample of size n, so that the ordered sample is $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$. The L-moments of a probability distribution are defined by

$$\lambda_1 = \mathrm{E}(X_{1:1}), \tag{1}$$
$$\lambda_2 = \tfrac{1}{2}\,\mathrm{E}(X_{2:2} - X_{1:2}), \tag{2}$$
$$\lambda_3 = \tfrac{1}{3}\,\mathrm{E}(X_{3:3} - 2X_{2:3} + X_{1:3}), \tag{3}$$
$$\lambda_4 = \tfrac{1}{4}\,\mathrm{E}(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}), \tag{4}$$

and in general

$$\lambda_r = r^{-1} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} \mathrm{E}(X_{r-j:r}). \tag{6}$$

L-moment ratios are the dimensionless quantities $\tau_r = \lambda_r/\lambda_2$, $r = 3, 4, \ldots$. They measure the shape of

This is an invited paper presented at the Forum "Math-for-Industry" 2011.

a distribution independently of its location and scale. The L-CV $\tau = \lambda_2/\lambda_1$ measures the distribution's dispersion relative to its mean, and is the L-moment analog of the coefficient of variation.

Given an ordered sample of data $x_{1:n}, \ldots, x_{n:n}$, an unbiased estimator of λ_r is the sample L-moment ℓ_r defined by

$$\ell_r = n^{-1} \sum_{j=1}^{n} w_{j:n}^{(r)} \, x_{j:n} \, . \tag{7}$$

The weight $w_{j:n}^{(r)}$ is, in the terminology of [15], the discrete Legendre polynomial $(-1)^r P_{r-1}(j-1, n-1)$. The weights can be computed as $w_{j:n}^{(1)} = 1$, $w_{j:n}^{(2)} = 2(j-1)/(n-1) - 1$, and recursively for $r \geq 3$ by

$$w_{j:n}^{(r)} = \frac{(2r-3)(2j-n-1)w_{j:n}^{(r-1)} - (r-2)(n+r-2)w_{j:n}^{(r-2)}}{(r-1)(n-r+1)} \, . \tag{8}$$

The sample L-CV is defined as $t = \ell_2/\ell_1$ and the sample L-moment ratios as $t_r = \ell_r/\ell_2$, $r = 3, 4, ...$; these estimators are unbiased only asymptotically as $n \to \infty$.

L-moments are analogous to ordinary moments, but have several advantages that make them suitable for use with the heavy-tailed distributions that often occur in geophysical and environmental data. Although moment ratios can be arbitrarily large, sample moment ratios have algebraic bounds; sample L-moment ratios can take any values that the corresponding population quantities can. L-moments are less sensitive than ordinary moments to outlying data values.

L-moments can be used to estimate parameters of distributions by equating sample and population L-moments. Typically, a frequency distribution is specified in terms of a set of parameters $\theta_1, \ldots, \theta_p$ and its L-moments can be expressed as functions of these parameters. By inverting this relation we can express the parameters in terms of the L-moments; when applied to the sample L-moments this yields estimators of the parameters. For example, the normal distribution is typically expressed in terms of its mean μ and standard deviation σ; its probability density function is $(2\pi\sigma^2)^{-1/2} \exp\{\frac{1}{2}(x-\mu)^2/\sigma^2\}$. Its first two L-moments are $\lambda_1 = \mu$, $\lambda_2 = \pi^{-1/2}\sigma$. Inverting this relation we have $\mu = \lambda_1$, $\sigma = \pi^{1/2}\lambda_2$; parameter estimates are then given by $\hat{\mu} = \ell_1$, $\hat{\sigma} = \pi^{1/2}\ell_2$. Estimators obtained by this procedure are generally consistent and in large samples have an asymptotically normal distribution. For some heavy-tailed distributions frequently used in extreme value analysis, estimation using L-moments is more accurate than maximum likelihood estimation in small and moderate samples.

Citations for proofs of the above properties are in [8].

3. Regional frequency analysis

Suppose that data are available at N sites, with site i having sample size n_i and observed data Q_{ij}, $j = 1, \ldots, n_i$. Let $Q_i(F)$, $0 < F < 1$, be the quantile function of the frequency distribution at site i. The key assumption is that the sites form a *homogeneous region*, i.e. that the frequency distributions of the N sites are identical apart from a site-specific scaling factor. We may then write

$$Q_i(F) = \mu_i q(F), \qquad i = 1, \ldots, N. \tag{9}$$

Here μ_i is the site-specific scaling factor, known in hydrology as the "index flood" following the usage in [4]. In this paper it will be termed the "index event". Typically it is chosen to be the mean of the site's frequency distribution, and is estimated by $\hat{\mu}_i = \bar{Q}_i$, the sample mean of the data at site i. Other location estimators such as the median or a quantile with specified exceedance probability could be used instead.

The remaining factor in (9), $q(F)$, is the *regional growth curve*, a dimensionless quantile function common to every site. It is the quantile function of the *regional frequency distribution*, the common distribution of the Q_{ij}/μ_i. The dimensionless rescaled data $q_{ij} = Q_{ij}/\hat{\mu}_i$, $j = 1, \ldots, n_i$, $i = 1, \ldots, N$, are the basis for estimating $q(F)$, $0 < F < 1$. It is usually assumed that the form of $q(F)$ is known apart from p undetermined parameters $\theta_1, \ldots, \theta_p$, so we write $q(F)$ as $q(F; \theta_1, \ldots, \theta_p)$. For example, these parameters may be the coefficient of variation and the skewness of the distribution, or the L-moment ratios τ_r defined above. The mean of the regional frequency distribution is not an unknown parameter: taking μ_i in (9) to be the mean of the frequency distribution at site i ensures that the regional frequency distribution has mean 1. The parameters are estimated separately at each site, the site-i estimate of θ_k being denoted by $\hat{\theta}_k^{(i)}$, and the estimates are combined to give regional estimates:

$$\hat{\theta}_k^{\mathrm{R}} = \sum_{i=1}^{N} n_i \hat{\theta}_k^{(i)} \Big/ \sum_{i=1}^{N} n_i \, . \tag{10}$$

Substituting these estimates into $q(F)$ gives the estimated regional growth curve $\hat{q}(F) = q(F; \hat{\theta}_1^{\mathrm{R}}, \ldots, \hat{\theta}_p^{\mathrm{R}})$. Quantile estimates at site i are obtained by combining the estimates of μ_i and $q(F)$:

$$\hat{Q}_i(F) = \hat{\mu}_i \hat{q}(F) \, . \tag{11}$$

When using regional frequency analysis with L-moments and the mean event magnitude as the index event, the sample L-moment ratios t and t_r, $r = 3, \ldots, p$, are computed for each site; we denote the values for site i by $t^{(i)}$ and $t_r^{(i)}$, $r = 3, \ldots, p$. The regional average L-moment ratios are then computed, as in (10), by

$$t^{\mathrm{R}} = \sum_{i=1}^{N} n_i t^{(i)} \Big/ \sum_{i=1}^{N} n_i \tag{12}$$

and

$$t_r^{\mathrm{R}} = \sum_{i=1}^{N} n_i t_r^{(i)} \Big/ \sum_{i=1}^{N} n_i \, , \quad r = 3, \ldots, p. \tag{13}$$

The regional frequency distribution is chosen to have mean 1 (since the mean is the index event) and its

L-moments and L-moment ratios are estimated by $\hat{\lambda}_1 = 1$, $\hat{\lambda}_2 = t^{\mathrm{R}}$, and $\hat{\lambda}_r = t_r^{\mathrm{R}}$, $r = 3, \ldots, p$. These estimates of the distribution's L-moments and L-moment ratios are used to obtain estimates of its parameters, as described in Section 2.

In regional frequency analysis with L-moments when the index event is not the mean event magnitude, it is still convenient to compute the regional frequency distribution as in the previous paragraph. The regional frequency distribution must be scaled appropriately to obtain quantile estimates at individual sites. For example if the index event is the median event magnitude and is estimated at site i by m_i, the appropriate scaling is by $m_i/\hat{q}(1/2)$, the ratio of the site-i median to the median of the regional frequency distribution. Thus in this case expression (14) is modified to

$$\hat{Q}_i(F) = m_i \hat{q}(F)/\hat{q}(1/2)\,. \tag{14}$$

This form will be used in the analysis of tsunami data in Section 4 below.

The practical implementation of regional frequency analysis involves deciding whether a group of sites constitutes a homogeneous region and making an appropriate choice of a regional frequency distribution. These stages typically involve some subjective judgement, but objective support for these decisions can be obtained from statistical measures defined in [10, chaps. 3–5]. These include a heterogeneity measure that compares the between-site variation in the L-moment ratios with what would be expected for a homogeneous region, and goodness-of-fit statistics that compare the regional average value of the L-kurtosis t_4 with its distribution when the proposed region is homogeneous and has some specified regional frequency distribution.

4. Analysis of worldwide tsunami data

Application of regional frequency analysis to tsunami data is not straightforward, since there are few data sets that contain consistent records of tsunami observations over a long period at multiple locations. As a measure of tsunami magnitudes we use runup height, defined as "the maximum height of the water observed above a reference sea level" [14].

We obtained data from the NGDC/WDC Global Historical Tsunami Database [13]. This appears to be one of the most complete sources of tsunami data. It contains data for 857 events, and a total of 12850 runup observations. The introduction to this database [14] cautions that "the reporting of large or destructive earthquakes is not homogeneous in space or time, particularly for periods prior to the 1900s" and says that this "introduce[s] uncertainties in the earthquake and tsunami databases for events prior to the late 1800s". We therefore used data from the database only for events since 1900: there are 637 such events, with a total of 12008 runup observations.

Observations of runup height are available at many sites. There are issues with irregular observations, when different tsunami events are recorded at inconsistent sets of sites.

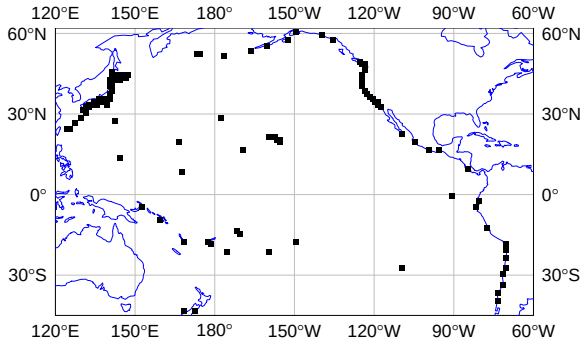

Figure 1: 1° grid squares with runup height observations for at least 10 tsunami events since 1900.

To overcome this, we combine adjacent sites into a "location" and use a single representative measurement for each tsunami event that affected the sites constituting a given location. Specifically, we combine all observations within each 1° square of latitude and longitude. For each event, we use the largest runup height recorded in each grid square. We judge that 10 is the smallest sample size that should be used; in smaller samples there may be serious biases in estimates of the L-moment ratios that are needed for regional frequency analysis of tsunami data. We use data for the 115 grid squares that have observations of at least 10 events. These locations are shown on Figure 1. They are all on or near the Pacific Ocean. No location on the Atlantic or Indian Oceans had sufficient observations to reach the 10-event threshold, which unfortunately limits the global applicability of the results of this analysis.

Sample L-moment ratios for the grid-square runup data are shown in Figure 2. The values of L-skewness and L-kurtosis cover a wider range and many of them are large compared to other kinds of environmental data such as streamflow and precipitation. This reflects the heavy-tailed nature of distributions of runup height. Figure 3 shows a typical example, the runup heights for grid square 34°N, 119°W. The data are plotted on a scale on which an exponential distribution would plot as a straight line: the marked convexity of the plotted points indicates that the distribution has a much heavier than exponential upper tail.

The discordancy measure D of [10, sec. 3.2] is a useful indicator of locations with L-moment ratios that are markedly different from those of the other locations in a regional data set; values $D > 3$ may indicate that there are problems with a location's data. For the tsunami data the grid square 7°N, 151°E was flagged as discordant, with $D = 16.3$. The data for this square look odd: the range of the observations is from 0.03m to 0.37m, but 10 of the 15 measured runup heights are either 0.09m or 0.10m. Though it is not clear whether the data are incorrect, we thought it best to exclude this site from further analysis. After this location was excluded, the discordancy statistic was calculated for the other 114 locations. Five of these had D values greater than 3, but in each case the data showed no major irregularities, and none of the sites was excluded.

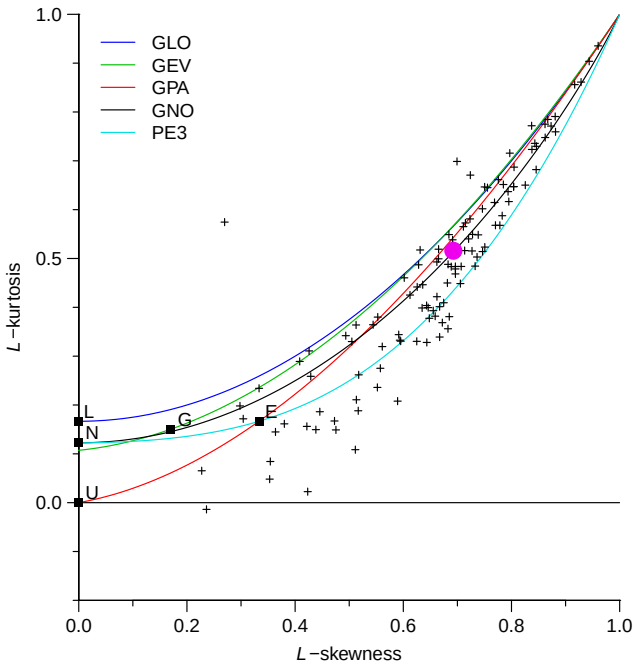

Figure 2: L-moment ratio diagram for tsunami data. "+" indicates sample L-moment ratios of tsunami data from different locations; the large dot is the average of these values. Squares show the L-skewness–L-kurtosis values for two-parameter distributions: exponential (E), Gumbel (G), logistic (L), normal (N), and uniform (U). Lines show the L-skewness–L-kurtosis relations for three-parameter distributions: generalized logistic (GLO), generalized extreme-value (GEV), generalized Pareto (GPA), generalized normal (GNO), and Pearson type III (PE3).

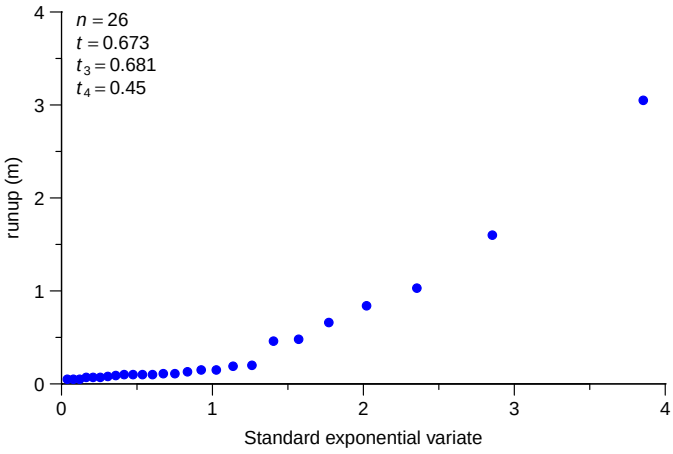

Figure 3: Runup heights for grid square 34°N, 119°W.

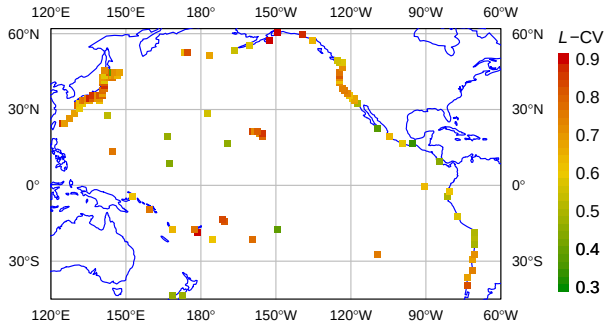

Figure 4: Tsunami data L-CV by grid square.

The heterogeneity measure H of [10, sec. 4.3] indicates whether the quantile functions for a set of locations satisfy the relation (9). It essentially measures the dispersion of the sample L-moment ratios at the different locations, relative to the amount of dispersion that would be expected if the locations constitute a homogeneous region. According to [10, sec. 4.3.3], a region should be regarded as "acceptably homogeneous" if $H < 1$, "possibly heterogeneous" if $1 \le H < 2$, and "definitely heterogeneous" if $H \ge 2$. For the 114 grid squares in the tsunami data, the computed value of H is 2.91, indicating that the locations as a whole are heterogeneous. It is therefore worthwhile to try to partition the locations into smaller sets that may be homogeneous, or at least less heterogeneous.

Identification of homogeneous regions within a regional data set is discussed in [10, sec. 4.1] and has been the subject of much research (e.g., [20, 3, 17, 12, 6]). Ideally it should be based on physical understanding. In this case, where we are looking for groups of sites that have a similar distribution of tsunami magnitude, we might consider, for example, the exposure of different segments of coastline to different kinds of tsunami events and the patterns of travel of tsunamis from their source locations to the coast. For the present work this information was not available and potential regions were identified based on distances between locations and on geographic patterns in the variation of the sample L-CV of runup heights. Sample L-CV is a natural quantity to use for this purpose, since in a homogeneous region all sites have the same L-CV and between-site variation in L-CV is a major contributor to heterogeneity. We acknowledge that the use of sample L-CV in both identification of regions and in the subsequent computation of the heterogeneity measure H compromises the integrity of H as a test of homogeneity [10, sec. 4.1.2]. However, in this case we feel that it is the best available approach, and we have tried to ensure that the regions that are identified have geographical coherence and plausibility rather than being unduly influenced by local variations in the sample L-CV statistics that may arise merely from the sampling variability of the statistics.

For the tsunami data the L-CV values for the data in each grid square are shown in Figure 4, with the area near Japan shown in greater detail in Figure 5. In the Japan area, L-CV values seem relatively high on the island of Honshu east of 135°E and relatively low on the island of Hokkaido and west of 135°E. There is also a group of high-L-CV sites southwest of 30°N, 135°E. We accordingly identify four potential regions, as shown in Figure 5. In the main body of the Pacific Ocean we identify three potential regions: the northwest, with relatively low L-CV; the

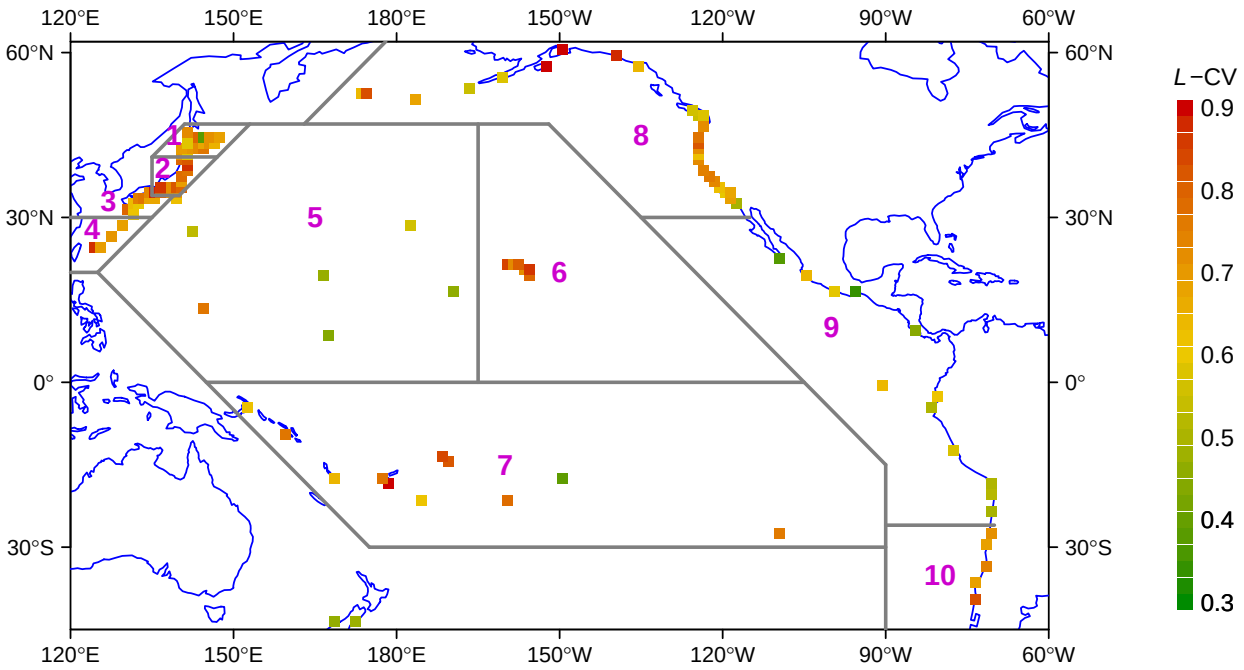

Figure 6: Tsunami region definitions.

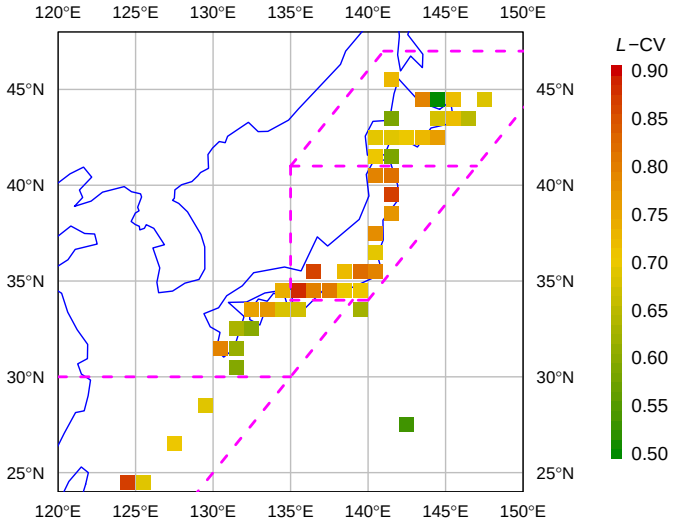

Figure 5: Tsunami region definitions for Japan.

Table 1: Heterogeneity measures for tsunami regions.

Region number	Region name	H
–	All data	2.63
1	Japan North	0.24
2	Japan Central	–1.49
3	Japan South	–1.36
4	E China Sea	–0.85
5	NW Pacific	0.42
6	Hawaii	–0.99
7	S Pacific	1.39
8	America North	1.19
9	America Central	0.23
10	America South	–1.25

Hawaii area with high L-CV; and the southern Pacific with mostly high L-CV. On the west coast of the Americas we identify three potential regions: a central region with relatively low L-CV and areas to the north and south with relatively high L-CV. All the regions are shown, with numerical labels, in Figure 6. The heterogeneity measure H was computed for each region; the values are given in Table 1. Regions 7 and 8 are classified as possibly heterogeneous; the others are acceptably homogeneous. Though it would be possible to refine the definition of Regions 7 and 8, we consider this set of regions to be adequate as a first attempt, and we use them in the subsequent analysis.

The regional average L-skewness and L-kurtosis of the ten regions are shown on Figure 7. All the regions lie close to the plotted lines for the generalized Pareto (GPA) and generalized normal (GNO) distributions. This is largely confirmed by the goodness-of-fit measure of [10, sec. 5.2], which deems the generalized Pareto and generalized normal fits acceptable for 7 and 6 regions respectively. We prefer to fit a generalized Pareto distribution, and we impose a lower bound of zero on the fitted distribution: this accords well with the nature of runup height observations, which are positive but with many values near zero. Each region is therefore fitted by the generalized Pareto distribution, with quantile function

$$q(F) = \alpha\{1 - (1 - F)^k\}/k\,. \qquad (15)$$

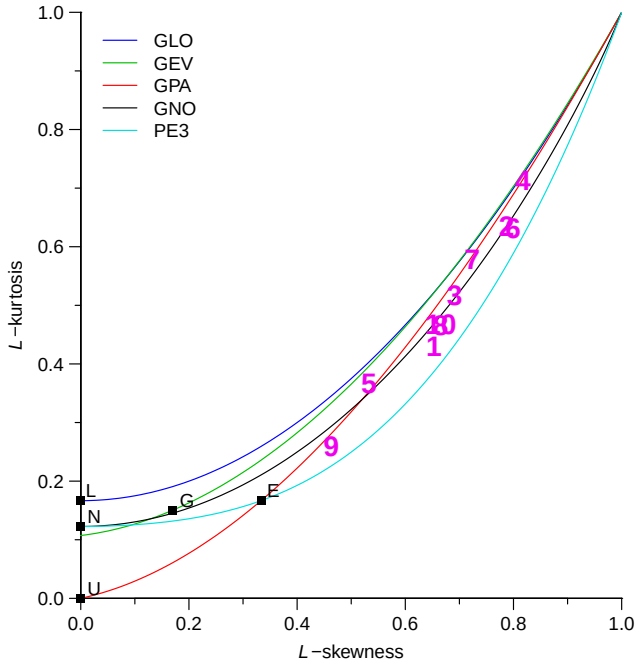

Figure 7: L-moment ratio diagram showing regional averages for tsunami data,plotted as symbols 1 to 10. Some points are hard to distinguish: L-moment ratios are almost identical for Regions 8 and 10, and are very similar for Regions 2 and 6.

The fitting procedure is as described in Section 3, and for this 2-parameter generalized Pareto distribution is very straightforward: the L-moments of the distribution are given by $\lambda_1 = \alpha/(1+k)$, $\lambda_2 = \alpha/\{(1+k)(2+k)\}$; the parameters are given in terms of the L-moment ratios by $k = 1/\tau - 2$, $\alpha = 1+k$; we therefore evaluate the regional average L-CV t^{R} as in (12) and use it as our estimate of τ to obtain estimates of k and α. Estimated parameters and quantiles of the fitted regional frequency distributions are given in Table 2. The regional frequency distributions are shown in Figure 8.

Estimation of the frequency distribution of runup height at a particular site requires an estimate of the index event, the scaling factor that is to be combined with the regional frequency distribution as in (14). For grid squares with a sufficiency of runup height observations the observed data can provide an estimate of the index event. Because the distribution of runup height is heavy-tailed and many grid squares have few observations, estimation of the mean may be unreliable and we prefer to use the median runup height as the index event.

Consider a grid square with observations of n tsunami events over a period of T years, and denote by h_{med} the median runup height of these events. The return period of the median event is $T/(\frac{1}{2}n)$ years, and in general the return period of the event with nonexceedance probability F is $R = T/\{n(1-F)\}$. Inverting this relation, the nonexceedance probability of the event with return period R years is $F = 1 - T/(Rn)$. The ratio of the magnitude of the event with return period R years to the magnitude of

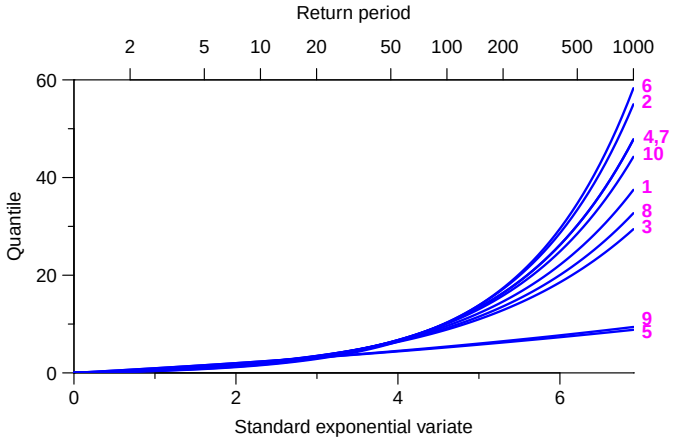

Figure 8: Regional frequency distributions for tsunami data.

Table 2: Parameters and quantiles of regional frequency distributions for tsunami data.

	Parameters		Quantiles			
Region	α	k	0.5	0.9	0.99	0.999
1	0.434	–0.566	0.37	2.06	9.62	37.48
2	0.287	–0.713	0.26	1.68	10.33	55.00
3	0.513	–0.487	0.42	2.18	8.87	29.43
4	0.347	–0.653	0.30	1.86	10.22	47.83
5	0.906	–0.094	0.65	2.33	5.22	8.82
6	0.257	–0.743	0.23	1.57	10.24	58.29
7	0.347	–0.653	0.30	1.86	10.22	47.79
8	0.479	–0.521	0.40	2.13	9.21	32.71
9	0.882	–0.118	0.64	2.33	5.39	9.40
10	0.376	–0.624	0.33	1.93	10.06	44.24

the median event is given by the regional frequency distribution as $q(1-T/(Rn))/q(1/2)$, where $q(\,.\,)$ is the quantile function of the regional frequency distribution. Thus the estimate of the event magnitude Q_R of return period R years for a specific site within the grid square is

$$Q_R = h_{\mathrm{med}}\; q(1-T/(Rn))/q(1/2)\,. \qquad (16)$$

We illustrate these calculations for the grid square 21°N, 158°W, which contains Honolulu, Hawaii. The tsunami record for this grid square is shown in Figure 9. In the $T = 111$ years starting in 1900 there have been $n = 95$ events, with a median runup height of $h_{\mathrm{med}} = 0.10$m. Honolulu is in Region 6, and the regional frequency distribution is a generalized Pareto distribution with quantile function (15) and parameters (from Table 2) α=0.257, k=−0.743. Event magnitudes (runup heights) for specified return periods can be obtained from (16); some examples are given in Table 3. The estimates obtained by this procedure are at best a rough approximation and do not make use of local knowledge: in particular we would obtain the same estimates for all sites within a grid square. For particular sites within a grid square it may be possible to improve

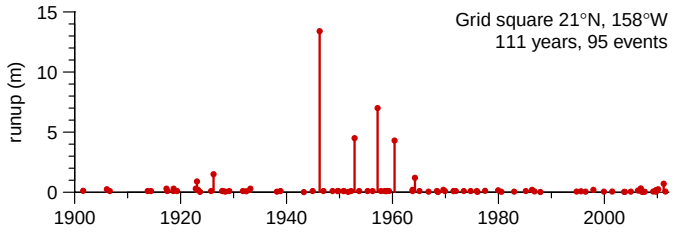

Figure 9: Tsunami record for the grid square containing Honolulu, Hawaii.

Table 3: Estimates of extreme tsunami magnitudes for Honolulu, Hawaii.

Return period (years)	10	100	1000
Event magnitude (m)	0.57	3.82	21.84

the estimates of the index event using knowledge of local conditions such as the underwater topography of the sea bed near the coast. We leave this as an area for future research.

Event magnitude estimates may also be required for a site with no or few events in the historical record. This is possible if the site can be assigned to a region and an independent estimate of the median event magnitude can be obtained. For example, on the west coast of the Americas regions have been identified, and for sites in grid squares with insufficient historical data the median event magnitude could be estimated by interpolation of the median event magnitudes in the nearest grid squares that have sufficient number of events in their historical records. Though a physically-based estimate might be preferred if it was available, this purely statistical interpolation-based approach may be satisfactory as an initial estimate.

5. Concluding remarks

Section 4 has shown how regional frequency analysis and L-moments can be used to estimate event magnitudes at tsunami-prone sites. It should be emphasized, however, that this is only a preliminary analysis and is subject to several qualifications.

The data set used has some weaknesses, since it lacks guarantees of uniformity of data collection procedures across time and space. For example, the cutoff at year 1900 is somewhat arbitrary. Magnitude-8 earthquakes occurred off northern Chile in 1868 and 1877, generating tsunamis with observed runup heights up to 18m and 24m respectively. If a reliable record that included these events could have been used in the analysis, L-CV values for grid squares in northern Chile (the southern part of Region 9) might have been much higher than the relatively low values computed from the post-1900 data. The lack of complete spatial coverage is also unfortunate. Had the quantity of data been sufficient, it would have been of interest to test whether the regional distribution of tsunami magnitudes is similar in the Atlantic and Indian Oceans to what we have found in different parts of the Pacific.

The runup height observations themselves are also not entirely satisfactory. Many of the observed runup heights are small, 0.10m or less, and it is not clear how consistently these have been observed at different sites and at different times. The L-moment statistics that are combined by regional averaging may therefore not be directly comparable: this can lead to bias and excess variability in the final estimates. It may be preferable to consider only observed runup heights that exceed some threshold of importance, since these can more reliably be compared across different sites.

As a related issue, the choice of index event needs further consideration. The median runup height is influenced by how many very small events are recorded at each site. It may be preferable to use an event of specified return period, such as the runup height with return period 10 years, which could be estimated more consistently at different sites.

Amalgamation of data by grid square needs further examination and justification. Though combination of data from neighbouring sites is effective in increasing the available sample sizes, more care is needed to ensure that data are combined only from sites with essentially the same frequency distribution.

The runup height distributions have very high skewness and heavy tails. For 7 of the 10 regions the fitted generalized Pareto distribution has $k < -0.5$, indicating that the distribution has infinite variance. There is not much evidence about the reliability of L-moment methods on distributions as skewed as this. Alternative approaches, such as the use of trimmed L-moments [5, 9] (which are designed to accommodate distributions with very heavy tails) or logarithmic transformation of the data (which reduces the skewness) may be worthy of consideration.

The region definitions obtained in this analysis seem reasonable, but could benefit from further refinement. Ideally this would make use of physical information about the tsunami susceptibility of different locations.

To summarize, we have presented a procedure for estimation of tsunami event magnitudes of specified return periods. In principle event magnitudes can be estimated for any site that may be affected by tsunamis, though in practice the available extent of historical data limited the analysis to sites in the Pacific basin. The procedure is entirely statistical and makes no use of physical understanding of the causes and propagation of tsunamis. This is arguably a weakness of the method that limits its ability, for example, to distinguish between risks at sites that are geographically close but have different exposure to tsunamis as a result of differences in the local topography of the offshore seabed; on the other hand, the statistical approach introduces no biases arising from choosing which kinds of earthquakes can generate tsunamis or from inaccurate modelling of tsunami propagation and landfall. It is plausible that a combination of the statistical and physically-based approaches may yield the most reliable estimates of tsunami risk.

Finally we note that a completely different regional

approach to tsunami event magnitude could be considered. An indirect approach would first obtain a frequency distribution for magnitudes of tsunami-generating earthquakes and convert this into a frequency distribution for tsunami magnitude. Thompson et al. [18] have presented a regional frequency analysis of global earthquake magnitudes; a similar approach could be used for the subset of tsunami-generating earthquakes. A spatial distribution of earthquakes could be obtained from the empirical distribution (possibly smoothed) of historical tsunami-generating earthquakes. Combining these distributions yields a joint distribution of tsunami frequency and location, and thence, for any site of interest, a joint distribution of earthquake magnitude M and distance R from the site. This can be converted into a distribution of tsunami runup height H using Abe's [1] relation $\log H = M - \log R - 5.80$.

Acknowledgements

The author thanks Prof. Masato Wakayama for his invitation to speak at the FMI-2011 conference, and all the conference organizers and support staff for the excellence of the conference arrangements. The final manuscript has benefited from helpful discussions with FMI-2011 participants, particularly Jane Sexton (Geophysics Australia) and Alejandro Jofré (Universidad de Chile), and the suggestions of an anonymous reviewer.

References

[1] Abe, K.: Estimate of tsunami run-up heights from earthquake magnitudes, in *Tsunami: Progress in prediction, disaster prevention and warning*, ed. Y. Tsuchita and N. Shuto, pp. 21–35, Kluwer Academic Publishers, Dordrecht, Netherlands, 1995.

[2] Bonnin, G. M., et al.: *NOAA Atlas 14. Precipitation-Frequency Atlas of the United States*, National Oceanic and Atmospheric Administration, Silver Spring, Md., 2002–2011.

[3] Burn, D. H.: Evaluation of regional flood frequency analysis with a region of influence approach, *Water Resour. Res.*, **26** (1990), 2257–2265.

[4] Dalrymple, T.: Flood frequency analyses, *Water Supply Paper 1543-A*, U.S. Geological Survey, Reston, Va., 1960.

[5] Elamir, E. A. H., and Seheult, A. H.: Trimmed L-Moments, *Comput. Statist. Data Anal.*, **43** (2003), 299–314.

[6] Gaal, L., Szolgay, J., Lapin, M., and Faško, P.: Hybrid approach to delineation of homogeneous regions for regional precipitation frequency analysis, *J. Hydrol. Hydromechan.*, **57** (2009), 226–249.

[7] Hosking, J. R. M.: *L*-moments: analysis and estimation of distributions using linear combinations of order statistics, *J. R. Statist. Soc. B*, **52** (1990), 105–124.

[8] Hosking, J. R. M.: *L*-moments, in *Encyclopedia of statistical sciences, update vol. 2*, ed. S. Kotz, C. Read, and D. L. Banks, 357–362, Wiley, New York, 1998.

[9] Hosking, J. R. M.: Some theory and practical uses of trimmed *L*-moments, *J. Statist. Plann. Inf.*, **137** (2007), 3024–3039.

[10] Hosking, J. R. M., and Wallis, J. R.: *Regional frequency analysis: an approach based on L-moments*, Cambridge University Press, Cambridge, U.K., 1997.

[11] Institute of Hydrology: *Flood Estimation Handbook*, 5 vols, Institute of Hydrology, Wallingford, Oxon., England, 1999.

[12] Lin, G.F., and Chen, L.-H.: Identification of homogeneous regions for regional frequency analysis using the self-organizing map, *J. Hydrol.*, **324** (2006), 1–9.

[13] National Geophysical Data Center / World Data Center (NGDC/WDC) Historical Tsunami Database, Boulder, Colo., U.S.A. Available at `http://www.ngdc.noaa.gov/hazard/tsu_db.shtml`.

[14] National Geophysical Data Center, World-Wide Tsunamis 2000 B.C. – to the present. Available at `http://www.ngdc.noaa.gov/nndc/struts/results?t=102564&s=207&d=207`.

[15] Neuman, C. P., and Schonbach, D. I.: Discrete (Legendre) orthogonal polynomials: a survey, *Int. J. Num. Meth. Eng.*, **8** (1974), 743–770.

[16] Pearson, C. P.: Regional frequency analysis of low flows in New Zealand rivers, *J. Hydrol. (N.Z.)*, **33** (1995), 94–122.

[17] Schaefer, M. G.: Regional analyses of precipitation annual maxima in Washington State, *Water Resour. Res.*, **26** (1990), 119–131.

[18] Thompson, E. M., Baise, L. G., and Vogel, R. M.: A global index earthquake approach to probabilistic assessment of extremes, *J. Geophys. Res.*, **112** (2007), B06314, doi:10.1029/2006JB00454.

[19] Willeke, G. E., Hosking, J. R. M., Wallis, J. R., and Guttman, N. B.: The National Drought Atlas (draft), *IWR Report 94-NDS-4*, U.S. Army Corps of Engineers, Fort Belvoir, Va., 1995. Also available online at `http://www.iwr.usace.army.mil/docs/atlas/Atlasintro.htm`.

[20] Wiltshire, S. E.: Identification of homogeneous regions for flood frequency analysis, *J. Hydrol.*, **84** (1986), 287–302.

Jonathan R. M. Hosking
IBM T. J. Watson Research Center, Yorktown Heights, New York, U.S.A.
E-mail: hosking(at)us.ibm.com