10

Forecasting US insured hurricane losses

THOMAS H. JAGGER, JAMES B. ELSNER, AND MARK A. SAUNDERS

Condensed summary

Coastal hurricanes generate huge financial losses within the insurance industry. The relative infrequency of severe coastal hurricanes implies that empirical probability estimates of the next big loss will be unreliable. Hurricane climatologists have recently developed statistical models to forecast the level of coastal hurricane activity based on climate conditions prior to the season. Motivated by the usefulness of such models, in this chapter we analyze and model a catalog of normalized insured losses caused by hurricanes affecting the United States. The catalog of losses dates back through the twentieth century. The purpose of this work is to develop a preseason forecast tool that can be used for insurance applications. Although wind speed is directly related to damage potential, the amount of damage depends on both storm intensity and storm size. As anticipated, we found that climate conditions prior to a hurricane season provide information about possible future insured hurricane losses. The models exploit this information to predict the distribution of likely annual losses and the distribution of a worst-case catastrophic loss aggregated over the entire US coast.

10.1 Introduction

Coastal hurricanes are a serious social and economic concern for the United States. Strong winds, heavy rainfall, and storm surge kill people and destroy property. The destructive power of hurricanes rivals that of earthquakes. On August 28, 2005, Hurricane Katrina's winds reached 78 meters per second ($\mathrm{m\,s}^{-1}$) in the central Gulf of Mexico, making it one of the strongest Atlantic hurricanes ever recorded. Early morning on the next day, Katrina struck Plaquemines Parish, Louisiana, with winds estimated near 65 $\mathrm{m\,s}^{-1}$. Katrina caused an estimated US$38 billion (bn) in insured losses as it roared across Louisiana, Mississippi, and Alabama.

It is important to know the return periods for losses incurred from storms of Katrina's magnitude or stronger and how the return periods vary when the climate fluctuates or changes (Elsner *et al*., 2006a). It is also valuable to be able to forecast the probability of a large loss before the hurricane season. Skillful forecasts of insured losses at lead times (forecast horizons) of 6 months or more would certainly benefit risk managers and others who are interested in acting on these forecasts. The rarity of severe hurricanes implies that empirical estimates of return periods likely will be unreliable. Fortunately, extreme value theory provides models for rare events and a justification for extrapolating to levels that are much greater than have already been observed. Moreover, statistical theory combined with knowledge of climate variability and its connection to regional storminess allows forecasts of seasonal hurricane activity.

Probability estimates of extreme hurricanes are available in the literature (Darling, 1991; Rupp and Lander, 1996; Heckert *et al*., 1998; Chu and Wang, 1998), but these studies do not address the question of how hurricane probabilities change with climate. This is done in Jagger *et al*. (2001), but the focus is on the probability of hurricanes of any intensity and not on the probability of the most extreme winds. Jagger and Elsner (2006) model the most extreme hurricane winds along the US coast and show how the probability of winds exceeding extreme thresholds changes with climate factors, including the North Atlantic Oscillation (NAO) and the El Niño/Southern Oscillation (ENSO).

Predictions of basin-wide Atlantic hurricane activity have been around since the middle 1980s (Gray, 1984). Studies focusing on climate factors that influence hurricane frequency regionally (Lehmiller *et al*., 1997; Bove *et al*., 1998; Maloney and Hartmann, 2000; Elsner *et al*., 2000a; Murnane *et al*., 2000; Saunders *et al*., 2000; Jagger *et al*., 2001; Larson *et al*., 2005) are more recent. Insights into climate conditions that affect regional hurricane activity are used to help predict landfall activity (Lehmiller *et al*., 1997; Elsner and Jagger, 2004, 2006; Saunders and Lea, 2005). Preseason forecasts of the number of hurricanes expected to affect the coast are useful especially if they are issued with significant lead time.

Saunders and Lea (2005) were the first to link predictions of US hurricane activity to skillful seasonal forecasts of loss. Here we present forecast models that can be used to directly predict the probability of a significant US financial loss from July 1. The models combine the strategy of Jagger and Elsner (2006) to estimate return periods with the strategy of Elsner and Jagger (2006) to forecast US hurricane activity before the start of the hurricane season. We begin with an examination of the normalized insured loss data and the data

associated with climate fluctuations. We then describe our modeling strategy and show results from a preseason model that predicts the annual expected loss and a model that predicts the worst-case scenario over a 100-year time horizon.

10.2 Normalized insured losses: 1900–2005

The work presented in this chapter was motivated by Katz (2002), who modeled total annual economic damage associated with hurricanes with a compound Poisson process. The process is compound since the total number of damaging hurricanes per year is fitted with a Poisson distribution, while the monetary amount of damage for individual hurricanes is fitted by the log-normal distribution. Damage totals are thus represented as a "random sum," with variations in total damage being decomposed into two sources, one attributable to variations in the frequency of events and another to variations in the damage from individual events. Results from Katz (2002) indicate a dependence of both hurricane occurrence and damage amount on the state of ENSO. Our idea is similar but with the following differences. First, we use preseason covariates to represent the climate rather than a contemporaneous above/below normal factor. Second, we use a threshold for dividing the loss data into small and large loss events, and third, we use simulation (random samples) to generate the distribution of losses.

We obtained insured loss data from Collins and Lowe (2001), who have produced a normalized record of insured losses for all hurricanes affecting the United States between 1900 and 1999. The normalization adjusts the damage from each hurricane to match what it would be if the storm had struck in the year 2000. This normalization is achieved by allowing for changes in inflation, wealth, and population, plus an additional factor, which represents a change in the number of housing units that exceeds population growth between the year of the loss and 2000. We extend the original Collins and Lowe (2001) data to 2005 using insured losses provided by the US Property Claims Service and inflate all losses to reflect 2005 US dollar values. The insured loss data for 1900–2005 comprise 178 loss events. The Collins and Lowe (2001) insured loss dataset is similar to the loss dataset of Pielke and Landsea (1998), who estimated total economic losses attributable to hurricanes since 1900. The rank correlation between the two annual hurricane loss time series is high, at 0.90 (1900–99).

Figure 10.1 shows the distribution and time series of insured losses over the period 1900–2005. The histogram bars indicate the percentage of events with losses in groups of US$1bn. The distribution is highly skewed, with 34% of the events having losses exceeding US$1bn and 19% of the events having losses

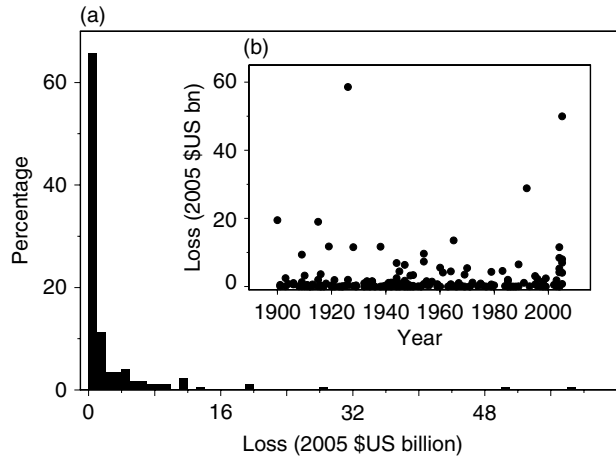192   *T. H. Jagger and others*

Figure 10.1. (a) Distribution of insured losses from hurricanes in the United States (excluding Hawaii). The distribution is highly skewed, with a few events generating very large losses. (b) Time sequence of the losses. Individual years may have more than one loss event.

exceeding US$3bn. The worst loss occurred with the 1926 hurricane that struck southeast Florida, creating an estimated insured loss adjusted to 2005 dollars of US$58.5bn. Hurricane Katrina in 2005 comes in second, with an estimated total loss of US$38.1bn. The time series of event losses is shown as an insert to Figure 10.1. Years with more than one loss have more than one dot. The data display large year-to-year variability but no obvious long-term trend, although here the data are not disaggregated into loss amount and number of loss events. The insured loss exceedances are shown in Table 10.1. Of the 178 loss events since 1900, 113 exceeded US$100 million (mn) in losses and 10 of these exceeded US$10bn. The geographic distribution of losses is shown in Figure 10.2. Plots are made for losses in four sizes, ranging from less than US$100 mn to more than US$10 bn. There does not appear to be a large geographic variation in loss locations with loss amount, with the exception of the largest loss amounts confined to southern exposures.

Because of the large skewness in loss values, we transform the data by using logarithms. A logarithmic transformation of the loss data is also used in Katz (2002). Here we use the base 10 logarithm for ease of interpretation. The logarithm to base 10 of a US$1bn loss is equal to 9. Figure 10.3 shows the logarithm of insured losses. The time series of log transformed annual losses shows no significant trend, although two of the highest yearly totals occurred in 2004 and 2005. The distribution of the logarithm of annual losses approximates a normal distribution, although there is some asymmetry in the tails. A quantile–quantile plot of the logarithm of losses against a normal distribution

Table 10.1. *Insured loss exceedances (US dollars adjusted to 2005)*

Values are the number of events exceeding various loss thresholds.

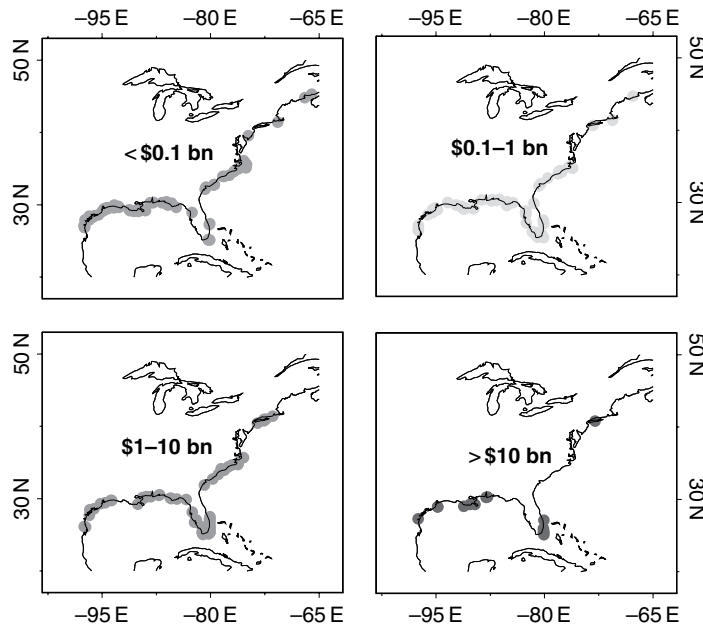|	Exceedance number	
US$ (2005)		Events
1 mn[a]		177
10 mn		172
100 mn		113
1 bn[a]		61
10 bn		10

[a] Abbreviations: mn, million; bn, billion.

Figure 10.2. Geographic distribution of normalized insured losses from hurricanes striking the United States between 1900 and 2005. See also Plate 22.

indicates a reasonable fit and provides evidence that the distribution of individual losses is log normal. Figure 10.3 also shows the annual number of loss events and the distribution. Again we see no obvious trend over time. There were three years with six loss events, with the most recent being 2004. The observed mean rate of loss events is 1.68, with a variance that is nearly equal, at
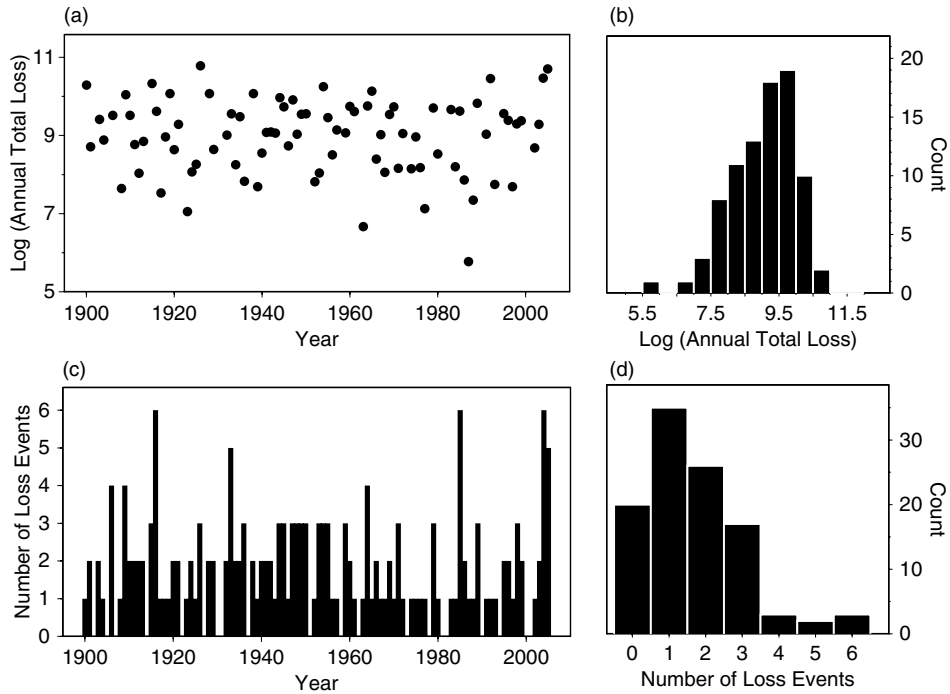
194 *T. H. Jagger and others*

Figure 10.3. (a) Time series and (b) distribution of the logarithm of annual total insured losses and the (c) time series and (d) distribution of the number of loss events.

1.896, consistent with the property of a Poisson distribution. A formal χ^2 test indicates that there is no reason to question a Poisson distribution for the annual number of loss events.

By examining the conditional variance, Katz (2002) estimated that about 17% of the variation in total annual damage is attributable to fluctuations in the annual number of storms. Thus we would expect that a climate variable that explains a portion of the fluctuation in annual number of events could be used help predict annual losses.

10.3 Climate variations

We argue that the annual distribution of insured hurricane losses depends to some extent on preseason climate factors. This conclusion is reasonable given that statistical relationships between US hurricane activity and climate are well established (Bove *et al*., 1998; Elsner and Kara, 1999; Elsner *et al*., 1999; Saunders *et al*., 2000; Elsner *et al*., 2000a, b, 2001; Elsner, 2003; Elsner *et al*., 2004; Saunders and Lea, 2005). More importantly for the present work, Jagger

et al. (2001) and Jagger and Elsner (2006) modeled the wind speeds of hurricanes at or near landfall and showed that the exceedance probabilities (e.g., wind speeds in excess of 100 knots) vary appreciably with the phase of the ENSO, the NAO, and Atlantic sea surface temperatures (SSTs). Similarly, Murnane *et al.* (2000) modeled the probability of coastal hurricanes conditioned on ENSO. A study by Goldenberg *et al.* (2001) suggests that the number and strength of Atlantic hurricanes follow a multidecadal cycle of changes in North Atlantic Ocean currents. This cycle, called the Atlantic Multidecadal Oscillation (AMO), might be related to changes in radiative forcing and/or changes in the thermohaline circulation.

The ENSO is characterized by basin-scale fluctuations in sea level pressure (SLP) between Tahiti and Darwin. Although noisier than equatorial Pacific SSTs, pressure values are available back to 1900. The Southern Oscillation Index (SOI) is defined as the normalized sea level pressure difference between Tahiti and Darwin. The SOI is strongly anti-correlated with equatorial Pacific SST, with an El Niño warming event associated with negative SOI values. Units are standard deviations. The relationship between ENSO and hurricane activity is strongest during the hurricane season, but we are interested in a predictive relationship, so we use a May–June average of the SOI as our predictor. The monthly SOI values (Ropelewski and Jones 1997) are obtained from the Climatic Research Unit (CRU) of the University of East Anglia, UK.

The NAO is characterized by fluctuations in sea level pressure differences. Index values for the NAO (NAOI) are calculated as the difference between the SLP for Gibraltar and for a station over southwest Iceland, and are obtained from the CRU (Jones *et al.*, 1997). The values are averaged over the pre- and early-hurricane season months of May and June (Elsner *et al.*, 2001). We speculate that the relationship might result from a communication between the middle latitudes and the tropics (Tsonis and Elsner, 1996) whereby below normal values of the NAO during the spring lead to dry conditions over the continents and to a tendency for greater summer/fall middle tropospheric ridging (enhancing the dry conditions). In turn, tropospheric ridging over the eastern and western sides of the North Atlantic basin during the hurricane season tends to keep the middle tropospheric trough of low pressure, responsible for hurricane recurvature, farther to the north and away from the westward tracking tropical cyclones (Elsner and Jagger, 2006).

The AMO is characterized by fluctuations in SST over the North Atlantic Ocean. Modeled SST and US National Oceanic and Atmospheric Administration (NOAA) optimal interpolated SST datasets were used to compute Atlantic SST anomalies north of the equator (Enfield *et al.*, 2001). Anomalies (in degrees centigrade, °C) are computed by month for the base

period 1951–2000. Data are obtained from the NOAA–Cooperative Institute for Research in Environmental Sciences' Climate Diagnostics Center (NOAA–CIRES CDC) back to 1871. For this study we average the Atlantic SST anomalies over the hurricane preseason months of May and June.

In summary, the distribution of US insured losses from hurricane winds is statistically modeled by using covariate (predictor) data for the period 1900–2005. We are interested in the preseason values (May–June averaged) of SOI, NAO, and Atlantic SST as predictors for the distribution of likely losses during the U.S. hurricane season, which runs principally from July through October. Figure 10.4 shows time series of the covariate values used in the model. All three series display large variability from year to year with a distinct nonlinear trend in the late springtime values of Atlantic SST.

The upper and lower quartile values of the SOI are 0.60 and −0.75 standard deviation (s.d.), respectively, with a median (mean) value of −0.16 (−0.10) s.d. Years of below (above) normal SOI correspond to El Niño (La Niña) events and thus to a lower (higher) probability of hurricanes. The upper and lower quartile values of the NAO are 0.42 and −1.08 s.d., respectively, with a median (mean) value of −0.39 (−0.32) s.d. Years of below (above) normal values of the NAO correspond to a weak (strong) NAO phase and thus to a higher (lower)
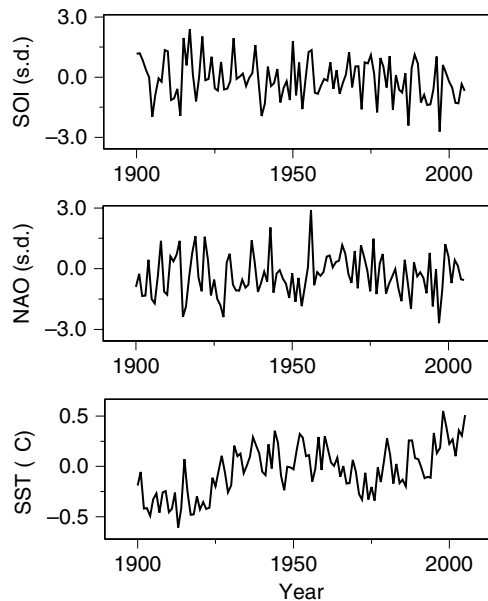

Figure 10.4. Time series of the three covariates used to predict insured wind losses from hurricanes before the start of the hurricane season. The values are averaged over the months of May and June. A linear detrended version of the Atlantic SST is sometimes referred to as the AMO.

probability of US hurricanes. The upper and lower quartile values of the Atlantic SST are 0.13 °C and −0.23 °C, respectively, with a median (mean) value of −0.04 (−0.04) °C. Years of above (below) normal values of SST correspond to a higher (lower) probability of hurricane activity. The linear correlation between the SOI and the NAO (SST) is a negligible + 0.03 (−0.04). The linear correlation between the NAO and Atlantic SST is a marginally significant value of −0.21.

10.4 Large and small losses

The total amount of insured losses calibrated to 2005 US dollars from the 178 events (1900–2005) is estimated at US$421bn. The large skewness in the insured losses per event and per annum suggests that it might be a good strategy to separate large losses from small losses for the purpose of prediction. It is often quoted that 80% of the total damage from all hurricanes is caused by the top 20% strongest storms. Figure 10.5 shows that the distribution of loss data is even a bit more skewed than that. In fact, we find that the top 30 loss events (less than 17% of the total number of loss events) account for more than 80% of the total loss amount.

The relative infrequency of the largest loss events argues for a split that favors including more data for modeling. Here we use a cutoff of US$100 mn and find that 113 of the 178 events (63.5%) exceeded this threshold. The
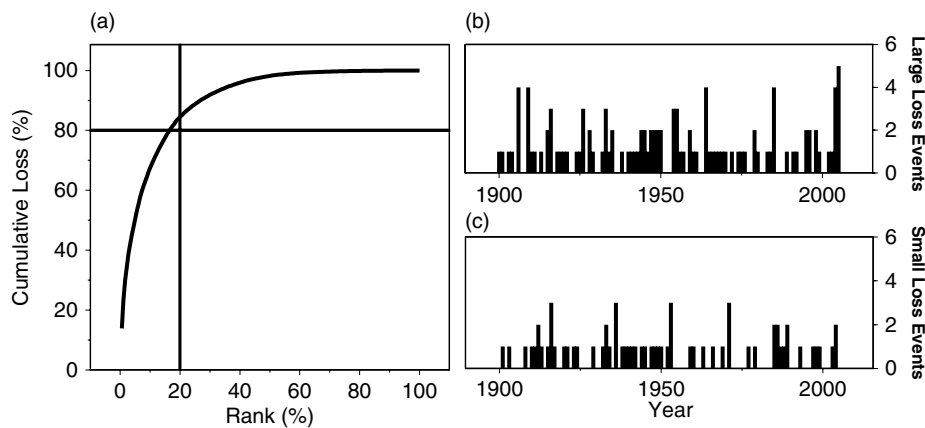

Figure 10.5. (a) Cumulative percent of total losses as a function of percent ranking. The reference lines indicate the oft-cited 80%/20% relationship, whereby the top 20% strongest hurricanes account for 80% of the losses. A split of the event counts into (b) large loss events and (c) small loss events based on losses exceeding US$100 mn is shown as annual time series. For reference, 2004 experienced four large and two small loss events.

remaining 65 events (36.5%) account for only 0.6% of the total losses. Thus it might be reasonable to assume that the small loss events are at the "noise" level. Time series of the annual number of large and small loss events are shown in Figure 10.5. The rank correlation between the two series is a negligible 0.06.

Next we examine the influence of the covariates, discussed in the previous section, on both the magnitude of annual loss and the number of annual loss events. For the number of loss events, we consider small and large loss events separately. Using the preseason Atlantic SST, we are able to explain 13% of the variation in the logarithm of loss values exceeding US$100 mn using an ordinary least squares regression model. The relationship is positive, indicating that warmer Atlantic SSTs are associated with larger losses as expected. The rank correlation between the amount of loss (exceeding US$100 mn) and the May–June Atlantic SST is $+0.31$ (p-value $= 0.0086$) over all years in the dataset and is $+0.37$ (p-value $= 0.0267$) over the shorter 1950–2005 period.

We also examine models for the number of loss events using the covariates. We find that the NAO is useful in predicting both the number of large loss events and the number of small loss events. The relationship is negative, indicating that when the preseason value of the NAO decreases, the probability of a loss event increases. The rank correlation between the total number of loss events and the preseason NAO is –0.29 (p-value $= 0.0032$) over all years and is -0.12 (p-value $= 0.3812$) over the shorter 1950–2005 period. Interestingly, we find no significant preseason relationship between event counts and SST or the SOI.

The analysis confirms that it is reasonable to model small and large loss events separately. However, it should be noted that it might be more appropriate to add measurement error to the data so as to reduce the weight of the smaller measurements rather than separate the data as is done here. Our final strategy combines a model for the loss amount with two models for the number of loss events: one for large losses and the other for small losses. We use the NAO for predicting the number of loss events (both large and small) and the SST for predicting the amount of damage given a loss event. We find that including the preseason SOI covariate does not help in forecasting the upcoming season's losses either for the amount of loss or for the number of loss events. This result is consistent with those for the models developed in Elsner and Jagger (2006) and Elsner *et al.* (2006b) for predicting coastal hurricane activity based on preseason data. Since it is well known that ENSO has an influence on shear during the hurricane season, it might be advantageous to include a predicted value of the SOI for the hurricane season rather than a preseason value as is done here.

10.5 Predicting annual losses

Results from the previous section provide the needed background for building a preseason model capable of predicting the annual expected loss. The model uses a hierarchical Bayesian specification. The final form of the model was based on comparison of the deviance information criterion (DIC) using several different models involving the three covariates. The DIC is a generalization of the Akaike information criterion (AIC) and Bayesian information criterion (BIC). It is useful in Bayesian model selection where the posterior distributions of the models are obtained by Markov chain Monte Carlo (MCMC) simulation. Like the AIC and BIC, it is an asymptotic approximation as the sample size becomes large. It is only valid when the posterior distribution is close to multivariate normal. We chose the model with the lowest value of DIC.

A schematic of the hierarchical model is shown in Figure 10.6. The predicted annual loss (TL) is the sum of the individual loss amounts (both large [LL_L] and small [LL_S] amounts) multiplied by the respective number of large (N_L) and small (N_S) loss counts. Given the mean (μ_L) and standard deviation (σ_L) of the logarithm of large losses, the logarithm of large loss follows a truncated normal distribution. Small loss amounts are also specified by using a truncated normal distribution, although the mean is not a function of any of
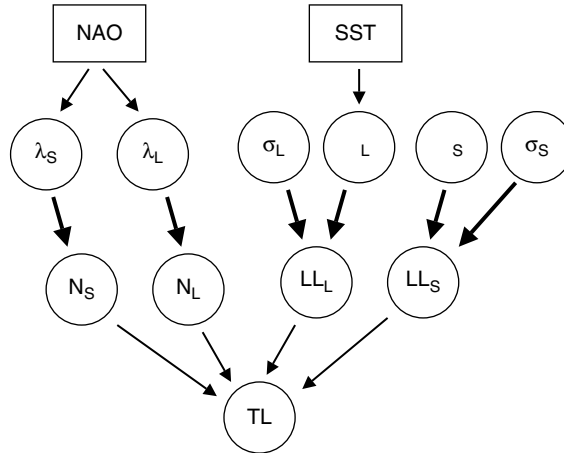

Figure 10.6. Hierarchical graph illustrating our strategy for simulating annual insured losses based on preseason values of the NAO and Atlantic SST. The connection between nodes is either stochastic (thick arrow) or logical (thin arrow). Node λ_L (λ_S) is the mean annual rate of large (small) losses, N_L (N_S) is the annual count of large (small) loss events, μ_L (μ_S) is the mean amount of large (small) loss on a log scale, σ_L (σ_S) is the standard deviation of large (small) loss amounts, LL_L (LL_S) is the logarithm of large (small) loss amount, and TL is the total loss.

the covariates. Given a mean annual rate of large losses (λ_L), the annual number of large losses follows a Poisson distribution with the natural logarithm of the rate given as a linear function of the NAO. Similarly, given a mean annual rate of small losses (λ_S), the annual number of small losses follows a Poisson distribution with the natural logarithm of the rate given as a separate function of the NAO.

Samples of the annual losses are generated by using the WinBUGS (Windows version of Bayesian inference Using Gibbs Sampling) developed at the Medical Research Council in the United Kingdom (Gilks *et al.*, 1996; Spiegelhalter *et al.*, 1996). WinBUGS chooses an appropriate MCMC sampling algorithm based on the model structure. In this way, annual losses are sampled conditional on the model coefficients and the observed values of the covariates. The cost associated with a Bayesian approach is the requirement to formally specify prior beliefs. Here we take the standard route and assume noninformative priors, which as the name implies provide little information about the parameters of interest. Markov chain Monte Carlo analysis, in particular Gibbs sampling, is used to sample the parameters given the data, since no closed form distribution exists for the truncated normal (or for the generalized Pareto distribution [GPD] used in the next section).

We check for mixing and convergence by examining successive samples of the parameters. Samples from the posterior distributions of the parameters indicate relatively good mixing and quick settling as two different sets of initial conditions produce sample values that fluctuate around a fixed mean. Based on these diagnostics, we discard the first 10,000 samples and analyze the output from the next 10,000 samples. The utility of the Bayesian approach for modeling the mean number of coastal hurricanes is described in Elsner and Jagger (2004).

Figure 10.7 shows the predictive posterior distributions of losses for two different climate scenarios. The first scenario is characterized by preseason conditions featuring a combination of high NAO values and low SST values. To offer a strong contrast, we set the values to their maximum and minimum, respectively, over the 106-year period (1900–2005; NAO = +2.9 s.d. and SST = −0.61 °C). This situation is unfavorable for hurricane activity along the US coast (Elsner and Jagger, 2006). Simulation results show that the probability of no loss (47%) is close to the probability of at least some loss (53%). This result contrasts with those from the second scenario, which is characterized by conditions favorable for hurricane activity (NAO = −2.7 s.d. and SST = +0.55 °C). Here the probability of at least some loss is 94%.

Perhaps more useful is the predictive distribution of losses, given that at least some loss occurs. Here the distributions are shown for the logarithm
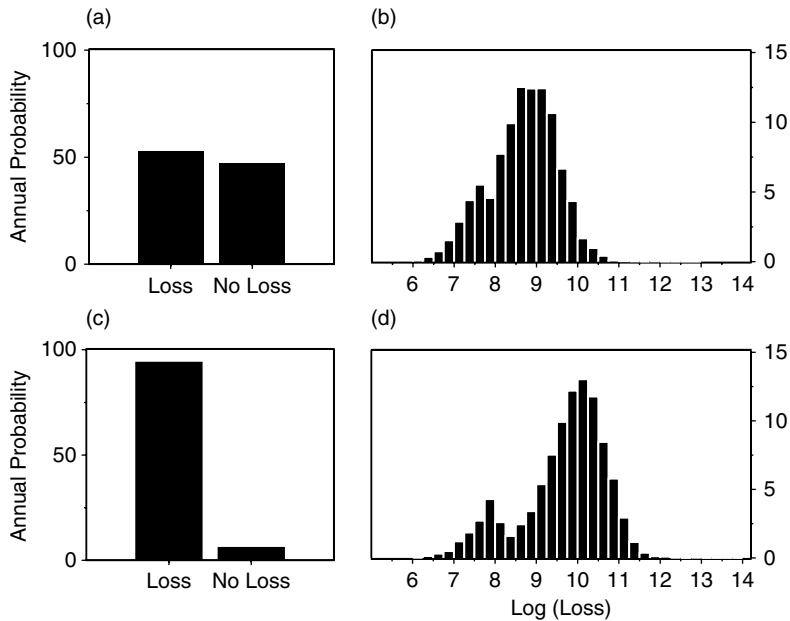
Figure 10.7. Simulated annual losses for two different climate scenarios. (a)
The probability of at least some loss and (b) the probability distribution of
loss amounts given at least one loss event under preseason climate conditions
foreshadowing an inactive US landfall season. Plots (c) and (d) are the same
as plots (a) and (b), respectively, except that they are based on preseason
conditions foreshadowing an active US landfall season.

of total annual loss from both scenarios and, as expected, the results are
divergent. In the case of favorable late springtime conditions for hurricane
activity, the loss distribution is shifted toward substantially higher loss
amounts relative to the case of unfavorable conditions. Converting to 2005
dollars, the expected yearly loss in a year with at least one loss when
conditions are favorable for hurricanes is estimated at US$25.2bn. This
compares with US$2.1bn when conditions are unfavorable. The overall
expected loss (taking into account the non-zero probability of no losses) is
US$23.7bn under favorable climate conditions and US$1.1bn under unfa-
vorable conditions. Therefore, assuming the model is correct and the future
will be the same as the past, the model is useful in portending the amount
of insured losses before the start of the season. The interesting side hump in
the distribution of losses is likely an artifact of using a truncated normal
distribution. Both the 2004 and 2005 hurricane seasons featured late spring-
time negative NAO values and above normal Atlantic SST values, which
combined to produce a forecast (hindcast) of above normal insured loss
probabilities.

10.6 Predicting extreme losses

While the modeling strategy described above makes sense for forecasting the distribution of likely losses associated with climate conditions before the start of the season, for financial planning it might be of greater interest to know the maximum possible loss. In this case, the normal distribution is replaced by an extreme value distribution for the logarithm of losses. For example, the family of generalized Pareto distributions describes the behavior of individual extreme events. Consider observations from a collection of random variables in which only those observations that exceed a fixed value are kept. As the magnitude of this value increases, the GPD family represents the limiting behavior of each new collection of random variables. This property makes the family of GPDs a good choice for modeling extreme events involving large insured losses. The choice of threshold, above which we treat the values as extreme, is a compromise between retaining enough observations to properly estimate the distributional parameters (scale and shape), but few enough that the observations follow a GPD family. A negative shape parameter implies an upper limit to the maximum possible loss.

The GPD describes the distribution of losses that exceed a threshold u but not the frequency of losses at that threshold. As we did with the annual loss model, we specify that, given a rate of loss events above the threshold, the number of loss events follows a Poisson distribution. Here there is no need to consider small loss events, as we are interested only in the large ones. Combining the GPD for the distribution of large loss amounts with the Poisson distribution for the frequency of loss events allows us to obtain return periods for given levels of annual losses.

We determine the particular threshold value for the set of insured losses by examining the plots shown in Figure 10.8. The mean residual life (MRL) plot shows the value of the mean excess as a function of threshold. The MRL plot is produced by averaging the difference (residual) in the observed logarithm of loss above a threshold as a function of the threshold. For example, at a log loss of 9 we subtract 9 from each observed log loss and average only the positive values (excesses). We repeat this operation for all thresholds. The mean excess is the expected value of the amount that the observations exceed the threshold. The standard errors on the mean excess allow us to compute confidence levels for the estimates. A nearly straight-line negative relationship between the mean excess and the loss above some threshold indicates the set of extreme losses. In other words, if extreme values follow a GPD, then their expected value is a linear function of the threshold. From the plot we see that a straight-line relationship is noted for losses at about 9 (US$1bn). The other two plots
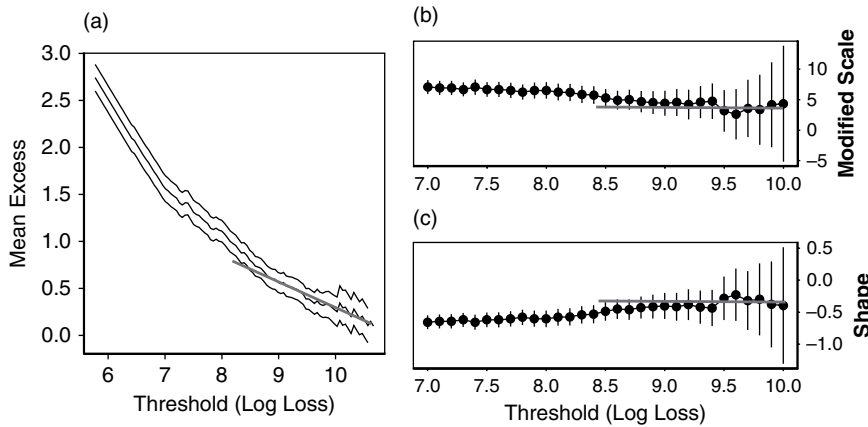
Figure 10.8. (a) Mean residual life plot for the logarithm of insured hurricane losses. The outside lines are the 95% confidence limits. An approximate linear decrease of the mean excess occurs after a threshold of about 9. The value of the (b) scale and (c) shape parameters from the GPD at various thresholds. The systematic variation is not detectable for thresholds exceeding about 9.

show the GPD parameters as a function of threshold. The systematic variation in the scale and shape parameters with threshold appears to end at a threshold value between 8.5 and 9, suggesting that only events with losses exceeding this level are extreme. Taken together, the diagnostic plots suggest that a threshold value is US\$1bn in losses.

As with the annual loss model, we use a Bayesian hierarchical specification for the model of extreme losses. Markov chain Monte Carlo samples are used to generate posterior predictive distributions. Here we are interested in the return level as a function of return period. A schematic of the hierarchical model is shown in Figure 10.9. The annual return level (RL_y) is determined by the return level of individual extreme events (RL_E) and the annual frequency of such events above a threshold rate (λ). The annual number of extreme events follows a Poisson distribution, with the natural logarithm of the rate specified as a linear function of the NAO. Given values for the scale (σ) and shape (ξ) parameters, the return level of individual extreme events follows a GPD. The logarithm of the scale parameter is a linear function of the NAO, and the shape parameter is a linear function of the SOI.

As before, samples of the return levels are generated by using WinBUGS and we use non-informative prior distributions. Samples from the posterior distribution of the model parameters indicate good mixing and convergence properties. We discard the first 10,000 samples and analyze the output from the next 10,000 samples. Applications of Bayesian extremal analysis are
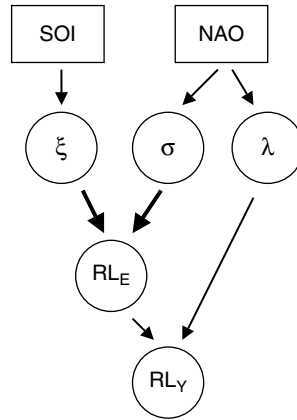
Figure 10.9. Hierarchical graph illustrating our strategy for simulating return levels for extreme losses conditional on the preseason values of the NAO and SOI. Nodes ξ and σ are the shape and scale parameters of the GPD, respectively; λ is the mean rate of extreme losses; RL_E is the return level for a particular loss event; and RL_Y is the return level for total losses over the year.
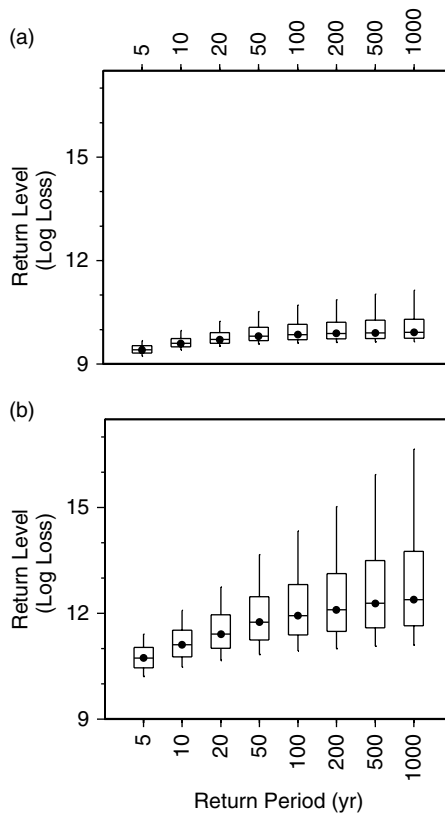

Figure 10.10. Simulated extreme losses for two different climate scenarios. (a) The distribution of return levels in the logarithm of insured losses for the case of unfavorable conditions for US hurricane landfalls, and (b) the distribution of return levels for the case of favorable conditions for US hurricane landfalls.

relatively rare (Coles and Tawn, 1996; Katz *et al.*, 2002; Coles *et al.*, 2003). In the context of local hurricane winds, Casson and Coles (1999) used a Bayesian analysis to estimate parameters of spatial regression models. They showed that including the spatial characteristics of extremes provides a substantial reduction in the confidence intervals for high quantiles. Bayesian approaches to modeling extreme wind behavior are given in Walshaw (2000) and Jagger and Elsner (2006).

Figure 10.10 shows the predictive posterior distributions of extreme losses for two different climate scenarios. The first scenario is characterized by preseason conditions featuring a combination of high NAO and high SOI values. Again, to offer a strong contrast, we set the values to their maximum and minimum over the 106-year period (1900–2005). Box and whisker plots are used to illustrate the variation in simulated extreme loss amounts for increasing return periods.

Results show the clear difference in expected extreme losses for the different climate conditions. Under the unfavorable scenario for US hurricanes, we find the expected return level of a 50-year extreme event at less than US$10bn; this compares with a return level of a 50-year extreme event loss of approximately US$630bn under favorable scenarios for US hurricanes. Thus the model is also useful for projecting extreme losses over longer time horizons given the preseason values of the climate covariates.

10.7 Summary

Coastal hurricanes are capable of generating large financial losses for the insurance industry. The rarity of large losses in the historical record implies that empirical estimates of next year's loss will have large errors. Annual loss totals are directly related to the size and number of hurricanes affecting the coast. Since some skill exists in forecasts of coastal hurricane activity, it makes sense to investigate the potential of predicting losses directly. This chapter demonstrates a strategy for making forecasts of annual insured losses by July 1 using preseason values for the NAO, Atlantic SST, and the SOI. Models are specified by using hierarchical Bayesian technology, and predictive posterior distributions are generated by using MCMC sampling. Markov chain Monte Carlo sampling provides a method of generating future loss projections. According to the model of expected annual loss, the probability of incurring a loss is higher compared to the climatological average when the NAO is negative. Also, the amount of loss is greater when Atlantic SST is above normal. Both conditions were met before the 2004 and 2005 hurricane seasons. While we did not perform an out-of-sample test of model skill, a similar

hierarchical Bayesian model for US hurricane activity using the same covari-
ates is cross-validated and shown to have skill above climatology in Elsner and
Jagger (2006).

These results are consistent with current understanding of hurricane climate
variability. Forecasts of extreme loss amounts are also possible using a some-
what different model specification and including a preseason value of the
SOI. Return level loss amounts exceed those of climatology under conditions
characterized by a negative NAO. It might be possible to develop a similar
model using data from as early as February 1 (see Elsner *et al.*, 2006b). While
the models here are developed from aggregate loss data for the entire United
States due to Atlantic hurricanes, it would be possible to apply the techni-
ques to model data representing a subset of losses, capturing, for example,
a particular reinsurance portfolio. Moreover, since the models use MCMC
sampling, they can be extended quite easily to include measurement error
as well as missing data. The models and data are available on our web site
(Google key words: hurricane climate).

References

Bove, M. C., Elsner, J. B., Landsea, C. W., Niu, X., and O'Brien, J. J. (1998). Effect of
    El Niño on U.S. landfalling hurricanes, revisited. *Bulletin of the American
    Meteorological Society*, **79**, 2477–82.
Casson, E., and Coles, S. (1999). Spatial regression models for extremes. *Extremes*, **1**,
    449–68.
Chu, P. -S., and Wang, J. (1998). Modeling return periods of tropical cyclone
    intensities in the vicinity of Hawaii. *Journal of Applied Meteorology*, **37**, 951–60.
Coles, S., and Tawn, J. A. (1996). A Bayesian analysis of extreme rainfall data. *Applied
    Statistics*, **45**, 463–78.
Coles, S., Pericchi, L. R., and Sisson, S. (2003). A fully probabilistic approach to
    extreme rainfall modeling. *Journal of Hydrology*, **273**, 35–50.
Collins, D. J., and Lowe, S. P. (2001). A macro validation dataset for U.S. hurricane
    models. *Casualty Actuarial Society, Winter Forum*, pp. 217–52.

Darling, R. W. R. (1991). Estimating probabilities of hurricane wind speeds using a large-scale empirical model. *Journal of Climate*, **4**, 1035–46.

Elsner, J. B. (2003). Tracking hurricanes. *Bulletin of the American Meteorological Society*, **84**, 353–6.

Elsner, J. B., and Jagger, T. H. (2004). A hierarchical Bayesian approach to seasonal hurricane modeling. *Journal of Climate*, **17**, 2813–27.

Elsner, J. B., and Jagger, T. H. (2006). Prediction models for annual U.S. hurricane counts. *Journal of Climate*, **19**, 2935–52.

Elsner, J. B., and Kara, A. B. (1999). *Hurricanes of the North Atlantic: Climate and Society*. New York: Oxford University Press.

Elsner, J. B., Bossak, B. H., and Niu, X.-F. (2001). Secular changes to the ENSO-U.S. hurricane relationship. *Geophysical Research Letters*, **28**, 4123–6.

Elsner, J. B., Jagger, T. H., and Niu, X. (2000a). Shifts in the rates of major hurricane activity over the North Atlantic during the twentieth century. *Geophysical Research Letters*, **27**, 1743–6.

Elsner, J. B., Jagger, T. H., and Tsonis, A. A. (2006a). Estimated return periods for Hurricane Katrina. *Geophysical Research Letters*, **33**, L08704, doi:10.1029/2005GL025452.

Elsner, J. B., Kara, A. B., and Owens, M. A. (1999). Fluctuations in North Atlantic hurricanes. *Journal of Climate*, **12**, 427–37.

Elsner, J. B., Liu, K.-b., and Kocher, B. (2000b). Spatial variations in major U.S. hurricane activity: statistics and a physical mechanism. *Journal of Climate*, **13**, 2293–305.

Elsner, J. B., Murnane, R. J., and Jagger, T. H. (2006b). Forecasting U.S. hurricanes 6 months in advance. *Geophysical Research Letters*, **33**, L10704, doi:10.1029/2006GL025693.

Elsner, J. B., Niu, X. -F., and Jagger, T. H. (2004). Detecting shifts in hurricane rates using a Markov chain Monte Carlo approach. *Journal of Climate*, **17**, 2652–66.

Enfield, D. B., Mestas-Nuñez, A. M., and Trimble, P. J. (2001). The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. *Geophysical Research Letters*, **28**, 2077–80.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Boca Raton, FL: Chapman & Hall/CRC.

Goldenberg, S. B., Landsea, C. W., Mestas-Nuñez, A. M., and Gray, W. M. (2001). The recent increase in Atlantic hurricane activity: causes and implications. *Science*, **293**, 474–9.

Gray, W. M. (1984). Atlantic hurricane frequency. Part I: El Niño and 30 mb quasi-biennial oscillation influences. *Monthly Weather Review*, **112**, 1649–68.

Heckert, N. A., Simiu, E., and Whalen, T. (1998). Estimates of hurricane wind speeds by "peaks over threshold" method. *Journal of Structural Engineering*, ASCE, **124**, 445–9.

Jagger, T. H., and Elsner, J. B. (2006). Climatology models for extreme hurricane winds near the United States. *Journal of Climate*, **19**, 3220–36.

Jagger, T. H., Elsner, J. B., and Niu, X. (2001). A dynamic probability model of hurricane winds in coastal counties of the United States. *Journal of Applied Meteorology*, **40**, 853–63.

Jones, P. D., Jónsson, T., and Wheeler, D. (1997). Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and Southwest Iceland. *International Journal of Climatology*, **17**, 1433–50.

Katz, R. W. (2002). Stochastic modeling of hurricane damage. *Journal of Applied Meteorology*, **41**, 754–62.

Katz, R. W., Parlange, M. B., and Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, **25**, 1287–304.

Larson, J., Zhou, Y., and Higgins, R. W. (2005). Characteristics of landfalling tropical cyclones in the United States and Mexico: climatology and interannual variability. *Journal of Climate*, **18**, 1247–62.

Lehmiller, G. S., Kimberlain, T. B., and Elsner, J. B. (1997). Seasonal prediction models for North Atlantic basin hurricane location. *Monthly Weather Review*, **125**, 1780–91.

Maloney, E. D., and Hartmann, D. L. (2000). Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian Oscillation. *Science*, **287**, 2002–4.

Murnane, R. J., and coauthors (2000). Model estimates hurricane wind speed probabilities. *Eos, Transactions of the American Geophysical Union*, **81**, 433–8.

Pielke, R. A., Jr., and Landsea, C. W. (1998). Normalized hurricane damages in the United States, 1925–95. *Weather Forecasting*, **13**, 621–31.

Ropelewski, C. F., and Jones, P. D. (1997). An extension of the Tahiti-Darwin Southern Oscillation Index. *Monthly Weather Review* 115, 2161–5.

Rupp, J. A., and Lander, M. A. (1996). A technique for estimating recurrence intervals of tropical cyclone–related high winds in the tropics: results for Guam. *Journal of Applied Meteorology*, **35**, 627–37.

Saunders, M. A., and Lea, A. S. (2005). Seasonal prediction of hurricane activity reaching the coast of the United States. *Nature*, **434**, 1005–8.

Saunders, M. A., Chandler, R. E., Merchant, C. J., and Roberts, F. P. (2000). Atlantic hurricanes and Northwest Pacific typhoons: ENSO spatial impacts on occurrence and landfall. *Geophysical Research Letters*, **27**, 1147–50.

Spiegelhalter, D. J., Best, N. G., Gilks, W. R., and Inskip, H. (1996). Hepatitis B: a case study in MCMC methods. In *Markov Chain Monte Carlo in Practice*, ed. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. London: Chapman & Hall/ CRC, pp. 45–58.

Tsonis, A. A., and Elsner, J. B. (1996). Mapping the channels of communication between the tropics and midlatitudes. *Physica D*, **92**, 237–44.

Walshaw, D. (2000). Modelling extreme wind speeds in regions prone to hurricanes. *Applied Statistics*, **49**, 51–62.