Modeling Severe Thunderstorm Risk in the United States

Shiraj Khan[1], Greta Ljung[2], Vineet Kumar Jain[3]

[1]Senior Engineer, AIR Worldwide Corp., Boston, MA USA, skhan@air-worldwide.com
[2]Principal Scientist, AIR Worldwide Corp., Boston, MA USA, gljung@air-worldwide.com
[3]Principal Engineer, AIR Worldwide Corp., Boston, MA USA, vjain@air-worldwide.com

ABSTRACT

The annual aggregate losses from severe thunderstorms have accounted for approximately half of all U.S. catastrophic insured losses since 1990, according to Insurance Services Office's Property Claim Services (PCS). In 2006, 2007, and 2008, nearly 90%, 55%, and 40%, respectively, of total catastrophic insured losses in the U.S. were directly related to severe thunderstorms. AIR developed the first probabilistic severe thunderstorm model for the insurance industry to manage severe thunderstorm risk proactively. The model simulates severe thunderstorm events representing large atmospheric systems which can spawn hundreds of tornadoes, hailstorms, and straight-line windstorms in a span of one to several days. The model is parameterized using historical data available in a severe thunderstorm database maintained by NOAA's Storm Prediction Center (SPC). Underreporting and population biases present in the historical data are overcome using a combined approach of smoothing and data augmentation. The engineering component of the model relates event's intensity to physical damage in terms of damage functions, which provide an estimate of mean damage by wind speeds or hail impact energy. The financial module translates damage into financial loss. Ground-up losses are calculated by applying the appropriate damage function to the replacement value of the insured property. Insured losses are determined by applying the policy conditions to the ground-up loss estimates. The hazards component of the model has been validated against NOAA data and other published information. The modeled losses are in good agreement with historical losses reported by PCS. The AIR model accounts for the policy conditions specific to the United States and has been used widely in the insurance and reinsurance industries to meet the wide spectrum of risk management requirements.

INTRODUCTION

According to the National Weather Service (NWS) definition, a thunderstorm is called "severe" if it produces a tornado, winds of at least 58 mph (50 knots), and/or hail measuring a minimum of 0.75 inches in diameter.

Tornadoes, spawned from severe thunderstorms, are some of the most destructive forces in nature. The United States has the most tornado development of any country in the world [1]. They cause an average of 70 deaths, 1,500 injuries, and an estimated USD 1.1 billion in property damages per year in the United States. Hailstorms typically receive less attention than very damaging tornadoes [2]. However, on an annual basis, hailstorms cause more than USD 1 billion in property damage in the United States. The level of damage from hailstorms depends on a number of factors including hailstone size, the rate at which the hailstones fall, the storm duration, the size of the storm area, and the wind speeds that accompany the hailstorm. Significant property damage usually occurs when hailstones are golf-ball size and larger. High winds can increase the kinetic energy of hailstones and blow them at angles significantly off the vertical, thus increasing the likelihood of broken windows and cladding damage. Larger hailstones generally create more damage than smaller stones because they can fall at higher speeds (up to 100 mph). Straight-line winds exhibit no circular motion because they lack a central vortex but can have windspeeds that exceed 100 mph. The most intense, longer-lived, and large scale windstorms are called "derecho". These storms can be as significant as tornadoes and hurricanes in terms of damage and casualties [3]. The damage path of severe straight-line winds can extend across hundreds of miles and may include toppled trees, downed power lines, and destroyed homes and automobiles.

While losses from individual severe thunderstorms are generally not as large as those from individual hurricanes or earthquakes, on an annual aggregate basis, the losses from severe thunderstorms have accounted for approximately 50% of all U.S. catastrophic insured losses since 1990, according to Insurance Services Office's Property Claim Services (PCS). In 2006, 2007, and 2008, nearly 90%, 55%, and 40%, respectively, of total catastrophic insured losses in the U.S. were directly related to severe thunderstorms. In terms of state rankings of natural catastrophe loss potential, Texas severe thunderstorm risk follows only California earthquake risk and Florida hurricane risk.

AIR developed the first probabilistic severe thunderstorm model for the insurance industry to help companies manage severe thunderstorm risk proactively. The model can help insurers with underwriting decisions, to strategically price their policies, and to determine various reinsurance and risk transfer strategies. The AIR U.S. Severe Thunderstorm Model broadly comprises of three components, i.e., hazard, engineering, and financial as shown in Figure 1.

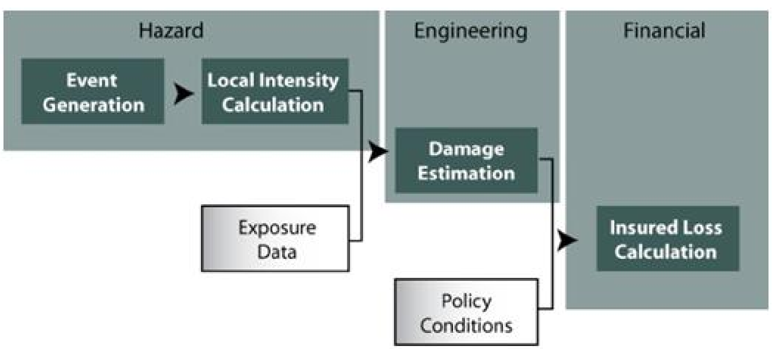

Figure 1. Components of AIR's Severe Thunderstorm Model

The paper is organized as follows: Section 1 provides additional details about severe thunderstorm risk in the United States. The hazard component of the AIR model is described in Section 2 while Section 3 presents the engineering aspect of the model. The financial module of the model is discussed in Section 4. The validation of AIR modeled losses is presented in Section 5. The paper ends with the discussion in Section 6.

1. SEVERE THUNDERSTORMS RISKS IN THE UNITED STATES

In the United States, severe thunderstorms occur nearly everywhere, but thunderstorm activity is particularly common in the Southeast, Great Plains, and the Midwest. This activity is in part due to the proximity of these regions to the Atlantic Ocean and the Gulf of Mexico. Also, a lack of major east-west mountain ranges causes the cool continental air mass to funnel down the Rocky

Mountains without impediment and collide with moist air from the Gulf of Mexico. Severe thunderstorm activity most frequently occurs during the spring and summer months in North America, although in southern California storms are more common during the winter and spring as a result of temperature inversions.

Tornadoes are common in an area termed “Tornado Alley”, which encompasses the lowland areas of the Missouri, Mississippi, and Ohio River Valleys and includes regions of Texas, Nebraska, Kansas, and Oklahoma [4]. Tornadoes have struck in all 50 states, however, and are more common in Florida than they are in Oklahoma. Tornadoes in Florida are generally weak, while those that have hit Oklahoma are some of the most violent on record. The southeastern states can experience tornadic activity during the winter months, but this trend shifts to the central U.S. during the months of March-May and to the Midwest during the warmer summer months. Storms producing hail occur most frequently where Colorado, Nebraska, and Wyoming convene. The most damaging hailstorm in the United States is the tri-state hailstorm in April 2001 which deposited 25-75 mm hailstones along a 585-km path resulting in insured losses of USD 2.2 billion at the time of catastrophe [5]. During the late spring and summer, severe windstorms are most common from the upper Mississippi to Ohio River valleys and from the mid-Mississippi River valley to the Southern Plains [6]. During September-April, severe windstorms are seen most frequently from eastern Texas to the southeastern U.S.

2. HAZARD

Severe thunderstorms are complex events, and their characterization involves an array of scientific variables. To model severe thunderstorms, AIR develops probability distributions for critical storm parameters such as annual frequency, event intensity, and size of the storm areas, testing them for goodness-of-fit and robustness. These probability distributions are then used in combination with the SPC data to produce a “seed set” of thunderstorm events that form the basis for the stochastic simulation. The simulation generates scenario years of event activity that represent a potential year of catastrophe experience. The AIR model allows for a variable number of events in each simulated year, just as might be observed in an actual year. Thousands of these scenario years are generated producing a large stochastic catalog that includes a complete range of potential annual incidents and full coverage of extreme events.

2.1 HISTORICAL DATA AND EVENT DEFINITIONS

Historical data on tornadoes, hailstorms, and straight-line windstorms are available from the severe thunderstorm database maintained by NOAA’s Storm Prediction Center (SPC) [7]. This database includes information on more than 47,500 tornadoes, 210,000 hailstorms, and 240,000 straight-line windstorms for the period 1955 to 2006. The database provides information about individual tornadoes, hailstorms, or straight-line winds that have been reported within each year. These individual occurrences are referred to as microevents by AIR. However, a severe thunderstorm event from the perspective of insured loss is a collection of tornadoes, hailstorms, or straight-line windstorms occurring over the course of one or more days and resulting from the same atmospheric event or frontal system. These events or storm clusters, referred to as macroevents by AIR, are identified by analyzing the geographical and temporal patterns of individual microevents using GIS mapping capabilities as well as clustering techniques.

2.2 DATA AUGMENTATION AND SMOOTHING

Despite the abundance of historical data, significant underreporting of microevents makes it difficult to fully characterize severe thunderstorm risk. The underreporting is evident from Figure 2 that shows the number of reported hailstorms, straight-line windstorms, and tornadoes for the period 1955 to 2002. The most severe underreporting exists for hail and straight-line windstorms, and both show an almost exponential growth in report numbers over time.

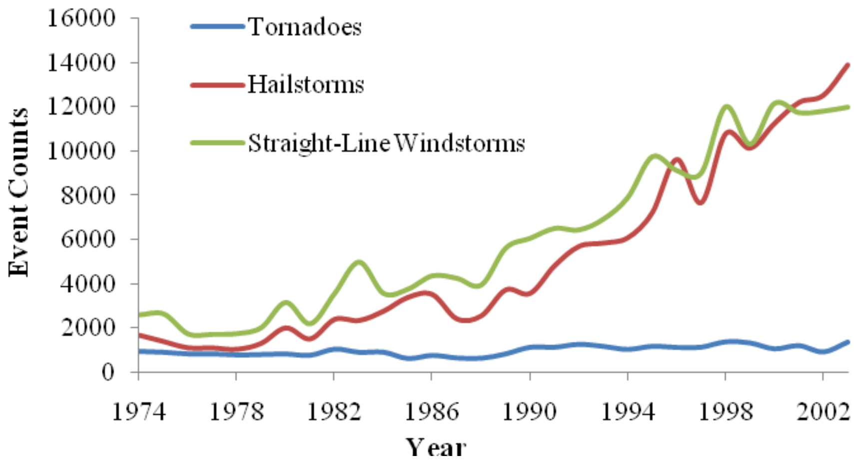

Figure 2. Number of Reported Microevents, 1974 - 2003

In addition to inconsistencies over time, there are also geographical differences in the level of reporting. In particular, there is a positive correlation between the number of reported events and population density, suggesting that many events may have gone undetected in areas where the population is low.

In order to compensate for underreporting, AIR employs a combination of kernel smoothing and data augmentation techniques to get a more realistic assessment of the true occurrence rates of tornadoes, hailstorms, and straight-line windstorms. By statistically smoothing the historical data, the model allows simulated storms to occur where no historical storm has been recorded. However, since the smoothing tends to lower the peak frequencies in areas where the reporting is believed to be complete, a statistical data augmentation technique is used to restore the peak frequencies. The smoothing and data augmentation are performed iteratively and continues until a credible probability surface is obtained.

2.3 DATA SIMULATION AND EXAMPLES

The grouping of tornadoes, hail, and straight-line wind micro events into severe thunderstorm event, followed by the measures undertaken to correct the data for underreporting, produces a large set of historical events that is used as a seed set for the stochastic simulation. Simulated events are generated using a smooth bootstrap procedure [8] that randomly samples from this seed set and spatially perturbs each selected event. The procedure preserves the seasonal and spatial distribution of simulated events and produces an accurate depiction of severe thunderstorm risk.

A simulated severe thunderstorm event and its underlying historical event before data augmentation and smoothing are shown in Figure 3. The left-hand side of Figure 3 shows a historical event that occurred from April 16 through April 19, 1995. The figure on the right shows the same event augmented and then perturbed to produce a potential future severe thunderstorm event.

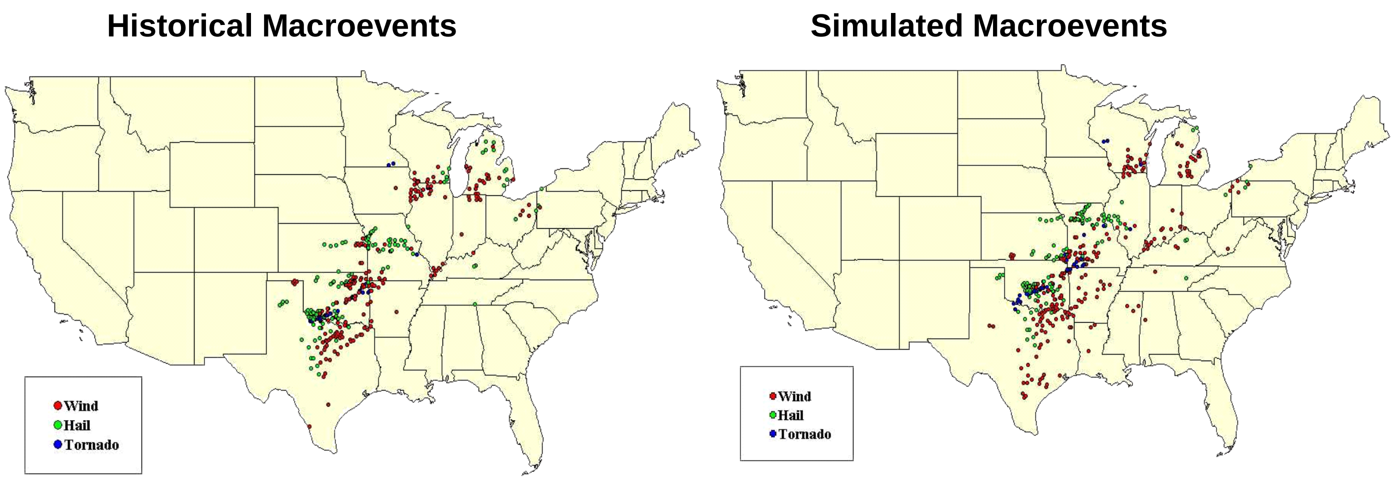

Figure 3. Historical vs. Simulated severe thunderstorm events, April 16-19, 1995

3. DAMAGE ESTIMATION

Potential losses are estimated using damage functions, which relate event intensity to physical damage in terms of the mean damage ratio. The intensity variables are wind speed for tornadoes and straight-line windstorms, and hail impact energy for hailstorms. The mean damage ratio is the ratio of repair costs of a property to its replacement value. As tornadoes, hailstorms, and straight-line windstorms inflict damage differently, separate damage functions are developed for each peril. Damage functions are available for building, contents, and time-element coverage losses. Because different structural types experience different degrees of damage during an event, damageability relationships vary according to construction materials and occupancy.

3.1 TORNADO DAMAGE FUNCTIONS

In the AIR model, tornado damageability is a function of the fastest quarter-mile wind speed. Due to variations in potential damage, the model uses separate functions for each combination of construction and occupancy classes. Roofs are often the first part of a building to be damaged by tornado winds, as once a single shingle is removed, neighboring shingles can easily be penetrated and lifted. Tornado winds can peel off unsecured slates, roll metal roofs, and damage windward overhangs and eaves. Failure of the roof system may weaken lateral support of walls, contributing to their collapse. Figure 4 shows relativities for AIR tornado damage functions for commercial buildings of various constructions, including buildings of unknown construction type.

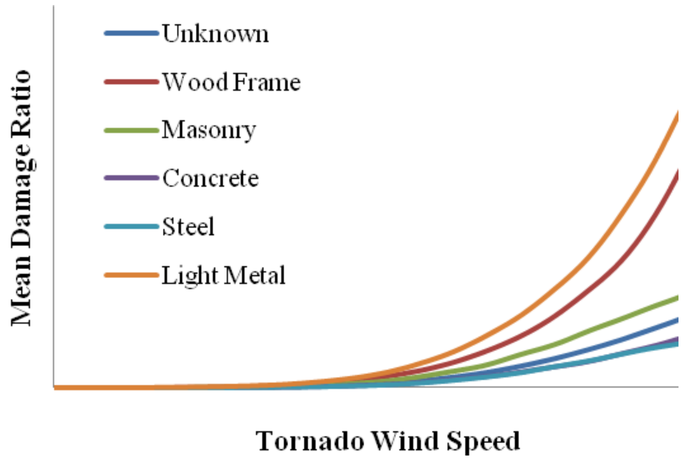

Figure 4. Building Damage Functions for Select Construction Types

3.2 Hail Damage Functions

The severity of hail damage depends on hailstone size, accompanying wind speed, the spacing between hailstones, the hailstorm area, and storm duration. Building roofs start to experience damage when hailstones approach three inches in diameter, and significant property damage usually occurs when hailstones are of golf-ball size and larger. High winds can increase the kinetic energy of hailstones and blow them at angles significantly off the vertical, thus increasing the likelihood of broken windows and cladding damage. The AIR hail damage function is a function of hail impact energy which represents the energy with which a hailstone strikes an object. Different construction and occupancy types behave differently with respect to hail impact energy; therefore, AIR develops different damage functions for different construction and occupancy types including automobiles which are generally more vulnerable than buildings to hailstorm damage.

3.3 Straight-line Winds Damage Functions

In the AIR model, damageability for straight-line windstorms is a function of 3-second gust speed and storm duration. Building damage from wind originates at the weak point in a structural system, leading to a sequential pattern of structural failures. As each connector or fastener is pulled from uplift, loads are transferred to the next point of vulnerability. The longer the duration of high winds, the longer this process will continue and the greater the resulting damage will be.

3.4 Contents Damage Functions

The AIR model considers damages to contents separately, with the damage ratio being the dollar loss to the contents divided by the replacement value of the contents. Contents vulnerability is a function of building construction, building damage, occupancy, and wind speed; occupancy in particular is crucial to the development of damageability relationships as it provides information on the contents likely present and their potential vulnerability.

3.5 Damage Functions for Time-Element Coverage

For time-element coverage, the damage ratio represents the per-diem expenses or business interruption losses associated with the expected number of days that the building is uninhabitable (residential structures) or unusable (commercial structures). In the AIR model, time-element damageability is a function of the mean building damage, the time required for repair or reconstruction, and occupancy. The functional relationship between building damage and loss of use is based upon published construction and restoration data and expert engineering judgment.

3.6 Uncertainty around the Mean Damage Ratio

Separate damage functions for each of building and contents provide estimates of the *mean*, or expected, damage ratio. However, these estimates are in fact sampled from full probability

distributions that allow for non-zero probabilities of zero percent and one hundred percent

damage. As is commonly seen in the course of damage surveys conducted in the aftermath of actual events, there can be a wide range of damage to structures of similar construction for the same wind speed or hail impact energy, including buildings that are untouched, as well as buildings that are completely destroyed. This variation in building damage can arise due to the inherent randomness in building response, or to differences in building characteristics, construction materials or workmanship. The uncertainty in damage is captured in the model by these probability distributions around the mean damage ratio.

4. Financial

The damage functions along with exposure are used to produce, for each event, a distribution of ground-up loss by location and coverage. Limits, deductibles and reinsurance are applied in the financial module to the ground-up loss distribution to produce gross and net loss estimates. The distributions are applicable to the analysis of a single exposure and usually have a high degree of uncertainty. The individual distributions are combined to obtain the portfolio distribution, where the uncertainty is lower. Convolution is a mathematical procedure that can be used to derive the probability distribution of the sum of multiple loss distributions. The convolution of two random variables F and G with density functions f(x) and g(x) respectively is represented by the equation,

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-t)g(t)dt$$

where, t is a dummy variable.

The AIR model employs an efficient and accurate numerical algorithm for “convolving” any number of non-parametric loss distributions. Extreme care has to be taken to deal with distributions with differing size of loss. This technique allows the correct representation of the shape of the loss distributions throughout the financial loss estimation process. Preserving the right shape is particularly important when insurance terms apply to the “tails” of the distributions.

The AIR model output may be customized by line of business, including construction class and coverage, and to any desired degree of geographical resolution, down to location level. The model also provides summary reports of exposures, comparisons of exposures and losses by geographical area, and detailed information on potential large losses caused by extreme tail events.

5. Validating Modeled Losses

The losses produced by the AIR model are extensively validated against observed loss estimates reported by PCS and actual claims data provided by clients. For a given event, the loss estimate issued by PCS represents the total losses from all event perils; the AIR model similarly reports losses for all three perils combined.

In Figure 5, AIR losses are compared to actual losses reported by PCS by line of business, i.e., personal, commercial and auto. PCS losses are adjusted to 2008 dollars. It should be noted that PCS losses have been reported by line of business since 1998.

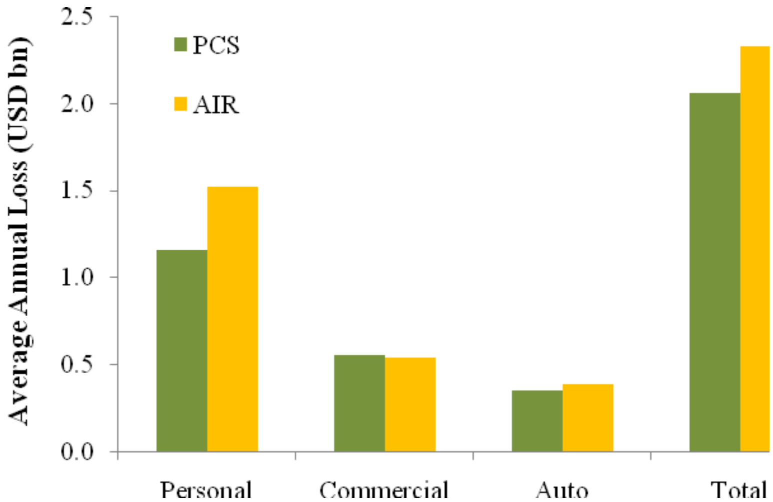

Figure 5. PCS vs. AIR Average Annual Occurrence Losses, 1998-2008

6. Discussion

Severe thunderstorms, often characterized by tornadoes, hailstorms, straight-line winds, are capable of causing substantial loss to life and property resulting in large insured losses. Since 1990, the annual aggregate losses from severe thunderstorms account for approximately 50% of all U.S. catastrophic insured losses. AIR developed the first probabilistic model for the insurance industry to manage severe thunderstorm risk proactively. Each component of the model is developed using published studies and internal research and validated against SPC and published data. The modeled losses are in good agreement with the actual losses reported by PCS. Despite the frequent nature of severe thunderstorms and the magnitude of potential losses, many insurers still treat them simply as a cost of doing business. But high resolution models now exist that can capture the risk at a highly localized level, enabling effective and proactive risk management practice.

Acknowledgement

The authors acknowledge help and support provided by the research division at AIR Worldwide Corporation.

REFERENCES

[1] C. A. Doswel III, Societal impacts of severe thunderstorms and tornadoes: lessons learned and implications for Europe, Atmospheric Research, vol. 67-68, 2003. p. 135-152.

[2] G. M. Morgan, Jr., P. W. Summers, Hailstorms and hailstorm characteristics, Thunderstorms: A Social, Scientific and Technological Documentary, Vol. 12, Thunderstorm Morphology and Dynamics, E. Kessler, U.S. Govt. Printing Office, Washington DS, 1982, p. 363-408.

[3] W. S. Ashley, T. L. Mote, Derecho hazards in the Unites States, Bull. Amer. Meteor. Soc. vol. 86, 2003, p. 1577-1592.

[4] H. E. Brooks, C. A. Doswell III, M. P. Kay, Climatological estimates of local daily tornado probability for the Unites States, Weather and Forecasting, vol. 18, 2003, p. 626-640.

[5] S. A. Changnon, J. Burroughs, The tristate hailstorm: The most costly on record, Monthly Weather Review, vol. 131,2003. p. 1734-1739.

[6] R. H. Johns, W. D. Hirt, Derechos: Widespread convectively induced windstorms, Weather and Forecasting, vol. 2, 1987, p. 32-49.

[7] J. T. Schaefer, R. Edwards, The SPC tornado/severe thunderstorm database, 11th Conf. on Applied Climatology, Dallas, TX, Amer. Meteor. Soc., 1999, p. 603-606.

[8] B. Efron, R. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall/CRC, 1994.