RiskLink®
RiskBrowser®
11.0

Principles of Model Validation: United States Hurricane Model

September 17, 2012

Risk Management Solutions, Inc.

7575 Gateway Boulevard, Newark, CA 94560 USA
http://www.rms.com

Warranty Disclaimer and Limitation of Liability

RiskLink® RiskBrowser®11.0 Principles of Model Validation: United States Hurricane Model.

September 17, 2012. Printed in the U.S.A.

Licenses and Trademarks

Contents

Introduction

Catastrophe models integrate scientific knowledge of the underlying phenomena, engineering principles governing the performance of buildings and other elements at risk, and financial and actuarial models, to derive the potential financial, human and economic consequences to different affected groups.

As catastrophe model output has grown to reside at the very core of insurance and reinsurance capital management and transactions, so has the importance of understanding the uncertainties inherent in catastrophe risk and models, and an increasing recognition that learning is ongoing. The industry is calling for more openness into model assumptions to help insurers and reinsurers better understand the decisions made relative to those assumptions, adapt faster to new information, and take ownership of their view of risk. RMS is committed to providing the tools and model transparency necessary to help users establish more resilient risk management strategies based on a full understanding of all aspects of catastrophe risk, and explicit consideration of the implications of model uncertainty on their portfolios.

In keeping with this principle, RMS encourages catastrophe model users to conduct their own validation of catastrophe model output, with the awareness that not all model validation metrics available in the market are scientifically valid. This paper discusses the primary limitations of some erroneous validation metrics, in order to help catastrophe model stakeholders better understand various validation practices.

Using examples from the version 11.0 RMS® U.S. Hurricane Model, released in 2011, this paper also demonstrates the principles underlying robust validation techniques, and raises awareness of potential pitfalls that can occur during the model validation process. While this document equips catastrophe model users with a framework they can use to evaluate the validation metrics of a catastrophe model, it is not designed to provide a comprehensive validation of the U.S. Hurricane Model itself. Licensed users of RMS models can access a full suite of regional validation documents for more information on specific RMS models.

Validating the RMS U.S. Hurricane Model

The RMS U.S. Hurricane Model comprises three individual models:

- The baseline model, which is calibrated to the long-term historical event record (the North Atlantic hurricane database or HURDAT)
- The storm surge model, which calculates the expected surge damage and loss from each stochastic event in the baseline model
- The medium-term rates forecast model, which provides an additional set of output based on forecasts of hurricane event frequency levels over the next five years

The choice of benchmark data used for validating model output must be carefully made, depending on which of the above three models is being validated. The 2011 upgrade of the RMS U.S. Hurricane Model (version 11.0) includes updates to the baseline hurricane hazard model, which are the predominant drivers of changes in hurricane risk for many insurance portfolios. Appropriate historical event data can be used to validate the baseline model.

The baseline model can be further broken down into "components" that are individually validated. Although examples of this process are described in section Component Validation on page 10, it can be challenging for model users to perform component-level validations, and therefore RMS provides licensed users with detailed documentation of this process.

However, model users can conduct loss validation tests that compare for both the industry as a whole, and individual insurance company portfolios, modeled and reported losses from historical events. This paper provides guidance in performing such overall loss comparisons in the section Overall Loss Validation on page13. Such tests include both insurance company portfolio losses, described in the section Portfolio Loss Validation on page 13, and industry loss comparisons, described in the section Industry Loss on page 14.

In addition to the baseline model updates in the version 11.0 U.S. Hurricane Model, the medium-term rates forecast model was also updated, to reflect the potential impact of sea surface temperature changes on average hurricane activity levels over the next five years. However, although the medium-term rates forecast methodology can be compared with historical data, by its nature, this is a forecasting model, and results are therefore are forward-looking and not designed to represent the past. Such comparisons are therefore not considered in this paper.

Validation Considerations

To validate a model, RMS performs a number of different validation tests that compare model components and output with comparable benchmarks, taking care to compare "apples-to-apples". The significance of each test is then carefully evaluated to avoid drawing false conclusions.

The following five points should be considered when evaluating validation techniques.

1. Catastrophe models are designed to extrapolate beyond a limited historical record in a rational and consistent manner.

Catastrophe models provide a representation of complex physical processes. They consist of multiple components (e.g., stochastic, hazard, and vulnerability), each characterizing a unique aspect of the overall process, calibrated independently, as well as together. The models are designed to produce a range of possible outcomes beyond those indicated by the historical record, which in the case of hurricanes is limited to approximately 110 years. Catastrophe models are particularly applicable to modeling low-frequency and high-severity catastrophe losses, as the historical record for such extreme events is often quite limited and potentially flawed. However, in spite of the lack of completeness associated with historical losses, it is always necessary to validate catastrophe models against the historical record to ensure consistency with actual observations. RMS carefully validates all of its models with reference to all available data - both by component and overall - before considering any model to be acceptable for release.

2. Calibration of model components must balance assumed physical relationships with signals and patterns found in observation records.

While catastrophe models should be consistent with known physical principles related to the underlying hazard, such principles may be countered by strong signals in the observation record; such signals cannot be ignored when the underlying physical theory is uncertain or incomplete. For example, intuitively we expect the frequency of stronger storms to be lower than weaker storms, but the 110 years of historical records in the North Atlantic hurricane database (HURDAT) show that in the southeast quadrant of Florida, the frequency of Category 2, 3, and 4 storms is nearly constant and does not exhibit a trend of decreasing frequency with increasing intensity. Unlike other model techniques based on localized statistical distributions, RMS' unique basin-wide track modeling process allows the U.S. Hurricane Model to reflect this historical pattern while also adhering to physical relationships and expectations, to ensure consistency with both physical principles and the historical record.

3. Hurricane model enhancements are based on long-term research programs designed to improve the modeling process.

RMS invests in substantial research to supplement analysis based on the historical record. For example, in 2008 RMS initiated a three-year project into post-landfall weakening rates that produced additional insights into factors that affect how quickly storms weaken (or fill) after landfall. As described in section Example 2: Long-Term

Research Improves Modeling of Inland Filling on page 11, RMS partnered with various academic centers of excellence[1], to publish the validated conclusions in a peer-reviewed academic paper. RMS does not adjust models in reaction to single events without additional research, thus it can take some time after events to release model updates.

4. Catastrophe models should be validated against appropriate measures of historical experience at both the component level and complete model level.

Comparisons to industry losses are a vital component of catastrophe model validation. Three primary industry loss experience benchmarks are commonly used to validate overall model performance: historical event losses, industry average annual loss (AAL), and the exceedance probability (EP) curve.

Each of these overall model validation benchmarks has its limitations, and taken individually cannot be used to conclude that a model is reasonable. However, collectively they form part of the overall validation framework.

Because historical industry loss experience is not comprehensive enough to conclusively validate the model as a whole, RMS spends a significant amount of time and resources to validate model components individually, using the science and data specific to each component. RMS publishes the results of its component and industry level validations to its client base with full transparency.

5. Model users should be aware of the potential for inappropriate model validation metrics.

While RMS encourages users of catastrophe models to conduct their own validation of catastrophe model output, RMS has observed the use of some model validation metrics that may be incomplete or misleading. These include real-time validation, incomplete comparison to public datasets, and erroneous conclusions derived from statistics based on the limited historical record. Some limitations of these other types of validation metrics are discussed in Misleading Validation Comparisons on page 23.

[1] RMS partnered with Tim Hall - Senior Scientist from NASA Goddard Institute for Space Studies and Dave Nolan, Associate Professor at University of Miami

Types of Model Validation

During the model development process, RMS uses a range of tests to validate model output at the individual component (e.g., the stochastic event, hazard, and vulnerability components) and overall model level. RMS models are thus designed both to capture individual processes accurately, such as how a hurricane strengthens and weakens over its lifecycle; and produce losses consistent with historical experience. RMS validation tests range from individual component-level validation to overall model validation tests on loss statistics. This section discusses the general types of validation tests RMS employs, and the limitations associated with each type.[2]

Component validation focuses on ensuring that model components provide reasonable results given specific inputs. At the component level, validation tests are typically run on the variables output by the component (e.g., wind speed from the hazard component), whereas at the overall model level, the model's loss output is tested.

As shown in Table 1, validation tests can be further characterized into those that test specific values for individual events, and those that test distributions of values across a set of events. In the latter case, it is generally the distributions resulting from the historical event set that are tested against the distributions resulting from the model's stochastic event set. The purpose of each type of validation test is described in the following table.

Each test provides an additional degree of confidence in the model. RMS considers a peril model reasonable when it passes all six validation tests.

For the v11.0 U.S. Hurricane Model, RMS model developers deliberately place more emphasis on the first three types of validation—the two component validation tests and the portfolio validation. The last three types of validation based on industry losses are primarily used to confirm that the model is generally reasonable, given the uncertainty in these methods. Relying too heavily on industry loss validations opens the process up to errors, due to uncertainties in producing credible industry loss estimates for events far back in time. The following sections explore the six validation tests in the context of the U.S. Hurricane Model.

[2] RMS documents detailed descriptions of all model validation tests performed. Licensed clients can access this documentation on the client portion of the RMS website.

Table 1: Types of Validation Tests for Catastrophe Models

Validation Level	Validated Element	Description
Component Validation	1. Event-Based Component Output Variables	Ensures that variables predicted by a component match observed values for specific historical events. For example, a model component that simulates storm filling rates is compared to rates calculated independently from HURDAT observations for specific storms (see page 10).
	2. Historical Distributions of Component Output Variables	Validates the distribution of component output variables against historical observations or other independent analyses. For example, the output of the hazard model represented as 100-year return period wind speeds is validated against the 100-year design wind speed maps used in U.S. building codes.
Overall Validation	3. Portfolio Validation (Historical Portfolio Event Loss)	Validates the overall model losses by comparing insurance--company claims data with modeled loss estimates for the underlying insurance-company portfolios. The historical event losses are modeled using reconstructed hazard footprints and the modeled losses are compared to the actual claims. The strength of this validation test lies in the large number of data points (many companies for each event) and the match of exposures underlying the companies' loss experience (see page 13).
	4. Historical Industry Event Loss	Compares modeled industry losses produced using an RMS industry exposure database (IED) to trended, reported market-wide losses to validate that the overall industry losses are well represented by the modeled loss. The industry loss observations are trended to reflect changes in exposure concentrations over time in the footprint of the event. Due to uncertainties in both trending and reported loss values, this test is limited to events from the last 20 to 25 years (see page 15).
	5. Historical Industry Average Annual Loss (AAL)	If the model passes both overall historical and industry loss validation tests above, it can be assumed that model losses based on the IED and reconstructed event footprints represent a good proxy for industry losses. Using reconstructed footprints for the last 110 years, we can thus establish a good proxy for industry losses for the entire HURDAT event history. This historical industry loss proxy can then be used to validate the model's stochastic event set by comparing the AAL of the historical industry loss proxy with the AAL of the model (see page 17).
	6. Historical Industry Exceedance Probability (EP)	Similar to the historical industry AAL, the proxy for industry losses above can also be compiled into an exceedance probability (EP) curve and compared against the model output (see page 22).

Component Validation

Catastrophe models are developed as a series of independent event, hazard, vulnerability, and loss components that together represent physical relationships and financial impacts of events on insured exposure. Each component is calibrated and validated independently and tested against real experience. The full scope of component-level validation is described in detail in RMS documentation and made available to licensed clients. In addition, for the U.S. Hurricane Model, material provided by RMS in its submission to the Florida Commission on Hurricane Loss Projection Methodology (FCHLPM) [1] illustrates a number of tests used to validate hurricane model components.

RMS uses component validation extensively throughout the model development process. This section uses two examples to highlight some of the methodological choices made in the development and validation of certain hurricane model components.

The first example demonstrates the process by which developers decide whether patterns observed in the historical record are significant enough to be included in the baseline model or should be discarded as statistical noise, using the example of landfall rates in south-east Florida..In the second example, RMS shows how new research can improve an existing model, using the example of a research project into inland weakening (filling) rates.

Example 1: Baseline Model Landfall Rates—Localized Versus Regional Patterns

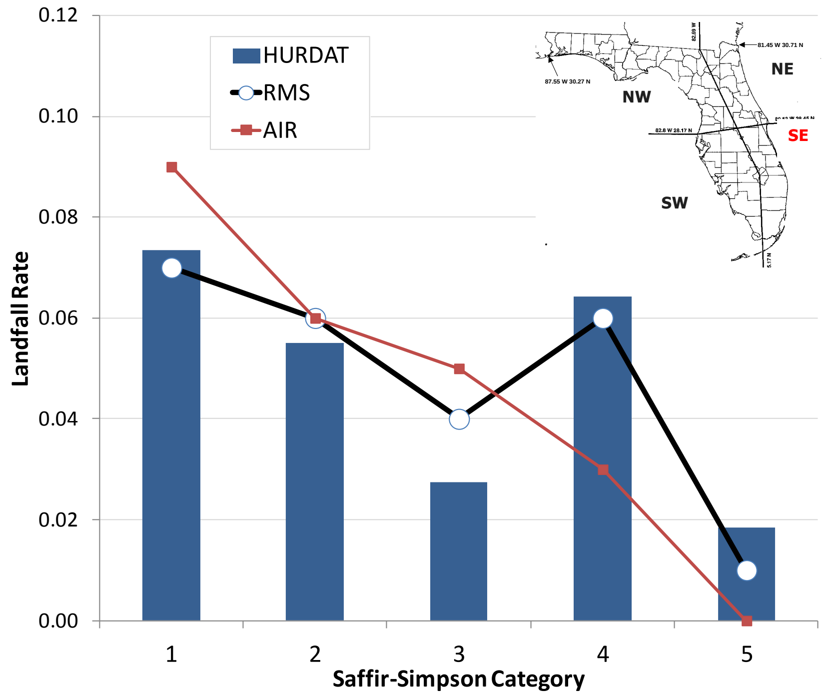
Catastrophe models balance physical theory with significant observational trends. Occasionally the observational trends are at odds with physical theory, and in such cases modelers must make finely judged decisions that weight the observation versus the theory; such decisions may significantly affect the model outcome.

For example, the v11.0 U.S. Hurricane Model development process balanced consistency with known regional physical principles, such as decreasing hurricane frequency with increasing latitude, and rates that are inversely proportional to intensity, with localized patterns observed in historical datasets.

The following example illustrates how this process has material consequences for modeled hurricane landfall rates in south-eastern Florida. Although basic theoretical considerations lead us to assume that hurricane rates should be inversely proportional to intensity, in south eastern Florida, the reverse is observed; the historical HURDAT record shows that the Category 4 landfall rate is higher than for Category 2 and 3 hurricanes.

In a region with a low frequency of hurricanes, it could be reasonable to attribute such a signal to "sparseness of data," thus discrediting the pattern. However, south-eastern Florida has the most extensive historical dataset in the U.S., due to its high frequency of hurricane landfalls. Rather than discount the observed pattern, RMS examined the physical relationships in more detail in this region, and discovered several factors that support the presence of a greater proportion of Category 4 storms landfalling in southeast Florida, giving further credibility to the historical observations. Specifically, published research [5] indicates that in this region, water depth and the presence of land masses (the Greater Antilles and Bahamas) dividing the Caribbean and Atlantic Oceans could alter the intensity of certain types of storms approaching this quadrant of Florida and thus affect hurricane intensity distributions. Thus, the effect of these air-ocean interactions on relative landfall frequency in southeast Florida may provide a physical explanation for the observed pattern.

Figure 1: Comparison of RMS and AIR landfall rates with HURDAT rates for the southeast quadrant of Florida (as reported in RMS and AIR submissions under 2009 FCHLPM standards) [1,2]



Given that these regional physical considerations provide a theoretical framework for the observed landfalling rates in Florida, RMS concluded for v11 that this signal is robust enough to be considered a real phenomenon and should not be treated as a statistical anomaly. Figure 1 shows how the RMS model thus follows the observed patterns in this quadrant of Florida unlike the AIR model. Notably, the same pattern is not present further north, despite a similar amount of data.

For portfolios with risk in this region, RMS concludes that an accurate adherence to historical patterns is prudent to avoid a potential underestimation of the probabilities of Category 4 or higher storms.

Example 2: Long-Term Research Improves Modeling of Inland Filling

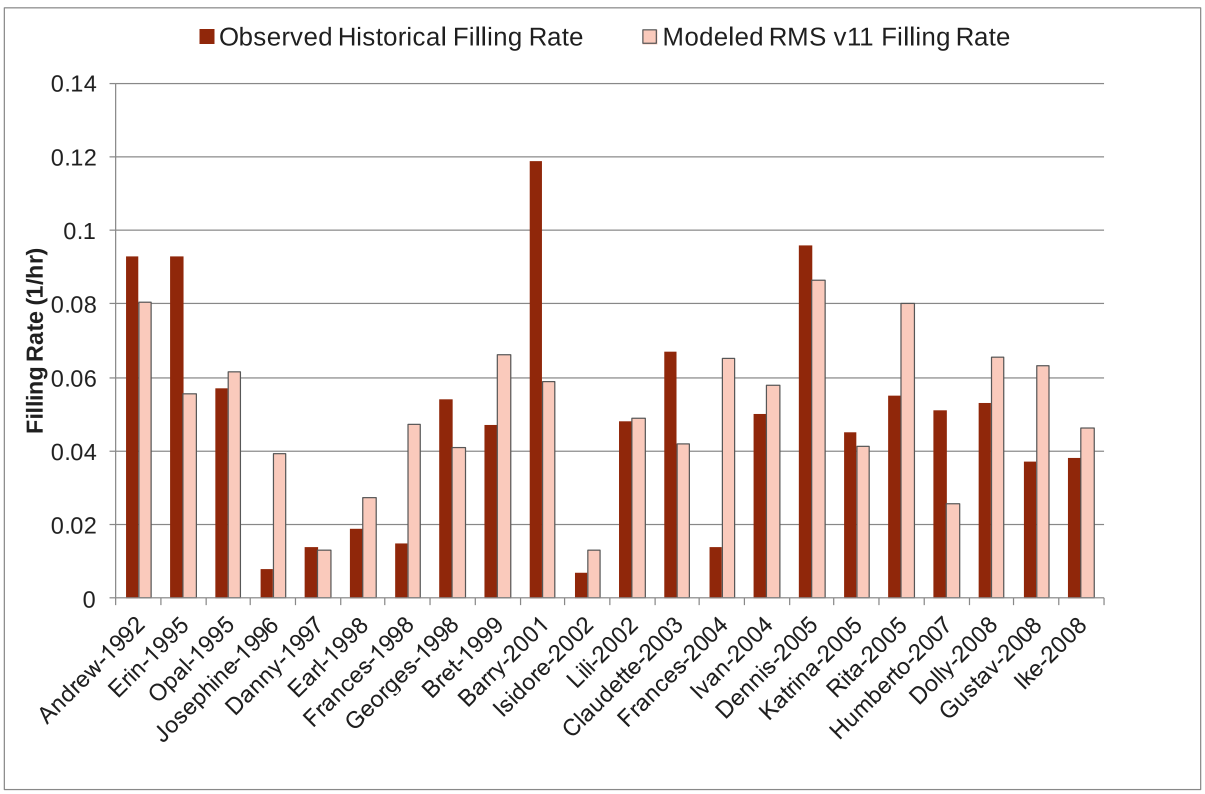
One of the key components updated in the v11.0 U.S. Hurricane Model involved the representation of how hurricane pressures "fill in" and wind speeds decay after they make landfall. After landfall, the storm loses its source of energy from the ocean, and in most cases the central pressure begins to increase or "fill." As the storm weakens, wind speeds decrease. While the wind speed at any location is also influenced by the surface roughness of that location and upwind locations for up to 50 miles, a significant determinant of wind speeds is the rate of this inland filling after landfall.

As recorded data on the factors that influence the inland filling process were very limited, in early 2008, RMS launched a fundamentally new study of the phenomena of inland decay rates—a process that involved just over three years of research and development, working closely with the academic community. The results of this study have been published in a peer-reviewed article in Monthly Weather Review [6]. Soon after initiating the project, Hurricane Ike caused unexpected levels of loss in Texas, reinforcing the already present need to increase knowledge about inland filling.

The RMS study [6] uses numerical weather modeling to analyze how storm decay relates both to characteristics of the hurricane's structure and to the influence of different terrains, and shows that up to eight parameters influence the rate of inland filling. Storms that predominantly travel over rough terrain will, on average, fill faster than those that travel over smooth terrain. Regional differences in terrain require different filling models to capture rates of hurricane decay more accurately, even if the storm's landfall characteristics are the same. Thus, the version 11.0 hurricane model includes different filling rate formulations for different regions to allow, for example, Florida east coast hurricanes to fill differently than Gulf Coast hurricanes.

Figure 2 illustrates how RMS validated the inland filling component against a selection of specific historical events from which the observed filling rate can be derived. This example demonstrates how this new component has the ability to reproduce a variety of filling rates for different hurricanes without systematic bias, and can reproduce historical filling rates across a range of hurricanes more accurately than hurricane models that do not use this new information.

Figure 2: Component validation of version 11.0 inland filling rates against specific key historical events in Florida from which the observed filling rate can be derived.



Overall Loss Validation

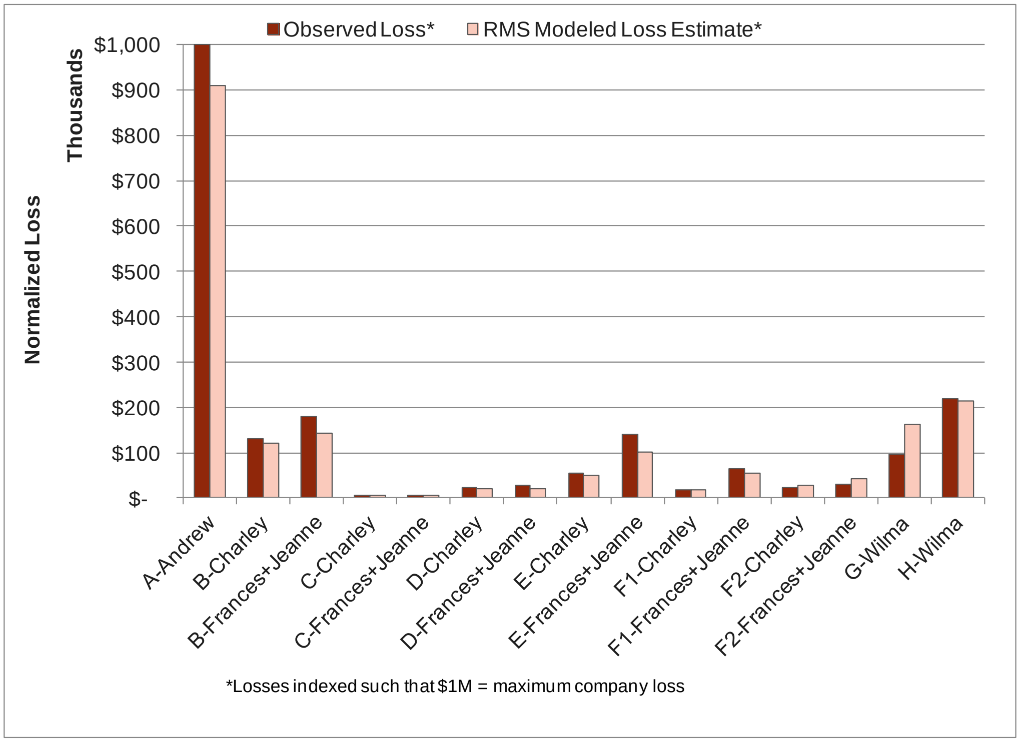
Catastrophe models are designed to estimate expected losses, and therefore a critical part of any model development is validation of the output of the whole model (losses) against various benchmarks (items 3 to 6 in Table 1). This process of overall model loss validation is generally split into two types—portfolio loss validation and industry loss validation.

Portfolio Loss Validation

A key validation of overall model performance is obtained by comparing insurance company portfolio claims data with modeled losses from historical events (item 3 in Table 1). This overall portfolio loss validation test confirms that the model's historical footprints and vulnerability components are collectively able to recreate observed losses from actual insurance portfolios. An example of an overall validation test for historical event portfolio losses is shown in Figure 3. In the figure, each individual insurance company is identified with a letter, while each storm is identified by name. Thus company F has provided data for four events, while losses from hurricanes Frances and Jeanne are shown for five companies. This figure demonstrates that modeled portfolio losses for hurricane events compare well with reported claim values for those same events.

In portfolio loss validation tests, individual insurance company exposure databases are used as model input to produce model losses comparable to the company's actual event loss experiences. The strength of this test is that it provides multiple data points for each event, as many companies are able to provide such data for a number of events; and that all modeled losses are based on exposures at the time of the event, eliminating the need to make trending adjustments to the exposure or actual loss. The model passes this validation test if the modeled losses are well correlated with actual losses without systematic bias to over or under-predict actual losses.

Figure 3: Comparison of actual company losses with modeled losses for historical hurricane events



Industry Loss Validation

Comparison of model output to recent incurred losses at the industry level is another validation method, although it should be noted that industry loss validations have greater uncertainty than component and portfolio level validation. It is important for model users to ensure that they are comparing relevant information, as several potential pitfalls exist with this test.

RMS' suggested best practice for validating a model against industry losses involves comparing model loss output to three benchmarks (items 4 to 6 in Table 1):

- Historical Industry Event Loss Validation: Adjusted industry loss observations for individual events over the last 20-25 years (item 4 in Table 1)
- Historical Industry Average Annual Loss Validation: The historical AAL derived from a proxy for 110 years of industry event losses, based on modeled output from an industry exposure database and reconstructed historical event footprints (item 5 in Table 1)
- Historical Industry Exceedance Probability (EP) Validation: The implied historical exceedance probability curve based on the proxy for 110 years of industry event losses and a simple event frequency assumption (item 6 in Table 1)

Historical Industry Event Loss Validation

The first step is to compare reported industry losses, adjusted for trends in exposure density, to reconstructed event model losses. Individual storm comparisons may be variable, but overall across a wide variety of storms such comparisons are expected to show that the model has no discernible biases. Because the number of homes and values exposed to storms has changed over time, the reported losses must be adjusted or trended to be on an "apples-to-apples" basis before making comparisons to the model output, which is based on a present-day representation of insured exposure. As these adjustments are subject to increasing uncertainties for events further back in time, RMS advises that no more than the past 20-25 years should be used when comparing modeled with reported industry losses.

Industry Loss Estimation Errors and Exposure Trend Adjustments

A standard source of historical loss data used for industry loss validation in the U.S. is provided by ISO's Property Claim Service (PCS), which has collected industry losses since about 1950. However, it should be noted that PCS does not collect data on the exposures (e.g., number of buildings) associated with the industry loss observations. For this reason, it is very difficult to create retrospective exposure datasets that can be used to derive representative model losses for the time of the event. In order to use these historical loss observations, RMS instead applies a trending model to the reported losses to derive what would be expected given today's exposures, and uses the current RMS industry exposure database (IED) to produce comparable model losses. Trending methodologies must account for both changes in inflation but also the growth in the number of buildings and their relative unit value over time. RMS has developed its own trending factors and methodologies, which builds on and advances the methodology used in Pielke et al., 2008 [3]. The RMS trending factors are more directly applicable to the insurance industry, and thus to catastrophe model output.

Even with this advanced methodology, uncertainties in trended losses will still be greater than the portfolio validation methods, especially for older storms. This is due to uncertainty in the trending factors as well as uncertainty around the reported event loss in the year that it occurred.

To better understand the uncertainty in reported losses, it helps to review the way in which PCS estimates are compiled. The PCS methodology involves sampling loss amounts from participating ISO insurers and extrapolating these based on market share or other proprietary data. PCS does not report the proportion of market share upon which they base this extrapolation, nor do they publish any information to establish credible ranges around the estimates. These ranges can be established by comparing PCS to other sources. Following major events, a state department of insurance may conduct a data call to understand the overall industry loss. The advantage of a department of insurance data call is that participation from all insurance companies will cover 90% of the industry or more. Table 2 compares PCS and Florida Office of Insurance (FL-OIR) estimates for some of the Florida storms that occurred in 2004 and 2005. Note that the FL-OIR reported losses are sometimes several billion dollars higher or lower than the reported PCS estimates, indicating that PCS estimates may be under or overestimated by as much as 50%. RMS developers place a higher degree of credibility on the FL-OIR benchmarks; however, this type of data is not routinely available for all events, and is often limited to only the admitted (i.e., residential) market.

Table 2: Comparison of Florida residential industry loss estimates from ISO's Property Claim Service (PCS) versus data calls conducted by the Florida Office of Insurance (FL-OIR).

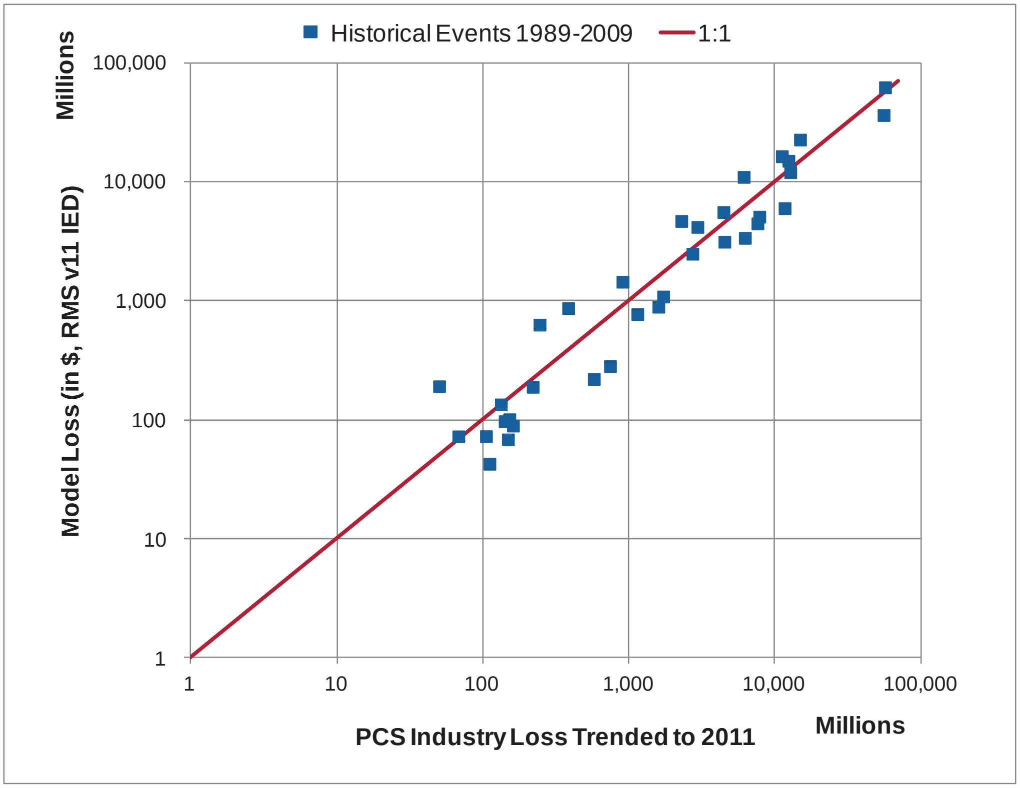
Hurricane	Year	PCS ($B)	FL-OIR ($B)	Percent Difference from PCS
Charley	2004	7.65	10.24	+34%
Ivan	2004	5.04	2.66	-47%
Jeanne+Frances*	2004	8.69	13.78	+59%
Wilma	2005	10.91	7.70	-29%
Katrina	2005	0.59	0.56	-5%
Dennis	2005	0.79	0.24	-70%

*Given overlapping impacts of Jeanne and Frances, losses for these two storms have been combined.

These observational uncertainties are amplified when combined with the large trending factors required for historical storms older than about 30 years. For example, for the 1926 Miami Hurricane, Pielke [3] estimated a trending factor of 1,495 to scale the economic loss at the time of the event ($105 million) to a 2005 equivalent loss of $157 billion. If the original loss estimate had been under or overestimated by 50%, then the possible range of loss in 2005 dollars could reach from $78 to 314 billion. Given the large trending factor, the loss in 2005 dollars is sensitive to errors in the reported loss at the time of the event.

For this reason, RMS validation against PCS data focuses on only the last 20-25 years of data. Figure 4 compares version 11.0 model losses to trended PCS losses for all hurricane events affecting the U.S. from 1989 to 2009. Although there is scatter above and below the 1:1 line, the data points group around the line thus demonstrating that the RMS model is unbiased, and confirming overall consistency with adjusted industry loss estimates. Note that individual variations for single events are expected due to the uncertainties in the original reported losses and trending methodologies, as well as model uncertainties.

Figure 4: Comparison of version 11.0 model losses to trended PCS losses between 1989 and 2009, demonstrating that results are unbiased and reflect overall industry patterns



Historical Industry Average Annual Loss Validation

Once the individual historical loss comparisons have been performed, the historical losses can be used for further model validation tests. For example, a further validation test involves checking whether summary statistics based on historical loss observations are consistent with the model statistics. The previous section establishes that there is diminishing confidence in industry loss reporting and trending going backwards in time. Although in basic statistics sample sizes of 30 from a 'normal' random process can be considered to be an 'infinitely-large' sample and thus can be used to establish accurate averages, hurricane loss distributions are not statistically normal - they are highly skewed. It is therefore inappropriate to calculate the historical average annual loss (AAL) for the past 30 years and use the outcome to check the baseline hurricane event rates. The following analysis illustrates why, for a highly skewed distribution like hurricane occurrence, 30 years of loss data is inadequate for validating models.

In this analysis RMS uses historical reconstructions, validated in the previous step, as a proxy for the PCS observations. Historical reconstructions are event-specific loss estimates for the present day industry exposure that reflect the loss if each historical event were to occur today, and are used because observations from PCS are non-existent prior to 1950, and because significant uncertainties exist in older PCS loss observations and exposure trend adjustment factors. This historical reconstruction proxy provides a means to extend the observed historical benchmark back 110 years to represent the full credible portion of the HURDAT storms database from 1900 to present.

Figure 5 shows (a) the 110 years of reconstructed industry loss proxy (red bars), and (b) a 30-year moving average of the same reconstructed losses (blue line).

The 30-year moving average (b) ranges from $4.1 billion to $20.3 billion. The graph shows that even after averaging over 30 years there is a significant amount of volatility in possible AAL estimates, depending on the time period chosen. The variability in these results stems from the inherent variability in the peril itself. Any conclusion drawn from a limited period is very much subject to which averaging period chosen. RMS concludes that an average of the last 30 years is not long enough to appropriately validate the model's long term AAL. The next line of investigation seeks to establish the required time period for model AAL validation.

Figure 5: Historical Annualized Loss reconstructions for the past 110 years and the variability of the 30-year moving based on the data.

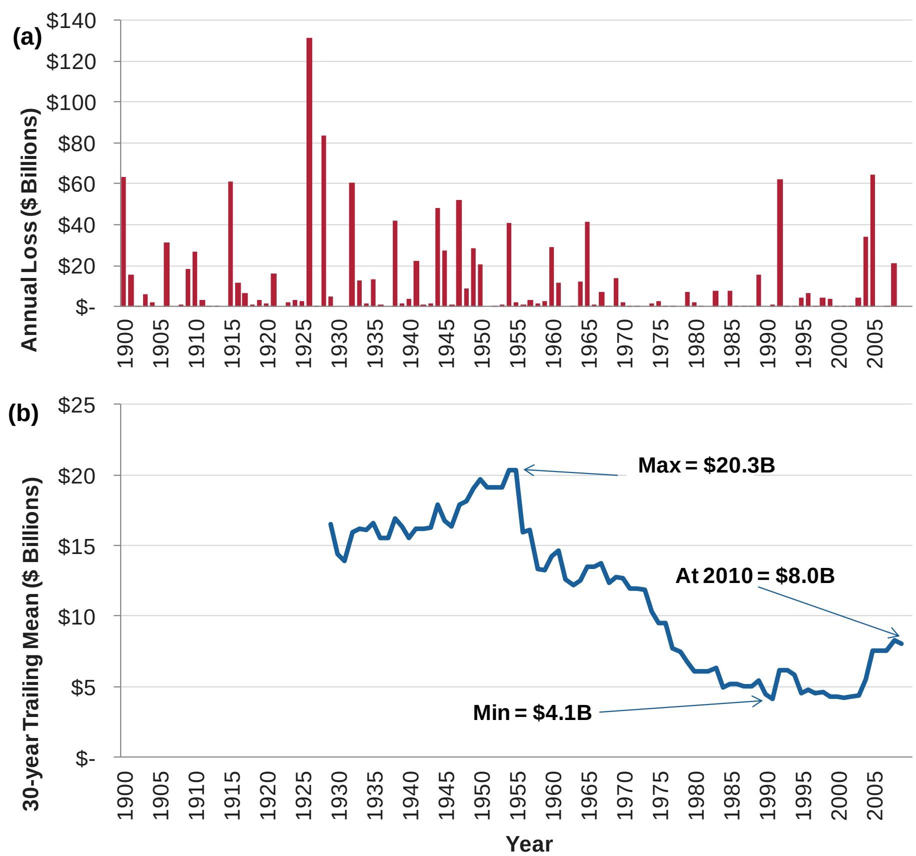
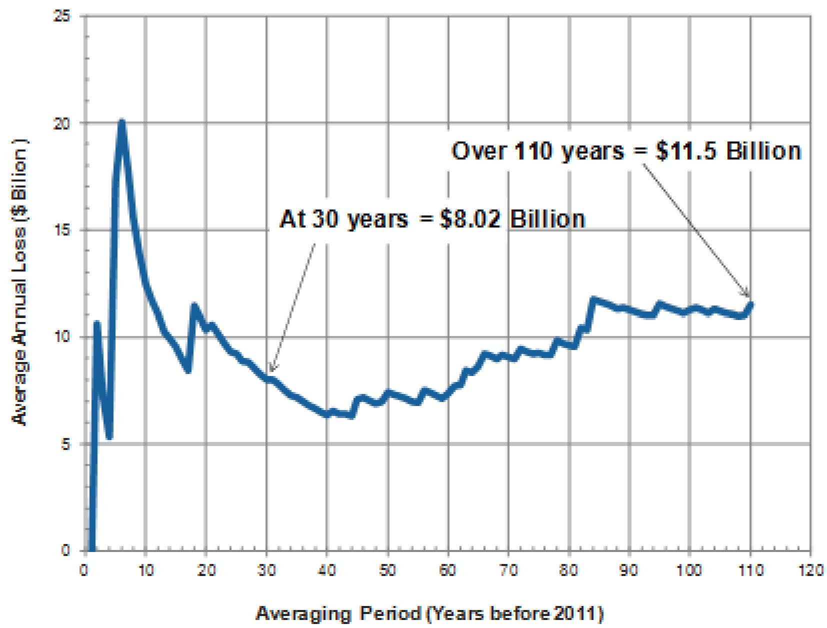


Figure 6 shows the same data as Figure 5, plotted instead as AAL estimates for different lengths of time periods starting from 2011 and reaching back in time. At each data point, an additional year of historical experience is added, and the corresponding change in AAL is measured. This graph reinforces the previous problem with the 30-year AAL metric, as significant fluctuations in the compiled AAL are present with the addition of new data starting at 30 years and increasing to as many as 90 years back. The AAL estimates based on the historical loss proxy appear to stabilize for time periods greater than 90 years, reaching $11.5 billion for the full 110-year period.

In theory, one can use this $11.5 billion AAL estimate to validate the model's industry AAL of $13.1 billion, but we should still have questions about the stability of the 110-year average, given the large variation of possible yearly losses. For example, if one more year with a loss of $130 billion (corresponding to the 1926 yearly loss) were added to the historical proxy data, then the AAL would change by about 10% from $11.5 billion to $12.6 billion. Rather than waiting for more history to occur, a different method of quantifying the uncertainty of the historical AAL is necessary.

Figure 6: Comparison of historical averages compiled over increasingly longer time windows within the 110 year long-term historical record, up to 2011.



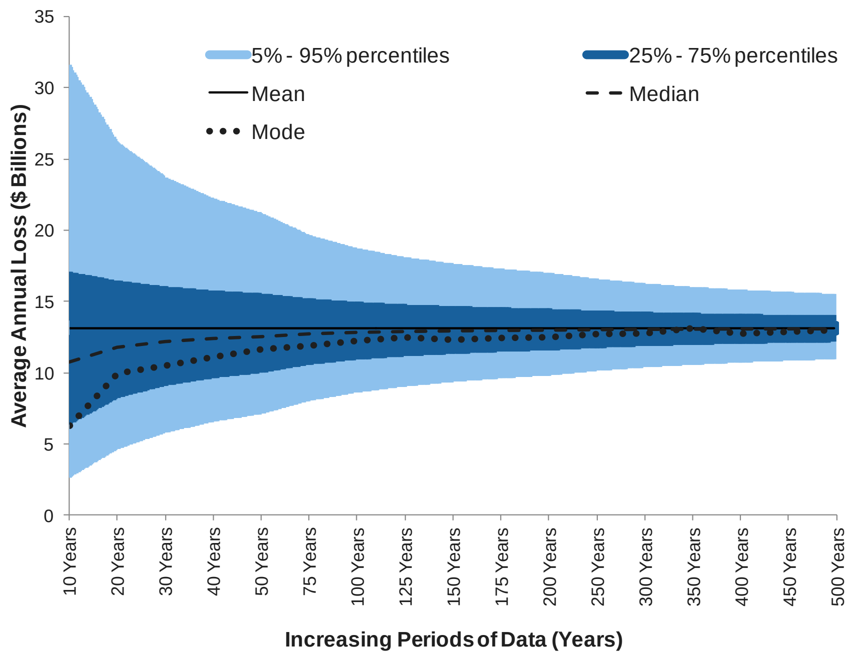
A Perspective on Annual Hurricane Loss Distributions

We have now established that a 30-year period is too short a timeframe to evaluate whether the implied AAL of a hurricane model is reasonable. We also know that hurricane losses exhibit great variability and have highly skewed distributions of loss. It appears that an average from 110 years of data is converging to a stable value, but is 110 years enough? In other words, if we compare the historical proxy AAL of $11.5 billion to the model AAL of $13.1 billion, does this demonstrate that the model is consistent with history?

To understand the uncertainty of the historical AAL, we need to know the shape of the distribution of many such 110-year AAL estimates. One way to approximate such a distribution is to randomly sample many 110-year periods from the model's implied annual industry loss distribution, calculate the 110-year average loss from each sample, and then plot the distributions of these simulated 110-year samples.

Figure 7 shows the output from precisely such an exercise for 110 years as well as other experience periods (number of years averaged together), plotting the trend in various distribution statistics as the experience period is increased from a 10-year to 500-year period. The black line represents the $13.1 billion industry AAL of our underlying baseline model. The dark blue area represents the range within which one such experience period's AAL estimate would fall into 50% of the time. The pale blue area is the equivalent for a 90% probability. The trends in the median (dashed) and mode (dotted) statistics are also displayed

Figure 7: Trend in uncertainty range and statistical metrics of simulated AAL estimates[3]



That both the median and the mode are below the mean is an expected result given the skewed nature of the underlying annual loss distribution [4]. Highly skewed distributions need more samples to accurately establish an average, and also have the feature that "small" samples are more likely to fall below the true mean. Note more likely does not mean always. This is especially true for shorter experience periods. Stated another way, the simulations in this graph indicate that the most likely

[3] Simulations based on 1 million random samples from the industry annual loss distribution of the baseline model. Note that even after simulating 1 million iterations the results do not demonstrate full convergence (which is why the lines are not perfectly smooth at shorter experience periods), but support the point that the sample average for skewed distributions is likely to be below the true mean

sample average (mode) will only lie within 10% of the true mean after about 100 years. For the U.S. Atlantic Basin as a whole, 100 years of data just barely reaches the threshold of an adequately stable sample.

When considering a regional subset of data for the northeast, for example, where the historical record is even more skewed, 100 years of data is still considered a “small” statistical sample. Hence, the fact that the AAL estimate derived from the historical record ($11.5 billion) is below the RMS model AAL ($13.1 billion) is not unexpected when considering the skewed nature. Overall, these considerations of the impact of skewness mean that RMS considers the AAL estimate based on the 110-year industry loss experience proxy of $11.5 billion to be consistent with the $13.1 billion industry AAL of the RMS hurricane baseline model.

Figure 7 also demonstrates that, even for longer experience periods, AAL estimates will continue to have large uncertainties. For example, with an experience period of 110 years, an AAL estimate has only a 50% probability of being inside the $11-15 billion range (red area), given a true underlying AAL (mean) of $13.1 billion. In other words, there is only a 50% probability that the experience-based estimate is contained in the $4 billion range around the true mean. Hence there is also a 50% probability that it could fall outside of this range. If the experience period was extended to 500 years, the width of this range would only reduce by 50%, to $2 billion ($12-14 billion).

The tendency for the historically derived AAL to fall below a catastrophe model's AAL is observed in other vendor models as well. Table 3 shows AAL estimates for Florida hurricanes from all commercial catastrophe model vendors that submit their models to the Florida Commission on Hurricane Loss Projection Methodology (FCHLPM). Comparing the model-estimated AAL with the historical AAL (108 years) for each vendor, all of the model losses are 20–30% higher than history. As just mentioned, when looking at sub-regional statistics, the skewed nature of the historical loss distributions becomes more pronounced than on a basin-wide basis. The ratio of modeled to historical loss should not be automatically interpreted as bias or model miss, but may be the result of the skewed distribution of historical results. Differences of 20–30% should not indicate that the model is invalid, but are consistent with the limitations of working with historical data.

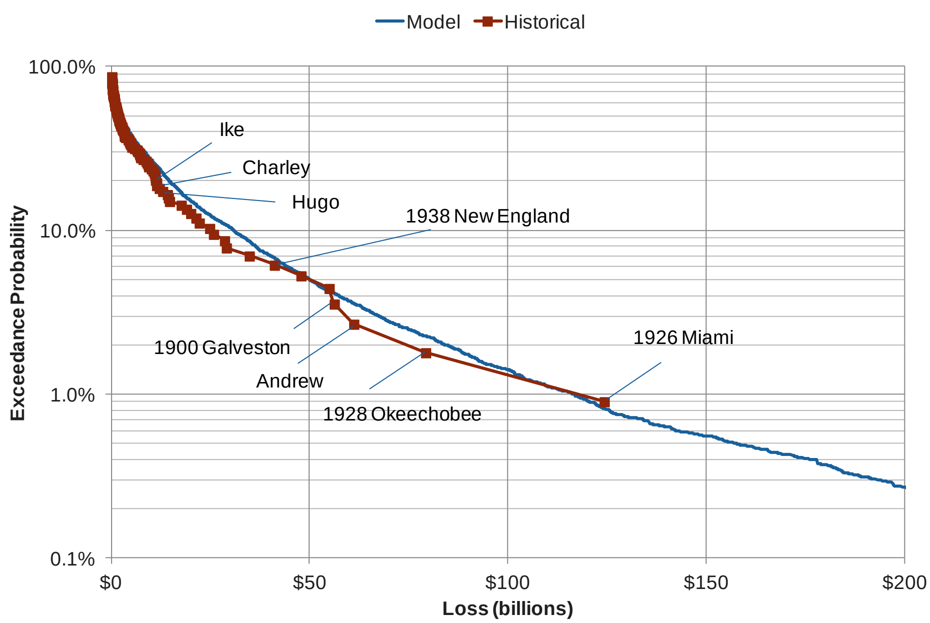
Table 3: Comparison of Historical and Modeled Average Annual Loss reported in submissions to the FCHLPM under the 2009 standards.

Model Vendor	Historical AAL ($B)	Model AAL ($ B)	Model/History
RMS	2.67	3.47	1.30
AIR	2.84	3.62	1.27
EQE	3.26	3.99	1.22
ARA	4.13	5.28	1.28

Historical Industry Exceedance Probability (EP) Validation

The final industry loss validation benchmark is the implied historical exceedance probability (EP) curve, which is based on the proxy for 110 years of industry event losses and a simple event frequency assumption. As an example, Figure 8 compares the RMS U.S. wind-only model EP curve with the historical EP curve, illustrating good agreement between the historical EP and the stochastic EP from version 11.0 wind-only losses. The historical EP curve is based on model reconstructions for all events since 1900, assuming a rate of 1/110 for each historical event, which represents the number of years over which exactly one occurrence of each historical event has been observed. This is a very simplistic statistical rate assignation, which gives every event an identical rate of occurrence, and does not consider whether one event is more or less likely to occur than another.

Figure 8: Comparison of Historical and Modeled EP curves (billions) for hurricane risk (wind only) in the U.S.



Misleading Validation Comparisons

RMS welcomes outside scrutiny of the version 11.0 U.S. Hurricane Model and all RMS models, and firmly supports model users undertaking their own catastrophe model validations. However, some validation techniques are not scientifically valid and others can be misleading. This section discusses some of the potential pitfalls in model validation, and some of the errors that may occur as a result.

Comparison to Real-Time Events

While a catastrophe model's ability to produce loss estimates for real-time events could potentially be considered a form of model validation, it is necessary to bear in mind that catastrophe models are not designed to be predictive forecast tools of single events, but instead to reproduce a range of possible events that can be used to "extrapolate" the historical record.

The art of predicting losses for a real-time event is a matter of selecting input parameters in light of highly uncertain and unverified hazard observations and estimates of affected exposure within days of the event's occurrence. Thus, it is more effective to create accurate historical reconstructions and use the validation techniques described in this paper, than to rely on real-time loss estimates as a validation technique.

Comparison to Public Domain Publications

One of the few public domain references for calculating annual hurricane losses is the paper by Pielke et al. [3] published in 2008, where economic hurricane losses reported in newspaper and weather service reports from 1900 to 2005 are converted to losses in 2005 dollars. The methodology uses three trending factors: inflation, number of building units, and changes in the building quality or wealth per capita. The study finds the estimated average annual losses in 2005 dollars to be approximately $10 billion.

While the Pielke framework is similar to the RMS Historical Industry AAL validation benchmark (see item 5 in Table 1), it is important to convert the loss estimates to the same time period. Pielke trends losses to the year 2005, whereas the v11.0 RMS hurricane model represents exposures and values reported for the year 2011. Pielke employs an average trending factor of 1.5 to convert losses from 1999 to 2005—a period of six years. Applying a similar factor to bring losses from 2005 to 2011 would increase the Pielke estimate from $10 billion to $15 billion. When compared to the model AAL of $13.1 billion, it could be suggested that the model does not overestimate the industry loss, but underestimates this loss, when compared on a consistent timeframe.

As previously mentioned, there is considerably more uncertainty in the inflated losses of events occurring greater than 20–25 years in the past than for more recent storms, both in terms of:

- The applicability of the trending data and methodology used to conduct the trending to the insurance industry.
- Consistency of the originally reported losses and applicability to the insurance industry.

Additionally, the Pielke study derived insured losses from economic losses, which include infrastructure and other public domain works in addition to privately insured losses. As a general rule, Pielke used a factor of 2 to convert between economic and insured losses. The difference between economic and insured losses varies by event but is typically greatest for large surge events that result in flood damage to buildings and infrastructure not covered by private insurance.

While the Pielke estimate corrected to 2011 vintage broadly validates the version 11.0 RMS U.S. Hurricane Model, the factors and data from Pielke cannot be used directly. However, the framework for trending, accompanied by other insurance-specific datasets, can be applied to the industry validation problem as just shown.

PCS-Derived Averages from Limited Time Periods

As discussed earlier, RMS advises that model users should be very careful when estimating average annual losses based on limited industry losses, such as the last 20–25 years of data reported in sources such as Property Claim Services (PCS). Event-based comparisons are appropriate, but averages derived from such brief observational periods do not consider how hurricane cycles (active verses inactive periods), or the limitations of statistics derived from highly skewed distributions can affect the results as demonstrated in the section Historical Industry Average Annual Loss Validation on page 17. Thus, comparisons of average annual loss based on only 30 years of history or less are commonly misleading. Instead, users should consider validating the model against the historical industry proxy technique described earlier.

Conclusion

Catastrophe models have become an integral component of insurance risk management strategies, and are valuable tools for dealing with high-severity, low-probability events like hurricanes and earthquakes. The component-based manner in which these types of models are created allows them to extrapolate beyond a limited historical record. They are specifically designed to overcome the limitations of working only with highly skewed historical loss distributions.

RMS models are validated in a number of different ways ranging from individual components of hazard and vulnerability to overall loss output from a wide variety of sources. This paper has described some of the validation methods employed at RMS using the version 11.0 U.S. Hurricane Model to illustrate these methods in practice.

Model users are encouraged to investigate the uncertainties inherent in catastrophe modeling. By outlining in a transparent manner the methods used by RMS staff to validate the hurricane model, users are provided with the tools needed to take ownership of their view of risk. This paper provides a framework through which model users can validate models on their own, and understand the strengths and limitations of the different validation methods available.

As the science of hurricane modeling continues to evolve, RMS model users can expect further model enhancements to be rigorously validated in a similar manner to the methods described in this paper.

References

1. RMS, "RMS North Atlantic Hurricane Model RiskLink 11.0.SP2: Submitted in compliance with the 2009 Standards of the Florida Commission on Hurricane Loss Projection Methodology ", Newark, CA, Sept 2011. (http://www.sbafla.com/method/portals/methodology/ModelSubmissions/2009/FinalRMS09StandardsDisclosures_20110920.pdf)
2. AIR Worldwide, "The AIR Hurricane Model: AIR Atlantic Tropical Cyclone Model V12.0.1: Submitted in compliance with the 2009 Standards of the Florida Commission on Hurricane Loss Projection Methodology", Boston, MA, Nov 2010. (http://www.sbafla.com/method/portals/methodology/ModelSubmissions/2009/20101103_AIR_2009_FCHLPM_acceptedrevisions.pdf)
3. Pielke, Jr., R. A., Gratz, J., Landsea, C. W., Collins, D., Saunders, M. A., and Musulin, R., "Normalized Hurricane Damages in the United States: 1900-2005", Natural Hazards Review, Vol. 9 / Issue 1, pp. 29-42, 2008. Supporting data is available at: (http://sciencepolicy.colorado.edu/publications/special/normalized_hurricane_damages.html)
4. Fleming, K., " Yep, We're Skewed ", Variance: Advancing the Science of Risk. Casualty Actuarial Society, Vol. 2 / Issue 2, pp. 179-183, 2009.
5. Cubukcu, N., R. L. Pfeffer, D. E. Dietrich, "Simulation of the Effects of Bathymetry and Land–Sea Contrasts on Hurricane Development Using a Coupled Ocean–Atmosphere Model". Journal of the Atmospheric Sciences, American Meteorological Society. 2000.
6. Collete, L., Daniel, V., Bellone, E., Nolan, D., "Using Mesoscale Simulations to Train Statistical Models of Tropical Cyclone Intensity Over Land". Monthly Weather Review. Vol. 138, No. 6. 2010.

Contacting RMS

When you contact RMS for technical support, provide the following information:

- Hardware and configuration details of the system you are using, including network details, for example desktop or client/server configuration.
- Details on the difficulty you are encountering, for example, the product and version you are using.

The RMS website features a password-protected Client Resources area for licensees of RMS catastrophe products. It includes a Documentation Library of software technical documentation and Model Methodology documents for RMS peril models. It also offers online access to catastrophe updates.

You can access the RMS website at www.rms.com. If you cannot access the site or do not have a user name and password, contact your RMS Service Representative or email info@rms.com.

RMS Global Product Support

Phone: 1-877-767-6266 (U.S., 1-877-RMS-O-COM)

1-877-767-0266 (U.S., 1-877-RMS-0-COM)

+44 207 444 7777 (International)

Email: support@rms.com

Additional RMS Information:

Phone: 1-510-505-2500

8:00 a.m. – 5:00 p.m., Pacific Time Zone

Fax: 1-510-505-2501

Write: Risk Management Solutions, Inc.

7575 Gateway Boulevard

Newark, CA 94560

USA

Email: info@rms.com

Web: http://www.rms.com

Send Us Your Feedback

Risk Management Solutions, Inc. appreciates your feedback on the quality of this document. Your suggestions are an integral part in enhancing and improving the usability of our documentation. Please send us your comments on any of the following topics:

- Did this document contain all the information you needed?
- Did the index help you access topics you were looking for?
- Did you find any technical errors in this document?
- What did you like most about this document?
- Do you have any suggestions for improving this document?

You can send feedback by email. Address your comments to Documentation@rms.com, subject: Documentation Feedback. Please include your contact information (name, company, and email address) if you want us to follow up with you.

Note: For product related questions or issues, contact your RMS Service Representative.