Catastrophe Ratemaking Revisited
(Use of Computer Models to Estimate Loss Costs)

by Michael A. Walters, FCAS, and
François Morin, FCAS

Catastrophe Ratemaking Revisited
(Use of Computer Models to Estimate Loss Costs)

By Michael A. Walters, FCAS, MAAA
and François Morin, FCAS, MAAA

Abstract

Recent developments in computer technology have significantly altered the way the insurance business functions. Easy access to large quantities of data has rendered some traditional ratemaking limitations obsolete. The emergence of catastrophe simulation using computer modeling has helped actuaries develop new methods for measuring catastrophe risk and providing for it in insurance rates. This paper addresses issues associated with these methods and provides actuaries, underwriters and regulators with an understanding of the features and benefits of computer modeling for catastrophe ratemaking.

Biographies

Michael A. Walters is a principal and consulting actuary with Tillinghast – Towers Perrin in Parsippany, New Jersey. A graduate of Fordham University, he has a master's degree in mathematics from the University of Notre Dame. He is a Fellow and past president of the Casualty Actuarial Society and a past vice president of the American Academy of Actuaries and past chairman of its Casualty Practice Council. He has authored two papers previously included in the CAS Syllabus of Examinations: "Homeowners Insurance Ratemaking" and "Risk Classification Standards." The latter was awarded the 1981 CAS Dorweiler prize.

François Morin is a consulting actuary with Tillinghast – Towers Perrin in Hartford, Connecticut. Mr. Morin has a bachelor of science degree in actuarial science from Laval University in Quebec City. He is a Fellow of the Casualty Actuarial Society and a Member of the American Academy of Actuaries. Mr. Morin provides catastrophe modeling consulting services to the insurance industry and has contributed to the development of ToPCAT, Tillinghast's catastrophe model.

Acknowledgment

The authors wish to acknowledge the input to the risk load analysis by Stephen Philbrick (FCAS) and the overall review efforts of Ronald Kozlowski (FCAS), Stuart Mathewson (FCAS), David Sommer (FCAS), and Tom Warthen (FCAS).

Catastrophe Ratemaking Revisited
(Use of Computer Models to Estimate Loss Costs)

WHY MODELING?

According to the CAS Principles of Ratemaking, a rate "is an estimate of the expected value of future costs, provides for all costs associated with the transfer of risk, and provides for the costs associated with an individual risk."

Traditionally, ratemaking has been regarded as the art of projecting scientifically measured past experience into valid conclusions about the future. However, for lines of business with catastrophe potential, questions always arise as to how much past insurance experience is necessary to accurately represent possible future outcomes and how much weight should be assigned to each year's experience. For instance, if a 1954 hurricane was the last severe event in a given state, may one assume that the return period for an event of the same severity is 40 years? What if historical records show that more severe storms occurred in the 1930s, before the advent of homeowners coverage? If the same storm happened today, would it affect the same properties? What level of damage would occur, given that the distribution of insureds had shifted to coastal communities and that the insured values at risk have trended at a pace that has exceeded inflation?

For these rare event calamities, reliance on actual insured experience does not allow accurate measurement of future expected loss. Therefore, one must use a much longer experience period, especially for event frequency. Computer simulation of events to obtain current insured losses has replaced traditional methods based exclusively on reported loss experience. These new methods can now be used not only to measure expected losses, but also to develop risk loadings to compensate for the variance in outcomes, compared to lower-risk insurance products.

The need for catastrophe modeling has existed for some time to aid in reinsurance purchase decisions as well as in insurance ratemaking. However, computer limitations on the amount of data that could be manipulated to develop a catastrophe model had usually rendered the concept impractical. But computer capacity has improved dramatically, which now makes catastrophe simulation feasible. It has also enabled scientists to expand their research and produce better simulations through a better understanding of catastrophic events.

WHAT TO MODEL

A state's most recent past may not be indicative of its true catastrophe potential because what happens in a given year is only a sample of what could have happened. The goal is to build a model to simulate what could realistically happen, based on information relevant to that state and to all refined geographic areas within the state.

Doing this with a computer model requires that the estimation process be separated between frequency and severity. For the frequency of hurricanes, there is a long history (more than 100 years) of recorded information to gauge the relative likelihood of landfall in a given state.

For severity of loss, however, the actual insured damage may not have been recorded. Certainly, the extent of loss if that same storm occurred again would depend on *today's* insured values, coverage and level of windstorm resistant structures. This is the first area utilizing computer simulation — taking the characteristics of a storm and replicating the windspeeds at various locations and times over its course after landfall. Next, the damage to buildings and the effect on insured values flow from the windfield created by the storm. Validation of the model examines actual storms over the recent past, so that the full range of possible storms is the basis for expected loss calculations as well as risk analysis on the possibility of adverse outcomes in any given year.

HOW TO MODEL FOR SEVERITY

The severity component of catastrophe modeling generally comprises three distinct modules with three separate skills required:

- event simulation (science)
- damageability of insured properties (engineering)
- loss effect on exposures (insurance).

The event simulation module is designed to reproduce natural phenomena. For a hurricane model, this involves predicting wind speeds at every ZIP code affected over the course of a single storm. The damageability module estimates the damage sustained by a given property exposed to the simulated event. The majority of the damage functions used in a catastrophe model are developed by engineers who better understand the physics of natural phenomena and can test the resistance of various materials to high windspeeds. (The results of the studies are also used to develop new materials and to implement new building codes to limit the damage from catastrophes.) The insured loss effect module incorporates the results of the first two modules and adjusts them for such factors as deductibles, co-insurance, insurance to value, and reinsurance. This is generally the only company-specific module because it includes all the factors that describe the in-force company book of business.

This part of catastrophe modeling is known as deterministic, because it allows the simulation of a predetermined event with known characteristics. The computer could duplicate this event, if it occurred today, with the resulting effects on the insured exposures calculated. Appendix A provides a detailed description of the process involved in developing and validating the severity component of a catastrophe model.

HOW TO MODEL FOR FREQUENCY

Once the deterministic model has been created, calibrated and validated, the modeler must analyze historical meteorological records and develop a probabilistic facet to the catastrophe model. The first step involves generating distributions for each of the parameters required as input to the hurricane model. A hurricane model may be dependent on a variety of factors, such as the radius of maximum speed, forward moving speed and pressure differential at the eye of the storm. Considerable effort must be spent in constructing these distributions so that accurate representations of realizable events can be obtained by combining the variables. For example, an analysis of the radius of maximum winds of historical events yields a conclusion that they are normally distributed ($N(\mu,\sigma)$), (with parameters of 16.840 and 10.567 in South Florida). Similarly, the forward moving speed of these events follows a lognormal distribution ($\Phi(\mu,\sigma)$) (with parameters of 2.304 and 0.283, respectively in South Florida.) Similar distributions must be built for each of the parameters that are hurricane specific in each geographic zone. One can obtain the historical data from National Oceanic and Atmospheric Administration (NOAA) publications.

The modeler then uses sampling techniques to randomly select the parameters from each distribution. Most catastrophe models rely on a Monte Carlo approach, a stratified sampling approach or a combination of both. Although Monte Carlo is easier to use and to explain to a nonstatistical audience, it does not have the sampling power of a stratified approach. Therefore, the modeler should consider both methods before generating the probabilistic database.

In conjunction with storm intensity distributions, conditional probabilities, storm paths, and landfall locations must be developed for each storm modeled. These parameters are based on actual storm paths of historical events over the last hundred years. The storm probabilities depend on the type of sampling utilized in selecting parameters for storm intensities. By nature, Monte Carlo sampling requires that all storms have the same probability, whereas stratified sampling can be done in such a way that probabilities are not all equally likely.

After selecting the storm intensity parameters and deriving the probabilities, one combines the two. The end result is the probabilistic library, which comprises a large enough number of events (in excess of 5,000) to represent all likely scenarios. For example, the database should include Category 5 storms making landfall in Maine (if they are at all possible) so that the damage associated with such an event can be calculated. (Stratified sampling allows a more efficient handling of this issue because it can cover all possibilities with fewer storms.) Because each event has an associated probability that is conditional on a hurricane making landfall, the sum of all probabilities will, by definition, add up to one. The modeler will then use these probabilities to derive annual expected loss costs.

BASIC OUTPUT OF MODEL

Once the probabilistic database is complete, one can proceed to calculate expected loss costs by ZIP code. To accomplish this, the modeler should run the entire event library against a set of exposures that assumes a constant value (e.g., $100,000 of Coverage A amount for homeowners) in each ZIP code. It is important to ensure that all exposed amounts are included in the simulation. For homeowners, it is customary to increase Coverage A amounts by 10% for appurtenant structures, 50% for contents and 20% for additional living expense (i.e., loss of use). Annual expected loss costs for a given ZIP code are then developed by multiplying the sum of the probability weighted simulated results across all storms by an annual hurricane frequency. The average annual frequency of hurricanes making landfall in the U.S. has been approximately 1.3.

For a given line of business, the expected losses by ZIP code are then:

$$EL_{ZIP} = F \times \sum_{storm} (P_{storm} \times E_{ZIP} \times DF_{storm})$$

Where EL_{ZIP} = Expected Losses for ZIP code for base class
F = Annual Hurricane Frequency
P_{storm} = Probability of storm
E_{ZIP} = Total exposure amount (Base class constant for all ZIP codes)
DF_{storm} = Damage factor for base class by ZIP code by storm

These expected losses represent insured losses for a base class amount of insurance, construction type and deductible. These may be selected as frame building with $250 deductible, with $100,000 Coverage A (building), $10,000 Coverage B (appurtenant structures), $50,000 Coverage C (contents) and $20,000 Coverage D (additional living expense). To convert this to a loss cost expressed as a rate per $1,000 of Coverage A amount requires division by the exposure base times 1,000.

$$ELC_{ZIP} = \frac{EL_{ZIP}}{COVA_{ZIP}} \times 1{,}000$$

Where ELC_{ZIP} = Expected Loss Cost for ZIP code
$COVA_{ZIP}$ = Base class Coverage A amount in ZIP code

A major feature of this calculation is its independence of an individual company's actual loss experience and of its exposure distribution. Being independent of individual company data, it is, in fact, appropriate for each insurer.

The next step is to average the loss costs by ZIP code over the insurer's exposure distribution within the territory structure it selects.

$$ELC_{terr} = \frac{\sum\limits_{ZIP} (ELC_{ZIP} \times COVA_{ZIP})}{\sum\limits_{ZIP} COVA_{ZIP}}$$

Where ELC_{terr} = Expected Loss Cost for territory

In Exhibit 1, the ZIP code loss costs per $1,000 of Coverage A amount for homeowners are averaged to a given territory structure to derive the territorial loss costs for hurricane coverage. It is likely that the more representative territory structure for hurricane will differ from regular homeowners territories. Because the latter evolved over time to respond to homogeneity considerations in setting rates for the perils of fire and theft, a company may wish to create new territories to reflect differences in hurricane loss potential.

ATTRIBUTES OF LOSS COSTS VIA COMPUTER MODELING

Credibility

Through computer simulation and stratified sampling, the most remote cells have complete credibility in the traditional sense. That is, the measurements can be taken at full value, without having to ballast them with actual results on a statewide basis, or on last year's results. One substitutes the random variation of low frequency actual storms with the set of all possible storms via the model. Moreover, the probabilities are assigned by the selection of the input parameters. This solves the problem of low credibility of actual results and the attempt to refine actual statewide data to territory.

While full credibility can be assigned in cell detail from computer simulation, this only means that random statistical variation can be resolved to eliminate the process risk from a ratemaking standpoint. However, there is still parameter risk in the selection of the key variables. It is possible that the event frequencies of the past 100 years are not representative of the next 100 years. This is especially true in the case of earthquake simulation, where the physics of shake intensity are not understood well enough by earthquake experts to generate fully reliable parameters of frequency and severity.

With full credibility in ZIP code detail, one can calculate statewide averages by averaging over ZIP code and territory. This is in stark contrast with the usual homeowners indicated loss costs, which first are developed statewide, and then must be distributed to the different class and territory cells with appropriate credibility weightings. This stems from the experience loss ratio method used to derive the result — actual insured experience that is a sample taken from what might have occurred over time. In contrast, hurricane loss costs are derived from the set of all possible events as constructed in the computer model. Using a hurricane model to produce loss costs is truly a pure premium method of ratemaking, versus the loss ratio method usually used in traditional ratemaking with historical insurance data.

Frequency of Review

Hurricane loss costs derived from modeling do not need frequent updates for two reasons. First, with more than 100 years of actual event characteristics shaping the model design, another year or two of actual results are unlikely to change model parameters much. In the early stages of model building, with each new hurricane to landfall, the potential exists to update some of the damage factors and the estimated effect of deductibles or other class factors. Also, when new class variables are developed, one can refine initial estimates with the loss experience of subsequent actual storms. For example, new kinds of shutters will have been tested, and it would be possible to incorporate their effect in the model.

Secondly, once adequate rate levels are achieved, annual updates are also not critical because the exposure base ($1,000 of Coverage A) is inflation sensitive. For the average territory loss costs, in the early years of implementation, it may be well to test for changing ZIP code distributions, as insureds and insurers react to some high loss costs in certain coastal areas.

Risk Variations

Non-hurricane homeowners loss costs vary significantly by fire protection class, reflecting the large portion of the coverage represented by the fire peril. Yet, the hurricane peril is obviously independent of protection class.

Policy form relativities basically increase as additional perils are covered. In Forms 1 and 2, the perils are specified, while Form 3 gives essentially all risk coverage on the building, but not on contents. Form 5 provides all risk coverage on contents. Thus, the wind coverage is identical in all the homeowners policy forms. Hence, if the hurricane loss costs are a material portion of total homeowners costs, the policy form relativities would have to vary substantially by territory or even by ZIP code if applied to an indivisible homeowners premium.

For construction class, a frame house can be almost as hurricane resistant as one made of brick or stone. For large hurricanes, the key is to protect the envelope of the building from penetration — i.e., the windows and the roof. Hence, the relative fire resistance of the construction is irrelevant for the hurricane peril.

Hurricane (and other catastrophes) ultimately may need a separate class plan because of different risk variation from the traditional covers. For example, for hurricanes, new rating factors will likely emerge for shuttering and for roof type (e.g., gable versus hip roof). Local enforcement of building codes is an early rating distinction that is implementable. Redoing all the traditional homeowners class relativities to meld with the new catastrophe classes would be very cumbersome. Perhaps the traditional homeowners territories could be retained, with a separate set of territory definitions for the hurricane rate.

A possible class plan with sample surcharges and discounts follows:

Category	*Criteria*	*Sample Factor*
Hurricane Shutters	None	+ 0.20
	Add-On	- 0.20
	Built-In	- 0.40
Roof Type	Hip	- 0.25
	Gable	+ 0.30
Location	Shielded by buildings	- 0.20
	Subject to projectiles	+ 0.20
	Beach front or subject to surge	+ 0.10
Town Building Code	Not enforced	+ 0.15
	Enforced; not inspected	- 0.10
	House inspected; within code	- 0.25

FORM OF RATING

If the hurricane peril does not vary by class the same way as the non-hurricane perils, should the hurricane rate be split out from the heretofore indivisible premium for homeowners? Should it have its own class plan? The answer to both questions is yes.

Basically, one can have the best of both worlds. The indivisible premium formerly simplified the review of loss experience and the rating of the homeowners policy, as well as lowering the cost of the monoline coverages, knowing that all the major perils were essentially compulsory. Virtually all of the advantages of the indivisible premium can be kept by still keeping hurricane *coverage mandatory. Yet, it is the very difficulty of the experience review that suggests the* segregation of it for ratemaking — using the pure premium method for hurricane ratemaking and allowing a loss ratio approach for the other perils.

Computer modeling could also be used for other catastrophe perils (e.g., earthquake, tornado and winter storm) such that the remaining non-catastrophe perils in homeowners would use the more traditional methods of ratemaking. Computer modeling for catastrophe perils actually makes ratemaking for the other perils much easier, because of less fluctuating results. With loss costs supplied by modeling and with a separate rate for each catastrophe peril, the actual catastrophe losses only need to be removed from the experience period and nothing need be loaded back to the normal homeowners losses. This means that catastrophe serial numbers ought to be retained for loss coding.

The overwhelming advantages of separate catastrophe rates are the simplification of the normal coverage rating and ratemaking as well as the better class and territory rating of the catastrophe coverages.

This does mean an extra rating step for the catastrophe coverages, but there already are so many endorsements in homeowners that this should not be much of a burden. Furthermore, if hurricane loss costs are left in the indivisible premium, the homeowners classes will become much more complicated to rate. The class relativities will have to vary greatly by hurricane zone, and the actuarial calculation of relativity indications will also be much more complex.

Another simplification via separate hurricane rating is not having to calculate a complicated set of statewide indications including hurricane. Instead, the indications can be produced, and actual rates selected, separately. Ostensibly, this creates a problem in rate filings, where tradition has called for a combined statewide average indicated rate change as well as a filed rate level change. However, this is mere custom, and not strictly required by the rating laws — which usually call for *rates* to be filed, not *rate changes*. In other words, statutory requirements are for *rates* to be reasonable, not excessive, inadequate or unfairly discriminatory. Filed measures of *rate changes* have merely been a convenient way for regulators to monitor reasonableness.

This is not to suggest that a rate filing should repress the estimate of statewide rate change. But given the different ways of calculating the appropriate rates (via a pure premium approach for hurricanes and a loss ratio method for other coverages), the statewide indication does not as readily come out of the ratemaking method as, for example, it does for auto insurance. Hence, other reasonable ways of estimating changes will need to be developed, instead of directly from the ratemaking method.

EXPENSE LOAD CONSIDERATIONS

If the hurricane peril is reinsured in a reasonable fashion, then the primary insurer ought to be able to pass those costs through to the policyholder. The reinsurance premium can be expressed as a function of the primary layer and added to the equation.

Then, the total expected hurricane loss costs would be adjusted to exclude the reinsured portion by having the hurricane computer model simulate the reinsurance layer. This is done by running all probabilistic storms against the insurer's exposure base by ZIP code and line of business. Each storm's losses in the reinsurance layer (1) are then allocated to line and ZIP code in proportion to total losses for that storm (2). Then each storm's probability is multiplied by the losses in the layer and accumulated (3). This produces the expected losses in the reinsurance layer.

(1) $$L_{XS} = \text{MIN}\left(\text{MAX}\left(\left(\sum_{ZIP} E_{ZIP} \times DF_{storm}\right) - \text{RET}, 0\right), \text{LIM}\right)$$

Where
L_{XS} = Total Losses in Layer for each storm
RET = Reinsurance Retention
LIM = Reinsurance Layer Size

(2) $$L_{XS,ZIP} = L_{TOT,ZIP} \times L_{XS} \div L_{TOT}$$

Where
$L_{XS,ZIP}$ = Excess Losses by zip code for each storm
L_{TOT} = Total Ground-Up Losses for each storm
$L_{TOT,ZIP}$ = Ground-Up Losses by zip code for each storm

(3) $$EL_{XS,ZIP} = F \times \sum_{storm} P_{storm} \times L_{XS,ZIP}$$

Where $EL_{XS,ZIP}$ = Expected Losses in Layer by Zip Code

The reinsurance premium can then be allocated to line of business and ZIP code in proportion to the expected excess losses in the reinsurance layer. Those premiums are then ratioed to the primary premium by line and ZIP code to get a factor to add to the indicated rate by line and ZIP code.

The remaining expected loss costs outside the reinsurance layer (above and below) would then be loaded for risk margin and expenses. The reinsurance pass-through would already have included the expenses and risk margin of the reinsurer.

RISK LOAD CONSIDERATIONS

Splitting the homeowners premium into a catastrophe and non-catastrophe component also allows for a separate calculation of a risk margin. As a result, the non-catastrophe component becomes easier to price, with less variability and a lower margin needed for profit. This makes it closer to a line of business like automobile physical damage in its target total rate of return and total target operating margin needed, which can be expressed as a percentage of premium.

Once a target margin is selected for the non-catastrophe component, the margin for the catastrophe piece can be calculated as a multiple of the non-catastrophe component, using some basic assumptions. One assumption is that profit should be proportional to the standard deviation of the losses. (Some actuarial theorists argue that risk load should be proportional to variance. It is important to note that these arguments apply to individual risks. The assumption that the required risk load for an entire portfolio is related to the standard deviation is not inconsistent with a variance based risk margin for individual risks. In addition, the high correlation of losses exposed to the risk of a catastrophe as well as the large contribution of parameter risk to the total risk load requirement provides additional arguments in favor of a standard deviation basis for risk load.)

The calculation of the risk load should be performed on a basis net of reinsurance since the reinsurance premium is being built back into the rates separately. However, calculating the risk load both gross and net of reinsurance may be an important exercise for an insurer analyzing retention levels. By doing so, the insurer may be able to evaluate its reinsurance protection by considering the total risk load required.

In the table below, one starts with a homeowners non-catastrophe pretax operating profit margin of 3%. At a 2.5 to 1 premium to surplus ratio, this is equivalent to about a 9.4% after-tax return on surplus (((2.5 x 3 + 7) x .65) = 9.4), assuming surplus can be invested at 7% pre-tax.

Calculation of the Hurricane Risk Margin as a Function of the Non-Catastrophe Risk Margin

(1)	*% of Loss (2)*	*Coefficient of Variation (3)*	*Standard Deviation (4)=(2)x(3)*	*Relativity (5)*	*Risk Margin (% of Mean) (6)*	*Dollar Return (7)*
Non-Catastrophe	80%	0.08	0.064	1.00	3%	0.0240
Hurricane	20%	3.50	0.700	10.94	131%	0.2625

Next, assume that the total pure premium can be split 80% non-catastrophe and 20% catastrophe. (This split is expected to be state-specific as the hurricane loss cost in hurricane-prone states will represent a greater proportion of the total loss cost.) Based on homeowners industry data adjusted to eliminate catastrophes, the coefficient of variation of non-catastrophe loss ratios has been about 8% over the past 40 years. The corresponding coefficient of variation for hurricane losses, based on computer models, might be 350%, for example. This implies that the standard deviation of hurricane catastrophe losses would be 10.94 times the standard deviation of non-catastrophe losses.

If a 3% operating margin for non-catastrophe homeowners produces a \$2.40 operating profit on an \$80 pure premium, then the operating profit for the hurricane pure premium should be 10.94 times that, or \$26.25 (10.94 x 2.40 = 26.25). Expressed as a percentage of the pure premium, this would result in a risk margin of 131% on top of the expected hurricane loss costs. (These operating margins would include investment income from policyholder-supplied funds, and therefore need to have that quantity subtracted to derive an underwriting profit margin to be applied to loss costs.)

One can actually convert the risk margin to be a direct function of the ratio of CV's , as the risk margin incorporates the ratio of the dollar profit to the mean:

$$\text{Risk Margin}_{CAT} = \text{Risk Margin}_{NON\text{-}CAT} \times CV_{CAT} \div CV_{NON\text{-}CAT}$$

RATE FILING ISSUES

The approval of computer models as the source of expected catastrophe loss and risk margin can be a lengthy process because it changes the way regulators can verify the calculations. Under traditional filings, basic data are included with the filing, and the underlying source data are often part of statistical plan information that has been implicitly approved by the regulators in the past.

With catastrophe modeling, the frequency of events is often taken from published information tracking 100 or more years of event history. For the key catastrophe event simulation, (a hurricane or an earthquake, for example), the source is usually a scientific paper describing the ability of various equations to simulate the event. For the probabilistic model generating expected losses, often thousands of events are used, each with a specific probability derived from past distributions of input parameters.

This presents a dimensionally different approach to the regulatory approval process. It lends itself to a separate evaluation of each independent modeler — to pre-clear each model before an actual rate filing is made utilizing that model's calculation of expected loss costs. This preclearing process can take several months' time, depending on the level of due diligence needed and on the amount of rate level increase implied by the use of models to replace the old ratemaking system.

Once the independent modelers have been approved, the resulting set of indicated loss costs can provide a range of reasonable answers within which to evaluate specific company filings if the insurer has built its own model. If that company-specific model has loss costs within the pre-cleared range, that is usually prima facie evidence of the overall reasonableness of the company model. Even if the insurer model has some results outside the range, that should not necessarily disqualify the result. It merely places an additional burden on the insurer to prove the result is reasonable based on its own assumptions and judgments.

The following steps can be considered in that regulatory approval process (the details of which are included in Appendix C):

- review general design of the model
- examine event simulation module
- test ability of module to simulate known past events
- check distributions of key input variables
- perform sensitivity checks on which inputs are most important
- verify damage and insurance relationship functions
- test output for hypothetical new events
- compare different modelers' results for loss costs
- conduct on-site due diligence and review of actual assumptions.

For independent modelers, and even for insurer specific models, it is important to preserve trade secret information during the approval process and afterwards. This will affect the likelihood of future innovations to know that research and development investments can be preserved.

The on-site due diligence of regulators should keep the inner workings of the models confidential, as long as the examining process is documented by the regulator, much in the same way a financial examination of an insurance company keeps key information confidential.

Even after the approval process of a model, the regulator can preserve the confidentiality of indicated loss costs by ZIP code by not publishing the ranges that it plans to use in reviewing other company filings. First, it is better policy not to disclose the high end of the range lest some insurers be tempted to file that answer rather than using a rigorous model. Second, publishing the rate may be tantamount to the regulator setting the rate instead of approving reasonable filed rates. And third, the regulator would not be receiving the direct public attention on why the rates are so high in certain areas.

FINAL PERSPECTIVE

In summary, computer models are now capable of simulating catastrophic events and creating probabilistic models of reality that can be used to generated expected loss costs for catastrophe perils. These same models also provide a means of including the reinsurance premiums in the primary pricing process and can help quantify the needed risk load in relation to profit margins required for the non-catastrophe perils.

The same model can also be used for insurer or corporate risk analysis including reinsurance purchase decisions, and for insurer marketing and underwriting strategies. These analyses are beyond the scope of this paper.

Use of computer models for ratemaking involves a different approach from the customary one in that it is a pure premium method in contrast to the usual loss ratio method involving past insured loss experience. But that carries advantages as well as challenges, as it attempts to deal with the true underlying probabilities of loss; not just with what appears in the last few years of actual insured loss experience - which is merely a sample of what could have occurred. The computer models attempt to measure what could have occurred.

Thus, the models rely heavily on computer simulations and other technical methods newly emerging as feasible because of the vast improvement in personal computer potential. This also requires a heavy investment not only in research and design, but in resources to have the model evaluated and accepted by regulators and others.

But it is worth the process, not only for the practical results in insurer ratemaking and planning, but also for the insights gained on these catastrophic events and the reduction in uncertainty for society in dealing with them.

Furthermore, the techniques developed in producing these computer models might ultimately be applied to other perils as well. After all, the essence of actuarial work is modeling reality to assess the present financial impact of future contingent events.

References

[1] Casualty Actuarial Society, "Statement of Principles Regarding Property and Casualty Insurance Ratemaking," as adopted May 1988.

[2] Walters, Michael A., "Homeowners Insurance Ratemaking," PCAS LXI, 1974, P. 15.

[3] Kozlowski, Ronald T. and Mathewson, Stuart B., "Measuring and Managing Catastrophe Risk," Casualty Actuarial Society, 1995 Discussion Paper Program.

Sample Insurance Company
State XYZ
Expected Loss Cost
Per $1,000 of Homeowners Coverage A

Exhibit 1

Base Class. Frame
Base Deductible: $250

Zip Code Loss Costs

Base Territory (1)	Zip Code (2)	Exposure in Coverage A Amount (3)	Expected Loss Cost (4)
A	2001	3,227,000	0.351
	2002	12,495,000	0.342
	2003	8,113,000	0.421
	2004	9,204,000	0.482
B	2005	1,198,000	1.232
	2006	3,254,000	1.425
	2007	6,681,000	1.647
	2008	11,341,000	1.552
C	2009	7,295,000	2.565
	2010	6,400,000	2.752
	2011	8,508,000	2.832
	2012	9,212,000	3.011
D	2013	17,346,000	3.742
	2014	15,212,000	3.953
	2015	13,900,000	4.032
	2016	6,573,000	4.211
	Total	139,959,000	2.464

Territory Loss Costs

Base Territory (1)	Exposure in Coverage A Amount (2)	Expected Loss Cost (3)
A	33,039,000	0.401
B	22,474,000	1.545
C	31,415,000	2.806
D	53,031,000	3.937
Total	139,959,000	2.464

Notes:
(2): In-force Coverage A amounts as of June 30, 1995.
(3): Expected Loss Costs derived from probabilistic hurricane modeling.

LOSSCOST.XLS 11/28/95

Appendix A

HOW TO CONSTRUCT A MODEL

The severity component of catastrophe models generally contain three modules which are initially built separately but eventually integrated. These modules are:

- event simulation (science)
- damagability of properties (engineering)
- loss effect on exposures (insurance)

Described below is the level of research and testing that must be performed to develop a catastrophe model before it can be used for ratemaking purposes.

Science Module

As a first step, the modeler must incorporate the physics of the natural phenomena in a module (also called the event generator module) that simulates as closely as possible the actual event. Examples of input for a hurricane model include the radius of maximum winds, pressure differential at the eye of the storm (ambient pressure minus central pressure), forward speed, angle of incidence, landfall location and directional path. For an earthquake model, factors such as magnitude, location of the epicenter, soil conditions, liquefaction potential and distance from the fault rupture are used to estimate the shaking intensity of the ground at a given location.

Complete testing of the event generator module must be performed to ensure that it can be used both for the reproduction of historical events and for the simulation of hypothetical or probabilistic events. As a first step, actual wind speed records for recent events should be compared to modeled results. Organizations such as the National Hurricane Center can provide actual recorded conditions for historical events. Second, the hurricane model should be used, and its accuracy tested, to predict wind speeds for hypothetical events along the Atlantic and Gulf coasts. Since one of the key drivers of a hurricane model is the terrain or roughness parameter, this testing will help evaluate the sensitivity of the model to this factor and will allow the modeler to perform the necessary refinements to the initial assumptions.

The predictive accuracy of the model is limited by the fact that some site-specific factors that affect the way an event behaves on a given property (e.g., topographic peculiarities that affect wind speeds, liquefaction propensity at a given location for earthquakes) cannot be captured and modeled. Therefore, one should not expect a model to exactly reproduce a single past event, but rather verify that it can simulate adequately hypothetical events with a given set of parameters. Thus, actual future events with other site differences do not require major modifications to the model, but rather provide additional information to further refine it. The two maps attached are modeled replications of Hurricane Hugo and of a simulated earthquake (of a 7.5 magnitude on the Richter scale) on the Newport-Inglewood fault in Southern California.

Engineering Module

Once the event generator has been developed, damageability functions are needed to estimate the damage to a property subject to an event of a given intensity. Input from various fields of the engineering profession, such as wind engineering and structural engineering, must be gathered to develop these functions. For damage by hurricane wind speeds, numerous studies have been performed that estimate these relationships. The functions should vary by line of business, region, construction, and coverage (building versus contents).

As was the case for the event generator module, accuracy of the damage functions is improved by analyzing actual past events. Actual loss experience of insurance companies should be compared to modeled losses in the most refined level of detail available. Whereas only aggregate loss amounts by catastrophe used to be collected by companies, it is now generally possible to see loss data at least by line of business and county (or even ZIP code).

Next, on-site visits to the locations of catastrophes can help assess the damageability of exposed structures. While not imperative, these visits provide additional insight to the modeler, especially in identifying future classification distinctions.

The refinement of the damage functions is an ongoing process that is dependent on input generally provided by the engineering community. Engineering studies and loss mitigation reports are constantly being published, and their conclusions should be adapted and incorporated into the damage functions being used in the catastrophe model.

Insurance Module

Once the science and engineering modules have been developed, they must be integrated with the insurance module to determine the resulting insured loss from a given event. Kozlowski and Mathewson [3] stress the importance of developing and maintaining a database of in-force exposures that captures the relevant factors that can be used in assessing the damage to a given risk. This database will not only include such factors as location, construction type, number of stories, age of building and coverage limits, but also replacement cost provisions, deductibles, co-insurance and reinsurance (both proportional and non-proportional).

Integration of Modules

The table below presents a sample calculation of the loss estimate generated by the model for a sample hurricane after integrating the three modules.

Sample Calculation of Hurricane Losses

Zip Code	*Exposure Amount*	*Deductible*	*Windspeed (mph)*	*Corresponding Damage Factor*	*Gross Resulting Loss*	*Net Resulting Loss*
2001	$180,000	$250	100	.15	27,000	26,750
2002	180,000	500	90	.08	14,400	13,900
2003	180,000	2%	80	.05	9,000	5,400

The example assumes that we have one single family dwelling in each zip code, each with a different deductible. Based on the parameters of the storm simulated, the event generator module calculates the average windspeed sustained by all structures within the zip code. In this case, the windspeeds decrease as the zip codes are further away from the coast.

The damageability module then predicts the damage sustained by each structure as a function of the windspeed. The damage factors generally vary based on factors such as construction type (e.g. frame versus wind-resistive), age of building and number of stories. The gross resulting loss is then calculated by multiplying the exposure amount by the damage factor. The estimate is then adjusted for insurance features such as deductibles and reinsurance. In this example, the gross loss is reduced by the deductible to derive the net resulting loss.

HOW TO VALIDATE

The final task in developing a catastrophe model lies in validating the simulated results. While intermediate levels of calibration are performed for each module, the modeler must verify how they interact by completing an overall analysis of the results.

Because the model purports to simulate reality, actual incurred loss experience is the obvious candidate to be used in testing modeled losses. It is important to realize that all comparisons are dependent on the quality of the data captured from the loss records of insurers. As described above, the modeler should gain access to various sets of insured loss data and verify that all relevant factors are reflected in the model. These would include line of business, construction class, coverage (e.g., building versus contents), and loss adjustment expense (LAE) as a percentage of loss.

One issue that is often raised when validating a catastrophe model is demand surge (or "price gouging"). Because this phenomenon is dependent on the time, size and location of the event, it should not be incorporated in the damage functions except to the extent it is "expected." For example, most models underestimated the actual losses from Hurricane Andrew. If the models were adjusted to exactly reproduce Andrew's losses, they would effectively include a provision for factors that were specific to Andrew and are not expected in the long run, for example:

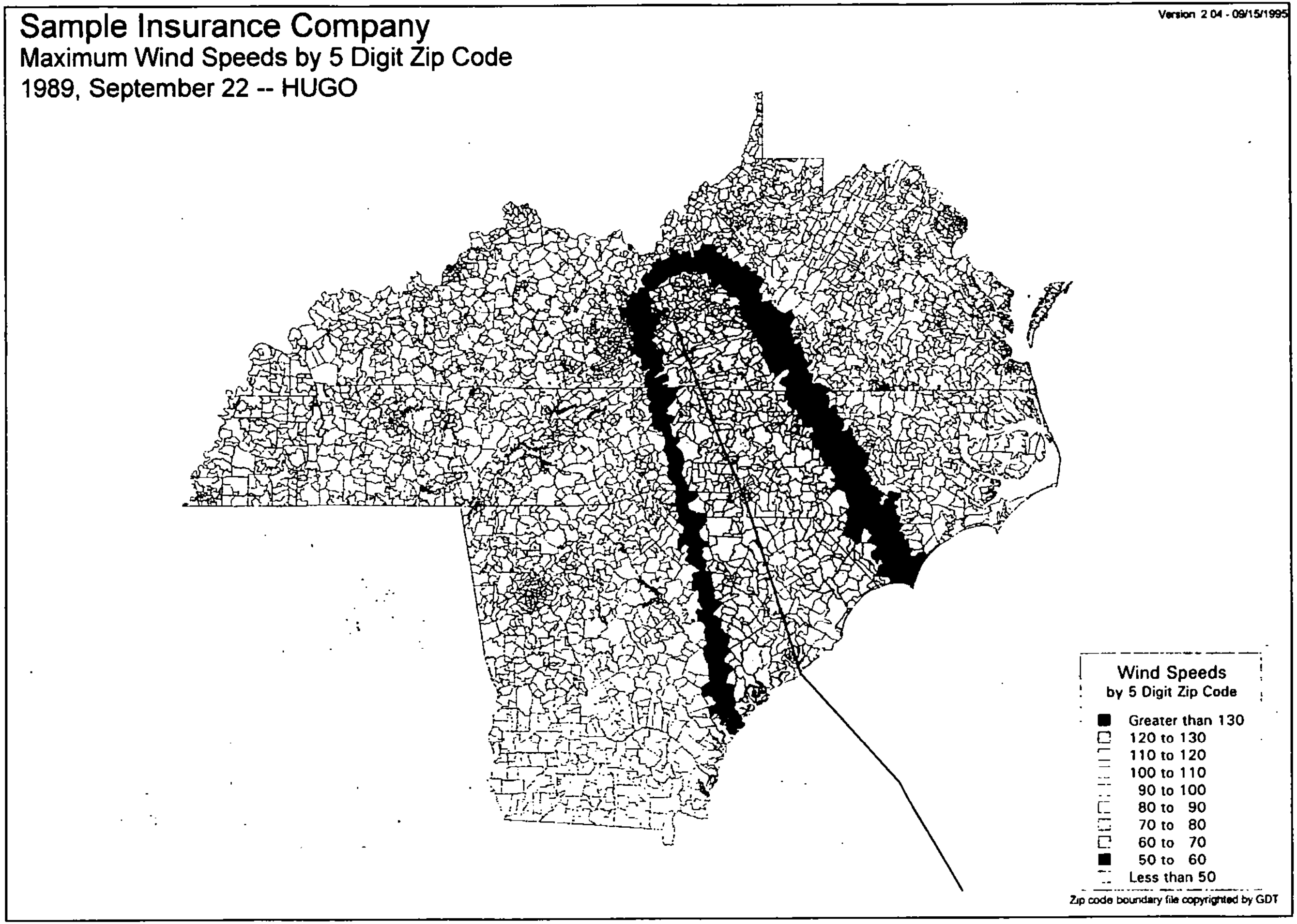
- inflation in reconstruction costs due to the excess of demand over supply
- excess claim settlements, as adjuster resources were overwhelmed by the volume of claims.

While these factors can be included separately in the reproduction of a single storm, they should not be part of the base model because they would inappropriately increase the expected level of future losses.

Another issue is storm surge from a hurricane, which as a flood loss is not officially covered by a homeowners policy. However, cynics expect that some adjustment of losses on houses affected will likely construe coverage from wind damage prior to the house being flooded. This can be handled with a small additional factor on those locales in low areas most susceptible to surge. However, from a ratemaking and rate filing standpoint, it is difficult to support much of an increase from a coverage that does not strictly apply to homeowners.

Appendix A

Map 1



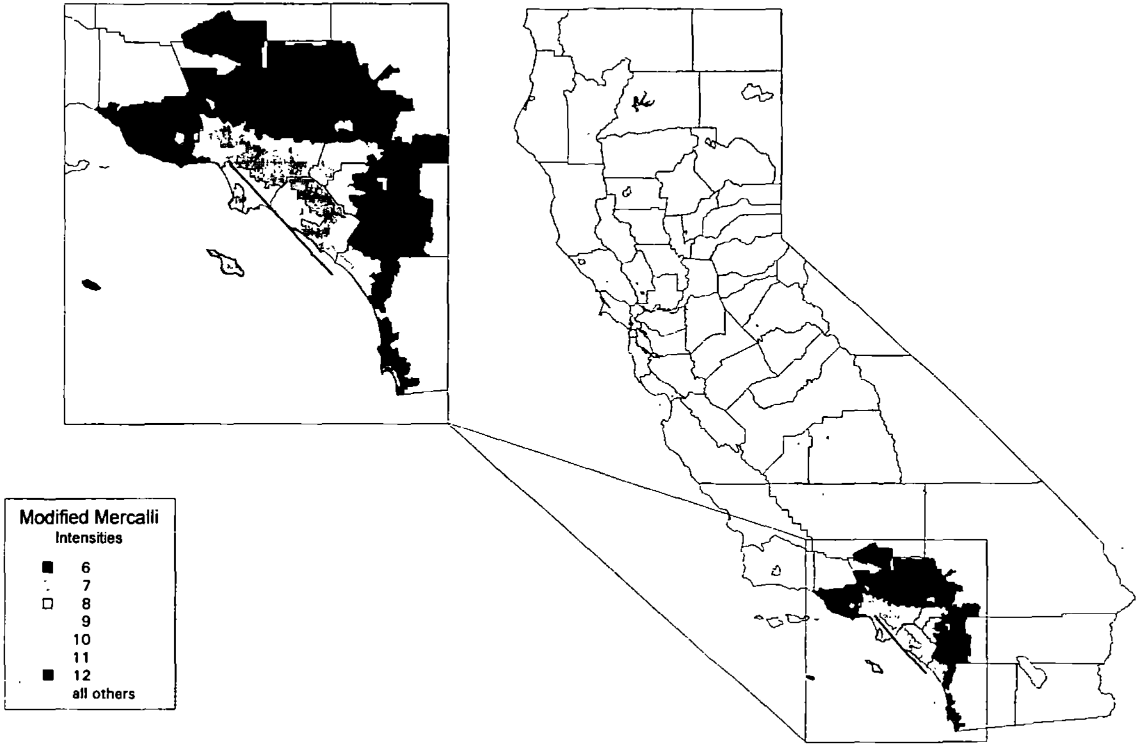
Sample Insurance Company
Maximum Wind Speeds by 5 Digit Zip Code
1989, September 22 -- HUGO

Map 2

Sample Insurance Company

Version: 2.04 - 08/10/1995

Shaking Intensities for California Region - Newport Inglewood Fault - Magnitude 7.50



Densities based on Zip Code, County Boundaries overlaid on top.

HOW OTHER PERILS ARE MODELED

Earthquake

Given that the library of historical earthquake events producing significant insured losses is scant compared to historical hurricane events, it is generally not expected that the level of precision of a computerized earthquake model will soon reach that of a hurricane model. Nevertheless, numerous models have been developed and a great amount of research has been done to define the various factors and relationships at play.

In the science module, the modeler attempts to reproduce the event by simulating shaking intensities in a ZIP code. As a starting point, the magnitude of an earthquake is generally expressed as a unit on the Richter scale. This implies a rupture length on a fault. Using other factors such as distance to the rupture, soil conditions and the liquefaction potential of the areas affected, the model estimates the shaking intensity for each ZIP code. The resulting shaking intensities are then usually converted to the Modified Mercalli Intensity (MMI) scale. This conversion is made necessary by the fact that most models use the ATC-13 damage functions as a starting point in their models.

The insurance module for an earthquake model is generally similar to a hurricane model. However, the use of percentage deductibles (which is not common on a standard homeowners policy) and separate coverage deductibles present a new twist to the equation. Hence, the model developed must have the capability of handling various deductible combinations. For instance, some earthquake policies apply a building deductible that is distinct from the contents deductible and the additional living expense deductible. A good model will apply the deductible credit separately for each coverage.

The insured loss data available to validate an earthquake model is more limited than for hurricanes. Also limiting is the fact that earthquakes are not all similar. For instance, most major faults in California have been of the strike-slip type. Yet the 1994 Northridge quake was a "blind" thrust-fault earthquake. These two types of earthquakes are by their nature very different and will cause a modeler to adjust the event generator model to reflect different shaking intensities.

Once the deterministic earthquake model has been developed, a probabilistic version must be generated. For earthquake modeling, a set of known faults is generally used as a starting point in building the library of events. Events of various strengths and locations are simulated for each fault. A probability is then assigned to each event in the library. These probabilities are generally expressed in a return time format such as 1 in 400 years. They can be obtained from geological sources such as the United States Geological Survey.

The Northridge event highlighted the fact that serious damage could be caused by earthquakes not located on well-known fault systems. This has implications for earthquake ratemaking because, while the frequency of these events is very much unknown at this time, inclusion of this type of event could increase the expected loss costs substantially. However, the modeler needs to take care that the long-run frequency of earthquakes remains reasonable.

Tornado and Hail

The actual loss experience of tornadoes and hailstorms is more readily available than for any other type of natural catastrophe. Given that there are roughly 1,000 tornadoes in the U.S. each year, the traditional way of developing a tornado/hail catastrophe loading in states with exposure to these perils has been to spread the actual loss experience over a number of years. However, this methodology does not get at the essence of why catastrophe modeling is the preferred approach, which is to estimate the current loss potential of a company given its distribution of exposures. Also implicit in any modeling approach is the simulation of events that have not occurred but are reasonably foreseeable given the historical database of events.

Tornadoes and hailstorms are typically generated by inland storms where moist, warm air masses collide with cooler, drier air masses. Such conditions are often present in the southwestern United States (northern Texas, Oklahoma) and the plains states (Iowa, Kansas, to cite some specific examples) where the Gulf of Mexico provides a continuous source of warm, moist air, and the Rocky Mountains create a source of cooler drier air as weather systems move over them. Tornadoes do, however, occur in all fifty states.

An inland storm capable of generating tornadoes may create one, or tens, of individual funnels over a widely dispersed area. A single funnel will produce damage over the portion of its track making contact with the earth. The length of that ground contact track can range from tens of feet to two hundred miles or more, and the width of the track within which damage can be produced by that funnel can range from tens of feet to a mile or more.

In order to accurately model the loss effects of a single funnel, it is therefore necessary to consider the small scale (nine digit zip code) location of exposures relative to the funnel path.

Because tornadoes and hailstorms are more sudden and unpredictable than hurricanes, most historical information has been the result of human observation. Current tornado databases generally consist of date and time, initial observed location, path width, path length and storm intensity for each event. Tornado intensity is generally measured on the basis of the Fujita scale, which translates an expected degree of damage to a range of windspeeds. For example, a tornado with a Fujita-scale intensity of F2 will be expected to tear roofs from frame houses. Engineering studies indicate that damage of this intensity can be generated by windspeeds of between 113 and 157 miles per hour.

Tornadoes do not behave like hurricanes. The spinning funnel-shaped updraft of a mature tornado is the most damaging windstorm produced by nature. Hence, the damage relationships at a given windspeed for a tornado are quite different from those of a hurricane. This indicates that the results of engineering and damage studies specific to tornadoes must be collected to develop a representative model.

The development of a hail model resembles that of a tornado model. However, difficulties lie in the definition of what is considered a hailstorm and which hailstorms are associated with tornadoes that are already included in a tornado database. The interpretation of the data present in the databases therefore has a significant impact on the overall frequency assumptions used in both models.

The validation of a tornado and/or a hail model against actual loss experience is dependent on the availability of loss data and on how much differentiation between the two perils is possible. (If this cannot be obtained, the modeler may have to calibrate the models on a combined basis. As a result, this would make the development and justification of territorial loss costs for all severe local storm perils easier.)

Winter Storm

Winter storm and freeze activity has been quite severe over the last few years. As a result, the need for better risk measurement and expected loss calculations has increased. Also, some of the same characteristics as hurricane prompt the use of a catastrophe model to simulate winter storm losses - changes in exposure and longer return periods than in an individual insurer's data base.

However, contrary to the other catastrophe perils, winter storms do not have a specific unit of measure that describes the intensity of a given event, and individual temperature is not the only factor that can describe these events. For example, wide temperature swings and absolute highs and lows over consecutive days have been identified as some of the factors that impact the intensity and duration of these events.

The damage functions associated with winter storms are also very different from those of the other perils. Because little of the damage is structural, damage functions are less severe than those of hurricanes, for example.

Similar to a hurricane model, the creation of a probabilistic database requires simulation of multiple events. While the parameters are different, each event is defined by a location (or landfall), size, intensity and duration.

Because individual winter storms have not been as surplus-threatening as hurricanes or earthquakes, the motivation to develop computer models has not been as high for risk analysis and development of PMLs. However, for ratemaking, this peril is equally as compelling as hurricane towards the use of computer modeling. Not only does it yield better expected loss estimates, but it allows the exclusion of past catastrophes from the normal homeowners ratemaking data base for better stability in rate level indications.

Appendix C

METHODS TO REVIEW CATASTROPHE MODELS IN REGULATORY PROCESS

1. Review general design of model

- Examine the credentials of the modeler
- What is the scientific basis for the key event simulation?
- What is the engineering support for the damage factors produced by each event severity?
- Are the insurance limitation features reasonable, e.g., deductibles, coinsurance and reinsurance calculations?

2. Examine event simulation module

- What are the credentials of the scientists who specified it?
- Has their work been published and/or peer reviewed?
- What special insights are they offering on the particular event to be simulated?

3. Test event generator's ability to simulate known past events

- Use published information from some critical events, such as Hurricanes Andrew and Hugo, the Loma Prieta earthquake (1989) or even the 1906 San Francisco earthquake
- Input some key parameters, such as central pressure, landfall, speed and radius of maximum wind, and examine the output windfield at various locations compared to published information on windspeeds. This can be done for any event, even if no current estimates of insured losses are available, as a test of the event simulation accuracy.

4. **Check key input distributions**
    - Compare the distributions of key input values among the different modelers, to see if there is any disparity in the key drivers of results. For hurricanes, a possible approach could be to look at the:
        - Distributions of central pressure at ten millibar intervals: 900-909, 910-919 etc.
        - Distributions of radius of maximum winds in five nautical mile ranges, and forward speeds in five knot ranges.
        - Probabilities of landfall for all storms affecting the state (direct hit and nearby landfalls).

5. **Conduct sensitivity checks**
    - Use a few sample events
    - Promulgate a sample exposure base statewide (e.g., 25 risks)
    - Vary the parameters one at a time, or perhaps a few in pairs
    - Observe changes is output (insured losses) for incremental changes in input
    - The goal is a rough measurement of the effect of changing inputs (e.g., central pressure, radius of maximum winds, forward speed)

6. **Verify damage and insurance relationship functions**
    - Examine the credentials of the engineers
    - Has the analysis been published and/or peer reviewed?
    - Analyze the damage curves (functions of increasing damage for increasing event intensity) separately for types of exposure, class and coverage
    - Review the insurance module for effects by deductible and reinsurance or coinsurance
    - Review the validation of the two components (damage and insurance effects) via multiple events over the past few years for multiple insurers; each event does not have to be replicated, but that they should average out over all events and all insurers.

7. Test output for hypothetical new events

- Select some new events defined by key parameters
- Use a sample database of exposures by ZIP code
- Compare results for different modelers and ask outside experts for their opinions on the reasonableness of these results.

8. Compare indicated loss costs for different modelers

- Select sample ZIP codes throughout the state
- Have modelers run all events with probabilities for those ZIP codes
- Use several base classes and coverages:
  - homeowners, $100,000 frame house, $250 deductible
  - tenants, $30,000 contents, masonry, $250 deductible
  - businessowners, $200,000, masonry, $1,000 deductible
- Compare modelers' loss costs per $1,000 of coverage by ZIP code
- Ask outliers to explain large differences from average.

9. Conduct on-site due diligence and review of key assumptions

- View a live running of the model, with actual input data
- Review input data sources — published and non-published
  - all key input parameters
  - frequency of events by location
  - key damage factors and sources
- Review output, including color coded maps showing ranges of expected loss costs.