

The Catastrophe of Dumbing Down Catastrophe Models

by Jayant Khadilkar

The use of catastrophe models has, over the last 20 years, become the norm for disaster risk management. Although some cat models existed earlier, it was in 1992 after Hurricane Andrew that the landscape of property catastrophe risk management changed and the importance of having a probabilistically simulated, event-based approach to evaluating risk was better understood.

Since Andrew, cat models have evolved to encompass the latest in scientific research for an increasing variety of perils. Nevertheless, many catastrophic events over the past decade (the multiple Florida hurricanes of 2004, Hurricane Katrina in 2005, Hurricane Ike in 2008, the swarm of tornados in 2011) have called attention to the fallibility of cat models and the insurance industry's overconfidence in their results. Cat models are unquestionably a valuable tool, but they are not the complete answer.

Cat models are built using limited data sources and scientific approximations and, as such, remain incomplete and have an abundance of built-in uncertainty. For example, Atlantic hurricane statistics are available back to 1851, but complete data did not become available until the advent of weather satellites in the 1950s. Given that U.S. hurricane models are built on this very limited historical data, how much confidence can a user have in a cat model's accuracy in determining what might be a 1-in-100-year event?

As another example, there are a number of scientifically valid methods to describe how hurricanes weaken after making landfall. Each of these methods is an approximation of reality that can produce significantly different results. Cat model developers choose one of these methods for their model and make many more choices and approximations before they arrive at a finished product. Therefore, the results produced by a single cat model have a lot of uncertainty around them. That is why different cat models produce different answers; the developers used different (though equally valid) methods and approximations for their models.

Despite these shortcomings, cat models have nevertheless represented a quantum leap for the insurance industry and how it considers catastrophic risk. Models provide a common framework for tying together hazard, vulnerability and exposure data. Models allow for the sharing of information in a consistent and recognized format. Risk takers can evaluate potential catastrophe scenarios and roughly estimate the probabilities of different sizes of loss. Where cat models excel is not the absolute measurement of risk, but in the relative evaluation of risk.

A Single Point of Failure

Virtually all cat models generate a loss distribution curve as part of their risk evaluation. A single point on the curve (though not always the same point), known as probable maximum loss (PML), has become the most commonly used measure of risk.

Many practitioners are justifiably uncomfortable with reducing the entire output of a cat model down to a single point. The credibility of loss estimates is significantly reduced as focus is narrowed down to either a single location or a single point on a loss distribution.

Yet the use of PML remains popular for a variety of reasons. It is a simple way to express the results of a complex model. The models themselves make it easier to produce single point estimates of PML. And regulators and rating agencies use it to calculate the financial strength of a risk taker.

If you ask people within the same organization for their understanding of PML, more likely than not you will get multiple interpretations. The definition of PML varies from business unit to business unit and from company to company. One might consider the 1-in-100-year occurrence loss to be the PML; another might look at the 1-in-250-year aggregate loss.

Yet the broader questions remain unanswered by the PML. How can decisions be

made based on a single point on the distribution? By its very nature, a single point on a loss distribution is unstable. Using this measure for complex decision making, such as the marginal impact of underwriting new policies, is a dangerous proposition as it provides limited information about the new policy. PML, since it is a single point on the distribution, is not additive across multiple risks. Unless the portfolio PML used for marginal calculation is refreshed every time a risk is added, this marginal impact on PML can lead to an undesirable portfolio. Meanwhile, the usual industry practice is to refresh the portfolio every month-or sometimes less frequently.

Another industry practice is to use PML to optimize a defined portfolio. This can lead to a hypothetical portfolio that performs well in theory but not well in practice. There are several reasons for this mismatch of on-paper and in-practice outcomes. Given that the optimization was conducted using either a single point, or in some cases a few points around the PML, if the exact resulting portfolio is not achieved in practice then its PML will not match the hypothetical portfolio PML. A portfolio just cannot be represented by such a narrow range of potential outcomes.

Because the PML has great uncertainty surrounding it and cannot measure all changes in the portfolio, the industry should become less reliant on it. Perhaps it would be more accurate to refer to PML as "probable meaningless loss" instead.

This does not, however, mean that cat models are without merit. Cat models are extremely valuable tools that are meant to provide what-if scenarios for multiple potential outcomes. Indeed, cat models can provide the basis for the development of a better risk measure, one that leverages their best parts and compensates for their failings.

Creating the Right Model

When we think about creating a more meaningful risk measure, there are a number of desirable features that it should include. It absolutely must be transparent, easy to understand and relatively stable. It should also be a consistent "yardstick" to measure risk across the entire organization and able to support day-to-day risk decision making. And it needs to be capable of taking into account the "right" part of the potential cat loss distribution.

This proper distribution will strongly depend on a company's risk tolerance and risk appetite. A mutual insurance company in the Midwest, for example, might be solely concerned about severe surplus erosion from a single event, and, as such, a measure capturing the extreme tail might be appropriate. A publicly traded company, for which quarterly earnings are an additional concern, might want to adopt a measure capturing the near-term volatility of risk as well as the tail.

The insurance industry needs to move away from depending on the PML and adopt a more company-specific risk measure. In order for this to happen, all constituents (risk takers, regulators, rating agencies, brokers) need to accept the fact that no two companies are similar when it comes to their catastrophe risk.

Cat models are and will remain a critical part of understanding, measuring and pricing catastrophe risk. But we must be wary of over-reliance and over-simplification of model results. To use cat models most effectively, we should always be conscious of their strengths and weaknesses. One clear step for increasing the quality of our decisions is to reduce our reliance on the PML and adopt other risk measures that consider all the events that may put stress on a company.

Jayant Khadilkar is a partner at TigerRisk, a Stamford, Connecticut-based reinsurance broking and risk/capital management firm