

The RMS Blog

TAG ARCHIVES: MODEL BLENDING

Will it Blend? (...And Now What?)

In previous posts on multi-modeling, Claire Souch and I discussed the importance of [validating models](#) and the [principles of model blending](#). Today we consider more practical issues: how do you blend, and how does it affect pricing, rollup, and loss investigation?

No approach to blending is superior. There are trade-offs between granularity of blending ratios allowed, ease of implementation, mathematical correctness, and downstream functionality. For this reason, several methods should be examined.

Hybrid Year Loss Table (YLT)

The years, events, and losses of a hybrid YLT are selected from source models in proportion to their assigned blending weights. For example, a 70/30 blend of two 10,000-year YLTs is composed of 7,000 years selected at random from Model A and 3,000 years selected at random from Model B.

Hybrid YLTs preserve whole years and events from their source models, and drilling down and rolling up still function. But the blending could have unexpected results, as users are working with YLTs instead of directly altering EP curves. Blending weights cannot vary by return period, which may be a requirement for some users.

Simple Blending

Simple blending produces a weighted average EP curve across the common frequencies or severities of source EP curves. Users might take a weighted average of the loss at each return period (severity blending) or a weighted average of the frequency of loss thresholds (frequency blending). [Aspen](#) has produced [an excellent](#)

[comparison](#) of the pros and cons of blending frequencies vs. severities.

Simple blending is appealing because it is easy to use and understand. Any model with an EP curve can be used—it need not be event-based—and weights can vary by return period or loss threshold. It is also more intuitive: instead of modifying YLTs to change the resulting EP curve in a potentially unexpected way, users operate on the same base curves. Underwriters may prefer it because they can explicitly factor in their own experience at lower return periods.

While this is useful for considering multiple views at an aggregate level, the result is a dead end. Users can't investigate loss drivers because a blended EP curve cannot be drilled into, and there is no YLT to roll up.

Scaled Baseline: Transfer Function

Finally, blending with a transfer function adjusts the baseline model's YLT with the resulting EP curve in mind. Event losses are scaled so that the resulting EP curve looks more like the second model's curve, to a degree specified by the user.

Unlike the hybrid YLT and simple blending methods, the baseline event set is maintained, and the new YLT and EP curve can be used downstream for pricing and roll-up. Blending weights can also vary by return period.

However, because it alters event losses, some drill-downs become meaningless. For example, say Model A's 250-year PML for US hurricane is mostly driven by Florida, while Model B's is mostly driven by Texas. If we adjust the Model A event losses so that its EP curve is closer to Model B's, state-level breakouts will not be what either model intended.

Alternatives

Given the downstream complexities of blending, it may be preferable to adjust the baseline model to look more like an alternate model, without explicitly blending them. This could be a simple scaling of the baseline event losses, so that the pure premium matches loss experience or another model. Or with more sophisticated software, users could modify the timeline of simulated events, hazard outputs, or vulnerability curves to match experience or mimic components of other models.

Where does that leave us?

Over the past month, we've explored why and how we can develop a multi-model view of risk. [Claire](#) pointed out that the foundation to a multi-model approach first requires validation of the individual models. Then I discussed the motivations for weighting one model over another. Finally, we turned to how we might blend models, and discovered its good, bad, and ugly implications.

Multi-modeling can come in many forms: blending results, adjusting baseline models, or simply keeping tabs on other opinions. Whatever form you choose, we're committed to building a complete set of tools to help you understand, take ownership of, and implement [your view of risk](#).

This entry was posted in Risk Modeling, RMS(one) and tagged Model Blending on November 11, 2013 [http://www.rms.com/blog/2013/11/11/blending-now-what/] by Meghan Purdy.

A Weight On Your Mind?

My colleague [Claire Souch](#) recently discussed the most important step in model blending: [individual model validation](#). Once models are found suitable—capable of modeling the risks and contracts you underwrite, suited to your claims history and business operations, and well supported by good science and clear documentation—why might you blend their output?

Blending Theory

In climate modeling, the use of multiple models in “[ensembles](#)” is common. No single model provides the absolute truth, but individual models’ biases and eccentricities can be partly canceled out by blending their outputs.

This same logic has been applied to modeling catastrophe risk. As [Alan Calder, Andrew Couper, and Joseph Lo of Aspen Re](#) note, blending is most valid when there are “wide legitimate disagreements between modeling assumptions.” While blending can’t reduce the uncertainty from relying on a common limited historical dataset or the uncertainty associated with randomness, it can reduce the uncertainty from making different assumptions and using other input data.

Caution is necessary, however. The forecasting world benefits from many models that are widely accepted and adopted; by the law of large numbers, the error is reduced by blending. Conversely, in the catastrophe modeling world, fewer points of view are available and easily accessible. There is a greater risk of a blended view being skewed by an outlier, so users must validate models and choose their weights carefully.

Blending Weights

Users have four basic choices for using multiple valid models:

1. Blend models with equal weightings, without determining if unequal weights would be superior
2. Blend models with unequal weightings, with higher weights on models that match claims data better
3. Blend models with unequal weightings, with higher weights on models with individual components that are deemed more trustworthy
4. Use one model, optionally retaining other models for reference points

On the surface, equal weightings might seem like the least biased approach; the user is making no judgment as to which model is “better.” But reasoning out each model’s strengths is precisely what should occur in the validation process. If the models match claims data equally well and seem equally robust, equal weights are justified. However, blindly averaging losses does not automatically improve results, particularly with so few models available.

Users could determine weights based on the historical accuracy of the model. In weather forecasting, this is referred to as “[hindcasting](#).” RMS’ medium-term rate model, for example, is actually a weighted average of thirteen scientific models, with higher weights given to models demonstrating more skill in forecasting the historical record.

Similarly, cat model users can [compare the modeled loss from an event with the losses actually incurred](#). This requires detailed claims data and users with a strong statistical background, but does not require a deep understanding of the models. An event-by-event approach can find weaknesses in the hazard and vulnerability modules. However, even longstanding companies lack a long history of reliable, detailed claims data to test a model’s event set and frequencies.

Weights could also differ because of the perceived strengths of model components. Using modelers’ published methodologies and model runs on reference exposures, expert users can score individual model components and aggregate them to score the model’s trustworthiness. This requires strong scientific understanding, but weights can be consistently applied across the company, as a model’s credibility is independent of the exposure.

Finally, users may simply choose not to blend, and to instead occasionally run a second or third model to prompt investigations when results are materially different from the primary model.

So what to do?

Ultimately, each risk carrier must consider its personal risk appetite and resources when choosing whether to blend multiple models. No approach is definitively superior. However, all users should recognize that blending affects modeled loss integrity; in our next blog, we’ll discuss why this happens, and how these effects vary by the chosen blending methodology.

This entry was posted in Risk Modeling, RMS(one) and tagged Model Blending, Open Modeling on October 8, 2013 [<http://www.rms.com/blog/2013/10/08/a-weight-on-your-mind/>] by Meghan Purdy.

Blend It Like Beckham?

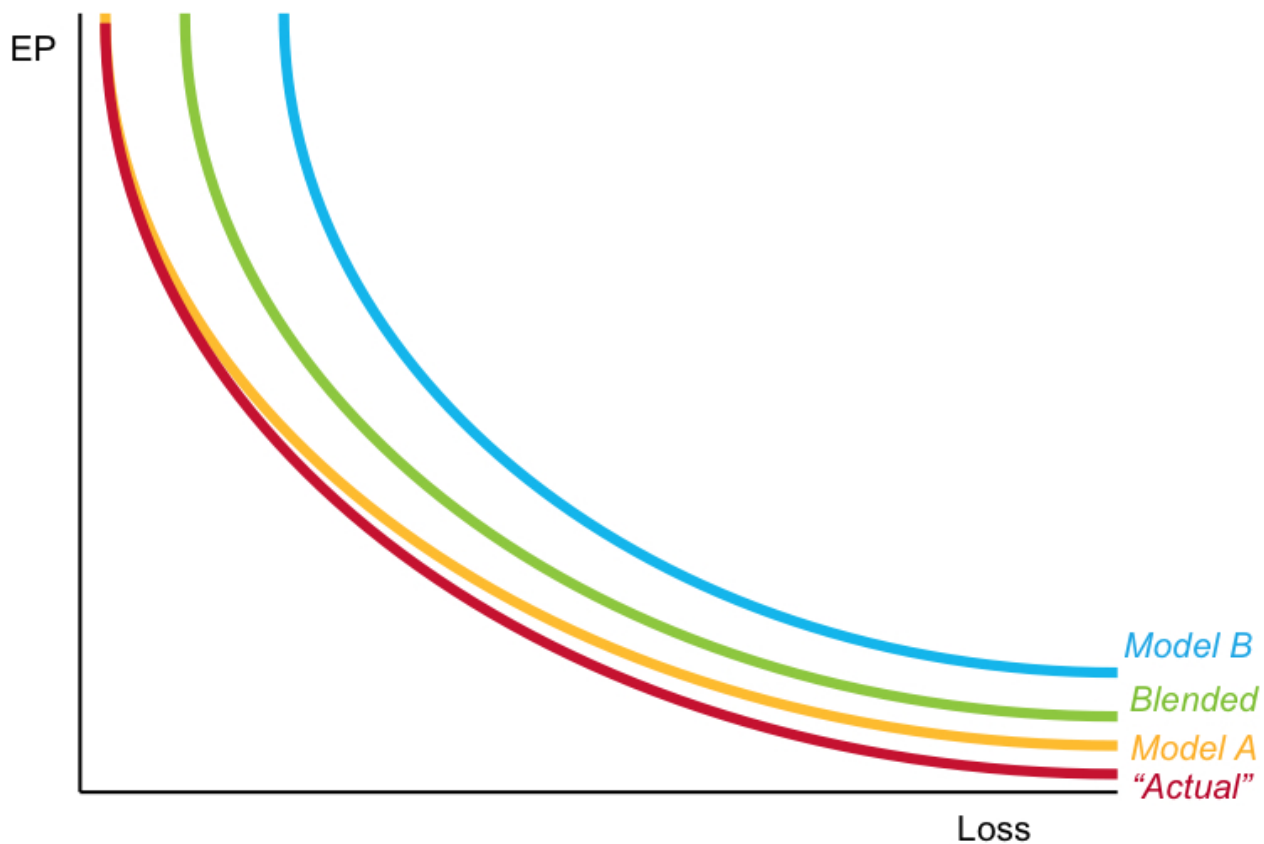
While model blending has become more common in recent years, there is still ongoing debate on its efficacy and, where it is used, how it should be done.

As RMS prepares to launch [RMS\(one\)](#), with its ability to run [non-RMS models](#) and blend results, the discussion around multi-modeling and blending “best practice” is even more relevant.

- If there are multiple accepted models or versions of the same model, how valid is it to blend different points of view?
- How can the results of such blending be used appropriately, and for what business purposes?

In two upcoming posts, my colleague Meghan Purdy will be exploring and discussing these issues. But before we can discuss best practices for blending, we need to take a step back: any model must be validated before it is used (either on its own or blended with other models) for business decisions. Users might assume that blending more models will always reduce model error, but that is not the case.

As noted by the 2011 report, [Industry Good Practice for Catastrophe Modeling](#), “If the models represent risk poorly, then the use of multiple models can compound this risk or lead to a lesser understanding of uncertainty.”



Blending Model A with a poor model, such as Model B, won't necessarily improve the accuracy of modeled losses

The fundamental starting point to model selection, including answering the question of whether to blend or not, is model validation: models must clear several hurdles before meriting consideration.

1. The first hurdle is that **the model must be appropriate for the book of business, and for the company's resources and materiality of risk**. This initial validation is done to determine each model's appropriateness for the business, and is a process that should preferably be owned by in-house experts. If outsourced to third parties, companies must still demonstrate active ownership and understanding of the process.
2. The second hurdle involves **validating against claims data and assessing how well the model can represent each line of business**. Some models may require adjustments (to the model, or model output as a proxy) to, for example, match claims experience for a specific line, or reflect validated alternative research or scientific expertise.
3. Finally, the expert user might then look at how much data was used to build the models, and the methodology and expertise used in the development of the model, in order **to discern which might provide the most appropriate view of risk for that company to use**.

Among the three major modeling companies' models, it would not be surprising if some validate better than others.

After nearly 25 years of probabilistic catastrophe modeling, it remains the case that all models are not equal; very different results and output can arise from differences in:

- Historical data records (different lengths, different sources)
- Input data, for example the resolution of [land use-land cover data](#), or elevation data
- Amounts and resolution of [claims data](#) for calibration
- Assumptions and modeling methodologies
- The use of proprietary research to address unique catastrophe modeling issues and parameters

For themselves and their regulator or rating agency, risk carriers must ensure that the individuals conducting this validation and selection work have the sufficient qualifications, knowledge, and expertise. Increasing scientific knowledge and expertise within the insurance industry is part of the solution, and reflects the industry's increasing sophistication and resilience toward managing unexpected events—for example, in the face of record losses in 2011, a large proportion of which was outside the realm of catastrophe models.

There is no one-size-fits all “best” solution. But there is a best practice. Before blending models, companies must take a scientifically driven approach to assessing the available models' validity and appropriateness for use, before deciding if blending could be right for them.

This entry was posted in Risk Modeling, RMS(one) and tagged Model Blending, Open Modeling on September 25, 2013 [<http://www.rms.com/blog/2013/09/25/blend-it-like-beckham/>] by Claire Souch.
