OASIS LMF

Requirements for results processing in ORD

14.01.22

LM01.01.02
11013 Weston Street
London
SE1 3ER
United Kingdom

Introduction

There are many uses of catastrophe model output. Most of the focus in developing ORD so far has been on creating a standard that can accommodate the output from different catastrophe model platforms to facilitate interoperability in sharing results across the insurance market.

In this discussion paper we turn our attention to the downstream use of the ORD results and what further requirements we need to consider as part of the standard to facilitate the processing of results. To do that, we present a structure for categorizing uses of catastrophe model output and a set of specific use cases that we use to develop a set of requirements and proposed solutions against.

Results processing

We use the following terminology with respect to use of catastrophe loss model results as discussed in Section 5 Developing a View of Risk of Natural Catastrophe Risk Management and Modelling: A Practitioners Guide.

Results "adjusting"- making changes to the results output from a single model to develop a view of risk. This can include input exposure scaling, severity (event loss) scaling, frequency (event rate) scaling, uncertainty adjustments and model component adjustments.

Results "blending" – to blend output from multiple models (on the same or overlapping input exposures) to form a composite view of risk. The forms this can take are severity blending, frequency blending and rank matching.

Model "fusion" – the combining of components from different models to created a fused or blended model.

Results "roll-up" – a commonly used term by practitioners, meaning combining multiple sets of results together into one consolidated result, for the purposes of, for example, producing enterprise wide EP curves for regulatory reporting.

In addition, we use the follow term which is not explicitly defined elsewhere;

Results "alignment" - transforming one format of output into another for the purpose of getting data from multiple sources into a common format. This is often required for import into downstream modelling tools, which have specific import formats.

Use cases

Category	Number	Use case
Results adjusting	1	Applying post-analysis loss amplification factors
Results adjusting	2	Combine results for two portfolios, ran separately against the model because post-loss adjustment factors have been applied to Region A but not Region B
Results adjusting / Roll-up	3	User-defined allocation of event losses to reporting zones. Based on where an event causes the most loss, all of the loss for the event is assigned to a 'Rating Zone'. This cleanly allocates each event to one zone. The detailed losses for each Rating Zone are then used to produce high level (AAL/EP) results. The definition of Rating Zones vary by peril. For example, hail losses in Florida might be assigned to a different Rating Zone than Hurricane losses in Florida.
Roll-up / alignment	4	Preparation for import into a roll-up tool. Roll-up tool only supports events that occur once and a single sample. Event ids and sample numbers need to be reindexed so that each event id only occurs once and there is only one sample.
Roll-up / alignment	5	PLT output needs to be identifiable across perils and vendors. There is an overlap in the use of Event Ids across different perils and vendors. In order to combine the results at the detailed event or year loss level, an id is required to uniquely identify events that belong to a particular vendor model.
Other	6	Local and international business has been run in separate analyses and the results must be combined. The analyses may have been run with different number of samples
Other	7	A large portfolio has been split up and ran in separate analyses against the same model due to its size, and the results must be combined. The analyses may have been run with different number of samples. Reasons for running with different numbers of samples might be that there may be very different convergence characteristics for different parts of the portfolio, such as requiring many more samples for commercial risks compared with residential.
Other	8	Results for two different portfolios ran separately must be combined in order to price reinsurance cover. The analyses may have been run with different number of samples.
Results adjusting	9	A portfolio is run twice against the same model with two different correlation assumptions, fully correlated and independent. A combined set of results is produced by weighting the individual results to reflect user-specified correlation factors

Requirements

1. Meta data for result adjustments

Use cases 1 and 2 highlights the need for explanatory data related to the post analysis adjustments that have been applied. What form this will take depends on the variety of adjustments that may have been made. The solution could range from a single field with explanatory text, to a json blob containing information about post-analysis adjustments with a variable nested structure.

2. Reporting Level definitions

Use case 3 highlights the issue of reporting level definitions, and how they may vary from company to company. In this particular case, the Rating Zones are defined differently for each peril, which means you cannot simply have a single field in the exposures which assigns a location to a Rating Zone, and generate reports for that summary level field.

Rating / reporting zones summary levels are used in portfolio roll-up, so there is also a requirement for consistency across analyses when combining results.

ODS has standard area codes representing the largest geographical division within a country (e.g. State code) and standard peril codes, so one approach could be to define in ORD a standard mapping table into a user defined reporting zone. For example;

AreaCode	CountryCode	PerilGroup	Reporting Zone
FL	US	WW1	Florida
FL	US	OO1	US South East
CA	US	QQ1	California
CA	US	OO1	Western US.
	UK	OO1	Rest of World
…	…	…	…

The mapping table format could be standardized, but the data inside it does not have to be. This would allow results originating from different sources with common standard reporting levels to be combined according to the user's reporting zone mappings in a post analysis adjustment process.

3. Support for grouped analyses

The uses cases 2 and 6-8 show a variety of operational reasons why analyses may be ran separately and combined which are not directly related to portfolio roll-up.

If combining of analyses must happen at the point of producing the results and before potentially sharing the results in the ORD package, it raises the question of whether ORD should be able to support 'grouped' analysis results. Currently the ORD package is designed to support multiple sets of single analysis results, but not a grouped set of results.

Grouped analysis results might require the following features;

- ELT/PLT data: An additional field or fields which uniquely identifies the Event, Model, Peril in event or period loss tables.

- Meta data: a list of analysis ids and associated model meta data associated with a grouped set of results, rather than single values that are associated with single results sets.
- Documentation: to produce standard methodologies for combining outputs and also general documentation to explain the differences between loss generation methodologies which give rise to different output formats.

Methodologies for combining output

There are different ways that output data can be combined. Generally it is best practice to combine results at a detailed level, i.e. combine event losses or period loss tables, and then use the same methodology as for a single analysis to produce high level results, that is, Average Annual Loss (ALT in ORD) and EP curves (EPT in ORD). In sample-based platforms, for example, it is the Sample Period Loss Table that is used to generate the ALT and EPT.

The reason that combining detailed results is preferable to combining high level results is so that correlation may be captured in the loss uncertainty distributions in the following situations;

- Per occurrence losses - exposures in two different sets of results that are affected by the same event.
- Annual aggregate losses – capture years in which there are multiple event occurrences from different perils which affects the aggregate annual loss distribution.

Depending on which detailed reports are populated, there are many ways to do this. Therefore it might be helpful to provide documentation on the various methods of combining detailed losses. Some methods for consideration are outlined below.

Sample loss tables (SELT/SPLT)

These reports store individual loss samples by Event or Event and Period.

Event or Event by Period loss tables from multiple analyses can be merged by rows. Losses can be summed across records where there is a common EventID, or EventId and PeriodNum, respectively.

This is simple if the number of samples and periods match, but needs a more careful approach otherwise, such as resampling.

If Sample by Sample losses are not available from the source analyses, and instead Moment or Quantile losses are being combined (for instance, if the detailed sample reports are too large to store on disk) it is possible to create a combined sample loss table from the combined moment or quantile loss table by random resampling.

Moment Loss tables (MELT/MPLT)

These reports provide the moments of the loss distribution, which are mean and standard deviation of loss.

Event or Event by Period mean and standard deviations from multiple analyses can be merged by rows. Mean losses can be summed across records where there is a common EventID, or EventId and Period, respectively.

A more careful approach is needed for summing standard deviations. If standard deviation is split into correlated and independent standard deviations (SDLossCor and SDLossInd are supported already in ORD), then the SDLossCor may be summed, and the SDLossInd may be squared and summed to capture the correlation correctly. Otherwise, in the absence of any other information, squaring the total standard deviation, SDLoss, and summing across records is the default approach, and assumes complete independence between losses.

Combined moment loss tables may be randomly resampled to create a combined sample loss table, if a parametric distribution for the losses can be assumed.

Quantile Loss tables (QELT/QPLT)

Quantiles are cut points dividing the range of a probability distribution into continuous intervals with equal probabilities, or dividing the observations in a sample set in the same way. For each Event, or Event and Period, the losses are provided for each probability interval.

Event or Event by Period loss quantiles from multiple analyses may be merged by rows, even if the set of loss quantile probabilities are different.

For rows where there is a common EventID, or EventId and Period, it is possible to combine the loss quantiles by convolution and create a combined quantile loss table, or by random resampling and create a combined sample loss table.

With both random resampling and convolution, an assumption has to be made about correlation. It is possible to convolve or randomly sample the loss quantiles with full independence or full correlation.

4. Support for downstream modelling tools

Use case 4 describes a results alignment process by which multiple samples of the same event are converted into new event occurrences because the modelling tool can only accept a single sample of each event.

There are no other specific examples of other formats that results are converted to for the purposes of ingestion into other systems, and there is an open question about the requirements of downstream modelling tools and to what extent they should be supported by ORD.