OASIS LMF

Combining results in ORD

28.08.25

Executive summary

This document provides a structured, high-level methodology for combining catastrophe loss model results in the Open Results Data (ORD) format, with a primary focus on the "roll-up" of results from multiple analyses to support group- and enterprise-level risk views. The framework outlined is designed to address practical industry needs such as regulatory reporting, portfolio management, and consolidated exceedance probability (EP) curve generation, by enabling flexible yet consistent aggregation of detailed model outputs.

Key Points

- **Terminology & Scope:** The proposal distinguishes between *results roll-up* (consolidating results across multiple analyses), *blending* (merging multiple model views), and *fusion* (integrating model components), with this methodology specifically focused on roll-up scenarios.
- **Method Overview:** Central to the approach is combining event or year losses at the most granular level, using a random, repeatable sampling method that preserves correlation structures and can transparently accommodate models from different suppliers with varying structures and outputs.
- **Technical Architecture:** The document proposes a unified *relational data model* for managing meta-data, analysis outputs, and groupings, which enables consistent processing, grouping configurations, detailed loss table generation, and supports complex grouping scenarios such as multiple risk portfolios and group-of-group aggregations.
- **Configuration & Flexibility:** The methodology allows configuration of grouping parameters (number of periods, random seeds, summary level alignment) and supports mean-only or secondary uncertainty-inclusive aggregation, addressing both mean and probabilistic views of group risk.
- **Outputs:** Resulting groupings enable generation of consolidated loss tables and risk metrics, such as Average Annual Loss (AAL) and EP curves, suitable for regulatory or internal reporting needs.

Benefits

This approach empowers risk practitioners to generate robust, reproducible, and transparent aggregated risk metrics even when combining disparate models, portfolios, or exposure summary levels. The framework emphasizes the importance of preserving uncertainty, handling different data formats, and ensuring statistical rigor in the grouping process. By outlining data structures, calculation logic, and configuration options, the document serves as a practical guide for implementing group risk aggregation workflows in the OASIS environment.

Related documents

- **ORD_requirements_for_results_processing_v0.2.pdf** – 2023 study on uses cases for results processing in ORD
- **OasisLMF_combining_results_20250806.pdf** – High level presentation of combing results methodology for Oasis TSG (Aug 2025)
- **Worked_example_combining_ORD_v1.xslx** – illustrative worked example of combining small data subsets of 4 analyses using this methodology.

Document versions

Version	Date	Author	Description
1	28th Aug 2025	Johanna Carter	First draft to share with Oasis TSG
2	23rd Sep 2025	Johanna Carter	Added acknowledgement

Acknowledgement

We gratefully acknowledge Guy Carpenter for their collaborative support and contributions to the technical work reflected in this document.

Contents

1. Introduction

This document presents a comprehensive technical proposal for combining catastrophe loss model results in the Open Results Data (ORD) format, focusing on methods for constructing consolidated views of risk from multiple analyses. The methodology outlined supports practitioners in catastrophe risk management and modelling, offering a robust, flexible framework for aggregating detailed model outputs—such as event loss tables (ELTs) and period loss tables (PLTs)—to generate enterprise-wide risk metrics including Average Annual Loss (AAL) and Exceedance Probability (EP) curves.

1.1 Purpose and Context

The approach detailed herein enables organizations to systematically roll up results from diverse models and analyses, preserving important correlations and uncertainties at the event and period level. This is essential for accurate group risk assessment, regulatory reporting, and strategic decision-making in the insurance and reinsurance sectors.

1.2 Scope of Methodology

By addressing practical challenges like inconsistent summary levels, varying period counts, and differing model structures, the proposed random sampling-based technique ensures repeatable, unbiased aggregation of loss data across multiple outputs. The document also and provides a high-level blueprint for the underlying relational data model required for implementation within ORD.

It is expected that the methodology and data model design described in this document will need to be further developed or refined through implementation.

2. Terminology

We refer to 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.

3. Business and Technical Context

In today's insurance and reinsurance landscape, organizations routinely utilize a diverse set of catastrophe models, platforms, and analytical approaches to inform critical risk decisions, regulatory submissions, and portfolio management.

This diversity creates the need for a robust, consistent methodology to combine disparate model outputs in a way that preserves key risk characteristics—including dependence, tail risk, and uncertainty—essential for enterprise and group-level reporting.

Effective aggregation in this context requires addressing not only business requirements for transparency, auditability, and comparability, but also overcoming significant technical challenges such as reconciling differing event definitions, period counts, model formats, and exposure summary levels across multiple suppliers and systems.

3.1 Business drivers

The high level business drivers that this methodology is designed to address are;

- Portfolio aggregation across models and geographies
- Regulatory submission of consolidated EP curves
- Enterprise risk management, including roll-up across subsidiaries and lines of business

A more detailed set of use cases with more emphasis on operational practices and challenges is outlined in *ORD_requirements_for_results_processing_v0.2.pdf*

3.2 Technical challenges

The challenges that this methodology is designed to address are;

- Different suppliers, model versions, event set definitions
- Inconsistent period counts and summary levels
- Different number of samples per analysis
- Mixed ELT formats (Moment, Quantile, Sample)

- Lack of support for multiple/grouped analyses in Open Data Standards

4. Overview of Proposed Methodology

4.1 Step-by-Step Process

The proposed methodology will combine results at a detailed level, i.e. combine event losses (MELT,SELT,QELT) or period loss tables (MPLT, SPLT, QPLT), 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).

The proposed method is a generalized random sampling method that can work consistently for the different types of detailed loss reports.

High-Level Process

- **Select Analyses and Output Levels:** Identify analyses and summary levels eligible for grouping based on compatible metadata fields.
- **Identify GroupEventSets**: Group analyses by shared event/occurrence set definitions.
- **Assign Events to Group Periods:** Use random, repeatable sampling (without replacement) to assign analysis periods to grouped periods, breaking model-driven event patterns.
- **Sample Losses per Period:** For each event, sample a loss from the appropriate severity distribution (mean, quantile, or sample), with the method chosen in the configuration.
- **Compute Grouped Reports and Metrics:** Produce detailed tables as well as high-level metrics, including Average Annual Loss (AAL) and Exceedance Probability (EP) curves.

4.2 Rationale and Challenges

- **'Bottom-up' combining from detailed event losses:** 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.

- **Use of random sampling:** Random sampling is a flexible approach which can deal with the following specific challenges of combining ORD results:
    - Results from different model suppliers / platforms
    - Different number of periods across models
    - Different number of samples across analyses
    - Mixed ELT formats (Moment/Quantile/Sample)

  The single loss per period approach is a well-established approach for rolling up results from multiple analyses in aggregation tools used in the industry.
- **Non-parametric event frequencies:** Using an occurrence file to express event frequency rather than event rates ensures the model provider's views on event clustering and temporal distribution are fully captured in the grouped result.
- **Challenge of approach:** The main challenge of a random sampling approach is sampling error and not achieving convergence in the grouped result. However, measuring the sampling error of single loss per period is straightforward compared with more complex simulation frameworks with multiple or nested stochastic variables.

5. Scope and Assumptions

- Types of combining: Roll-up
    - Out of scope: blending of multiple views of the same risk(s).
- Number of analyses to group
    - Normal case 3-5 analyses, stress case 20 analyses
- Grouping features
    - Multiple summary levels e.g. All risks and by LOB
    - Perspectives: GU/IL/RI. 'Fill' perspective if missing
    - Support grouping of groups. In practice, this means preserving enough information about the provenance of each detailed loss in a grouped result to be able to append the losses of new analyses rather than restarting the grouping from the original analyses.
    - Out of scope: aggregation across summary levels. Each summary level required at a grouped level must be present at the individual analysis
- Event definitions
    - Uniquely identity events within a grouped analysis

- Source analysis data
    - Format: ORD. This methodology does not include combining results in different formats. If the source analysis data is in a different format, it must first be converted into ORD format to be grouped with other ORD format analyses.
    - Event frequency: The methodology assumes all analyses have either PLTs, or ELTs and occurrence files describing event frequency.
- Outputs
    - Grouped event loss tables and year loss tables
    - EPT (Mean/with Secondary uncertainty) and ALT

6. Data and Meta Data Architecture

In ORD v2, a data package is a directory containing different sorts of files. Results data tables are stored in csv or parquet format, and analysis meta data is stored in multiple json files. One ORD v2 data package is designed to store the results data for a single analysis.

6.1 ORD Meta Data Structure

An analysis in ORD has the following types of meta data described in a nested json structure;

Analysis: A unique identifier of the analysis run and associated information including analysis description, status, currency, portfolio name, user name, date and time.

- **Settings**: Information about what the general settings of the run were, such as which model was run, which reports and perspectives were requested, and the number of samples.
    - **Model Settings**: this contains the model file identifiers that were used, such as event set and event occurrence set, with descriptions.
        - **Correlation Settings**: The correlation factors that were applied in the analysis, if any.
    - **Output Sets**: which 'summaries' or exposure summary levels were requested, for each financial perspective.
        - **Summary Info:** for each Output Set, a list of SummaryIds representing a particular group of exposures, say 'Residential', 'Commercial' and 'Industrial' for the ground up LOB Output Set with their descriptions and TIVs.
- **Exposure Files**: The filenames associated with the portfolio name.

- **Exposure Summary**: A high level summary of the total value and number of locations that were modelled and not modelled.
    - **Exposure Coverage:** A detailed breakdown of TIV and number of locations by coverage and peril that was modelled.
- **Versions:** containing platform and technical environment version information**.**

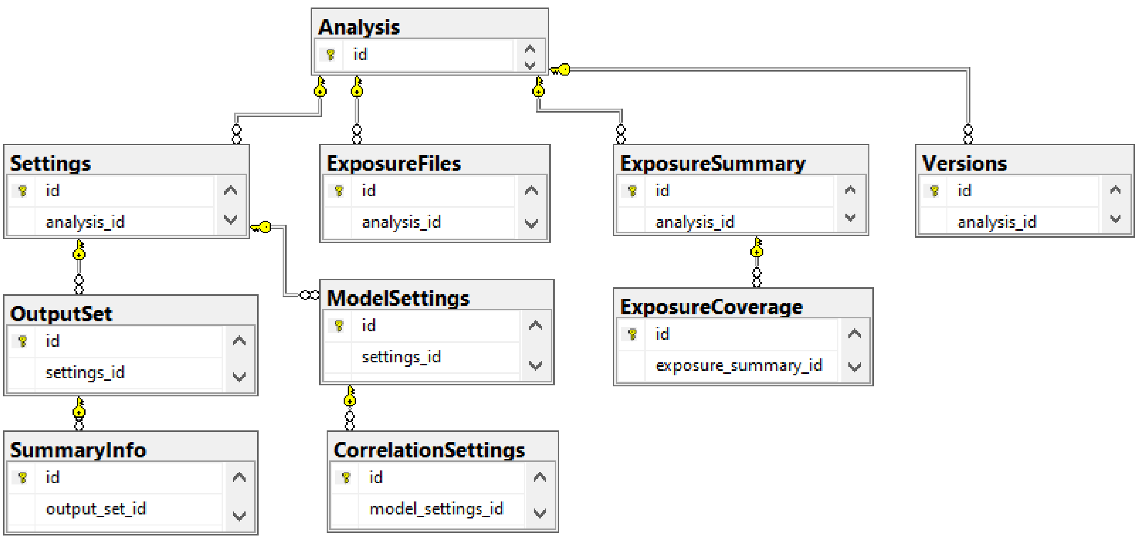
6.2 Proposed Relational Model

In order to identify the grouping possibilities across any set of separate analyses, it is necessary to bring the meta data together in an organized format.

To achieve this, the above structure can be represented in a relational data model. Each table object in the relational model can hold the meta data of multiple analyses. To maintain symmetry between json and the relational equivalent, each nested json object can be represented as a separate data table with a foreign key relationship to a table containing the parent level meta data.

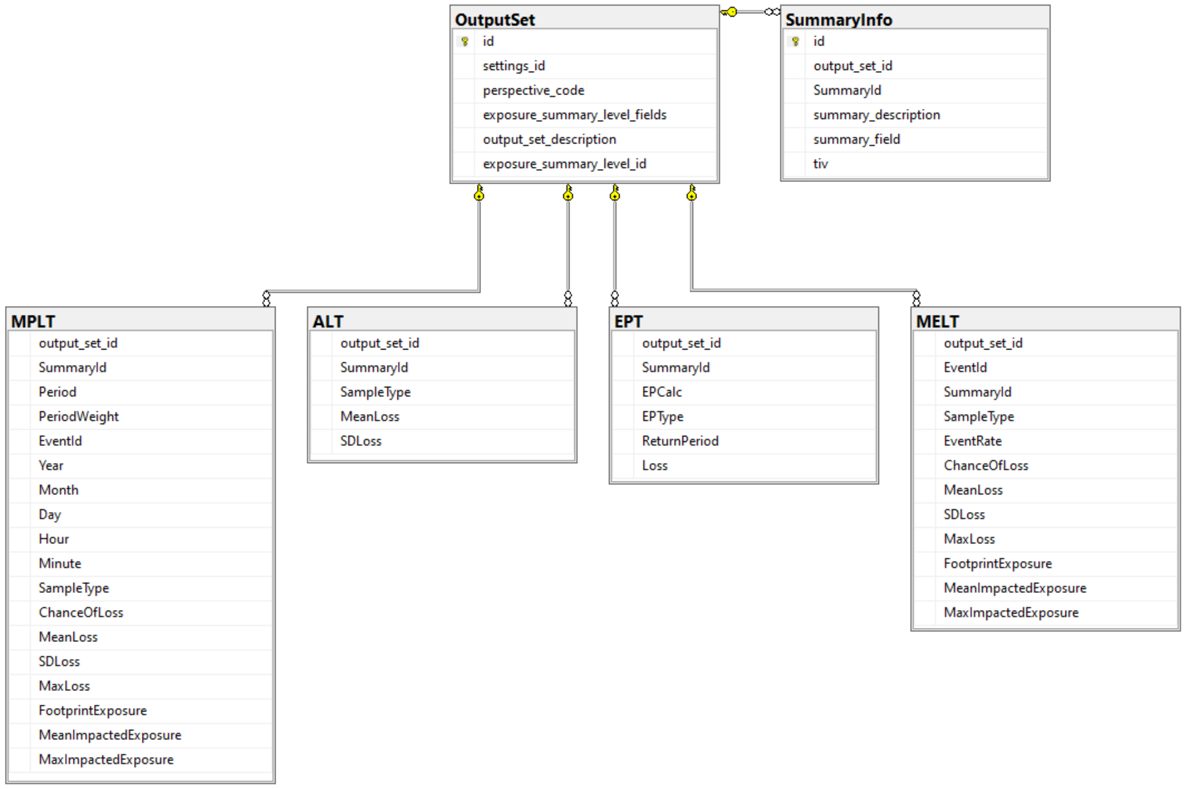
Figure 1 shows the high-level relational model, including suggested table names with their primary and foreign keys.

Figure 1: High level relational model for ORD meta data



OutputSet, from json 'summaries', lists each individual summary level result within an analysis and assigns a globally unique id. With each OutputSet uniquely defined across multiple analyses, the ORD results themselves can be stored in a single table per report type (MELT, ALT and EPT as examples shown in Figure2) with a foreign key reference to the **OutputSet** table.

Figure 2: Relationship of ORD results tables (four shown) to meta data



A mixture of case (PascalCase and snake_case) is proposed to easily distinguish fields that originate from ORD results files (such as SummaryId) from relational fields (such as the foreign key output_set_id that links result data to a parent record in parent table OutputSet(id)).

Relationally, the grouped analysis data can be structured as follows;

Group

Lists each group

id	Description
1	Illustrative group across multiple analyses and models
2	Illustrative group of groups and analyses

GroupAnalysis

Links groups to analyses

group_id	analysis_id
1	1
1	2
1	3
1	4

group_id	analysis_id
2	5

GroupGroup

To support groups of groups (group 2 includes all analyses in group 1 plus analysis 5)

parent_group_id	child_group_id
2	1

GroupSet

Lists the output summary levels available for each group

id	group_id	perspective_code	exposure_summary_level_fields	group_output_description
1	1	gr	[]	Gross - all risks
2	1	gr	["LOB"]	Gross – by LOB
3	1	ri	[]	Net - all risks

GroupOutputSet

Links the output summary levels selected for grouping to the required detailed data 'OutputSet's from the relational model in Figure 1

group_set_id	output_set_id
1	1
1	6
1	9
1	11
2	2
2	7
2	10
2	12
3	4
3	6
3	9
3	11

Figure 3: High level relationship of groups to analyses (not all tables shown)

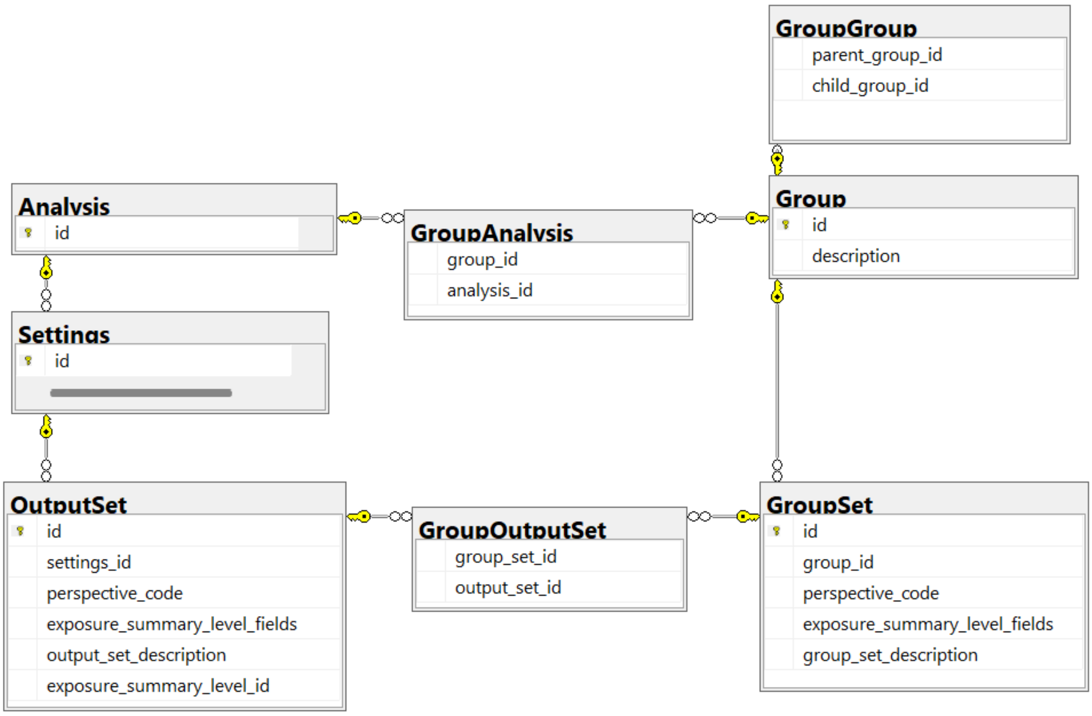
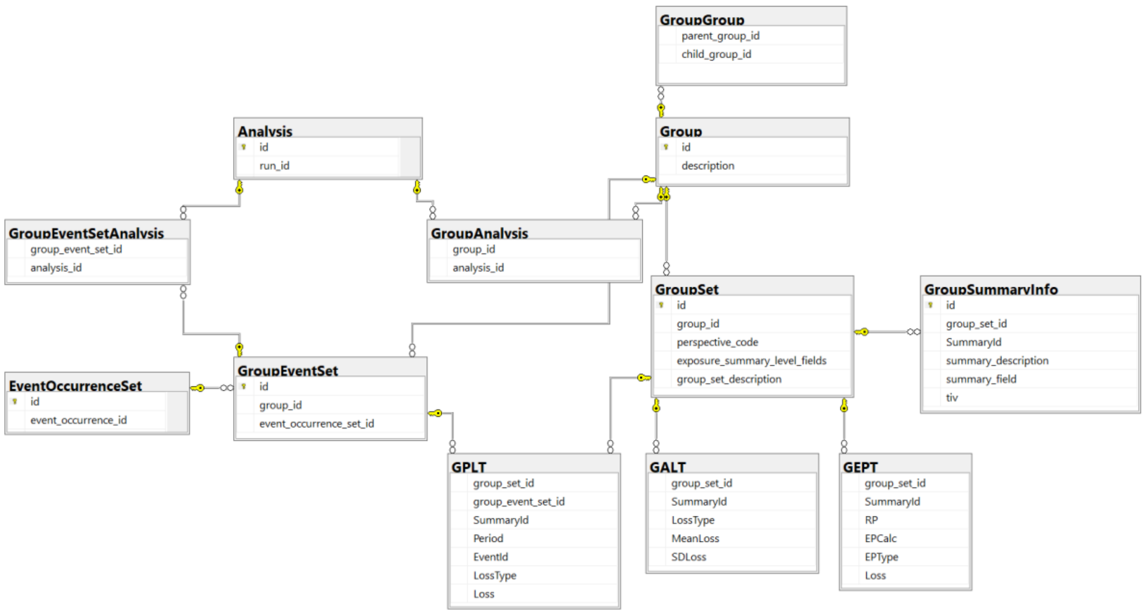


Figure 4: All group tables (Analysis child tables not shown)



7. Detailed Grouping Methodology

7.1 Analysis and Output Selection

The normal case will be that there will be multiple summary level output sets (for example, gross loss by portfolio, gross loss by line of business, e.tc.) available within each single analysis results package.

Summary level output sets within individual analyses that are compatible for grouping can be recognized by their meta data; principally the perspective and OED fields representing the exposure summary level.

This step involves bringing the meta data together in order to decide what can be grouped, and then based on the selections, creating meta-data for the grouped analysis.

Configuration

A configuration option to allow for 'filling' of financial perspectives to enable more grouping of output sets.

group_fill_perspectives: true/false

When set to true;

- a missing 'net' output set could be replaced with the 'gross' output set to produce a grouped 'net' output set, and
- a missing 'gross' output set could be filled with a 'ground up' output set to produce a grouped 'gross' output set.

Some example meta-data for four analyses arranged into relational tables follows;

Analysis table – example selections

id	run_id	description
1	29485	Analysis 1
2	29486	Analysis 2
3	29488	Analysis 3
4	29491	Analysis 4

OutputSet table for analysis selections

id	settings_ id	perspective_ code	output_set_ description	exposure_ summary_level_fields	exposure_ summary_level_id
1	1	gr	Gross, all risks	[]	1
2	1	gr	Gross, by LOB	["LOB"]	2

id	settings_ id	perspective_ code	output_set_ description	exposure_ summary_level_fields	exposure_ summary_level_id
3	1	gr	Gross, by LocName	["LocName"]	3
4	1	ri	Net, all risks	[]	1
5	1	ri	Net, by LocName	["LocName"]	2
6	2	gr	Gross, all risks	[]	1
7	2	gr	Gross, by LOB	["LOB"]	2
8	2	gr	Gross, by GeogName1	["GeogName1"]	3
9	3	gr	Gross, all risks	[]	1
10	3	gr	Gross, by LOB	["LOB"]	2
11	4	gr	Gross, all risks	[]	1
12	4	gr	Gross, by LOB	["LOB"]	2

Any consistent set of perspective_code and exposure_summary_level_fields across the selected analyses can be grouped. The value of exposure_summary_level_id (which is an internal identifier of each output set originating from the json file) is irrelevant here as the output sets may have inconsistent ids but are nonetheless compatible.

OutputSets that can be grouped for this example are;

Gross, all risks

id	settings_id	perspective_code	exposure_summary_level_fields
1	1	gr	[]
6	2	gr	[]
9	3	gr	[]
11	4	gr	[]

Gross, by LOB

id	settings_id	perspective_code	exposure_summary_level_fields
2	1	gr	["LOB"]
7	2	gr	["LOB"]
10	3	gr	["LOB"]
12	4	gr	["LOB"]

With the config setting ***group_fill_perspectives*** = ***true***, "Net – all risks" grouped result could also be created by 'filling' net results for analysis 2, 3 and 4 (not present) with gross losses, and grouping these with net results for analysis 1;

Net, all risks

id	settings_id	perspective_code	exposure_summary_level_fields
4	1	ri	[]
6	2	gr	[]
9	3	gr	[]
11	4	gr	[]

Relationally, the data for analysis groups can be structured as follows;

Group

Lists each group

id	Description
1	Illustrative group across multiple analyses and models
2	Illustrative group of groups and analyses

GroupAnalysis

Links groups to analyses

group_id	analysis_id
1	1
1	2
1	3
1	4
2	5

GroupGroup

To support groups of groups (group 2 includes all analyses in group 1 plus analysis 5)

parent_group_id	child_group_id
2	1

GroupSet

Lists the output summary levels available for each group

id	group_id	perspective_code	exposure_summary_level_fields	group_output_description
1	1	gr	[]	Gross - all risks
2	1	gr	["LOB"]	Gross – by LOB
3	1	ri	[]	Net - all risks

GroupOutputSet

Links the output summary levels selected for grouping to the required detailed data 'OutputSet's from the relational model in Figure 1

group_set_id	output_set_id
1	1
1	6
1	9
1	11
2	2
2	7
2	10
2	12
3	4
3	6
3	9
3	11

Alignment of Summary info

Within each group set, the SummaryId, which identifies each group of exposures, will generally not be consistent across output sets.

There needs to be a preparation step to recognize all of the distinct SummaryId description fields in the output sets and recreate the SummaryId index for the group set to ensure there is consistency in the meaning of these values before grouping.

To illustrate, the following two tables show SummaryId definitions for two compatible output sets, and the reindexed grouped SummaryId definitions for the group set.

Summary Info for compatible output sets

id	output_set_id	SummaryId	summary_description	summary_field	tiv
2	2	1	Res	LOB	850,997,000
3	2	2	Com	LOB	686,009,000
21	7	1	Com	LOB	236,851,000
22	7	2	Ind	LOB	550.224,000

Summary Info for group set

id	group_set_id	SummaryId	summary_description	summary_field	tiv
1	2	1	Res	LOB	850,997,000
2	2	2	Com	LOB	922,860,000
3	2	3	Ind	LOB	550.224,000

7.2 Identification of GroupEventSets

The second step involves assigning a GroupEventSet id to each analysis in the group by identifying common event definitions across the analyses in order to combine losses correctly in the grouped result.

A GroupEventSet links a subset of analyses which share common event ids and event occurrences. A grouped result may contain multiple GroupEventSets, each of which links to one or more analyses.

This step is facilitated by bringing together the meta-data of the selected analyses in the group and defining the condition that uniquely identifies event sets.

The attributes contained in ORD meta data regarding event definitions are proposed to be;

- **event_set_id**– a locally unique identifier of the event set file used in the analysis
- **event_set_description** – a description of the event set used in the analysis
- **event_occurrence_id**– a locally unique identifier of the occurrence set file used in the analysis
- **event_occurrence_description** – a description of the occurrence set used in the analysis.
- **event_occurrence_max_periods** – the overall timespan covered by the occurrence set used in the analysis, expressed as a number of periods (normally years).
- **model_supplier_id** – name of the model supplier

- **model_name_id** – name of the model
- **model_description** – description of the model
- **model_version** – version of the model

In a grouped analysis, unique event sets could be identified by one or more of the above meta data fields.

Configuration

Grouping configuration for recognizing common event definitions across analyses.

group_event_set_fields: [field list]

This config option is to provide flexibility in how event sets are uniquely defined, recognizing that not all meta data fields will be populated, or uniquely identify an event set on their own.

Example

group_event_set_fields: [model_supplier_id, event_set_id, occurrence_set_id]

Each EventId has the same definition across analyses to be grouped where model_supplier_id, event_set_id and occurrence_set_id share common values.

This means that losses would be combined into a single occurrence for the purposes of OEP calculation if they share a common EventId within a GroupEventSet.

Analysis meta data

id	model_ supplier_id	model_ name_id	model_ version	event_ set_id	event_ occurrence_id	event_occurrence_ max_periods
1	SupplierA	ATFL	2025_01	EUFL02	EUFL02_base	10,000
2	SupplierA	DEFL	2024_03	EUFL02	EUFL02_base	10,000
3	SupplierA	BEFL	2021_06	EUFL01	EUFL01	10,000
4	SupplierB	ITFL	2021_08	EUFL01	EUFL01	100,000

The relational tables for GroupEventSets are proposed as follows;

EventOccurrenceSet

Populated with the group event set fields specified in configuration (only) to create an index of distinct event occurrence sets.

id	model_ supplier_id	model_ name_id	model_ version	event_ set_id	event_ occurrence_id	event_occurrence_ max_periods
1	SupplierA	NULL	NULL	EUFL02	EUFL02_base	NULL
2	SupplierA	NULL	NULL	EUFL01	EUFL01	NULL
3	SupplierB	NULL	NULL	EUFL01	EUFL01	NULL

Note: keeping this table independent of any particular group of analyses within a database helps to support grouping of groups, because the ids can be reused if new analyses with matching group event set keys are to be added to existing groups.

GroupEventSet

Lists the GroupEventSets for each group and links to the EventOccurrenceSet

id	group_id	event_occurrence_set_id
1	1	1
2	1	2
3	1	3

With these GroupEventSet field settings, Analysis 1 and 2 belong to the same GroupEventSet. They would have maximum occurrence losses summed by EventId because they share the same GroupEventSet id and therefore, the same event definitions.

Analysis 3 and 4 each have their own GroupEventSetId assigned because they are from entirely different models (model_supplier_id differs).

This information is represented in the GroupEventSetAnalysis table.

GroupEventSetAnalysis

analysis_id	group_event_set_id
1	1
2	1
3	2
4	3

7.3 Period Sampling

Period sampling deals with assigning events from multiple analyses to a common group period. The sampling method of periods is proposed to be random, repeatable, without replacement:

- **Random:** This means that we will use random numbers to select a period number from 1 to the maximum period number for each GroupEventSet, and we will use a different sequence of random numbers for each GroupEventSet, so that there is no correlation between the choices of periods across analyses. This is to remove patterns in model-provided occurrence files which could introduce bias in grouped results (such as small events at the beginning of the timeline and large events at the end)
- **Repeatable:** This means that we will use the same sequence set of random numbers to select a period number from an analysis if we were to repeat the grouping (with a configurable seed to enable different random numbers to be used), leading to exactly the same numerical grouped results.
- **Without replacement:** This means that each time a period for each event set group is sampled, it cannot be selected again until all other periods have been sampled.
    - This ensures full representation of the period losses as long as the number of periods set for the grouped result is at least equal to the number of periods in each event set grouped analysis.
    - Essentially we are creating a random order of the given periods for each event set to align with the GroupPeriod 1 through to G. Each cycle of the full set of periods will use a different random order.

Example

GroupEventSet id 2 has 25 periods in total, of which periods 2,10,12,15, and 21 have loss-causing events, and the other periods have no events.

GroupEventSet id 2 Event Occurrences

Period	EventId	Year	Month	Day
2	17	2	1	1
10	2	10	1	1
12	1	12	1	1
15	3	15	1	1
21	42	21	1	1

The group number of periods is set to 100 periods. All of the periods with losses for GroupEventSet id 2 are assigned randomly to the group periods in full cycles before they repeat. Each period will be sampled exactly 4 times in the group period timeline.

Period Sampling of GroupEventSet id 2 across 100 GroupPeriods

Cycle	GroupPeriod	group_event_set_id	Period
	7	2	10
	8	2	12
1	9	2	15
	13	2	2
	16	2	21
	26	2	15
	40	2	12
2	46	2	10
	47	2	2
	49	2	21
	52	2	21
	56	2	2
3	59	2	12
	67	2	10
	72	2	15
	79	2	2
	86	2	10
4	88	2	12
	91	2	15
	96	2	21

Not all GroupPeriods will have a loss-causing event occurrence assigned to them.

Configuration

A configuration option will set the number of periods for grouping;

group_number_of_periods: {integer}

Default value: 1,000,000

The default value can be changed to be equal to the event set with the largest number of periods, or more as required.

Another configuration option seeds the random numbers;

group_period_seed:{integer}

Default value: 1

A change in the seed will lead to a different set of random numbers to be generated for period sampling for each GroupEventSet.

The output of this step is an intermediate calculation table of GroupPeriods with assigned periods for each GroupEventSet.

GroupPeriod table

group_event_set_id	GroupPeriod	Period
3	1	1
3	5	50
1	7	4
2	7	10
2	8	12
2	9	15
1	10	2
3	10	62
2	13	2

This table provides the link between the Group Periods and all of the event occurrences for the analyses in the group, by joining to the detailed period loss tables (MPLT, SPLT etc) or the models event occurrence file, where EventIds are assigned to Periods.

7.4 Loss Sampling

The period sampling determines which events from all GroupEventSets are assigned to GroupPeriod g. Loss sampling deals with generating grouped losses for each period g.

Two types of sampling can be performed:

- Mean only (no random numbers used)
- Full uncertainty sampling from each event's severity distribution (random numbers used)

Mean only is compatible with MELTs only. I.e. the MELT has to be present for all analyses in the group. With this option, we take the mean event loss as the GroupPeriod event loss every time it occurs.

Full uncertainty sampling can be used with MELT, QELT, SELT or a mixture.

The sampling method of losses for each period g for type 2) full uncertainty is proposed to be random, repeatable and independent.

- **Random:** This means that we will use random numbers to sample a loss for each GroupPeriod and event occurrence.
- **Repeatable:** This means that we will use the same random numbers to generate a loss for the same event and group period if we were to repeat the grouping process so that the sampled losses are repeatable (with a

configurable seed to force different random numbers to be used if required).

- **Independent:** If the event from which we sample a loss occurs in more than one GroupPeriod (because, for example, it occurs in multiple periods from the source analysis) a different random number should be used to sample a different and independent loss sample per occurrence.
    - Independent samples across occurrences of the same event ensures that we sample a range of losses from the event loss distribution across GroupPeriods, rather than using the same severity loss sample every time.
    - Independent samples across GroupPeriods also simplifies the statistical approach of measuring sampling error in annual loss metrics and ensures better convergence.

In the following calculation logic, g represents the group period number between 1 and G, the set value of group_number_of_periods.

Calculation Logic

1) Mean only

   Sum the event mean losses, grouping by *group_event_set_id, GroupPeriod, SummaryId, SampleType, EventId*

2) Full uncertainty

This involves generating random numbers which will be used to inverse-transform sample a loss quantile from the event loss cumulative distribution function. Different methods can be used depending on which format of ELT is available.

Generate the random numbers by creating a Group Period Quantile Table, an intermediate calculation table needed for loss sampling with secondary uncertainty.

Group Period Quantile Table

GroupPeriod	group_event_set_id	output_set_id	EventId	Quantile
1	3	4	1	0.817935
5	3	4	74	0.948615
7	1	3	2	0.016913
8	1	3	1	0.147552
9	1	3	3	0.465603
10	3	4	92	0.491842

This joins the GroupPeriods to all event occurrences based on the GroupPeriod table and generates a single random number to use as the severity quantile. All random numbers are independent (but see Correlation section for options).

Depending on which detailed loss table is available to draw a loss from, the process is as follows:

a) MELT

- Parameterise an event loss distribution for SampleType=2 using Mean, SD, MaximumImpactedExposure. (set parametric function in config) per EventId.
- Compute loss quantile from inverse transform of parametric distribution function per EventId using quantile probability from GPQT.

b) QELT

- Compute loss quantile from discrete inverse transform sampling of QELT distribution function per EventId using linear interpolation of quantile probability from the GPQT.

c) SELT

- Rank order the loss samples from low to high.
- Use inverse transform sampling of the uniform distribution using the quantile probability from GPQT to select a sample index.
- Select the loss associated with that sample index as the loss quantile.

Configuration

The options for sampling are:

- **group_mean**:{true/false}
  - Output the grouped mean results.
- **group_mean_type**:{int}
  - Set type 1 for analytical mean grouped result, or type 2 for sample mean
- **group_secondary_uncertainty**:{true/false}
  - Output the grouped results with secondary uncertainty (type 2).
- **group_parametric_distribution:**{beta/gamma}
  - The parameters will be estimated from the MeanLoss, SDLoss and MaxImpactedExposure value from the MELT.
- **group_format_priority:**{M,Q,S}
  - This controls which format of ELTs is used for loss severity sampling when there are multiple formats available per output set.

    M – Mean

Q – Quantile

S - Sample

{M} means that only the MELTs will be used for severity loss sampling.

{Q,M} means that QELT will be used as a priority format for severity loss sampling, falling back on the MELTs where QELT is missing. etc.

The output of the loss sampling process is the Group Period Loss Table.

Group Period Loss Table

group_event_set_id	group_set_id	GroupPeriod	SummaryId	EventId	LossType	Loss
3	1	1	1	1	2	3279925.5
3	1	5	1	74	2	3473524
1	1	7	1	6	2	78078.90137
2	1	7	1	2	2	23172.25977
2	1	8	1	1	2	19766.51953
2	1	9	1	3	2	19766.51953
1	1	10	1	1	2	201071.5039
1	1	10	1	4	2	67776.29883
3	1	10	1	92	2	3351865.75
2	1	13	1	17	2	210010.5938

The extra column 'LossType' allows for multiple loss type selections, with enumeration which is consistent with the equivalent 'SampleType' in individual loss tables.

LossType

1=Grouped Analytical Mean,

2=Grouped Sample Mean With Secondary Uncertainty,

3=Grouped Sample Mean

In this detailed group output format, there is a single loss per period per event (per loss type). This differs from the individual analysis report tables MPLT, QPLT and SPLT which all contain distributions of event losses per event and per period in different forms.

7.5 Generation of Grouped Metrics (AAL, EP Curves)

The Group Period Loss Tables is the starting point for calculation of the grouped metrics.

Calculation logic - Average Annual Loss

Sum the Loss, grouping by GroupPeriod, SummaryId, LossType to produce a Period Loss per SummaryId and LossType

- **Mean** calculation – grouping by SummaryId and LossType, sum the Period Loss and divide by G to produce the AAL per SummaryId and LossType
- **Standard deviation** calculation – grouping by SummaryId and LossType, square root of the sum of squared differences between each Period Loss and the AAL, divided by G-1, for each SummaryId and LossType,

Calculation logic – EP curves

Sum the Loss, grouping by group_event_set_id, EventId, GroupPeriod, SummaryId, LossType to produce an Event Occurrence Loss per group_event_set_id, EventId, GroupPeriod SummaryId and LossType.

- **OEP** calculation:
  - Group by GroupPeriod, SummaryId, LossType and take the maximum loss to produce a maximum occurrence loss per GroupPeriod
  - Rank the maximum occurrence loss per GroupPeriod, SummaryId and LossType from highest to lowest.
  - Rank/G = Occurrence Exceedance Probability
- **AEP** calculation:
  - Group by *GroupPeriod, SummaryId, LossType* and take the sum of the loss to produce the aggregate loss per GroupPeriod
  - Rank the aggregate loss per GroupPeriod, SummaryId and LossType from highest to lowest.
  - Rank/G = Aggregate Exceedance Probability

Configuration

- **group_alt**:{true/false}
  - Output the grouped ALT report.
- **group_ept**:{true/false}
  - Output the grouped EPT report.
- **group_plt:**{true/false}
  - Output the grouped PLT report. Although this is necessary for the calculation of GALT and GEPT, it will be very large in some cases and does not have to be persisted.

8. Output Formats and Reporting

8.1 Group Period Loss Table

Below is the proposed detailed output table for groups in the relational model, which provides full provenance of each loss back to each analysis and group (via the group_set_id foreign key) and the event definitions (via the group_event_set_id foreign key).

Relational GPLT

group_event_set_id	group_set_id	GroupPeriod	SummaryId	EventId	LossType	Loss
3	1	1	1	1	2	3279925.5
3	1	5	1	74	2	3473524
1	1	7	1	6	2	78078.90137
2	1	7	1	2	2	23172.25977
2	1	8	1	1	2	19766.51953
2	1	9	1	3	2	19766.51953
1	1	10	1	1	2	201071.5039
1	1	10	1	4	2	67776.29883
3	1	10	1	92	2	3351865.75
2	1	13	1	17	2	210010.5938

A file-based GPLT, following convention for single analyses, would not need a group_set_id as there is one set of result files per group_set with the id as part of the filename. However it would need the GroupEventSetId field (presented here in PascalCase format) in order to be able to identify which EventOccurrenceSet each EventId belongs to.

File-based GPLT

GroupEventSetId	GroupPeriod	SummaryId	EventId	LossType	Loss
3	1	1	1	2	3279925.5
3	5	1	74	2	3473524
1	7	1	6	2	78078.90137
2	7	1	2	2	23172.25977
2	8	1	1	2	19766.51953
2	9	1	3	2	19766.51953
1	10	1	1	2	201071.5039
1	10	1	4	2	67776.29883
3	10	1	92	2	3351865.75
2	13	1	17	2	210010.5938

9.2 Average Loss & EP Curves

The following tables are proposed as high-level statistical summaries of the grouped results.

Group Average Loss table (GALT)

This is essentially the same format as the ALT, other than re-defining the type of grouped loss. The relational version has a foreign key to the GroupSet table.

Relational GALT

group_output_id	SummaryId	LossType	Mean	Stdev
1	1	1	989,340	120,217

File-based GALT

SummaryId	LossType	Mean	Stdev
1	1	989,340	120,217

LossType

1=Grouped Analytical Mean,

2=Grouped Sample Mean With Secondary Uncertainty,

3=Grouped Sample Mean

Grouped Exceedance Probability Table (GEPT)

This is the same format as the EPT, other than redefining values for EPType to identify grouped loss types.

Relational GEPT

group_set_id	RP	EPCalc	EPType	Loss
	10	1	2	2,668,619
	5	1	2	1,334,718
	2	1	2	680,996
	10	3	2	3,993,822
	5	3	2	1,498,348
	2	3	2	851,245

File-based GEPT

RP	EPCalc	EPType	Loss
10	1	2	2668619
5	1	2	1334718
2	1	2	680996

RP	EPCalc	EPType	Loss
10	3	2	3993822
5	3	2	1498348
2	3	2	851245

EPCalc

(no difference to single analysis EPType)

1 - OEP

2 – OEP TVAR

3 - AEP

4 – AEP TVAR

EPType

(redefined for groups)

1=Grouped Analytical Mean

2=Grouped Sample Mean With Secondary Uncertainty,

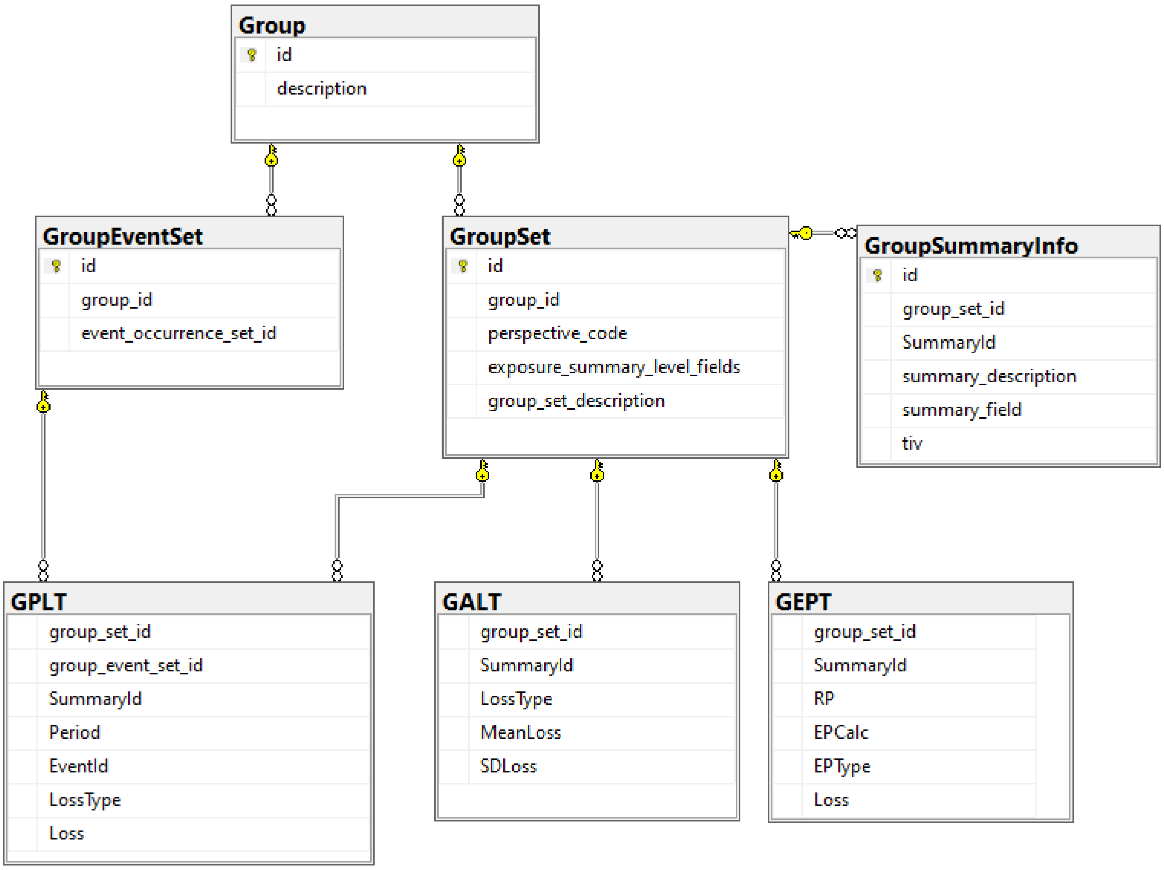
3=Grouped Sample Mean

9.3 Grouped results in the relational model

Since each grouped results table, e.g. GEPT, has a relationship to a different parent table (GroupSet) than individual analysis results tables (OutputSet), they would not be stored in the same table in the relational model as the individual results tables, even though they have the same fields.

The proposed relational data model for grouped results to meta data is as follows (not all Group tables shown);

Figure 5: Relationship of ORD group results tables to meta data



9. Correlation

The random numbers can be generated differently to capture correlation within a GroupEventSet. This is where there are multiple losses coming from the same EventId and it would be a better approximation of the combined exposure single-run result to assume some correlation.

Fully correlated random numbers can be generated within GroupEventSets by dropping the output_set_id from the seeding columns in the GPQT:

GPQT

GroupPeriod	group_event_set_id	output_set_id	EventId	quantile
1	3	4	1	0.817935
5	3	4	74	0.948615
7	1	3	2	0.016913
8	1	3	1	0.147552
9	1	3	3	0.465603
10	3	4	92	0.491842

Partial correlation could be implemented by using a Gaussian copula to generate correlated random numbers to apply to loss samples arising from different analyses for the same EventId (same GroupEventSet and EventId, correlated across different output_set_id).

The correlation factor can be driven by the Correlation Settings in model settings, which would tend to be the same factor within any given model, and therefore within each GroupEventSet;

Correlation_settings:{damage_correlation_factor: 0.125}

10. Conclusion

The methodology set out in this document delivers a rigorous, flexible solution for the group-level roll-up of catastrophe model results in the ORD format.

By systematically combining detailed losses across multiple analyses—using random, repeatable sampling and a unified relational data model—the framework preserves key elements of risk, including correlation structures, exposure summaries, and uncertainty across model types and suppliers. This enables the generation of robust, transparent enterprise metrics such as AAL and EP curves, supporting regulatory submissions and internal group risk management with statistical repeatability and full data lineage.

The configuration options and relational model provide the extensibility required for evolving operational needs, while the approach's grounding in industry best practices ensures both analytical soundness and practical application.

As risk modelling continues to advance in complexity, the methodology equips practitioners with a reliable foundation for scalable, auditable, and future-ready risk aggregation within the Oasis platform.