Operational risk scenario modelling

Technical document

May 2017

Contents

1.0 Introduction and Purpose

Operational risk is intrinsic in any business. It can give rise to loss through the actions of staff as well as manifesting itself through weaknesses in systems and processes and through external events such as cybercrime and physical asset damage. Operational risk is also inherently difficult to measure. Historic internal data may be sparse or aged and external data can have limited applicability. To provide a forward looking view that seeks to alleviate these data issues, organisations have adopted a process called 'Scenario Analysis'.

As part of its operational risk identification process, a firm will identify its most significant operational risks. Scenario analysis of these risks is then performed by business and risk experts. This seeks to explore the potential evolution of key risk scenarios over a twelve-month time horizon given the institution's current risk profile and control framework. Experts will consider the expected frequency of associated events and provide an estimate of the possible event severities over this time frame. The event severities are typically framed within the working lifetime of the contributors.

Statistical techniques are then used to extrapolate the scenario outcomes to quantify the low frequency, high impact nature of operational risk to confidence levels associated with holding sufficient capital. This end to end process enables businesses to consider appropriate levels of mitigation such as, investment in controls, insurance, credit facilities and capital buffers.

elanev have distilled these statistical techniques to provide a flexible and rapid modelling framework. The frequency and severity information and related correlations from each scenario analysed are aggregated in real time to estimate an organisation's operational risk loss profile at the required confidence level and time horizon.

The elanev approach ("the Model") is:

- **Regulatory cognisant –** The approach adopted when using the default options has passed regulator scrutiny at leading institutions.
- **Industry benchmark –** A range of industry standard distributions and modelling techniques associated with operational risk scenario modelling are available to the user.
- **Accessible –** A fully documented, readily understood approach accessed via a webmail based infrastructure hosted on elanev servers.
- **Repeatable** – An annual based service.
- **Flexible** – Multiple options for model parameterisation.
- **Convergent** – Up to five million simulations can be run for any one scenario with more available in the bespoke approach.
- **Fast –** Near real time output facilities, 'what-if' analysis, discussion and challenge.

The purpose of this document is to provide the reader with an understanding of the elanev approach, the Model inputs and outputs and enable the user to apply their inputs to the Model.

2.0 The elanev Approach

The elanev approach to modelling operational risk is based on standard statistical methods which provide a forward looking view of operational risk loss using the results of scenario analysis as inputs. These methods are commonly applied within banks, investment firms and insurers operating in the regulated financial markets of Europe, Asia, the Middle East, Africa and South America where they have gained regulatory approval. Within Europe these models are utilised by banks and investment firms to calculate Pillar 1 and Pillar 2 operational risk capital and by insurers for calculating Solvency 2 requirements. They are also being used in firm-wide stress tests such as the European Banking Authority's 2016 EU-wide stress test.

Institutions with a loss data history have looked to statistical methods to analyse operational risk loss. These methods typically sought to understand the probability of loss event frequency and loss event severity using statistical distributions. Initially historic in nature, the techniques were subsequently applied to the results of scenario analysis to provide a forward looking view of operational risk loss.

Within the elanev approach a firm's scenario based estimates are used to parameterise a frequency and a severity distribution for each key risk scenario. The distributions used, and the methods for parameterising them, are based on techniques common to these models. Similarly, the use of Monte-Carlo simulation to combine the frequency and severity distributions to identify a loss profile for each of the key risks are also commonly used techniques.

The elanev approach offers the following distributions to enable firms to model their loss profiles in a manner most appropriate to their risk scenarios:

- Poisson and Bernoulli distributions to model loss frequency
- Lognormal and Pareto distributions to model loss severity
- Multivariate Gaussian and Student's t-distributions for copula modelling

The Model outputs include:

- An undiversified loss which is calculated from the fully correlated individual loss profiles.
- A diversified loss calculated from correlated individual loss profiles.

The diversified loss utilises user specified correlations between the key risk scenarios. These are applied to standard industry multi-variable ('multivariate') statistical distribution to model the relationship between scenarios. Within the industry these multivariate distributions are referred to as copula models. A confidence level can be selected over a one year time horizon to which undiversified losses and diversified losses can be calculated.

The approach is summarised in the following schematic.

Inputs	Model	Outputs

Inputs

For each scenario:

- Expected frequency
- Two severities and associated frequencies:
  - Data point 1
  - Data point 2

Across scenarios:

- Correlations

1	0.5	0.25	0.5	0.5
0.5	1	0.75	0.25	0.75
0.25	0.75	1	0.5	0.75
0.5	0.25	0.5	1	0.25
0.5	0.75	0.75	0.25	1

Model

For each scenario loss is simulated from frequency and severity distributions

Loss profiles for each scenario are correlated using a copula model to give total loss

Outputs

- For each scenario the loss at the required confidence level
- The undiversified loss – the sum of the losses from each scenario at the required confidence level
- The diversified loss – the correlated scenario losses at the required confidence level

3.0 Model Inputs

The elanev approach uses the following scenario analysis derived data set as inputs to the Model:

- An expected loss frequency for each scenario
- Two estimates of loss severity, and associated frequencies, for each scenario
- An estimate of correlation between each scenario

Firms can select the frequency and severity distributions that best match their risk scenarios. Furthermore, firms can select the copula model that best matches the independency between their scenarios.

The frequency distribution is parameterised using the scenario based estimate of the expected loss frequency for each scenario.

The severity distribution is parameterised using scenario based estimates of the severity of loss at two frequency values. There is an additional option to specify one of these two points as the median, mode or mean of the severity distribution.

The copula model is parameterised using the correlations identified between the risk scenarios.

In addition to the flexibility offered in the distribution choice and the methods available for parameterisation, the user can additionally specify the number of scenarios and Monte-Carlo simulations required as well as selecting the confidence level and reporting currency.

Bespoke Service: Should additional distributions or associated flexibility be required, this can be made available under elanev's bespoke service offering.

4.0 Model Outputs

The core model outputs are:

- An undiversified loss which is calculated from the fully correlated individual loss profiles.
- A diversified loss calculated from correlated individual loss profiles.

These are calculated at the selected confidence level and provided for each model run. In addition a set of detailed outputs for each risk scenario and the diversified loss profile are outputs are provided as a pdf report ("the Report").

For each risk scenario the following are provided:

- A plot of the probability mass function of the frequency distribution selected.
- The value of the frequency parameter.
- A plot of the probability density function and the cumulative distribution function of the severity distribution selected.
- The values of shape and scale parameters calculated from the inputs.
- A plot of the cumulative distribution function of the simulated loss distribution.
- The simulated loss values at the following percentiles; 10% to 90% in steps of 10% and 95%, 99.5% and 99.9%.
- Values of mode, mean, median, standard deviation, skewness and kurtosis of the simulated loss profile.

For the diversified loss profile the following are provided:

- A plot of the cumulative distribution function of the simulated diversified loss distribution.
- The simulated diversified loss values at the following percentiles; 10% to 90% in steps of 10% and 95%, 99.5% and 99.9%.
- Values of mode, mean, median, standard deviation, skewness and kurtosis of the simulated diversified loss profile.
- The undiversified loss and diversified loss at the required confidence level and the corresponding diversification benefit [the difference between the undiversified loss and diversified loss as a percentage of the undiversified loss].
- The undiversified loss for each scenario and the required confidence level together with its relative ranking.

For the user's benefit the Report also provides commentary on the inputs and outputs.

5.0 User Guide

elanev has a user-friendly data interface which is currently in the form of a Microsoft® Excel 2016 workbook. This enables the user to input their data and set up the elanev model in line with their requirements. The workbook comprises three worksheets:

- Scenarios
- Correlations
- Configuration

Each worksheet has default settings. The default settings do not constitute elanev advice on the Model setup, and correspondingly, the user should be able to justify to their regulator the settings selected based on their understanding of the underlying operational risks being assessed.

In the subsequent subsections details of each worksheet are provided.

5.1 Scenarios

Within this worksheet the user can record their scenario estimates, as well as the number of scenarios, the required confidence level and their currency basis.

5.1.1 Number of Scenarios

The user can specify a number of scenarios between 2 and 20. This controls the number of scenarios that are modelled up and updates the workbook accordingly.

Bespoke Service: Some firms may require additional scenarios. This is covered by the elanev bespoke service.

5.1.2 Confidence

The user can use the "Confidence" drop down menu to select a required confidence level of 99.5% or 99.9% when calculating the scenario level and undiversified and diversified operational risk loss outputs.

Insight Tip: Banks and investment firms typically apply a confidence level of 99.9% (1 in 1000) and insurers use a 99.5% (1 in 200) confidence level.

5.1.3 Currency

The user can select the currency of the severity inputs and hence the corresponding loss outputs.

5.1.4 Scenario Name

Users can specify a name for each risk scenario of up to 100 characters. This is provided under the "Scenario Name" heading. A Scenario ID and Scenario Name are made available for each scenario up to the number of scenarios specified.

5.1.5 Expected Frequency

The user can enter the expected frequency values for each scenario using the input cells.

If using a Poisson distribution to model loss event frequency the user can vary the expected frequency between 0.1 and 50 events per annum.

If using a Bernoulli distribution to model loss event frequency the user can vary the expected frequency between 0.1 and 1 event per annum.

Details of frequency distribution selection are provided in the Frequency distribution subsection below.

Insight Tip: The Poisson distribution can be used to model the possibility of multiple events within the one-year time horizon. The Bernoulli distribution can be used to model events which only occur once over the time horizon.

5.1.6 Severity Values – Data Point 1 and Data Point 2

A minimum of two data points are required to parameterise a two parameter distribution. For this reason, the Model requires the user to specify two data points generated through their scenario analysis process for each scenario. A frequency estimate is provided along with an associated severity estimate for each data point. The first data point (data point 1) is more frequent but less severe than the second data point (data point 2) which has a lower frequency, but higher severity.

The user can enter the values for each scenario using the corresponding input cells for each data point. The frequency of data point 1 should be less than the lower value of 1 or corresponding expected frequency and greater than the corresponding data point 2 frequency. The frequency of data point 2 should be less than the corresponding data point 1 frequency and greater than 0.0001 (1 in 1000). The severity of data point 1 should be greater than 1000 but less than the severity of data point 2. Correspondingly the severity of data point 2 is greater than the severity of data point 1, but less than 1 billion.

The user should enter the full monetary amount e.g. 25,000 rather than an abbreviated amount i.e. 25.

Insight Tip: The estimates generated through the scenario analysis process are typically based on the working lifetime experiences of the experts involved (e.g. 1 in 10 years to 1 in 25 years for data point 1 and 1 in 30 years to 1 in 40 years for data point 2). The model extrapolates these frequency and severity inputs to the required confidence level; either the 99.5th (1 in 200) or to the 99.9th (1 in 1000) confidence level.

When viewed from this confidence level these input frequencies are relatively frequent and close together. Correspondingly, if the difference between the severity values is relatively large, then a significant number at the required confidence level can be expected.

The extrapolation means that the user should expect the Model output to be a number of times larger than the sum of the severities of associated with data point 2.

5.2 Correlations

Within the correlation worksheet the user can record the correlations between their scenarios.

5.2.1 Correlation Matrix

To recognise diversification between the operational risk scenarios the user will need to specify a correlation matrix. The length and width of the matrix is the same as the number of scenarios. The values in the diagonal of the matrix from the top left hand corner to the bottom right hand corner are set to one. This unit diagonal describes the correlations between a given scenario and itself.

The remaining cells of the matrix can be used to provide a correlation between zero and one to reflect the interdependency between two given scenarios. The user should record their correlations between scenarios in the lower triangle of the matrix provided.

A value of zero is associated with full independence i.e. no correlation between two scenarios; and a value of one describes full correlation between two scenarios. As correlation reflects the extent to which two variables have a linear relationship with each other, the correlation between scenario A and scenario B is the same as Scenario B and Scenario A. This results in symmetry in the unit diagonal; the values in the lower triangle of the matrix are the same as the values in the upper triangle of the matrix. Accordingly, the upper triangle will automatically update for these changes.

The model enforces a unit diagonal and symmetry by using the lower triangle of the matrix provided to over write the upper triangle.

Insight Tip: Although possible, negative correlations are not recognised in operational risk modelling which is why a lower zero bound is used in the correlation matrix.

5.3 Configuration

Within the configuration worksheet the user can select the number of simulations, the frequency distribution, the severity distribution and the copula model correlations and associated configuration.

5.3.1 Simulations

The model uses a Monte-Carlo process to generate a loss profile from the user specified frequency and severity distributions when correlating losses to provide a diversified loss profile. The Monte-Carlo process uses random numbers to simulate the loss distribution, which results in a degree of variability in the outputs. Increasing the number of simulations increases the number of losses and hence the granularity of the Model. Increasing the granularity decreases the variability. Using the "simulation" drop down menu the user can select whether to generate 500,000; 1,000,000 or 5,000,000 simulations of loss for each scenario.

Insight Tip: Increasing the number of simulated losses increases the runtime. Therefore consider using 1,000,000 simulations for initial Model runs and 5,000,000 simulations for final Model runs.

5.3.2 Severity Percentile

The severity distribution is a two parameter distribution and requires a minimum of two points to enable a value for each parameter to be calculated. Each point relates to the severity of an event and the corresponding probability, termed a percentile, of that severity. Each point can be calculated from the frequency and severity estimates recorded by the user in the *Scenarios* worksheet against Data point 1 and Data point 2.

To do this the user can elect to equate the frequency of each data point to a percentile directly (for example: 1 in 1000 years equates to 99.9th percentile) or by scaling the frequency of each data point by the corresponding expected loss frequency. When equating the frequency to the percentile directly without scaling the frequency is a probability.

Insight Tip: Scaling assumes that the loss event severity and loss event frequency are not independent. Scaling a 1/20 year frequency by the corresponding expected loss frequency of 0.5 events per annum would give a probability for the 1 in 20 year event of 10% i.e. if the expected frequency is 1 in 2 years and the severe case event is once every 20 years then 10 events occur over this period.

5.3.3 Correlate

Using the "Correlate" drop down menu the user can elect whether to correlate frequencies or simulated losses by applying correlations to the frequency or loss distributions respectively.

Insight Tip: Firms typically correlate the scenario loss profiles. When using a Bernoulli distribution to model frequency the user should strongly consider correlating at a loss level rather than at frequency level.

5.3.4 Copula Model

To calculate the diversified loss the Model uses a routine referred to as a 'copula model' to impose the correlations between the risk scenarios. A copula is a probability distribution that can be used to model the dependency between multiple variables. These distributions are referred to as multivariate probability distributions. Different types of distribution can be used to form copula models and two of the most common, utilised by the Model, are the Gaussian distribution and Student's t-distribution. The copula model can be selected using the "copula model" drop down menu.

The selection of the Student's t-distribution requires a parameter referred to as the degree of freedom parameter.

Insight Tip: The Gaussian copula generates a weak correlation in the tail when correlating fat tailed distributions. To compensate, larger values of correlations within the matrix can be used or a Student's t-distribution with a small degree of freedom parameter can be applied.

The Gaussian copula is the most commonly applied copula model within the industry as a result of its relative simplicity. However, the Student's t-distribution is finding increasing favour due to its relatively fatter tail.

Bespoke Service: Alternative copula are available as part of the elanev bespoke service.

5.3.5 Degree of Freedom Parameter

Selecting Student's t-distribution requires a parameter referred to as the degree of freedom parameter. This has a positive integer value which the user can choose to be between 2 and 1000. The smaller the value the fatter the tail of the distribution. At large values the Student's t-distribution tends towards a Gaussian distribution.

Insight Tip: Firms typically a degree of freedom parameters with integer values in the low single digits.

5.3.6 Random Seed

The random number generator is initialised using a number referred to as a "seed". The user can select to fix the seed keeping it constant or to allow it to vary.

Insight Tip: Fixing the seed will allow the user to view the effect of varying scenario input values.

5.3.7 Frequency Distribution

The user can specify whether to apply a Poisson or Bernoulli distribution to model the frequency of a given risk scenario using the "Frequency distribution" drop down menu.

Insight Tip: Poisson is the most commonly applied distribution within the industry to model the frequency of operational risk loss events.

Bespoke Service: Alternative frequency distribution models are available as part of the elanev bespoke service.

5.3.8 Severity Distribution

The user can specify whether to apply a Lognormal or Pareto (type 1) distribution to model the severity of a given risk scenario using the "Severity distribution" drop down menu. Both distributions provide a model of the probability of an event with a given severity and are fully described using two parameters. These two parameters are referred to as the scale and shape parameters.

Insight Tip: The lognormal is the most commonly applied distribution within the industry to model the severities associated with operational risk scenario analysis. However, the Pareto distribution is increasing being cited by regulators.

The scale parameter determines the scale, or statistical dispersion, of the probability distribution. The larger its value the more the distribution will be spread out, the smaller the more concentrated the distribution will be. The shape parameter affects the shape of the distribution. The larger its value, the larger the values are in the tail of the distribution.

Bespoke Service: Alternative severity distribution models are available as part of the elanev bespoke service.

5.3.9 Severity Fit

Using the "Severity Fit" drop down menu there is the option to parameterise the severity distribution using both frequency severity data points or by specifying the lower of the two severities as a mean, mode or median of the distribution and using this in conjunction with the higher frequency severity value.

6.0 Model Details

This section is aimed at providing the user with a broad understanding of the key modelling concepts within the elanev approach. Here we will elaborate on the distributions available, provide further detail on parameterisation and discuss the Monte-Carlo simulation of loss and the properties of the correlation matrix.

6.1 Distributions

The probability mass functions ("pmf") of the discrete Poisson and Bernoulli distributions can be examined below. These are used within the elanev approach to describe the frequency of a loss event. The probability density functions ("pdf") of the continuous lognormal, Pareto and multivariate Gaussian and Student t- distributions are also given. For the lognormal and Pareto distribution we also present the cumulative distribution functions ("cdf") together with the mean, mode and median as these are used to parameterise the two severity distributions. For each distribution we provide example plots of the corresponding pmfs or pdfs. Plots of the equivalent univariate pdfs are provided as a more reader friendly illustration of the multivariate distributions.

6.1.1 Poisson Distribution
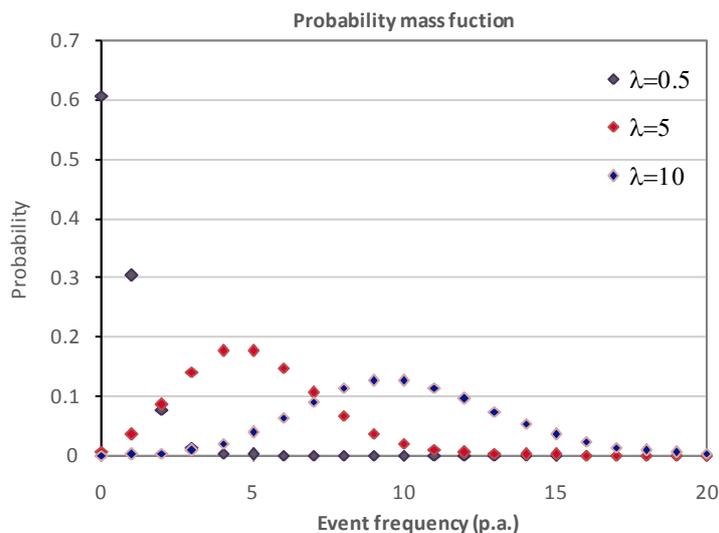
pmf:

$$f(k \mid \lambda) = \frac{\lambda^k}{k!} \exp(-\lambda)$$

mean: $\quad \lambda$

where $k \in \mathbb{Z}^+$ and $\lambda > 0$

6.1.2 Bernoulli Distribution

Probability mass fuction

p=0.8
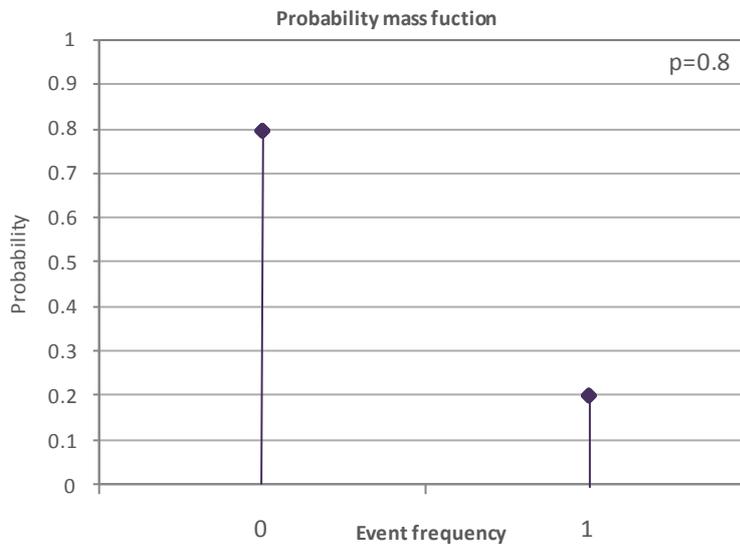
pmf:
$$f(k|p) = \begin{cases} 1 - p & k = 0 \\ p & k = 1 \end{cases}$$

mean: p

where $k \in \mathbb{Z}^+ x \in \{0,1\}$ and $p\,(0< p<1), p \in \mathbb{R}$

6.1.3 Lognormal Distribution

Probability density function
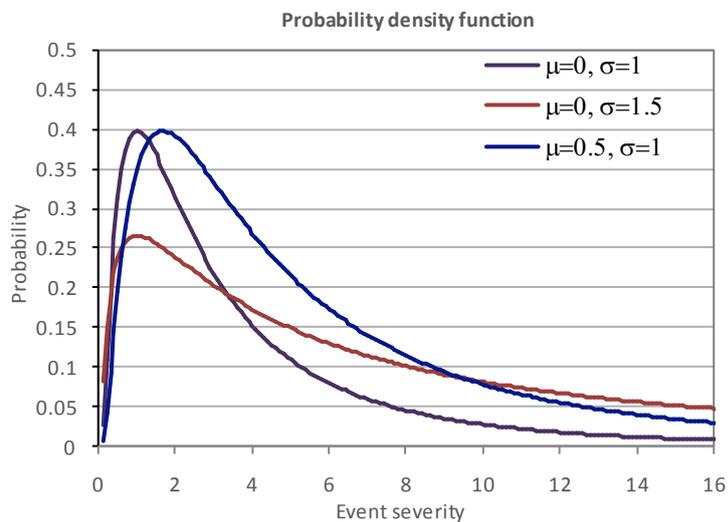
pdf:
$$f(x\,|\,\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(\frac{-\left(\ln x - \mu\right)^2}{2\sigma^2} \right)$$

cdf:
$$F(x\,|\,\mu, \sigma) = \frac{1}{2} + \frac{1}{2} erf\left(\frac{\ln x - \mu}{\sigma\sqrt{2}} \right)$$

mean: $e^{\mu + \frac{\sigma^2}{2}}$

mode: $e^{\mu - \sigma^2}$

median: e^{μ}

variance: $e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right)$

where x≥ 0, the scale, $\mu \in \mathbb{R}$ and shape $\sigma > 0$.

6.1.4 Pareto Distribution

Probability density function
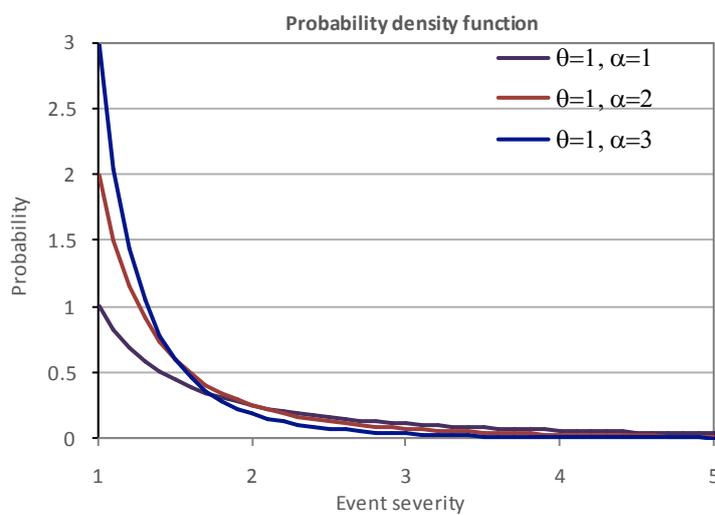
pdf: $f(x \mid \alpha, \lambda) = \dfrac{\alpha \lambda^{\alpha}}{x^{\alpha+1}}$

cdf: $F(x \mid \alpha, \lambda) = 1 - \left(\dfrac{\lambda}{x}\right)^{\alpha}$

mean: $\dfrac{\alpha\lambda}{\alpha-1}$ for $\alpha > 1$.

mode: λ

median: $\lambda \sqrt[\alpha]{2}$

variance: $\dfrac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$

where x≥λ, the scale, $\lambda > 0$ and shape $\alpha > 0$.

6.1.5 Multivariate Gaussian Distribution

pdf $\qquad f(x|\mu,\Sigma) = \frac{1}{\sqrt{|\Sigma|(2\pi)^d}} e^{\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}$

where Σ is the d-by-d correlation matrix, and μ is the mean of the distribution which is set to zero within the copula model. Σ is a d-by-d symmetric and positive semi-definite matrix.

6.1.6 Multivariate t Distribution

pdf $\qquad\qquad f(x|\Sigma,\nu) = \frac{1}{\sqrt{|\Sigma|(\nu x)^d}} \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x\prime\Sigma^{-1}x}{\nu}\right)^{-\left(\frac{\nu+d}{2}\right)}$

where Σ is the d-by-d symmetric and positive semi-definite correlation matrix, and ν is the degree of freedom parameter $\nu \in \mathbb{Z}^+$ and $\nu > 0$.
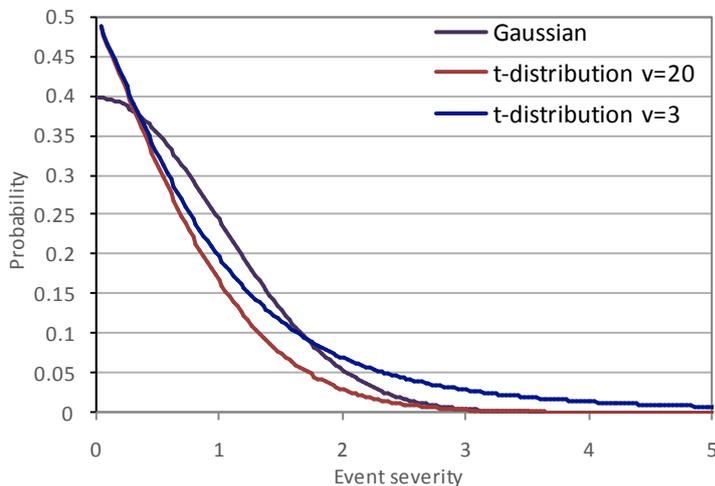
6.2 Distribution Parameters

The elanev approach, in line with the industry models, implicitly assumes that the user's loss profile can be described by the loss event frequency and loss event severity distributions selected by the user.

The elanev approach, in line with industry models, assumes that the severity distribution selected by the user passes through both scenario data points that they provide. This assumption permits the cumulative distribution function of the nominated distribution to be rearranged in terms of its shape and scale parameters and accordingly values for each to be calculated.

6.3 Loss Profile

To model the loss profile a Monte-Carlo process is used. This combines the frequency and severity distributions selected and parameterised by the user using random number generation. For each distribution selected there is a corresponding random number generator. With a sufficient large set of random numbers from each generator the corresponding distribution can be obtained.

The random number generator associated with the frequency distribution simulates the number of events that could occur within the one-year time-horizon. The random number generator associated with the severity distribution simulates the impact of a single event. Combining the two gives the total of the event severities within that time horizon; this is a single simulation step. To obtain a loss profile that simulates loss in the tail associated with the required confidence level a large number of simulation steps is required. Once the simulation is complete the losses are rank ordered to obtain the cumulative loss distribution function. The loss at the required confidence level can then be determined. For 1 million simulations the 5000th largest loss corresponds to the 99.5[th] confidence level and the 1000th largest loss corresponds to the 99.9[th] confidence level. Simply multiply these loss positions by 5 to obtain the corresponding values for 5 million simulations.

6.4 Correlation Matrix Properties

The elanev approach assumes that the interdependency between scenarios can be modelled using the multivariate distribution, or copula model, selected by the user. In order to apply the copula, the correlation matrix will need to be represented as two sub-components. These sub components are obtained through Cholesky decomposition. The decomposition can be achieved if the correlation matrix possesses a unit diagonal, is symmetric about that diagonal and is positive semi-definite.

Unlike the unit diagonal and its symmetry, the positive semi-definite property is not readily observed, being a mathematical construct. To ensure the correlation matrix possesses the positive semi-definite property, the elanev approach uses a routine referred to as 'spectral decomposition routine', to adjust a non-positive semi-definite matrix and provide a positive semi-definite matrix with similar, but slightly different correlation values.

6.5 Further Reading

Key modelling concepts detailed above are readily available within the Wikipedia and Wolfram online resources, however, additional information on the techniques used can be found in the following publications:

[1] A.M. Law and W. D. Kelton, *Simulation, Modelling and Analysis* McGraw Hill, 2000.

[2] P. Jackel. *Monte-Carlo Methods in Finance*, John Wiley & Sons, 2002.

[3] W. H. Press, S.A. Teukosky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, 1992.

7.0 Validation

To validate the Model elanev ran an exhaustive series of tests. A selection of these are presented below. The purpose of these tests are to ensure that the Model is parameterised as expected, that the Monte-Carlo convolution returns the expected result and that the copula model applies the correlation matrix as expected.

Test	Model	Tests	Result
1	Bernoulli - Lognormal - Gaussian	Parameterisation, convolution & correlation	Pass
2	Bernoulli - Lognormal - Student-t	Parameterisation, convolution & correlation	Pass
3	Poisson - Lognormal - Gaussian	Parameterisation, convolution & correlation	Pass
4	Poisson - Lognormal - Student-t	Parameterisation, convolution & correlation	Pass
5	Bernoulli - Pareto - Gaussian	Parameterisation, convolution & correlation	Pass
6	Bernoulli - Pareto - Student-t	Parameterisation, convolution & correlation	Pass
7	Poisson - Pareto - Gaussian	Parameterisation, convolution & correlation	Pass
8	Poisson - Pareto - Student-t	Parameterisation, convolution & correlation	Pass
9	Poisson - Lognormal - Student-t	Parameterisation, convolution & correlation	Pass
10	Poisson - Lognormal - Student-t	Convergence, parameterisation, convolution & correlation	Pass

Summary of the test batteries presented in this section.

Parameterisation

Test batteries 1 to 8 each contain four severity parameterisation tests one for each available method for parameterising the severity distribution:

1. Percentile method using data point 1 and 2
2. Median and data point
3. Mean and data point
4. Mode and data point

To perform these tests the inputs are derived from a predetermined scale and shape parameter. These are the compared to the Model output to test that the Model is correctly deriving the scale and shape parameters.

As an add-on product to the Model elanev can offer the facility to parameterise the severity distribution using the shape parameter in conjunction with the median, mean, mode or percentile as inputs. Test battery 9 contains four severity parameterisation tests one for each of these methods. The add-on also allows for the use of scale and shape parameters as direct model inputs.

The add-on product to the Model also contains a convergence testing facility. Test battery 10 provides a convergence test.

As a test of the frequency distribution parameterisation the expected frequency inputs to the Model are compared to the Model output to test that the Model is applying the correct frequency parameter. This test is performed in all test batteries 1 to 10.

Convolution

The frequency and severity distributions are joined using convolution to provide a loss profile for each scenario from which the loss at the required confidence level is calculated. To test that the Monte-Carlo convolution process provides the expected results the Model output is compared with an alternative model. This is performed for each of the four parameterisation tests on an individual loss basis and undiversified loss basis across all 10 test batteries for both a 99.5% and 99.9% confidence level.

Convolution involving the Bernoulli distribution can be solved analytically using the inverse of the cumulative distribution function of the corresponding severity distribution. This forms the alternative model for test batteries 1,2,5 and 6. For test batteries 3 and 4 and 7 to 10 which involve convolution with a Poisson distribution a recursive formula is used as an alternative model.

As the Model uses Monte – Carlo simulation an average value for the loss at both confidence levels together with the corresponding standard deviation is used in this comparison. This average value is calculated over 25 model runs for each test battery.

Correlation

The copula model used within the elanev model has been extensively tested through comparisons with copula model functionality available within common scripting languages. An additional test which is applied in the case of test batteries 1 to 10 is to calculate the correlations between the loss profiles within each test battery following the application of the copula model. These can then be compared to the Model inputs.

Some differences can be expected between the input and output correlations as:

1. The copula models used apply a Monte – Carlo process which as discussed introduces variation.
2. The Lognormal and Pareto distributions possess a tail and as copula models weakly correlate the tail, in comparison to the main body, this can reduce correlations exhibited between the loss profiles.

Convergence

To demonstrate convergence the Model applies five different simulation sizes: 50,000; 100,000; 500,000; 1,000,000 and 5,000,000 and for each of these 25 model runs are performed. From this the mean percentile loss at the required confidence and its standard deviation can be calculated for each of the five simulation sizes together with the maximum and minimum values. Test battery 10 provides a test of model convergence for the diversified loss at a confidence level of 99.9%.

7.1 Test Battery 1

Model inputs

No.	Description	Frequency		Severity					Correlation				Configuration	
		Expected Freq	Distribution	Frequency	Severity	Frequency	Severity	Distribution	1	2	3	4	Variable	Value
1	Data point 1 and 2 percentile	1	Bernoulli	0.05	50,133	0.025	58,687	Lognormal	1	0.25	0.5	0.75	Severity fit	Frequency
2	Median and data point 2	1	Bernoulli		22,026	0.025	58,687	Lognormal	0.25	1	0.25	0.5	Correlate	Losses
3	Mean and data point 2	1	Bernoulli		24,959	0.025	58,687	Lognormal	0.5	0.25	1	0.25	Copula	Gaussian
4	Mode and data point 2	1	Bernoulli		17,154	0.025	58,687	Lognormal	0.75	0.5	0.25	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency	Severity parameters		Loss	
		Parameter	Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	1	10	0.5	79,863 ± 0.1%	103,246 ± 0.2%
2	Median and data point 2	1	10	0.5	79,862 ± 0.1%	103,314 ± 0.3%
3	Mean and data point 2	1	10	0.5	79,852 ± 0.1%	103,269 ± 0.2%
4	Mode and data point 2	1	10	0.5	79,820 ± 0.1%	103,172 ± 0.2%
				Undiversified loss	319,397 ± 0.05%	413,001 ± 0.1%
				Diversified loss	247,321 ± 0.05%	301,542 ± 0.1%

Correlation

1	2	3	4
1	0.24	0.48	0.73
0.24	1	0.24	0.48
0.48	0.24	1	0.24
0.73	0.48	0.24	1

Model comparison

No.	Description	Loss		Difference(%) to Model outputs	
		Loss at 99.5%	Loss at 99.9%	99.5%	99.9%
1	Data point 1 and 2 percentile	79,851	103,271	0.01%	-0.02%
2	Median and data point 2	79,851	103,271	0.01%	0.04%
3	Mean and data point 2	79,851	103,271	0.00%	0.00%
4	Mode and data point 2	79,851	103,271	-0.04%	-0.10%
	Undiversified loss	319,403	413,086	0.00%	-0.02%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs
Four tests, one for each of the four available methods for parameterising the Lognormal severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 10 and 0.5.
In addition the application of the Bernoulli frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 1 and a scale and shape parameter of 10 and 0.5. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Bernoulli and Lognormal distributions can be solved analytically using the inverse of the Lognormal cumulative distribution function. This provides an alternative analytical model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.2 Test Battery 2

Model inputs

No.	Description	Frequency		Severity					Correlation				Configuration	
		Expected Freq	Distribution	Frequency	Severity	Frequency	Severity	Distribution	1	2	3	4	Variable	Value
1	Data point 1 and 2 percentile	0.5	Bernoulli	0.05	310,163	0.025	425,051	Lognormal	1	0.75	0.5	0.25	Severity fit	Probability
2	Median and data point 2	0.5	Bernoulli		59,874	0.025	425,051	Lognormal	0.75	1	0.75	0.5	Correlate	Losses
3	Mean and data point 2	0.5	Bernoulli		98,716	0.025	425,051	Lognormal	0.5	0.75	1	0.75	Copula	Student's t (DoF: 2)
4	Mode and data point 2	0.5	Bernoulli		22,026	0.025	425,051	Lognormal	0.25	0.5	0.75	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency	Severity parameters		Loss	
		Parameter	Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	0.5	11	1	613,456 ± 0.2%	1,064,165 ± 0.4%
2	Median and data point 2	0.5	11	1	612,980 ± 0.3%	1,064,765 ± 0.5%
3	Mean and data point 2	0.5	11	1	613,263 ± 0.3%	1,064,785 ± 0.6%
4	Mode and data point 2	0.5	11	1	612,614 ± 0.2%	1,062,609 ± 0.4%
				Undiversified loss	2,452,313 ± 0.1%	4,256,325 ± 0.2%
				Diversified loss	1,992,627 ± 0.1%	3,452,971 ± 0.3%

Correlation

1	2	3	4
1	0.72	0.47	0.23
0.72	1	0.74	0.49
0.47	0.74	1	0.74
0.23	0.49	0.74	1

Model comparison

No.	Description	Loss		Difference(%) to Model outputs	
		Loss at 99.5%	Loss at 99.9%	99.5%	99.9%
1	Data point 1 and 2 percentile	613,140	1,064,655	0.05%	-0.05%
2	Median and data point 2	613,140	1,064,655	-0.03%	0.01%
3	Mean and data point 2	613,140	1,064,655	0.02%	0.01%
4	Mode and data point 2	613,140	1,064,655	-0.09%	-0.19%
	Undiversified loss	2,452,558	4,258,622	-0.01%	-0.05%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs
Four tests, one for each of the four available methods for parameterising the Lognormal severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 11 and 1.
In addition the application of the Bernoulli frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 0.5 and a scale and shape parameter of 11 and 1. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Bernoulli and Lognormal distributions can be solved analytically using the inverse of the Lognormal cumulative distribution function. This provides an alternative analytical model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.3 Test Battery 3

Model inputs

No.	Description	Frequency Expected Freq	Frequency Distribution	Frequency	Severity	Severity Frequency	Severity	Severity Distribution	Correlation 1	2	3	4	Configuration Variable	Value
1	Data point 1 and 2 percentile	1	Poisson	0.05	368,016	0.025	473,530	Lognormal	1	0.5	0.25	0.75	Severity fit	Frequency
2	Median and data point 2	1	Poisson		98,716	0.025	473,530	Lognormal	0.5	1	0.5	0.25	Correlate	Losses
3	Mean and data point 2	1	Poisson		135,944	0.025	473,530	Lognormal	0.25	0.5	1	0.5	Copula	Gaussian
4	Mode and data point 2	1	Poisson		52,052	0.025	473,530	Lognormal	0.75	0.25	0.5	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency Parameter	Severity parameters Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	1	11.5	0.8	990,012 ± 0.1%	1,402,571 ± 0.3%
2	Median and data point 2	1	11.5	0.8	990,070 ± 0.2%	1,403,463 ± 0.4%
3	Mean and data point 2	1	11.5	0.8	989,919 ± 0.1%	1,402,241 ± 0.3%
4	Mode and data point 2	1	11.5	0.8	989,920 ± 0.1%	1,401,468 ± 0.2%
				Undiversified loss	3,959,921 ± 0.06%	5,609,743 ± 0.1%
				Diversified loss	2,879,641 ± 0.08%	3,852,407 ± 0.2%

Correlation

1	2	3	4
1	0.45	0.22	0.69
0.45	1	0.45	0.22
0.22	0.45	1	0.45
0.69	0.22	0.45	1

Model comparison

No.	Description	Loss Loss at 99.5%	Loss at 99.9%	Difference(%) to Model outputs 99.5%	99.9%
1	Data point 1 and 2 percentile	990,125	1,402,244	-0.01%	0.02%
2	Median and data point 2	990,125	1,402,244	-0.01%	0.09%
3	Mean and data point 2	990,125	1,402,244	-0.02%	0.00%
4	Mode and data point 2	990,125	1,402,244	-0.02%	-0.06%
	Undiversified loss	3,960,501	5,608,977	-0.01%	0.01%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs
Four tests, one for each of the four available methods for parameterising the Lognormal severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 11.5 and 0.8.
In addition the application of the Poisson frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 1 and a scale and shape parameter of 11.5 and 0.8. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Poisson and Lognormal distributions can be solved semi-analytically using a recursive formula. This provides an alternative model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.4 Test Battery 4

Model inputs

No.	Description	Frequency Expected Freq	Frequency Distribution	Frequency	Severity	Severity Frequency	Severity	Severity Distribution	Correlation 1	2	3	4	Configuration Variable	Value
1	Data point 1 and 2 percentile	2	Poisson	0.05	428,172	0.025	686,901	Lognormal	1	0.25	0.75	0.5	Severity fit	Probability
2	Median and data point 2	2	Poisson		36,316	0.025	686,901	Lognormal	0.25	1	0.25	0.75	Correlate	Losses
3	Mean and data point 2	2	Poisson		111,860	0.025	686,901	Lognormal	0.75	0.25	1	0.25	Copula	Student's t (DoF: 2)
4	Mode and data point 2	2	Poisson		3,828	0.025	686,901	Lognormal	0.5	0.75	0.25	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency Parameter	Severity parameters Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	2	10.5	1.5	2,696,159 ± 0.3%	5,319,812 ± 0.7%
2	Median and data point 2	2	10.5	1.5	2,697,645 ± 0.3%	5,324,045 ± 0.8%
3	Mean and data point 2	2	10.5	1.5	2,697,026 ± 0.3%	5,313,911 ± 0.5%
4	Mode and data point 2	2	10.5	1.5	2,694,953 ± 0.3%	5,313,032 ± 0.6%
				Undiversified loss	10,785,783 ± 0.1%	21,270,799 ± 0.3%
				Diversified loss	8,346,046 ± 0.1%	16,304,313 ± 0.3%

Correlation

1	2	3	4
1	0.26	0.73	0.47
0.26	1	0.25	0.73
0.73	0.25	1	0.35
0.47	0.73	0.35	1

Model comparison

No.	Description	Loss Loss at 99.5%	Loss at 99.9%	Difference(%) to Model outputs 99.5%	99.9%
1	Data point 1 and 2 percentile	2,697,354	5,318,298	-0.04%	0.03%
2	Median and data point 2	2,697,354	5,318,298	0.01%	0.11%
3	Mean and data point 2	2,697,354	5,318,298	-0.01%	-0.08%
4	Mode and data point 2	2,697,354	5,318,298	-0.09%	-0.10%
	Undiversified loss	10,789,416	21,273,192	-0.03%	-0.01%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs
Four tests, one for each of the four available methods for parameterising the Lognormal severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 10.5 and 1.5.
In addition the application of the Poisson frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 2 and a scale and shape parameter of 10.5 and 1.5. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Poisson and Lognormal distributions can be solved semi-analytically using a recursive formula. This provides an alternative model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.5 Test Battery 5

Model inputs

No.	Description	Expected Freq	Distribution	Frequency	Severity	Frequency	Severity	Distribution	1	2	3	4	Variable	Value
1	Data point 1 and 2 percentile	1	Bernoulli	0.05	110,784	0.025	164,624	Pareto	1	0.25	0.5	0.75	Severity fit	Frequency
2	Median and data point 2	1	Bernoulli		29,720	0.025	164,624	Pareto	0.25	1	0.5	0.5	Correlate	Losses
3	Mean and data point 2	1	Bernoulli		46,667	0.025	164,624	Pareto	0.5	0.5	1	0.25	Copula	Gaussian
4	Mode and data point 2	1	Bernoulli		20,000	0.025	164,624	Pareto	0.75	0.5	0.25	1	Simulations	5,000,000

(Header groups: Frequency | Severity | Correlation | Configuration)

Model outputs

No.	Description	Parameter	Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	1	20000	1.75	412,740 ± 0.4%	1,034,104 ± 0.8%
2	Median and data point 2	1	20000	1.75	413,262 ± 0.3%	1,037,034 ± 0.8%
3	Mean and data point 2	1	20000	1.75	412,967 ± 0.5%	1,035,740 ± 0.8%
4	Mode and data point 2	1	20000	1.75	413,146 ± 0.3%	1,037,067 ± 0.7%
			Undiversified loss		1,652,115 ± 0.21%	4,143,944 ± 0.3%
			Diversified loss		1,304,134 ± 0.17%	2,966,775 ± 0.5%

(Header groups: Frequency (Parameter) | Severity parameters (Scale, Shape) | Loss)

Correlation

1	2	3	4
1	0.24	0.48	0.73
0.24	1	0.48	0.48
0.48	0.48	1	0.24
0.73	0.48	0.24	1

Model comparison

No.	Description	Loss at 99.5%	Loss at 99.9%	Difference(%) 99.5%	Difference(%) 99.9%
1	Data point 1 and 2 percentile	412,956	1,035,895	-0.05%	-0.17%
2	Median and data point 2	412,956	1,035,895	0.07%	0.11%
3	Mean and data point 2	412,956	1,035,895	0.00%	-0.01%
4	Mode and data point 2	412,956	1,035,895	0.05%	0.11%
	Undiversified loss	1,651,826	4,143,580	0.02%	0.01%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs

Four tests, one for each of the four available methods for parameterising the Pareto severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 20000 and 1.75.
In addition the application of the Bernoulli frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs

For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 1 and a scale and shape parameter of 20000 and 1.75. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison

The convolution of the Bernoulli and Pareto distributions can be solved analytically using the inverse of the Pareto cumulative distribution function. This provides an alternative analytical model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.6 Test Battery 6

Model inputs

No.	Description	Expected Freq	Distribution	Frequency	Severity	Frequency	Severity	Distribution	1	2	3	4	Variable	Value
1	Data point 1 and 2 percentile	0.75	Bernoulli	0.05	134,164	0.025	189,737	Pareto	1	0.75	0.5	0.25	Severity fit	Probability
2	Median and data point 2	0.75	Bernoulli		42,426	0.025	189,737	Pareto	0.75	1	0.75	0.5	Correlate	Losses
3	Mean and data point 2	0.75	Bernoulli		60,000	0.025	189,737	Pareto	0.5	0.75	1	0.75	Copula	Student's t (DoF: 2)
4	Mode and data point 2	0.75	Bernoulli		30,000	0.025	189,737	Pareto	0.25	0.5	0.75	1	Simulations	5,000,000

(Header groups: Frequency | Severity | Correlation | Configuration)

Model outputs

No.	Description	Parameter	Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	0.75	30000	2	367,335 ± 0.3%	821,292 ± 0.7%
2	Median and data point 2	0.75	30000	2	367,665 ± 0.3%	821,904 ± 0.8%
3	Mean and data point 2	0.75	30000	2	367,338 ± 0.3%	821,335 ± 0.6%
4	Mode and data point 2	0.75	30000	2	367,322 ± 0.3%	821,875 ± 0.8%
			Undiversified loss		1,469,660 ± 0.2%	3,286,405 ± 0.3%
			Diversified loss		1,250,731 ± 0.1%	2,751,818 ± 0.3%

Correlation

1	2	3	4
1	0.69	0.45	0.23
0.69	1	0.72	0.48
0.45	0.72	1	0.73
0.23	0.48	0.73	1

Model comparison

No.	Description	Loss at 99.5%	Loss at 99.9%	Difference(%) 99.5%	Difference(%) 99.9%
1	Data point 1 and 2 percentile	367,423	821,584	-0.02%	-0.04%
2	Median and data point 2	367,423	821,584	0.07%	0.04%
3	Mean and data point 2	367,423	821,584	-0.02%	-0.03%
4	Mode and data point 2	367,423	821,584	-0.03%	0.04%
	Undiversified loss	1,469,694	3,286,335	0.00%	0.00%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs

Four tests, one for each of the four available methods for parameterising the Pareto severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 30000 and 2.
In addition the application of the Bernoulli frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs

For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 0.75 and a scale and shape parameter of 30000 and 2. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison

The convolution of the Bernoulli and Pareto distributions can be solved analytically using the inverse of the Pareto cumulative distribution function. This provides an alternative analytical model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.7 Test Battery 7

Model inputs

No.	Description	Frequency		Frequency	Severity	Frequency	Severity	Distribution	Correlation				Configuration	
		Expected Freq	Distribution						1	2	3	4	Variable	Value
1	Data point 1 and 2 percentile	1	Poisson	0.05	151,459	0.025	206,104	Pareto	1	0.75	0.5	0.25	Severity fit	Frequency
2	Median and data point 2	1	Poisson		54,432	0.025	206,104	Pareto	0.75	1	0.5	0.5	Correlate	Losses
3	Mean and data point 2	1	Poisson		72,000	0.025	206,104	Pareto	0.5	0.5	1	0.75	Copula	Gaussian
4	Mode and data point 2	1	Poisson		40,000	0.025	206,104	Pareto	0.25	0.5	0.75	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency	Severity parameters		Loss	
		Parameter	Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	1	40000	2.25	517,206 ± 0.2%	948,767 ± 0.5%
2	Median and data point 2	1	40000	2.25	517,819 ± 0.2%	950,297 ± 0.5%
3	Mean and data point 2	1	40000	2.25	517,375 ± 0.3%	949,584 ± 0.4%
4	Mode and data point 2	1	40000	2.25	517,560 ± 0.2%	950,516 ± 0.5%
	Undiversified loss				2,069,959 ± 0.11%	3,799,163 ± 0.2%
	Diversified loss				1,660,159 ± 0.10%	2,811,954 ± 0.3%

Correlation

1	2	3	4
1	0.69	0.45	0.22
0.69	1	0.45	0.45
0.45	0.45	1	0.69
0.22	0.45	0.69	1

Model comparison

No.	Description	Loss		Difference(%) to Model outputs	
		Loss at 99.5%	Loss at 99.9%	99.5%	99.9%
1	Data point 1 and 2 percentile	517,411	949,315	-0.04%	-0.06%
2	Median and data point 2	517,411	949,315	0.08%	0.10%
3	Mean and data point 2	517,411	949,315	-0.01%	0.03%
4	Mode and data point 2	517,411	949,315	0.03%	0.13%
	Undiversified loss	2,069,645	3,797,261	0.02%	0.05%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs
Four tests, one for each of the four available methods for parameterising the Pareto severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 40000 and 2.25.
In addition the application of the Poisson frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 1 and a scale and shape parameter of 40000 and 2.25. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Poisson and Pareto distributions can be solved semi-analytically using a recursive formula. This provides an alternative model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.8 Test Battery 8

Model inputs

No.	Description	Frequency		Frequency	Severity	Frequency	Severity	Distribution	Correlation				Configuration	
		Expected Freq	Distribution						1	2	3	4	Variable	Value
1	Data point 1 and 2 percentile	3	Poisson	0.05	165,723	0.025	218,672	Pareto	1	0.5	0.75	0.25	Severity fit	Probability
2	Median and data point 2	3	Poisson		65,975	0.025	218,672	Pareto	0.5	1	0.75	0.75	Correlate	Losses
3	Mean and data point 2	3	Poisson		83,333	0.025	218,672	Pareto	0.75	0.75	1	0.5	Copula	Student's t (DoF: 2)
4	Mode and data point 2	3	Poisson		50,000	0.025	218,672	Pareto	0.25	0.75	0.5	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency	Severity parameters		Loss	
		Parameter	Scale	Shape	Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	3	50000	2.5	978,218 ± 0.1%	1,531,494 ± 0.5%
2	Median and data point 2	3	50000	2.5	978,291 ± 0.1%	1,530,825 ± 0.4%
3	Mean and data point 2	3	50000	2.5	978,145 ± 0.1%	1,529,727 ± 0.4%
4	Mode and data point 2	3	50000	2.5	978,855 ± 0.2%	1,533,497 ± 0.5%
	Undiversified loss				3,913,510 ± 0.1%	6,125,543 ± 0.2%
	Diversified loss				3,422,831 ± 0.1%	5,262,303 ± 0.2%

Correlation

1	2	3	4
1	0.51	0.7	0.26
0.51	1	0.75	0.73
0.7	0.75	1	0.51
0.26	0.73	0.51	1

Model comparison

No.	Description	Loss		Difference(%) to Model outputs	
		Loss at 99.5%	Loss at 99.9%	99.5%	99.9%
1	Data point 1 and 2 percentile	978,019	1,530,331	0.02%	0.08%
2	Median and data point 2	978,019	1,530,331	0.03%	0.03%
3	Mean and data point 2	978,019	1,530,331	0.01%	-0.04%
4	Mode and data point 2	978,019	1,530,331	0.09%	0.21%
	Undiversified loss	3,912,076	6,121,322	0.04%	0.07%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

Inputs
Four tests, one for each of the four available methods for parameterising the Pareto severity distribution are applied. Accordingly, each of the severity input values are derived from scale and shape parameters of 50000 and 2.5.
In addition the application of the Poisson frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 3 and a scale and shape parameter of 50000 and 2.5. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Poisson and Pareto distributions can be solved semi-analytically using a recursive formula. This provides an alternative model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.

7.9 Test Battery 9

Model inputs

No.	Description	Frequency		Severity				Correlation				Configuration	
		Expected Freq	Distribution	Input Variable	Value	Shape	Distribution	1	2	3	4	Variable	Value
1	Scale and shape	2	Poisson	Scale:	10.5	1.5	Lognormal	1	0.25	0.5	0.75	Severity fit	Probability
2	Median and shape	2	Poisson	Median:	36,316	1.5	Lognormal	0.25	1	0.25	0.5	Correlate	Losses
3	Mean and shape	2	Poisson	Mean:	111,860	1.5	Lognormal	0.5	0.25	1	0.25	Copula	Student's t (DoF: 2)
4	Mode and shape	2	Poisson	Mode:	3,828	1.5	Lognormal	0.75	0.5	0.25	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency Parameter	Severity parameters Scale	Shape	Loss Loss at 99.5%	Loss at 99.9%
1	Scale and shape	2	10.5	1.5	2,696,159 ± 0.3%	5,319,812 ± 0.7%
2	Median and shape	2	10.5	1.5	2,697,645 ± 0.3%	5,324,045 ± 0.8%
3	Mean and shape	2	10.5	1.5	2,697,026 ± 0.3%	5,313,911 ± 0.5%
4	Mode and shape	2	10.5	1.5	2,694,953 ± 0.3%	5,313,032 ± 0.6%
	Undiversified loss				10,785,783 ± 0.15%	21,270,799 ± 0.3%
	Diversified loss				8,125,978 ± 0.14%	15,938,822 ± 0.3%

Correlation

1	2	3	4
1	0.26	0.49	0.72
0.26	1	0.26	0.5
0.49	0.26	1	0.36
0.72	0.5	0.36	1

Model comparison

No.	Description	Loss Loss at 99.5%	Loss at 99.9%	Difference(%) to Model outputs 99.5%	99.9%
1	Scale and shape	2,697,354	5,318,298	-0.04%	0.03%
2	Median and shape	2,697,354	5,318,298	0.01%	0.11%
3	Mean and shape	2,697,354	5,318,298	-0.01%	-0.08%
4	Mode and shape	2,697,354	5,318,298	-0.09%	-0.10%
	Undiversified loss	10,789,416	21,273,192	-0.03%	-0.01%

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.

As an add-on product to the Model elanev can offer the facility to parameterise the severity distribution using the scale and shape parameters directly or using the shape parameter in conjuction with the median, mean, mode or percentile. These are tested here.

Inputs
Four methods for parameterising the Lognormal severity distribution are tested. Accordingly, each of the severity input values are derived from scale and shape parameters of 10.5 and 1.5.
In addition the application of the Poisson frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four methods of parameterising the severity distribution the Model outputs a frequency parameter of 2 and a scale and shape parameter of 10.5 and 1.5. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Poisson and Lognormal distributions can be solved semi-analytically

7.10 Test Battery 10

Model inputs

No.	Description	Frequency		Severity					Correlation				Configuration	
		Expected Freq	Distribution	Frequency	Severity	Frequency	Severity	Distribution	1	2	3	4	Variable	Value
1	Data point 1 and 2 percentile	2	Poisson	0.05	50,133	0.025	58,687	Lognormal	1	0.5	0.75	0.25	Severity fit	Probability
2	Data point 1 and 2 percentile	0.5	Poisson	0.05	310,163	0.025	425,051	Lognormal	0.5	1	0.75	0.75	Correlate	Losses
3	Data point 1 and 2 percentile	1	Poisson	0.05	368,016	0.025	473,530	Lognormal	0.75	0.75	1	0.5	Copula	Student's t (DoF: 2)
4	Data point 1 and 2 percentile	0.5	Poisson	0.05	428,172	0.025	686,901	Lognormal	0.25	0.75	0.5	1	Simulations	5,000,000

Model outputs

No.	Description	Frequency Parameter	Severity parameters Scale	Shape	Loss Loss at 99.5%	Loss at 99.9%
1	Data point 1 and 2 percentile	2	10	0.5	188,563 ± 0.1%	228,593 ± 0.2%
2	Data point 1 and 2 percentile	0.5	11	1	677,211 ± 0.3%	1,137,143 ± 0.4%
3	Data point 1 and 2 percentile	1	11.5	0.8	989,919 ± 0.1%	1,402,241 ± 0.3%
4	Data point 1 and 2 percentile	0.5	10.5	1.5	1,242,078 ± 0.3%	2,779,188 ± 0.7%
	Undiversified loss				3,097,771 ± 0.1%	5,547,165 ± 0.4%
	Diversified loss				2,627,726 ± 0.1%	4,789,404 ± 0.3%

Correlation

1	2	3	4
1	0.45	0.65	0.23
0.45	1	0.66	0.63
0.65	0.66	1	0.44
0.23	0.63	0.44	1

Model comparison

No.	Description	Loss Loss at 99.5%	Loss at 99.9%	Difference(%) to Model outputs 99.5%	99.9%
1	Data point 1 and 2 percentile	188,560	228,593	0.00%	0.00%
2	Data point 1 and 2 percentile	677,649	1,136,614	-0.06%	0.05%
3	Data point 1 and 2 percentile	990,120	1,402,245	-0.02%	0.00%
4	Data point 1 and 2 percentile	1,243,470	2,786,880	-0.11%	-0.28%
	Undiversified loss	3,099,799	5,554,332	-0.07%	-0.13%

Model convergence

Decreasing variability with simulation size

Results

Test	Pass/Fail	Comment
Frequency parameterisation	Pass	Input and output values agree. Value consistent with loss outputs.
Severity parameterisation	Pass	Output values agree to those used to derive inputs.
Convolution	Pass	For the same inputs, Model outputs agree with comparison model.
Correlation	Pass	Correlations calculated from loss profiles agree with input values.
Convergence	Pass	Model is highly convergent.

As an add-on product to the Model, elanev can offer the facility to test its convergence. The Model's convergence is demonstrated here.

Inputs
Four tests which parameterise a Lognormal severity distribution using data point 1 and 2 percentiles are applied. The severity input values are derived from scale and shape parameters of 10 and 0.5, 11 and 1, 11.5 and 0.8 and 10.5 and 1.5.
In addition the application of the Poisson frequency distribution, the Monte-Carlo convolution and correlations applied using the copula model are tested.

Outputs
For each of the four scenarios the Model outputs a frequency parameter and a scale and shape parameter. These are the consistent with the Model's inputs.
The correlations between the loss profile are also calculated. These are similar to those used as inputs. Small differences are due to the variability inherent in Monte-Carlo techniques.
The losses at the 99.5th and 99.9th percentiles are calculated. This is done by completing 25 model runs and calculating the average value of each percentile for each of the four parameterisation methods considered. The standard deviation about this mean value is given as a percentage.

Comparison
The convolution of the Poisson and Lognormal distributions can be solved semi-analytically using a recursive formula. This provides an alternative model. Comparing the results from both approaches gives agreement to within a standard deviation of the Model's loss outputs.
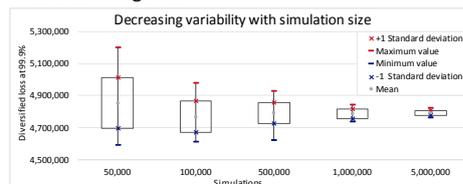
Convergence
To demonstrate convergence the Model applies five different simulation sizes: 50,000; 100,000; 500,000; 1,000,000 and 5,000,000 and for each of these 25 model runs are performed. From this the mean 99.9th percentile loss and its standard deviation can be calculated for each of the five simulation sizes together with the maximum and minimum values. These outputs show (see graph) that as simulation size increases variability in the output decreases

8.0 Conditions of Use

The provision of the elanev model ('the Model') and this document is prepared under the conditions herein.

We will not assess the validity or applicability of the Model inputs or model outputs to your business. We will not assess the accuracy of your model inputs or your model outputs or seek to identify or resolve errors in your model inputs and as a consequence in your model outputs. We will not review your model configuration nor the applicability of the corresponding assumptions to your business or any assumptions made by you in applying the Model.

When applying the Model to regulated entities it is reasonable to expect that the corresponding regulator will require you to be able to understand and justify your model inputs, your model outputs and your model configuration. Furthermore, it is reasonable to expect that your regulator will require you to have an understanding of the Model methodology as applied to your business.

The insights and methodologies presented in this document are based on the industry experience of elanev and are a distillation of this and accordingly this industry experience, whilst reasonably broad, does not extend to the entire industry. These insights and methodologies may be subsequently superseded as techniques develop.

This document is private and confidential and is prepared solely for your firm's internal use and must not be disclosed to any third parties without elanev's prior written approval.