Actuarial Perspectives on Fire Losses, Particularly for Heavy Timber Construction

Jérémie Barde
with Etienne Marceau and Hélène Cossette

59th Actuarial Research Conference
Middle Tennessee State University, Murfreesboro, TN

August 8, 2024

UNIVERSITÉ LAVAL
Faculté des sciences et de génie
École d'actuariat

CIMMUL

Quantact

CIRCERB

Agenda

Introduction

Introduction

Research context

This research has been conducted in collaboration with the *Chaire Industrielle de Recherche sur la Construction Écoresponsable du Bois* (CIRCERB).

- Industrial Research Chair on Eco-Responsible Wood Construction

What is CIRCERB?

- Multidisciplinary academic platform integrated with an industrial consortium.
- Works across the entire value creation network in the wood construction sector.

Objective: develop eco-friendly solutions that utilize wood to reduce the ecological footprint of buildings.

Introduction

Actuary role

Where can actuaries help fulfill CIRCERB's goal?

Issue discussed with experts in the heavy timber construction sector (e.g., Cecobois):

- Premiums for heavy timber construction sites can be seven times higher than for steel/concrete sites.

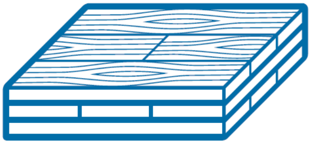
How to fairly rate the insurance premium for a construction site of a mass timber building, known as CLT ?

Introduction

Research objective

Provide an actuarial perspective on fire losses for different types of structure including heavy timber construction.

Illustration of CLT panel :



Cross-laminated timber (CLT)
is a wood panel consisting of several (usually 3, 5, or 7) layers of dimension lumber oriented at right angles to one another and glued together to form structural panels. CLT is used for floors, walls, and roofs.

Source : ISO A Verisk Business (Kahn, 2020)

Introduction

Arbora Project

Arbora project: one of the largest residential complexes constructed with mass timber in Montreal.

Source : (Levée et al., 2020)

Database

Data at disposition

Here are some data we had access to:

- Fire loss data from the city of Toronto;
- Fire loss data from the city of San Francisco;
- National Fire Information Data Base (NFID).

In this presentation we will only present the NFID:

- Put together by Len Garis, Director of Research for the National Indigenous Fire Safety Council and Adjunct Professor in the School of Culture, Media.
- Objective: gather as much data as possible on fire claims in Canada.
- Used in various works, recently in (Zheng et al., 2022).

National Fire Information Database (NFID)

Categories	Variables	Description
Accident information		
	YEAR :	Years of Incident
	MONTH :	Month of Incident
	DAY :	Day of Week of Incident
	RESPONSE :	Response Time of First Vehicle
Property information		
	GENCONST :	Type of Construction
	PROPCLAS :	Property Classification
	RISKVALA :	Value of Contents at Risk
Protection against fire		
	SPRINPRO :	Sprinkler Protection
	FIREDET :	Fire Detection Devices
Dollar Loss		
	DOLLOSSA :	Dollar Loss – Building/Vehicle
	DOLLOSSB :	Dollar Loss – Contents
Circumstances and contributing factors to the fire		
Discovery of fire and action taken		
Victim's information		

National Fire Information Database (NFID)

Presentation

Description of the NFID :

- 467 929 observations;
- 136 explanatory variables;
- From 2005 to 2015;
- Estimated losses.

Adjustment done to the `DOLLOSSA` variable :

- Removal of 360 181 unavailable observations;
- Removal of 5113 observations with no loss;
- Observations remaining: 102 635.

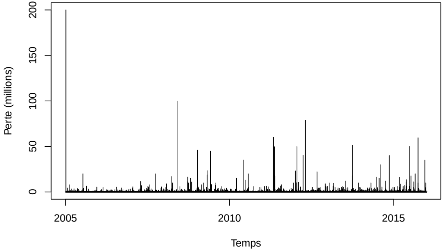
Creation of the `DATE` variable, with the `YEAR`, `MONTH` and `DAY` variables.

National Fire Information Database (NFID)

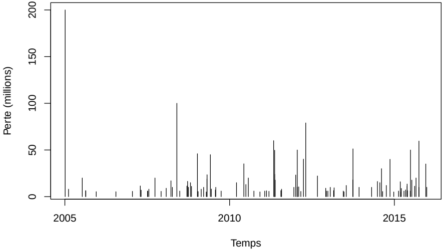
Descriptive statistics

Descriptive statistics for the loss amount.

Min.	$1st\ Qu.$	Median	Mean	$3rd\ Qu.$	Max.
1	1000	5000	75 800	30 000	200 000 000



(a) All losses



(b) Losses greater then 5 millions

UNIVERSITÉ LAVAL

National Fire Information Database (NFID)

Frequency of claims

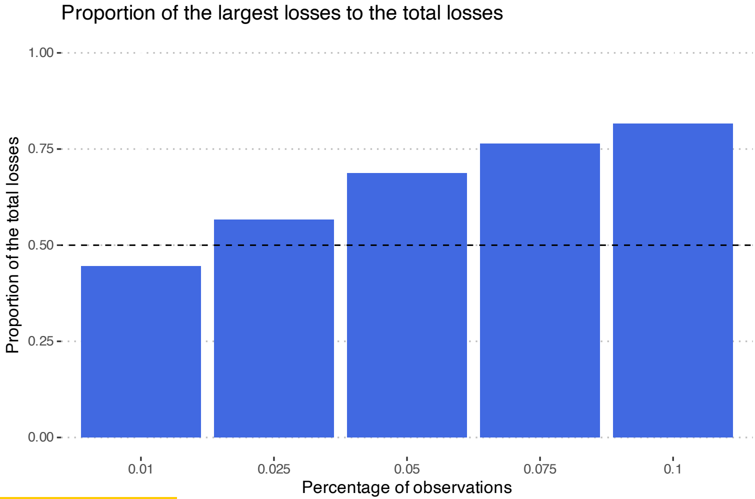
Intervals	Number of observations						
	2009	2010	2011	2012	2013	2014	2015
$[1, 50[$	7134	6869	7581	8083	7531	7234	7428
$[50, 500[$	1594	1431	2022	1902	1778	1811	1752
$[500, 1000[$	127	124	388	159	174	145	166
$[1000, 10\,000[$	78	58	107	98	89	83	89
$[10\,000, 50\,000[$	6	4	4	6	4	6	9
$[50\,000, \infty[$	0	0	1	2	1	0	2

Table: Intervals are given in multiple of 1000

Remark : It is normal to observe this kind of result for fire loss observations.

National Fire Information Database (NFID)

Contribution of the biggest losses



Collective risk model

Collective Risk Model

Definition

The temporal evolution of the total claim amount is modeled by an compound process $\underline{X} = \{X(t), t > 0\}$, defined as a random sum, i.e.,

$$X(t) = \begin{cases} \sum_{k=1}^{N(t)} B_k, & N(t) > 0 \\ 0, & N(t) = 0, \end{cases}$$

where $\underline{N} = \{N(t), t > 0\}$ is a counting process and $\underline{B} = \{B_k, k \in \mathbb{N}\}$ is a sequence of non-negative rvs.

For this project, the classic assumptions are used:

- $\underline{B}$ is a sequence of independent rvs;
- $B_1, B_2, \ldots$ follow the same distribution as B;
- the rv B and the counting process $\underline{N}$ are independent.

Collective Risk Model

Interpretation

Given these assumptions :

- We can easily calculate the expected total loss amount:

$$E[X(t)] = E[N(t)] \times E[B]$$

Interpretation: $E[X(t)]$ = pure premium of the insurance policy over the time $[0,t]$.

- We can also model $N(t)$ (frequency) and B (severity) separately.

Modelling

Severity Modelling

Tools

To model the severity, we will use two different tools:

- Extreme value theory: Peak-over-Threshold (POT) method;
- Splicing: combination of probability distributions (Brazauskas and Kleefeld, 2016).

Définition 1

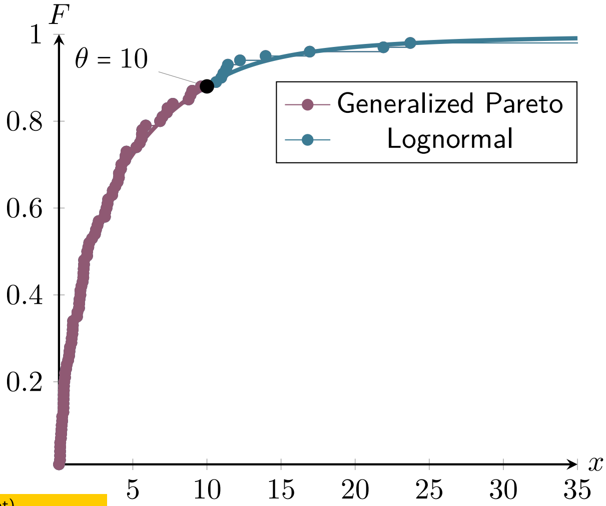
Let B be a continuous variable that follows a composite distribution, and let $\{D_j, j = 1, 2\}$ be variables that are not identically distributed. Then,

$$f_B(x) = \begin{cases} w_1 \frac{f_{D_1}(x)}{F_{D_1}(\theta)} & , x \leq \theta \\ & \\ w_2 \frac{f_{D_2}(x)}{1 - F_{D_2}(\theta)} & , x > \theta \end{cases}$$

where $w_i > 0$ and $\sum_{i=1}^{2} w_i = 1$.

Severity Modelling

Example



Severity Modelling

Distribution Tested

Splicing distributions tested:

- `LN-GPD`: Lognormal-generalized Pareto distribution;
- `We-GPD`: Weibull-generalized Pareto distribution;
- `GB2-GPD`: Generalized beta of the second kind-generalized Pareto distribution.

We based our choice on:

- QQplots;
- Cramer-Von Mises (W^2) and Anderson-Darling (A^2) statistics;
- AIC and BIC information criteria.

Severity Modelling

Chosen Distribution

Tests results:

Distribution	W^2	A^2	AIC	BIC
`LN-GPD`	13	83	-2 237 688	-2 237 640
`We-GPD`	68	428	-2 242 921	-2 242 874
`GB2-GPD`	12	74	-2 237 440	-2 237 374

Best splicing distribution for our data:

- `GB2-GPD`: Generalized beta of the second kind-generalized Pareto distribution;
- `LN-GPD`: Lognormal-generalized Pareto distribution.

Severity Modelling

Parameters and characteristics

We obtain the following estimated parameters for the GPD portion with threshold at $\theta = 350\ 000$

$\hat{\alpha}$	$\hat{\lambda}$
1.36	332 209

(a) LN-GPD distribution

$\hat{\alpha}$	$\hat{\lambda}$
1.40	343 089

(b) GB2-GPD distribution

The tail index α affects the riskiness of the distribution. In those cases the variance doesn't exist due to $\alpha \leq 2$.

In actuarial literature, $\alpha \approx 1.5$ is frequently observed for fire losses in commercial insurance, see Antal and Re (2007).

Severity Modelling

Expected Value and Variance

Expected value compared to empirical results:

Distribution	$E[B]$
Empirical	75 800
`LN-GPD`	76 274
`GB2-GPD`	73 187

UNIVERSITÉ LAVAL

Severity Modelling

Risk Measures VaR and $TVaR$

	κ				
Distribution	0.90	0.95	0.99	0.995	0.999
empirical	130 000	290 000	855 895	1 500 000	5 000 000
LN-GPD	132 343	277 560	903 834	1 494 471	4 852 422
GB2-GPD	125 968	273 232	901 700	1 476 447	4 657 006

(a) Values of $VaR_{\kappa}(B)$

	κ				
Distribution	0.90	0.95	0.99	0.995	0.999
empirical	619 893	1 042 588	3 378 455	5 881 809	18 180 293
LN-GPD	618 499	1 045 204	3 385 605	5 630 593	18 394 037
GB2-GPD	592 403	999 682	3 154 707	5 176 620	16 365 563

(b) Values of $TVaR_{\kappa}(B)$

Frequency modelling

We use a homogeneous Poisson process to model the frequency of the claims.

We use the maximum likelihood method to estimate the parameter of the Poisson process:

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{nt}.$$

Choosing time to be on a daily basis ($t = 1$), we obtain $\hat{\lambda} = 24.91$.

Interpretation: the estimated parameter $\hat{\lambda}$ = average number of fires per day.

Risk Sharing

Creation of Classes

Objective: calculate the contribution of each class to the total loss amount.

We create classes based on the 5 different structure types.

Class j	Construction	Number n_j of fires incidents
CC	Combustible - open wood joist	12 242
PCC	Protected combustible - wood protected by plaster	26 039
HTV	Heavy timber	738
NCC	Non-combustible - exposed steel	1797
PNCC	Protected non-combustible protected steel/concrete	3041

Descriptive Statistics of the Classes

Loss amount (in millions)	CC	PCC	HTC	NCC	PNCC
Median :					
	0.017	0.011	0.030	0.010	0.004
Mean :					
	0.110	0.109	0.380	0.218	0.154
Percentiles :					
0.99	1.19	1.00	2.32	3.59	2.06
0.999	5.84	6.00	63.15	15.28	20.00
Maximum losses :					
	60	60	100	46	45
Proportion of observations:					
	0.25	0.53	0.05	0.07	0.09

Greatest losses for each class

Combustible (CC)	
Loss (millions)	Sector
60.0	Diverse
20.0	Residential
19.0	Residential
16.2	Residential
12.0	Residential

Protected Combustible (PCC)	
Loss (millions)	Sector
59.6	Business
50.0	Manufacturing
35.2	Residential
35.0	Institutional
23.5	Residential

Heavy Timber (HTC)	
Loss (millions)	Sector
100	Business
50.0	Manufacturing
6.5	Residential
5.0	Residential
4.0	Residential

Non-Combustible (NCC)	
Loss (millions)	Sector
46.0	Manufacturing
16.4	Manufacturing
15.0	Manufacturing
13.0	Manufacturing
8.0	Storage

Protected Non-Combustible (NCC)	
Loss (millions)	Sector
45.0	Manufacturing
40.0	Residential
22.2	Residential
20.0	Assembly
20.0	Storage

Descriptive Statistics of the Classes

In this table we compare the median loss amount and the maximum loss amount to the median value of the building:

	Median Loss	Maximal Loss	Median Building Value
CC	16 700 $	60 millions $	150 000 $
PCC	10 500 $	59 millions $	300 000 $
HTC	30 000 $	100 millions $	215 950 $
NCC	10 000 $	46 millions $	1 000 000 $
PNCC	4000 $	45 millions $	2 500 000 $

Modelling Severity and Frequency

To calculate the contribution of each class, we will use the Collective risk model with

- Severity: `LN-GPD` distribution;
- Frequency: homogeneous Poisson process.

For the α parameter, we get

Class j	CC	PCC	HTC	NCC	PNCC
$\hat{\alpha}_j$	1.75	1.75	1.30	1.52	1.40

For the λ parameter with $t = 1$, we get

Class j	CC	PCC	HTC	NCC	PNCC
$\hat{\lambda}_j$	3.05	6.48	0.18	0.45	0.76

Risk Sharing Rules

Objective: Find a fair way to calculate each participant's contribution.

Let $S_d = X_1 + \cdots + X_5$ be the total loss amount. Then, the contribution of class j is given by C_j.

We consider two risk sharing rules:

1. Conditional mean: $C_j^{cm} = E[X_j|S]$;
2. Proportional mean: $C_j^{pm} = \frac{E[X_j]}{E[S]}S$.

To calculate C_j^{cm}, we use the approach presented in Blier-Wong et al. (2022).

We use the ordinary generating function of the expected allocations, which is defined as

$$\mathcal{P}_S^{[j]}(t) := \sum_{j=1}^{\infty} E[X_j \times 1_{\{S=k\}}]t^k, \quad j = 1, \ldots d.$$

Results

	Classes ($s = 500\ 000$)				
	CC	PCC	HTC	NCC	PNCC
(α)	(1.75)	(1.75)	(1.30)	(1.52)	(1.40)
Cond. Mean	29	60	2	4	6
Prop. Mean	26	52	4	8	10

Classes ($s = 1\ 000\ 000$)				
CC	PCC	HTC	NCC	PNCC
(1.75)	(1.75)	(1.30)	(1.52)	(1.40)
29	60	2	4	5
26	52	4	8	10

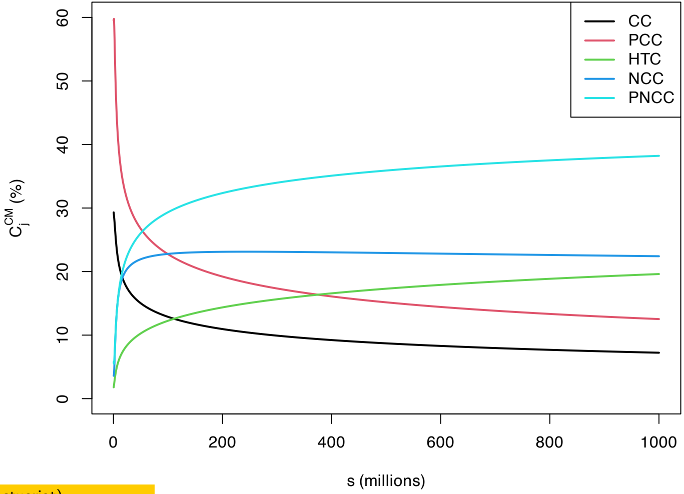
	Classes ($s = 30\ 000\ 000$)				
	CC	PCC	HTC	NCC	PNCC
(α)	(1.75)	(1.75)	(1.30)	(1.52)	(1.40)
Cond. Mean	24	47	4	12	12
Prop. Mean	26	52	4	8	10

Classes ($s = 117\ 540\ 000$)				
CC	PCC	HTC	NCC	PNCC
(1.75)	(1.75)	(1.30)	(1.52)	(1.40)
20	38	6	18	18
26	52	4	8	10

Table: Contribution (%) of each class given $S = s$

Results

Comparison of contributions to the total loss for the 5 classes



Conclusion

Concluding remarks

What we have accomplished:

- Investigate the National Fire Information Database;
- Estimate the distribution of the fire loss amount using the `LN-GPD` and `GB2-GPD` distributions.
- Using risk sharing, it has been determined that the type of construction has a significant impact on the total loss.

Future work:

- Use predictive model for the distribution below the threshold;
- Also add covariates for the generalized Pareto distribution portion,
- Work with real insurance data.

Acknowledgements

I sincerely thank La Chaire Industrielles de Recherche sur la Construction Écoresponsable du Bois (CIRCERB) for having largely funded my work, which allowed me to focus full-time on my studies.

I'm also thankful to La Chaire d'Actuariat for enabling me to come here and present today.

I would also like to thank the following people:

- Prof. Etienne Marceau
- Prof. Hélène Cossette
- My colleagues from the ACT&RISK Lab

Thank you for your attention!

Any questions?

References I

Antal, D. P. and Re, S. (2007). Mathematical methods in reinsurance. *Lecture Notes, ETH Zurich*.

Blier-Wong, C., Cossette, H., and Marceau, E. (2022). Generating function method for the efficient computation of expected allocations. *arXiv preprint arXiv:2207.02654*.

Brazauskas, V. and Kleefeld, A. (2016). Modeling severity and measuring tail risk of norwegian fire claims. *North American Actuarial Journal*, 20(1):1–16.

Kahn, D. (2020). The mass timber revolution. *White paper ISO A Verisk Business*.

Levée, V., Lavoie, K., and Bellavance, S. T. (2020). Arbora : Une structure en bois apparente dans un projet résidentiel de grande envergure. *Cecobois*, page 16.

Zheng, A., Garis, L., and Pike, I. (2022). Fire severity outcome comparison of apartment buildings constructed from combustible and non-combustible construction materials. *Fire technology*, 58(4):1815–1825.

Abstract

Fire losses represent a significant risk in property insurance. To comprehend and quantify this risk, we use two Canadian databases containing numerous fire incidents: one from the city of Toronto and another named the National Fire Information Database (NFID). We model the losses employing various parametric families and use risk measures to quantify the risk of the model. In the pursuit of developing a peer-to-peer insurance model for the heavy timber construction sector, we are interested in risk-sharing rules. Our aim is to comprehend the dynamics of risk sharing by examining a portfolio of buildings with diverse structural types.

Reserve

Analytical form of the characteristics

For a splicing distribution with GPD above the threshold, the characteristics are:

- Expected Value:

$$E[B] = w\frac{E\left[D_1 \times 1_{\{D_1 \leq \theta\}}\right]}{F_D(\theta)} + (1-w)\left(\frac{\sigma}{1-\xi} + u\right)$$

- Value-at-Risk (VaR):

$$c = VaR_\kappa(B) = \begin{cases} F_{D_1}^{-1}\left(\frac{\kappa F_{D_1}(\theta)}{w}\right), & \kappa \leq w \\ \frac{\sigma}{\xi}\left(\left(\frac{1-\kappa}{1-w}\right)^{-\xi} - 1\right) + \theta, & \kappa > w \end{cases}$$

- Tail-value at risk ($TVaR$):

$$TVaR_\kappa(B) = \begin{cases} \frac{w}{F_{X_1}(u)(1-\kappa)}\left(E[D_1 \times 1_{\{D_1 \leq \theta\}}] - E[D_1 \times 1_{\{D_1 \leq c\}}]\right) + \frac{1-w}{(1-\kappa)}E[D_2 \times 1_{\{D_2 > \theta\}}] & ,c \leq \theta \\ (1-w)\frac{\sigma}{\xi}\left(\frac{1}{1-\xi}(1-\kappa)^{-\xi} - 1\right) + \theta & ,c > \theta \end{cases}$$