Insurance Data Science Conference 19 - 20 June 2025

Programme and Abstract Booklet

Scientific Committee

2025-05-29

Contents

Conference sponsors **2**

Scientific Committee **3**

Programme **4**

Abstracts of contributed talks **9**

Index of presenters **68**

Conference sponsors

Gold sponsors

Posit: The open source data science company.

Ledger Investing: We are building an open insurance system to connect risk to capital.

Silver sponsors

Mirai Solutions: Smarter analytics - better decisions

Markel: Bold ideas. Honest actions.

Ki: Risk, simplified.

Scientific Committee

Academic	Industry
Andreas Tsanakas (Bayes Business School)	Can Baysal (Munich Re, Munich)
Arthur Charpentier (Université du Québec à Montréal)	Claudio Giancaterino (Towards Innovation Lab, Italy)
Bernard Wong (UNSW Sydney)	Davide de March (Markel, London)
Despoina Makariou (University of St.Gallen)	Dylan Liew (Bupa, London)
Filip Lindskog (Stockholm University)	Giorgio Alfredo Spedicato (Unipol Group, Milan)
Ioannis Kyriakou (Bayes Business School)	Gráinne McGuire (Optum)
Katrien Antonio (KU Leuven, University of Amsterdam)	Jürg Schelldorfer (Swiss Re, Zurich)
Mario Wüthrich (RiskLab, ETH Zurich)	Luca Baldassarre (Swiss Re, Zurich)
Mathias Lindholm (Stockholm University)	Marcela Granados (Databricks, New York)
Mike Ludkovski (UC Santa Barbara)	Markus Gesmann (ICMR, London)
Montserrat Guillén (Universitat de Barcelona)	Markus Senn (Partner Re, Zurich)
Munir Hiabu (University of Copenhagen)	Mick Cooney (Describe Data, Dublin)
Pietro Millossovich (Bayes Business School)	Ronald Richman (Johannesburg, South Africa)
Rui Zhu (Bayes Business School)	Wui Hua (John) Ng (Zurich Financial Services, Melbourne)
Salvatore Scognamiglio (University of Naples)	
Silvana Pesenti (University of Toronto)	
Stefan Weber (House of Insurance, University of Hannover)	

Keynotes

- **James David Long** (CTO, Palomar)
- **Magdalena Ramada** (Senior Director, WTW)
- **Johanna Ziegel** (Department of Mathematics, ETH Zurich)

Panel discussion

Effective Collaboration between Academia and Industry in Insurance Data Science

Chair: Dylan Liew, Bupa

Panellists:

- Debashish Dey, Aviva
- Tina Thomson, Gallagher Re
- George Tzougas, Heriot-Watt University
- Malgorzata Wasiewicz, UCL

Programme

Venue

- City, University of London, **Northampton Square**, London **EC1V 0HB**

19 June 2025

08:30 - 09:00 Registration

09:00 - 09:15 Welcome (Room B200)

09:15 - 10:15 Keynote 1 (Room B200) (Chair: Andreas Tsanakas)

Magdalena Ramada (WTW): Emerging Technologies - Current Impact on the Insurance Value Chain

10:15 - 11:15 Regular Session 1

Stream 1 Room B200: Fairness (Chair: Karol Gawlowski)

- **Charlotte Jamotton** (Université Catholique de Louvain): A multivariate energy distance approach to premium fairness adjustment
- **Mathias Lindholm** (Stockholm University): Sensitivity-based measures of discrimination in insurance pricing
- **Marie-Pier Côté** (Université Laval): A scalable toolbox for exposing indirect discrimination in insurance rates

Stream 2 Room B103: Data (Chair: Markus Gesmann)

- **Chris Halliwell** (Markel): NLP for data granularity improvement
- **Mick Cooney** (Describe Data): Analysing ship utilisation using AIS data
- **Manuel Caccone** (Italian Society of Actuaries (ISOA)): Application of NLP models in loss modeling for actuarial science

11:15 - 11:40 Coffee break

11:40 - 12:40 Lightning Session 1

Stream 1 Room B200: Climate (Chair: Mathias Lindholm)

- **Hirbod Assa** (University of Essex): Drought parametric insurances by a two-step machine learning approach under climate change scenarios
- **Nan Zhou** (Complutense University of Madrid): The influence of climate change on insurance sustainability: Evidence from Spanish agricultural insurance
- **Rasmussen Yubo** (Heriot-Watt University): A sub neural network approach for forecasting climate-related claim costs in property insurance
- **Despoina Makariou** (University of St Gallen): Mitigating systemic risk in catastrophe insurance: The role of human judgment in model diversification
- **Yushan Liu** (Institute Polytechnique de Paris): Meta-modelling paths of simple climate models using Neural Networks and Dirichlet polynomials: An application to DICE

Stream 2 Room B103: Natural Language Processing (Chair: Ioannis Kyriakou)

- **Amin Karbassi** (AXA XL): Track trending topics in insurance for emerging risk identification
- **Claudio Giorgio Giancaterino** (Towards Innovation Lab): Harnessing conditional generative models for synthetic non-life insurance premium data
- **Pratyush Singh** (Swiss Re): Enhancing contract wordings analysis with generative AI: A timeline of efficiency and accuracy
- **Matej Otcenas** (SwissRe): Advancing Claims Document Processing with LLMs: From PoC to Production
- **Malgorzata Smietanka** (UCL): LLMs for claims processing: A fully local and compliant solution
- **Daniel Jakobi** (Markel): Identifying similar insurance claims using text-based vector search

12:40 - 13:40 Lunch

13:40 - 14:40 Regular Session 2

Stream 1 Room B200: Mortality (Chair: Pietro Millossovich)

- **Carlos Arocha** (Arocha & Associates GmbH): Climate-enhanced pricing: Using gradient boosting machines to personalise life insurance rates
- **Jens Robben** (University of Amsterdam): Granular mortality modeling with temperature and epidemic shocks: A three- state regime-switching approach
- **Huiling Zheng** (UCL): Fine-grained mortality forecasting with deep learning

Stream 2 Room B103: Pricing Methods (Chair: Bernard Wong)

- **Robert Carruthers** (Ki Insurance): Predicting the full D&O insurance tower with securities class action data
- **Pierre-Olivier Goffard** (Université de Strasbourg): Market-based insurance ratemaking: Application to pet insurance
- **Salvatore Scognamiglio** (University of Naples Parthenope): The credibility transformer

14:40 - 15:40 Panel discussion (Room B200)

Effective Collaboration between Academia and Industry in Insurance Data Science (Chair: Dylan Liew)

- Debashish Dey, Aviva
- Tina Thomson, Gallagher Re
- George Tzougas, Heriot-Watt University
- Malgorzata Wasiewicz, UCL

15:40 - 16:00 Coffee break

16:00 - 17:00 Regular Session 3

Stream 1 Room B200: Portfolio Management (Chair: Mick Cooney)

- **Francesca Nuzzo** & **Leonardo Ruggieri** (Generali Italia): Strategic asset allocation for insurance product development: A machine learning approach
- **Markus Gesmann** (Insurance Capital Markets Research): Winning strategies: Predict relative performance
- **Freek Holvoet** (KU Leuven): Multi-view spatial embeddings for insurance portfolio analytics

Stream 2 Room B103: Explainability (Chair: Rui Zhu)

- **Karol Gawlowski** & **Patricia Wang** (EY): Ensembling GLM and XGB
- **Lucas Muzynoski** (Avenue Analytics): Fully transparent machine learning: Exact factor table representation of GBMs
- **Quentin Guibert** (Université Paris-Dauphine): Explainable boosting machine for predicting claim severity and frequency in car insurance

17:00 - 18:00 Keynote 2 (Room B200) (Chair: Gráinne McGuire)

James David Long (Palomar): Is Data Engineering the New Data Science?

19:00 Conference dinner

- **Ironmongers' Hall**, Shaftesbury Place, Barbican, London EC2Y 8AA

20 June 2025

09:15 - 10:15 Room B200: Keynote 3 (Chair: Mario Wüthrich)

Johanna Ziegel (ETH Zurich): Conformal calibration guarantees for reliable predictions

10:15 - 11:15 Lightning Session 2

Stream 1 Room B200: Risk Management & Theory (Chair: Davide De March)

- **Emilio L. Sáenz Guillén** (City St Geoge's): Non-parametric insurance loss modelling using variable-knot splines
- **Francesco Ungolo** (University of New South Wales): An augmented variable Dirichlet process mixture model for the analysis of dependent lifetimes
- **Brandon Schwab** (Leibniz University Hannover): Elevating trust in high-stakes decisions using glass-box models and robust feature selection
- **Filip Lindskog** (Stockholm University): Claims processing and costs under capacity constraints
- **Ran Xu** (Xi'an Jiaotong-Liverpool University): LSTM-based coherent mortality forecasting for developing countries
- **Giovanni Rabitti** (Heriot-Watt University): Analytical variable importance indices for generalized additive models

Stream 2 Room B103: AI/ML (Chair: Claudio Giancaterino)

- **Amaryllis Mouyiannou** (Swiss Re): Lost in translation: Is portfolio analytics an academic exercise or a strategic asset?
- **Anne van der Scheer** (Perunum Actuarieel Advies): From claim counts to interarrival times using a small neural framework
- **Yiannis Parizas** (Actuary and Open-Source Developer): Advancing non-life insurance modelling with NetSimR
- **Michael Ramati** (Earnix): Price leakage in demand models
- **Gwen Chan** (Ki Insurance): More data or more model: A framework for achieving predictive analytics objectives in Specialty insurance
- **Gustavo Martinez** (Mirai Solutions): Agentic AI applications in insurance

11:15 - 11:45 Coffe break

11:45 - 12:45 Regular Session 4

Stream 1 Room B200: Regression Methods (Chair: George Tzougas)

- **Mario Wüthrich** (ETH Zurich): Tests for auto-calibration
- **Yuval Ben Dror** (Earnix): A two-step regularization algorithm to cluster categories in GLMs
- **Tian Dong** (UNSW Sydney): Distributional regression for actuarial applications: distributional refinement network

Stream 2 Room B103: Climate Risks (Chair: Hirbod Assa)

- **José Garrido** (Concordia University): Catastrophic-risk-aware reinforcement learning with extreme-value-theory-based policy gradients
- **Olivier Lopez** (Ensae Institut Polytechnique de Paris): Design of parametric insurance via machine learning and optimal combination with traditional insurance

- **Mathias Valla** (France & Aix-Marseille University): Feature and quantile selection for the actuarial climate index: Everything, everywhere, all at once

12:45 - 13:45 Lunch

13:45 - 14:45 Lightning Session 3

Stream 1 Room B200: Bayes & Graphs (Chair: Despoina Makariou)

- **Mark Shoun** (Ledger Investing): Copula models of correlation in insurer loss reserves
- **Conor Goold** (Ledger Investing): Automating tail factor detection points using Bayesian hidden Markov models and latent change-point models
- **Paul Wilsens** (KU Leuven): Machine learning in an expectation-maximisation framework for nowcasting
- **Juan Ignacio de Oyarbide** (Addactis): Pricing workers compensation via Bayesian hierarchical modeling
- **George Tzougas** (Heriot-Watt University): Dynamic hierarchical graph neural networks for spatiotemporal prediction of flood-related claims
- **Aurélien Couloumy** (University Lyon 1): Causal knowledge graphs for risk interpretation using LMM: A new tool for insurers

Stream 2 Room B103: AI/ML (Chair: Filip Lindskog)

- **Tsz Chai Fung** (Georgia State University): Statistical learning of trade credit insurance network data with applications to ratemaking and reserving
- **Emanuele Fabbiani** (Xtream): From SHAP to EBM: How to explain gradient Boosting models
- **Meryem Schalck** (IPAG Business School): Auto insurance fraud detection: Machine learning and deep learning applications
- **Guangyuan Gao** (Renmin University of China): Additive tree latent variable models with applications to insurance loss prediction
- **Juan Yanez** (University of Barcelona): How does granularity affect motor insurance claim predictions in a telematics setting?
- **Samuel Gyamerah** (Toronto Metropolitan University): Metaheuristic-informed machine learning for optimizing strike temperatures in weather index insurance

14:45 - 14:55 Room B200: Closing comments (Markus Gesmann)

Abstracts of contributed talks

Climate-Enhanced Pricing: Using Gradient Boosting Machines to Personalise Life Insurance Rates

Carlos Arocha, Arocha & Associates GmbH (presenter)

Abstract: As climate change continues to impact global health and mortality rates, integrating climate risk into life insurance pricing is becoming increasingly important. In this session, I will present a practitioner's approach to using Gradient Boosting Machines (GBMs) for more accurate and climate-aware life insurance pricing. Unlike academic discussions, this session will focus on practical tips and techniques that can be directly applied to everyday pricing strategies.

I will demonstrate how to combine personal policyholder data with climate-related variables using GBMs to create a more comprehensive risk model. Attendees will learn how to preprocess data, engineer features, and implement GBMs in real-world scenarios, providing actionable insights for improving pricing accuracy. Session attendees wil walk away with clear, actionable strategies to enhance pricing models using machine learning, with a focus on the growing importance of climate risk.

Keywords: Machine Learning, Gradient Boosting Machines (GBM), Climate Risk, Life Insurance

References

1. **Büchi, G., & Bonhoeffer, S. (2020).**
   Climate Change and Life Insurance: Assessing the Impact of Climate Risk on the Insurance Industry.
   Journal of Environmental Economics and Management, **62**(3), 345-367.
2. **Yuan, Y., & Zhou, M. (2018).**
   Predicting Life Insurance Pricing with Machine Learning Algorithms.
   Journal of Risk and Insurance, **85**(2), 450-479.
3. **XGBoost Documentation. (2022).**
   XGBoost: A Scalable Tree Boosting System.
   Retrieved from https://xgboost.readthedocs.io/.

Contact details

- Email: ca@ArochaAndAssociates.ch
- Homepage: ArochaAndAssociates.ch
- Social media: LinkedIn
- LinkedIn: https://www.linkedin.com/in/carocha/

Drought parametric insurances by a Two-Step machine learning approach under climate change scenarios

Hirbod Assa (presenter)

University of Essex

Abstract: In this paper, we utilize data from the IPCC to develop a predictive model for the Palmer Drought Severity Index (PDSI). Our approach involves a two-step modeling process: initially applying a random forest regression, followed by a linear regression correction, achieving a forecasting accuracy exceeding 94%. This method aims to predict the drought index using a minimal set of climate indices, with projections extending to the year 2100 based on CMIP6 models. The analysis focuses on selected U.S. states, assessing the impacts of different climate scenarios and climate change under various SSP pathways. On that basis we designe a parametric insurance on drought index. If time permits, we will also explore the implications of these drought projections on supply chain risks in the beef commodity market.

Keywords: Drought, IPCC, CMIP6, Climate Change, Parametric insurance

Contact details

- Email: h.assa@essex.ac.uk
- Homepage: www.hirbod-assa.com
- Social media: https://www.linkedin.com/in/hirbod-assa-a8a75237/

Application of NLP Models in Loss Modeling for Actuarial Science

Manuel Caccone, Gruppo Unipol (presenter)

Abstract: This paper explores the application of Natural Language Processing (NLP) models to address misclassification issues in multi-peril insurance policies. Traditional loss modeling in actuarial science relies on frequency-severity models and compound models using Generalized Linear Models (GLMs), but these approaches struggle to capture the complexity of loss data and often result in significant noise.

Our research demonstrates how NLP techniques, particularly Zero-Shot Classification (ZSC), can be applied to loss adjuster evaluations and policyholder declarations to improve the classification of claims. Using a dataset of 78 health insurance claims across five perils (Dentistry, Visits, Mammography, Analysis, and Diagnostics), we identified a 69.23% misspecification ratio when comparing original classifications to those suggested by our NLP approach.

The implementation of ZSC significantly altered the empirical distribution functions of claims and revealed more reasonable correlation patterns between different perils. By recovering the true context of claims information, we observed a "spreading distribution" effect rather than the "narrowing distribution" that occurs with incorrect classification. This approach not only improves the accuracy of loss classification but also enhances the foundation for developing more accurate pricing models.

Keywords: Actuarial Science, Natural Language Processing, Zero-Shot Classification, Loss Modeling, Claim Misclassification

References

1. Boulieris, P., Pavlopoulos, J., Xenos, A., Vassalos, V. (2023). Fraud Detection with Natural Language Processing. Machine Learning, 1-22.
2. Devlin, J., Chang, M., Lee, K., Toutanova, K. (2018). BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. CoRR abs/1810.04805.
3. Pitkänen, P. (1975). Tariff Theory. ASTIN Bulletin 8(2), 204-228.
4. Chaudhary, A. (2020). Zero Shot Learning for Text Classification.
5. HugginFace. (2024). Zero-Shot Classification.

Contact details

- Email: manuel.caccone@gmail.com
- Affiliation: Italian Society of Actuaries (ISOA)
- LinkedIn: www.linkedin.com/in/manuel-caccone-42872141

Predicting the full D&O insurance tower with Securities Class Action data

Robert Carruthers, Ki Insurance (presenter)

Shane Murphy, Ki Insurance (presenter)

Abstract: Directors and Officers (D&O) is a type of liability insurance that covers individuals for claims made against them while serving on a board of directors and/or as an officer. The low claim frequency and long tail (3-5 years) of D&O make it notoriously difficult to predict. Claims can be up to hundreds of millions, and hence there is an active insurance market in Lloyd's of London for Excess as well as Primary layer coverage.

The market has traditionally relied on fixed Increased Limit Factor (ILF) curves to price the Excess tower. Practice varies widely, with curves tending to be of unclear derivation, out-of-date, and statically applied across very different insureds.

We have implemented a novel machine learning-based approach to D&O underwriting by combining live sources of financial data about our insureds with D&O-specific Securities Class Action data, building a model that covers the entire US Public D&O market. A key innovation in our model is the application of distributional regression with the NGBoost framework to predict the entire insurance tower at the policy level for a given insured.

This work has transformed our approach to D&O underwriting by offering unprecedented levels of sophistication and transparency. This talk will cover the full spectrum of the project from technical approach and modelling through to lessons learned from implementation in a full production context and release as part of an automated rating system.

Keywords: D&O insurance, insurance layer cost prediction, distributional regression, risk modelling, firmographic data, Securities Class Action

References

1. Tyralis, H. and Papacharalampous, G., 2024. A review of predictive uncertainty estimation with machine learning. Artificial Intelligence Review, 57(4), p.94.
2. Duan, T., Anand, A., Ding, D.Y., Thai, K.K., Basu, S., Ng, A. and Schuler, A., 2020, November. Ngboost: Natural gradient boosting for probabilistic prediction. In International conference on machine learning (pp. 2690-2700). PMLR.
3. Matheson, J.E. and Winkler, R.L., 1976. Scoring rules for continuous probability distributions. Management science, 22(10), pp.1087-1096.
4. Donelson, D.C., Hopkins, J.J. and Yust, C.G., 2015. The role of directors' and officers' insurance in securities fraud class action settlements. The Journal of Law and Economics, 58(4), pp.747-778.
5. Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E.K. and Suvas, A., 2017. Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. Journal of international financial management & accounting, 28(2), pp.131-171.

Contact details

- Email: robert.carruthers@ki-insurance.com, shane.murphy@ki-insurance.com
- Homepage: ki-insurance.com

More data or more model: A framework for achieving predictive analytics objectives in Specialty insurance

Gwen Chan, Ki Insurance (presenter)

Priyank Shah, Lane Clark & Peacock (presenter)

Abstract: What is the best route to achieving your predictive analytics objectives for a particular business case? Is it collecting more data, adding more breadth or more depth, or doing more feature engineering and investigating different model architectures? In this talk, we introduce a practical framework that helps organisations answer these questions by using: - Learning curves to estimate the likely improvements from additional data vs model improvements and assess the value of different potential data assets. - Heuristics which can be used in cases where the detailed modelling needed for learning curves is not possible. Drawing on two real-world case studies from the insurance industry, we showcase how this framework can be applied in different business contexts.

First, we examine how a Lloyd's syndicate evaluated the importance of different data assets for algorithmic underwriting performance. This case study shows how organisations can use the framework to identify where targeted data acquisition strategies will boost model performance, and where they won't.

Our second case study focuses on improving the performance of a model which predicts if there will be deteriorations in insurers' reserves. Attendees will learn a systematic approach for identifying whether additional data or model structure changes are the best route for improving predictive performance. This includes quantifying the marginal value of additional data assets, and aligning these insights with strategic business goals, specifically within the insurance industry.

Keywords: specialty insurance, predictive analytics, predictive performance

Contact details

- Email: gwen.chan@ki-insurance.com, priyank.shah@lcp.uk.com
- Homepage: ki-insurance.com, https://www.lcp.com

Analysing Ship Utilisation using AIS Data

Mick Cooney, Describe Data

Abstract: Automatic Identification System (AIS) is a maritime technical standard developed by the International Maritime Organisation (IMO).

AIS is a sophisticated radio technology which combines GPS, VHF and data processing technologies to enable the exchange of relevant information in a strictly defined format between different marine entities. This may be the simple exchange of position, course, speed and identity information between individual vessels or more sophisticated data exchanges between specialist shore and buoy located devices.

In this talk we discuss how we combined data engineering, data analysis, AIS data and domain knowledge to track the utilisation of vessels globally to understand market trends, risk factors, ship valuations and cargo trends for use in both finance and insurance, with a particular focus on smaller vessel sizes (less than 10,000 dead-weight-tonnage).

In particular, we discuss techniques to overcome issues of dealing with messy and inconsistent data and how to leverage additional data sources to enhance the insights produced from the analysis.

Keywords: Marine, AIS, Cargo

References

1. International Telecommunication Union (2010) Technical characteristics for an automatic identification system using time-division multiple access in the VHF maritime mobile band

Contact details

- Email: mcooney@describedata.com
- Repository: https://github.com/kaybenleroll/
- Social media: https://www.linkedin.com/in/mick-cooney/

A scalable toolbox for exposing indirect discrimination in insurance rates

Olivier Côté, Université Laval

Marie-Pier Côté, Université Laval (presenter)

Arthur Charpentier, Université du Québec à Montréal

Abstract: Insurance pricing involves grouping policyholders by risk to set adequate premiums. To this end, insurers rely on policyholder characteristics as proxies for the unobserved risk. This raises fairness concerns when sensitive personal information, such as socioeconomic or demographic factors, may (intentionally or inadvertently) affect pricing. When sensitive attributes are specified and collected, group fairness metrics are standard in the assessment of disparate impact, indirect discrimination, and unfair bias (see, e.g., Charpentier, 2024). However, these metrics fail to deliver segment-specific insights for actuaries. We revisit fairness as resting on three fundamental pillars: actuarial fairness, solidarity, and causality. We then translate those pillars as five ratemaking benchmarks covering the whole spectrum of fairness, spanning a best-estimate premium to a corrective premium as in Côté et al. (2025). Building on the policyholder-specific metrics of Lindholm et al. (2025), we define actuarially relevant local metrics prior to pricing: the risk spread, the proxy vulnerability, the fairness range, and the parity cost. We also define metrics to assess a given commercial ratemaking scheme (post-pricing): implied commercial loading, commercial burden, implied propensity and excess lift. We propose a partition-based scalable method to detect systematic unfairnesses: pre-pricing identification of proxy-vulnerable individuals and post-pricing detection of commercial loadings. Using a large-scale Canadian auto insurance dataset, we showcase the tools practicality for protecting financially vulnerable individuals. Our analysis suggests that proxy vulnerability is skewed: most policyholders may receive small rebates, while a vulnerable minority may face 15-30% overpricing due to proxy discrimination based on financial status.

Keywords: Bias, Fairness, Proxy discrimination

References

1. Côté, O., Côté, M.-P., Charpentier, A. (2025). A fair price to pay: Exploiting causal graphs for fairness in insurance. *Journal of Risk and Insurance* **92**, 33-75.
2. Charpentier, A. (2024). *Insurance, biases, discrimination and fairness.* Springer.
3. Lindholm, M., Richman, R., Tsanakas, A., Wuthrich, M.V., Sensitivity-Based Measures of Discrimination in Insurance Pricing (2024). *Available at SSRN: 4897265*.

Contact details

- Email: marie-pier.cote@act.ulaval.ca
- Homepage: https://marie-pier-cote.fsg.ulaval.ca/
- Social media: https://www.linkedin.com/in/marie-pier-cote-631848314/

Causal knowledge graphs for risk interpretation using LMM: a new tool for insurers

Aurélien COULOUMY, Dylogy, University Lyon 1 (presenter)

Christelle ROVETTA, Dylogy (presenter)

Idir HAREB, Dylogy

Julien CRESPY, Dylogy

Charlie CHAIDRON, Dylogy

Abstract: The current approach to knowledge management in the insurance sector has notable limitations, primarily due to the restricted utilization of available data. Although advances in artificial intelligence (AI) enable improved processing of unstructured information, their potential remains underexploited. In practice, data predominantly remains tabular, considered immutable, correlated simplistically or merely temporally, or even treated as independent and often arbitrarily aggregated. These limitations directly impact risk estimation, which significantly relies on the richness and structure of the underlying knowledge.

This presentation aims to introduce various methods that combine programmatic and statistical approaches, particularly through the use of generative AI for the automated creation of directed acyclic graphs (DAG) [1]. These techniques partly overcome difficulties related to data formats, the types of links considered, multidimensionality, and aggregation methods [2].

Illustratively, we will exploit claims reports from multiple lines of insurance businesses, such as transportation, energy, and medical liability. First, we will present an agent-based architecture integrating various large language model (LLM) services capable of automatically reconstructing a standardized causal graph knowledge base. We will detail the performance obtained and the methods used to evaluate the effectiveness of these approaches, leveraging established benchmarks [3].

Secondly, we will illustrate through three practical case studies the use of these causal graphs: the first linked to actuarial sciences and the improvement of pricing techniques, the second oriented towards underwriting and proactive development of insurance offerings, and finally, the third focused on prevention related to natural disasters.

Finally, we will conclude by addressing the challenges associated with these approaches, particularly concerning costs, initial data models, limitations of the context within the employed LLMs, the complexity of abstraction, technical scalability, and the essential change management required within teams.

Keywords: LMM, DAG, Risk management, Prevention, Pricing, Underwriting

References

1. Long, Stephanie, Tibor Schuster, and Alexandre Piché. “Can large language models build causal graphs?.” arXiv preprint arXiv:2303.05279 (2023).
2. Gopalakrishnan S, Chen VZ, Dou W, Hahn-Powell G, Nedunuri S, Zadrozny W. “Text to Causal Knowledge Graph: A Framework to Synthesize Knowledge from Unstructured Business Texts into Causal Graphs.” Information. (2023).
3. Jin, Zhijing, et al. “Can large language models infer causation from correlation?.” arXiv preprint arXiv:2306.05836 (2023).

Contact details

- Email: acouloumy@dylogy.com
- Homepage: https://www.dylogy.com/
- Repository: -
- Social media: https://www.linkedin.com/in/aur%C3%A9lien-couloumy-5aa778a8/

Pricing Workers Compensation via Bayesian Hierarchical Modeling

Juan de Oyarbide, MSc, IAP, Addactis

Abstract: Bayesian models have been studied in actuarial science since the early twentieth century (e.g., Whitney, 1918), serving as a theoretical foundation for many traditional credibility formulas, such as those by Bühlmann and Straub (1970). However, the adoption of full Bayesian hierarchical models was historically constrained by limited computational resources and the lack of accessible software (Klugman, 1991). With the advent of advanced hardware and user-friendly software such as PyMC (Abril-Pla et al., 2023)—providing faster sampling algorithms and flexible model specification—there is now a renewed opportunity to employ Bayesian methods in complex actuarial applications. This work presents a Bayesian hierarchical model applied to pricing workers compensation contracts, illustrating how to effectively represent heavy-tailed losses (Oyarbide & Paulo, 2020), borrow information among risk classes, and assess variation through posterior and predictive distributions.

Keywords: Credibility, Bayesian Inference, Hierarchical Modeling, Non-Life, Pricing, Risk-Analysis.

References

1. Whitney, A. W. (1918). The Theory of Experience Rating. *Proceedings of the Casualty Actuarial Society*, 4, 274–292.
2. Bühlmann, H., & Straub, E. (1970). Credibility procedures for insurance claims. *Bulletin of the Swiss Association of Actuaries*, 70(1), 111–133.
3. Klugman, S. (1991). *Bayesian Statistics in Actuarial Science: With Emphasis on Credibility.* Springer Netherlands.
4. Oyarbide, J., & Paulo, R. (2020). Collective risk assessment in affordable care act markets: A Bayesian hierarchical model. *SSRN*.
5. Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fonnesbeck, C. J., Kochurov, M., Kumar, R., Lao, J., Luhmann, C. C., Martin, O. A., et al. (2023). PyMC: A modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science*, 9, e1516.

Contact details

- Email: juan.deoyarbide@addactis.com

Distributional Regression for Actuarial Applications: Distributional Refinement Network

Benjamin Avanzi, University of Melbourne

Tian Dong, UNSW Sydney (presenter)

Patrick Laub, UNSW Sydney

Bernard Wong, UNSW Sydney

Abstract: Actuarial modeling often requires the ability to model the full distribution of losses, not just point estimates.
This talk introduces the Distributional Refinement Network (DRN; Avanzi et al., 2024) and covers two key aspects:
(i) the motivations behind DRN and its theoretical foundations, and
(ii) practical coding examples for actuarial applications.

The first part highlights the advantages of DRN for distributional regression, which combines an inherently interpretable baseline model (such as GLMs) with a flexible neural network – a modified Deep Distribution Regression (DDR; Li et al., 2021) method. Inspired by the Combined Actuarial Neural Network (CANN; Schelldorfer and Wuthrich, 2019), our approach flexibly refines the entire baseline distribution.

The second part demonstrates the use of the 'drn' package to train a DRN model, showcasing its versatility in solving a range of actuarial regression problems.

Keywords: Distributional Regression, Neural Networks, Actuarial Modeling, Uncertainty Quantification

References

1. Avanzi, B., Dong, T., Laub, P., Wong, B. (2024). *Distributional Refinement Network for Actuarial Applications.*

2. Li, R., Bondell, H.D., Reich, B.J. (2021). *Deep Distribution Regression. Journal of Machine Learning Research*, **22**(1), 1-33.

3. Schelldorfer, J., Wuthrich, M.V. (2019). *Combined Actuarial Neural Networks (CANN) for Risk Classification. ASTIN Bulletin*, **49**(2), 413-436.

Contact details

- Email: tian.dong@unsw.edu.au
- Repository: https://github.com/EricTianDong/drn
- Social media: LinkedIn

R2VF: A Two-Step Regularization Algorithm to Cluster Categories in GLMs

Yuval Ben Dror, Data Science Researcher, Earnix Ltd.

Abstract: Over recent decades, extensive research has focused on overcoming the restrictive assumptions of Generalized Linear Models (GLMs) to improve predictive accuracy and interpretability. Key developments include coefficient regularization (1), feature selection, and ordinal category clustering (2). However, efficiently grouping nominal categories in GLMs without excessive computational costs (3) remains a challenge. This presentation introduces Ranking to Variable Fusion (R2VF), a novel two-step method for fusing nominal and ordinal categories in GLMs. R2VF first transforms nominal features into an ordinal representation using regularized regression, then applies variable fusion to optimize category grouping. This approach balances model complexity and interpretability while mitigating overfitting. We compare R2VF against existing methods and demonstrate its effectiveness in handling high-cardinality categorical variables while maintaining model simplicity and explainability. This is particularly relevant for insurance pricing models, where categorical variables—such as policyholder attributes and vehicle classifications—are prevalent. The need for interpretable, computationally efficient models makes R2VF a strong candidate for actuarial applications. Recognizing this, R2VF has been implemented as part of Earnix's Auto-GLM (4) feature. A link to our paper: https://arxiv.org/abs/2503.01521

Keywords: GLM, Nominal Variables, Variable Fusion, Regularized Regression

References

1. Tibshirani, R. *Regression Shrinkage and Selection via the Lasso*. Journal of the Royal Statistical Society: Series B, 58(1), 267–288, 1996.
2. Land, S., & Friedman, J. *Variable Fusion: A New Method of Adaptive Signal Regression*. 1996.
3. Bondell, H. D., & Reich, B. J. *Simultaneous Factor Selection and Collapsing Levels in ANOVA*. Biometrics, 65, 169–177, 2009.
4. Fujita, S., et al. *AGLM: A Hybrid Modeling Method of GLM and Data Science Techniques*. 2020.

Contact details

- Email: yuval.bendror@earnix.com
- Social media: https://www.linkedin.com/in/yuval-ben-dror-881939109/

From SHAP to EBM: How to Explain Gradient Boosting Models

Emanuele Fabbiani, xtream & Catholic University of Milan

Abstract: The insurance industry relies heavily on data to make critical decisions, from assessing risk and detecting fraud to segmenting customers and predicting claims. **XGBoost** is considered a state-of-the-art model for regression and classification and is often the go-to solution for such tasks.

However, its complex inner workings often resemble a "black box," making it difficult to understand how it generates predictions. This lack of transparency is a **significant challenge** in the highly regulated insurance sector, where explainability is essential for compliance, trust, and ethical decision-making.

This presentation explores the world of Explainable AI (XAI) and highlights two powerful techniques: **SHapley Additive exPlanations** (SHAP) [1] and **Explainable Boosting Machine** (EBM) [2]. We'll explain the theory behind these methods and discuss how they unravel the decision-making processes of XGBoost models. Through real-world examples, we'll demonstrate how SHAP and EBM can reveal the contribution of individual features to a prediction, enabling insurers to:

- justify underwriting decisions,
- enhance fraud detection,
- improve customer segmentation, and
- meet regulatory requirements.

Attendees will gain **practical knowledge** of SHAP and EBM, including their strengths and limitations in insurance applications. They will also receive hands-on experience using Python libraries such as `shap` and `interpret-ml`.

Keywords: XGBoost, Explainability, SHAP

References

1. Scott Lundberg. *A unified approach to interpreting model predictions.* arXiv preprint arXiv:1705.07874. 2017.
2. Yin Lou, Rich Caruana, Johannes Gehrke, and Giles Hooker. *Accurate intelligible models with pairwise interactions*. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 623–631. 2013.

Contact details

- Email: emanuele.fabbiani@xtreamers.io
- Repository: https://github.com/donlelef
- Social media: https://www.linkedin.com/in/emanuelefabbiani/

Additive tree latent variable models with applications to insurance loss prediction

Zhihao Wang, Xinjiang University of Finance and Economics, Urumqi 830012, China.

Yanlin Shi, Macquarie University, Sydney, NSW 2109, Australia.

Guangyuan Gao, Renmin University of China, Beijing 100872, China. (presenter)

Abstract: We consider a specific class of regression models with discrete latent variables including mixture model, zero-inflated count model and Tweedie's compound Poisson model, which are commonly used in actuarial science and other fields. When fitting these parametric regression models, regression functions are estimated for both the observed response variable and the latent variable, respectively. Feature engineering, variable selection and model selection become challenging due to the involvement of multiple regression functions and latent variable. To address these challenges, we propose additive tree latent variable models. To calibrate these models, we introduce an iteratively re-weighted gradient boosting (IRGB) algorithm that combines the EM algorithm with the gradient boosting. In each iteration, the IRGB algorithm trains only one weak learner in a stagewise manner. Theoretical analysis demonstrates the monotonic behavior of the likelihood in the IRGB algorithm. We further illustrate the advantages of the proposed nonparametric methods through an empirical example of motor insurance claim counts and a case study on French motor third-party liability insurance pure premiums.

Keywords: Nonparametric regression; Additive models; Latent variable; Gradient boosting; Insurance loss prediction

References

1. Delong, Ł., Lindholm, M., Wüthrich, M. V. (2021). Making Tweedie's compound Poisson model more accessible. *European Actuarial Journal* **11/1**, 185-226.
2. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *The Annals of statistics* **29**, 1189-1232.
3. Gao, G. (2024). Fitting Tweedie's compound poisson model to pure premium with the EM algorithm. *Insurance: Mathematics and Economics* **114**, 29-42.
4. Lange, K. (1995). A gradient algorithm locally equivalent to the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* **57/2**, 425-437.
5. Yip, K. C., Yau, K. K. (2005). On modeling claim frequency data in general insurance with extra zeros. *Insurance: Mathematics and Economics* **36/2**, 153-163.

Contact details

Catastrophic-risk-aware reinforcement learning with extreme-value-theory-based policy gradients

Parisa Davar, Concordia University and Deloitte, Montreal, Canada

Frédéric Godin, Concordia University and Quantact Laboratory, CRM, Montreal, Canada

José Garrido, Concordia University, Montreal, Canada and Digital Insurance And Long term risk (DIALog) research chair, France (presenter)

Abstract: This paper tackles the problem of mitigating catastrophic risk (which is risk with very low frequency but very high severity) in the context of a sequential decision making process. This problem is particularly challenging due to the scarcity of observations in the far tail of the distribution of cumulative costs (negative rewards). A policy gradient algorithm is developed, that we call POTPG. It is based on approximations of the tail risk derived from extreme value theory. Numerical experiments highlight the out-performance of our method over common benchmarks, relying on the empirical distribution. An application to financial risk management, more precisely to the dynamic hedging of a financial option, is presented.

Keywords: Risk-aware reinforcement learning, catastrophic risk, extreme value theory, peaks-over-threshold (POT), hedging

References

1. Bader, B., Yan, J., Zhang, X. (2018). Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate. *The Annals of Applied Statistics* **12(1)**, 310–329.
2. Carbonneau, A., Godin, F. (2021). Equal risk pricing of derivatives with deep hedging. *Quantitative Finance* **21(4)**, 593–608.
3. Godin, F., Mayoral, S., Morales, M. (2012). Contingent claim pricing using a normal inverse Gaussian probability distortion operator. *Journal of Risk and Insurance* **79(3)**, 841–866.
4. Godin, F. (2016). Minimizing CVaR in global dynamic hedging with transaction costs. *Quantitative Finance* **16(3)**, 461–475.
5. Troop, D., Godin, F., Yu, J.Y. (2022). Best-arm identification using extreme value theory estimates of the CVaR. *Journal of Risk and Financial Management* **15(4)** 172.

Contact details

- Email: jose.garrido@concordia.ca
- Homepage: https://www.concordia.ca/faculty/jose-garrido.html
- Repository: https://www.researchgate.net/profile/Jose-Garrido-16

Ensembling GLM and XGB

Karol Gawlowski, Manager, EY

Patricia Wang, Director, Convex

Abstract: Bridging the gap between models' predictive power and interpretability is one of the key problems in modern predictive analytics specifically in insurance. Despite the availability of more performant ML tree based models less predictive GLMs are still a go-to method due to their explainable nature. We propose a multiplicative architecture of ensembling GLMs and GBMs, which leverages GBMs superior performance for residual modelling, while retaining the familiar linear formulaic representation of the resulting model.

Our proposition leverages the state of the art interpretability technique - SHAP, to arrive at a per-observation set of GLM coeffecient corrections. These corrections help modellers to understand how the ML model deviates from the underlying GLM while inproving its performance.

Keywords: SHAP, GLM, XGB, Predictive Modelling, XAI

References

1. Tianqi Chen, Carlos Guestrin (2016) XGBoost: A Scalable Tree Boosting System. *22nd SIGKDD Conference on Knowledge Discovery and Data Mining*
2. Scott M Lundberg, Su-In Lee (2017) A Unified Approach to Interpreting Model Predictions *Advances in Neural Information Processing Systems 30*
3. Noll, Alexander and Salzmann, Robert and Wuthrich, Mario V. (2020) Case Study: French Motor Third-Party Liability Claims *SSRN*

Contact details

- Email: karol.gawlowski1@uk.ey.com, Karol.Gawlowski@bayes.city.ac.uk
- Repository: github.com/Karol-Gawlowski/
- Social media: linkedin.com/in/karol-gawlowski/

Winning Strategies: Predict relative performance

Markus Gesmann, Insurance Capital Markets Research

Abstract: Competition is the nature of business and sport. Employing data and artificial intelligence (AI) to analyse competitors is increasingly crucial for maintaining a competitive edge.

While absolute metrics are indispensable for monitoring individual progress, relative performance ultimately determines success.

Furthermore, sustained relative outperformance drives higher valuations, manifested as price-to-book or price-to-earnings multiples in the corporate sector, or enhanced remuneration for athletes.

Insurance Capital Markets Research (ICMR) has observed strikingly similar patterns within the (re)insurance industry, particularly within the Lloyd's of London market. Notably, the relative performance of underwriting teams, or entire syndicates, when benchmarked against competitors, exhibits a significantly greater degree of consistency than their absolute loss ratio performance.

This presentation will outline a predictive model, developed by ICMR, designed to forecast future relative performance within the Lloyd's marketplace. The model leverages historical rank performance to predict future outcomes, drawing upon methodologies presented in [1, 2].

Keywords: Performance monitoring, Portfolio optimisation, Rank modelling, Plackett-Luce

References

1. Luce, R. Duncan. 1959. Individual Choice Behavior: A Theoretical Analysis. New York: Wiley.
2. Plackett, Robert L. 1975. "The Analysis of Permutations." Appl. Statist 24 (2): 193–202. https://doi.org/10.2307/2346567.

Contact details

- Email: markus.gesmann@insurancecapitalmarkets.com
- Homepage: https://insurancecapitalmarkets.com

Harnessing Conditional Generative Models for Synthetic Non-Life Insurance Premium Data

Claudio Giorgio Giancaterino, Towards Innovation Lab (presenter)

Abstract: This study is oriented to a synthetic non-life insurance premium dataset generated using several Conditional Generative Models. Unlike unconditional models, which create data freely without guidance, conditional models are directed by input conditions to produce targeted outputs. As a benchmark, has been employed a Conditional Gaussian Mixture Model. The validation of generated data followed several steps: data visualization comparison with univariate analysis, PCA and UMAP representations between trained data and samples generated. To evaluate the goodness of data generated has been employed in addition to these visual methods, consistecy check of data produced, and the statistical Kolmogorov–Smirnov test. Followed by predictive modelling of frequency and severity with Generalized Linear Models (GLMs) exploited by Tweedie distribution as a measure of the generated data's quality going ahead with the evidence of features importance. For further comparison, advanced Deep Learning architectures have been employed: Conditional Variational Autoencoders (CVAEs), CVAEs enhanced with a Transformer Decoder, and a Conditional Diffusion Model. The analysis assesses each model's ability to capture the underlying distributions, preserve complex dependencies, and maintain conditional relationships intrinsic to the premium data. These findings provide insightful directions for enhancing synthetic data generation in insurance, with potential applications in risk modelling, pricing strategies with data scarcity, and regulatory compliance.

Keywords: Conditional Variational Autoencoders, Conditional Gaussian Mixture Model, Conditional Diffusion Model, Conditional Variational Autoencoders with a Transformer Decoder, PCA, UMAP, GLMs

References

1. Ian Goodfellow and Yoshua Bengio and Aaron Courville, 2016, *Deep Learning*, MIT Press.
2. Mario V. Wuthrich, Ronald Richman, Benjamin Avanzi, Mathias Lindholm, Michael Mayer, Jürg Schelldorfer, Salvatore Scognamiglio, 2025, *AI Tools for Actuaries*, SSRN.
3. David Foster, 2023, *Generative Deep Learning*, 2nd Edition, O'Reilly.
4. Christopher Bishop, 2006, *Pattern Recognition and Machine Learning*, Springer.
5. Jamotton, Charlotte; Hainaut, Donatien, 2023, *Variational autoencoder for synthetic insurance data*, ISBA

Contact details

- Email: c.giancaterino@gmail.com
- Homepage: https://towardsinnovationlab.com/
- Repository: https://github.com/claudio1975/Generative_Modelling
- Social media: https://www.linkedin.com/in/claudiods/

Market-based insurance ratemaking: application to pet insurance.

Pierre-Olivier Goffard, Université de Strasbourg, Institut de Recherche Mathématique Avancée, Strasbourg, France (presenter)

Pierrick Piette, Univ Lyon, Université Claude Bernard Lyon 1, Institut de Science Financière et d'Assurances (ISFA), Laboratoire SAF EA2429, F-69366, Lyon, France

Gareth W. Peters, University of California Santa Barbara, Department of Statistics and Applied Probability, Santa Barbara CA 93106-3110, USA

Abstract: This talk introduces a method for pricing insurance policies using market data. The approach is designed for scenarios in which the insurance company seeks to enter a new market, in our case: pet insurance, lacking historical data. The methodology involves an iterative two-step process. First, a suitable parameter is proposed to characterize the underlying risk. Second, the resulting pure premium is linked to the observed commercial premium using an isotonic regression model. To validate the method, comprehensive testing is conducted on synthetic data, followed by its application to a dataset of actual pet insurance rates. To facilitate practical implementation, we have developed an R package called IsoPriceR. By addressing the challenge of pricing insurance policies in the absence of historical data, this method contributes to enhancing pricing strategies in emerging markets.

Keywords: Insurance Pricing, Bayesian Inference, Approximate Bayesian Computation, Isotonic Regression.

References

1. Goffard, P.-O., Piette, P., & Peters, G. W. (2023). *Market-based insurance ratemaking*. Working paper or preprint. https://hal.science/hal-04297811.
2. Sisson, S. A., Fan, Y., & Beaumont, M. (2018). *Handbook of Approximate Bayesian Computation*. Chapman and Hall/CRC.
3. Barlow, R. E., Brunk, H. D., Bartholomew, D. J., & Bremner, J. M. (1972). *Statistical inference under order restrictions (The theory and application of isotonic regression)*.

Contact details

- Email: goffard@unistra.fr
- Homepage: https://pierre-olivier.goffard.me/
- Repository: market_based_insurance_ratemaking
- Social media: LinkedIn

Automating tail factor detection points using Bayesian hidden Markov models and latent change-point models.

Conor Goold, Senior Data Scientist, Ledger Investing, Inc. (presenter)

Abstract: Typical loss development workflows fit loss and tail development models separately, using actuarial judgement and rules-of-thumb to choose tail development starting points, training data windows (i.e. experience periods and development lags) and tail factors. In this talk, I will share work on automating the loss and tail development fitting process using Bayesian hidden Markov models and Bayesian latent change point models. I will compare results to a more traditional two-step process on industry data using out-of-sample performance metrics and the calibration of forecasted ultimate loss ratio distributions. Advantages and disadvantages to these automated processes will be discussed in the context of the actuarial workflow and insurance data science more broadly, including computational and inferential considerations.

Keywords: loss development factors, mixture modelling, loss reserving

References

1. Goold, C. (preprint). Joint estimation of insurance loss development factors using Bayesian hidden Markov models. https://arxiv.org/abs/2406.19903

Contact details

- Email: conor@ledgerinvesting.com

Explainable Boosting Machine for predicting claim severity and frequency in car insurance

Markéta Krúpová, ADDACTIS France (presenter)

46 bis Chemin du Vieux Moulin, 69160 Tassin-la-Demi-Lune, France

Nabil Rachdi, ADDACTIS France

46 bis Chemin du Vieux Moulin, 69160 Tassin-la-Demi-Lune, France

Quentin Guibert, Université Paris-Dauphine (presenter)

CEREMADE, Université Paris-Dauphine, PSL University, CNRS, 75016 Paris, France

Abstract: In a context of constant increase in competition and heightened regulatory pressure, accuracy, actuarial precision, as well as transparency and understanding of the tariff, are key issues in non-life insurance. Traditionally used generalized linear models (GLM) result in a multiplicative tariff that favors interpretability. With the rapid development of machine learning and deep learning techniques, actuaries and the rest of the insurance industry have adopted these techniques widely. However, there is a need to associate them with interpretability techniques. In this paper, our study focuses on introducing an Explainable Boosting Machine (EBM) model that combines intrinsically interpretable characteristics and high prediction performance. This approach is described as a glass-box model and relies on the use of a Generalized Additive Model (GAM) and a cyclic gradient boosting algorithm. It accounts for univariate and pairwise interaction effects between features and provides naturally explanations on them. We implement this approach on car insurance frequency and severity data and extensively compare the performance of this approach with classical competitors: a GLM, a GAM, a CART model and an Extreme Gradient Boosting (XGB) algorithm. Finally, we examine the interpretability of these models to capture the main determinants of claim costs.

Keywords: car insurance, pricing, explainable boosting machine, generalized additive model, cyclic gradient boosting, interpretable machine learning, glass-box.

References

Lou, Y., Caruana, R., and Gehrke, J. (2012). Intelligible models for classification and re- gression. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. Beijing China: ACM, pp. 150–158. doi: 10.1145/2339530. 2339556.

Nori, H., Jenkins, S., Koch, P., and Caruana, R. (2019). InterpretML: A Unified Framework for Machine Learning Interpretability. doi: 10.48550/arXiv.1909.09223.

Wood, S. N., Goude, Y., and Fasiolo, M. (2022). Interpretability in Generalized Additive Models. Interpretability for Industry 4.0 : Statistical and Machine Learning Approaches. Springer, Cham, pp. 85–123. doi: 10.1007/978-3-031-12402-0_4.

Zakrisson, H. and Lindholm, M. (2025). A tree-based varying coefficient model. Computa- tional Statistics, pp. 1–30. doi: 10.1007/s00180-025-01603-8.

Contact details

- Email: guibert@ceremade.dauphine.fr
- Homepage: https://www.ceremade.dauphine.fr/~guibert/

- Repository: https://github.com/qguibert
- Social media: https://www.linkedin.com/in/quentin-guibert-74ab4033

Metaheuristic-Informed Machine Learning for Optimizing Strike Temperatures in Weather Index insurance

Samuel Asante Gyamerah, Toronto Metropolitan University, Toronto, Canada

Abstract: For potential buyers of weather index insurance, such as producers, the question of whether a weather index insurance is an efficient hedging instrument is just as important as the modelling and pricing of its underlying [1]. This has triggered several attempts to design effective and efficient modelling and pricing techniques for weather derivatives. However, "geographical basis risk," which is defined as the mismatch of temperature between the local farming location and the reference weather station, continues to undermine the efficiency of pricing models and the uptake of weather derivatives as a risk-hedging instrument. This poses a real economic risk to producers such as farmers and stakeholders in the agricultural sector. Minimizing the effect of basis risk is not just a technical exercise in improving spatial pricing models but also an important step toward building confidence and increasing the adoption of weather derivatives as practical, farmer-friendly weather risk hedging instruments. The project's goal is to develop an optimization framework that uses metaheuristic-informed machine learning techniques to determine optimal strike temperatures for weather index insurance for multiple weather stations. Specifically, the study aims to 1) reduce geographical basis risk by aligning the payoff of the local weather station with the payoff of the reference weather station; 2) penalize strike temperatures that do not account for climate extremes; and 3) ensure that stations with correlated weather patterns have similar strike temperatures, ensuring geographic consistency.

Keywords: risk management, climate change, agriculture insurance, metaheuristics

References

1. Ender, M., & Zhang, R. (2015). Efficiency of weather derivatives for Chinese agriculture industry. China Agricultural Economic Review, 7(1), 102-121.

Contact details

- Email: asante.gyamerah@torontomu.ca
- Homepage: https://www.torontomu.ca/math/our-people/samuel-asante-gyamerah/
- Social media: www.linkedin.com/in/samuelasantegyamerah

NLP For Data Granularity Improvement

Iftekhar Khaled, Data Scientist (Presenter)

Chris Halliwell, Head of Predictive Analytics (Backup Presenter)

Abstract: Insurance companies regularly collect free-text descriptions of claims—through emails, call logs, or digital forms—yet this valuable information is often stored in siloed systems and remains underutilized for risk segmentation and underwriting. We address this gap by developing a semantic classification pipeline to transform unstructured claim notes into structured, granular insights. Our approach begins with an ontology of main categories (e.g., Fire, Water) and subcategories (e.g., Wildfire, Pipe Leak, Storm Damage), which is then enriched with domain-specific embeddings derived from the Mixedbread.ai model [1]. Next, we employ a double-ranked voting algorithm that uses cosine similarity to map each claim note to its most probable main–subcategory pair. Compared to traditional keyword-based solutions, our method captures nuanced language contexts—such as "frozen pipes inside water heater"—leading to more accurate and interpretable classification. We demonstrate this pipeline on real-world property claims data, validating our results against a hand-labeled benchmark and illustrating how the system can approach or even exceed human-level assignments. This approach aligns with recent advancements in applying cosine similarity measures for text classification tasks [2]. By converting large volumes of text into a structured database of actionable features, insurers can refine risk-based pricing and implement targeted loss control measures. This presentation details our step-by-step methodology, including best practices for implementing contextual embeddings, the ranking algorithm, and the integration of these outputs into underwriting workflows. Our methodology is informed by comprehensive studies on predictive analytics in insurance claims [3] and the application of natural language processing in actuarial contexts [4]. References

Keywords: NLP, Claims, Vector Embedding, Mixedbread, Data Quality

References

1. Mixedbread AI. "Embedding Models." Available: https://www.mixedbread.ai/docs/embeddings/models
2. S. Sohangir and D. Wang, "Improved sqrt-cosine similarity measurement," Journal of Big Data, vol. 4, no. 1, p. 25, 2017. Available: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-017-0083-6
3. Milliman, "5 things to know about large language models in claims management," 2023. Available: https://www.milliman.com/en/insight/5-things-large-language-models-claims-management
4. A. Richman and J. Wüthrich, "Actuarial Applications of Natural Language Processing Using Transformers," 2022. Available: https://arxiv.org/pdf/2206.02014

Contact details

- Email: chris.halliwell@markel.com
- Phone: 07595924871

Multi-view spatial embeddings for insurance portfolio analytics

Freek Holvoet, KU Leuven (presenter)

Christopher Blier-Wong, University of Toronto

Katrien Antonio, KU Leuven & University of Amsterdam

Abstract: Accurate assessment of spatial risk factors, particularly those influenced by climate, weather, and demographic factors, is crucial for the insurance industry to improve underwriting precision and enhance risk management. However, spatial data is often unstructured or high-dimensional. Embedding methods are needed to convert spatial data into meaningful representations for modelling tasks. This paper introduces a novel approach for constructing spatial embeddings using contrastive learning in a multi-view setting. We combine two types of spatial data, satellite images and OpenStreetMap tags, to create a multi-view embedding. This embedding captures both visible geospatial features and detailed neighbourhood context. By linking the multi-view embedding to coordinates using contrastive learning, we can efficiently extract spatial embeddings based solely on latitude and longitude. This simplifies spatial data integration in predictive models, eliminating the need to repeatedly gather large amounts of spatial data whenever spatial context is required. In a case study on French real estate prices, we show that our spatial embeddings effectively capture geographic structures. This leads to improved predictive accuracy and provides valuable spatial insights. Our results show that our multi-view spatial embedding model allows us to incorporate spatial context in predictive tasks in an easy and flexible way.

Keywords: spatial data, embedding, machine learning, geospatial representations

References

1. Blier-Wong, C., Baillargeon, J.-T., Cossette, H., Lamontagne, L., & Marceau, É. (2020). Encoding neighbor information into geographical embeddings using convolutional neural networks. *The Florida AI Research Society*, 15–20.
2. Blier-Wong, C., Lamontagne, L., & Marceau, E. (2024). A representation-learning approach for insurance pricing with images. *ASTIN Bulletin: The Journal of the IAA*, 1–30.
3. Klemmer, K., Rolf, E., Robinson, C., Mackey, L., Rußwurm, M. (2023). SatCLIP: Global, general-purpose location embeddings with satellite imagery. *In arXiv [cs.CV]*. arXiv. http://arxiv.org/abs/2311.17179

Contact details

- Email: freek.holvoet@kuleuven.be

Identifying Similar Insurance Claims using Text-Based Vector Search

Daniel Jakobi, Markel Insurance SE, FOM Hochschule für Oekonomie & Management (presenter)

Davide De March, Markel International

Abstract: In a softening market, reducing operational expense and anticipate claims' costs is crucial to maintain profitability, particularly since increasing premium earnings is challenging due to limited or negative rate, stricter underwriting standards and reduced capacity.

The majority of costs stem from claims payouts and processing expenses and Markel has investigated AI-driven solutions to identify insurance claims similarity based on text data received from brokers, policyholders, and claimants.

To address this challenge, we analysed a large dataset of over 600,000 German email texts and attached documents. Our technical approach leveraged PySpark, Mosaic AI Vector Search, OpenAI Embeddings, and the Databricks analytics platform.

We utilized Reciprocal Rank Fusion to rank retrieved text pieces based on scores from Mosaic AI Vector Search and the frequency of text pieces related to specific claims. The use of multilingual large language models (LLMs) to generate embeddings enables easy expansion of the vector search to claims in other languages. Moreover, the concept of applying vector search and text embeddings to retrieve similar results can be readily adapted to other domains, such as healthcare.

Our results demonstrate the effectiveness of identifying similar claims based solely on text data. By filtering the results based on claim similarity, significantly improved the claims handling process, by collecting structured data from unstructured forms, and provide claim handler with similar case to assess the claim cost.

Keywords: vector search, claims, similarity, insurance, automation, llm

References

1. Jasper Xian, Tommaso Teofili, Ronak Pradeep, and Jimmy Lin. (2024). Vector Search with OpenAI Embeddings: Lucene Is All You Need. *In Proceedings of the 17th ACM International Conference on Web Search and Data Mining (WSDM '24). Association for Computing Machinery, New York, NY, USA*, 1090–1093.
2. Jan Rygl, Jan Pomikálek, Radim Řehůřek, Michal Růžička, Vít Novotný, Petr Sojka. (2017). Semantic Vector Encoding and Similarity Search Using Fulltext Search Engines. *Annual Meeting of the Association for Computational Linguistics 2017, Vancouver, Canada*.

Contact details

- Email: daniel.jakobi@markel.de, davide.demarch@markel.com

A Multivariate Energy Distance Approach to Premium Fairness Adjustment

Charlotte Jamotton, Université catholique de Louvain (presenter)

Donatien Hainaut, Université catholique de Louvain

Abstract: Fairness in insurance has become increasingly critical, particularly in light of regulations calling for non-discriminatory premium estimation, such as the EU Gender Directive (2012), which prohibits gender-based pricing in insurance. This research explores group fairness mechanisms designed to mitigate group-level disparities. We introduce a group fairness adjuster based on the energy distance. The multivariate nature of the energy distance allows for the simultaneous fairness adjustment of premium distributions across multiple sensitive demographic groups from different (non-binary) sensitive variables, such as gender and age. Our method incorporates an autocalibration technique to correct biases in the total number of claims after fairness adjustments, ensuring overall balance in the predictions. Additionally, our approach allows for the adjustment of estimated discriminatory premiums for new policyholders without requiring retraining of the fairness adjustment model. We demonstrate the effectiveness (i.e., significant reductions in group-level disparities while maintaining near-initial model accuracy) of our adjustments in the context of car insurance, known to rely on demographic factors such as age and gender to assess risk and determine premiums.

Keywords: fairness, demographic parity, autocalibration, actuarial pricing, energy distance

References

1. Côté, O., Côté, M.-P., & Charpentier, A. (2024). A fair price to pay: Exploiting causal graphs for fairness in insurance. Journal of Risk and Insurance.
2. Charpentier, A. (2023). Insurance, biases, discrimination and fairness.
3. Frees, E. W., & Huang, F. (2023). The discriminating (pricing) actuary. North American Actuarial Journal, 27(1), 2–24.

Contact details

- Email: charlotte.jamotton@uclouvain.be

Track Trending Topics in Insurance for Emerging Risk Identification

Amin Karbassi, Ph.D., Senior Risk Engineer, AXA XL (presenter)

Abstract: Understanding risk trends is crucial for identifying future exposures in the insurance sector, particularly in the context of emerging risks. Emerging risks can often appear in scientific research articles as topics of interest before they become recognized risks for the public and consequently for the insurance sector. While traditional clustering strategies can categorize articles into topics, these methods often overlook the evolution of such topics over specific timescales. Moreover, even when articles are grouped, the results can be extensive and challenging to navigate. To address these challenges, this presentation explores the application of a time-biased trend detection techniques to analyze peer-reviewed articles published in ScienceDirect over 25 years from 20005 to 2025 related to emerging contaminants topic.

For data normalization, hyperparameter tuning was performed to determine the optimal number of components for Latent Dirichlet Allocation (LDA) and Singular Value Decomposition (SVD). A term frequency vector was created using CountVectorizer, features were extracted with LDA, and the feature matrix was transformed into a normalized Term Frequency-Inverse Document Frequency (TF-IDF) representation. Additionally, the number of features was reduced using SVD, and the resulting Numpy array was normalized. A term matrix was built for the entire corpus to facilitate further analysis.

By employing a time-biased document clustering method, insights into well-defined trending topics within the field are provided. The methodology outlined is aimed at enabling practitioners to identify and quantify trends effectively. Various levels of temporal bias can be experimented with by users, thereby detecting both short- and long-timescale trends. This approach allows experts to delve into the identified topics, gaining a deeper understanding through relevant keywords and key articles. This work aims to assist professionals in the insurance industry in navigating the vast literature on any topic of interest (e.g., PFAS, Wildfire), to identify critical concepts and factors essential to managing emerging risks.

Keywords: Topic Modeling, Trend Detection, Latent Dirichlet Allocation, Emerging Risks

References

1. Behpour, S., Mohammadi, M., Albert, M. V., Alam, Z. S., Wang, L., & Xiao, T. (2021). Automatic trend detection: Time-biased document clustering. *Journal of Knowledge-Based Systems*, Vol. 220, 106907
2. Blei, David M., Ng, Andrew Y. and Jordan, Michael I. (2003). Latent dirichlet allocation. *Journal of Machine Learning* Res. 3, 993–1022.

Contact details

- Email: amin.karbassi@axaxl.com
- Repository: https://github.com/aminkarbassi
- Social media: https://www.linkedin.com/in/aminkarbassi

Sensitivity-based measures of discrimination in insurance pricing

Mathias Lindholm, Stockholm University

Abstract: Different notions of fairness and discrimination have been extensively discussed in the machine learning, operations research, and insurance pricing literatures. As not all fairness criteria can be concurrently satisfied, metrics are needed that allow assessing the materiality of discriminatory effects and the trade-offs between various criteria. Methods from sensitivity analysis have been deployed for the measurement of demographic unfairness, that is, the statistical dependence of risk predictions on protected attributes. We produce a sensitivity-based measure for the distinct phenomenon of proxy discrimination, referring to the implicit inference of protected attributes from other covariates. For this, we first define a set of admissible prices that avoid proxy discrimination. Then, the measure is defined as the normalised L2-distance of a price from the closest element in that set. We use arguments from variance-based sensitivity analysis, to attribute the proxy discrimination measure to individual (or subsets of) covariates and investigate how properties of the data generating process are reflected in those metrics. Furthermore, we build on the global (i.e., portfolio-wide) measures of demographic unfairness and proxy discrimination to propose local (i.e., instance- or policyholder-specific) measures, which allow a fine-grained understanding of discriminatory effects. Finally, we apply the methods developed in the paper to a real-world insurance dataset, where ethnicity is a protected variable. We observe substantial proxy-discriminatory effects for one ethnic group and identify the key variables driving this.

Keywords: Sensitivity analysis, proxy discrimination, demographic parity, insurance pricing, algorithmic fairness

References

1. Lindholm, M., Richman, R., Tsanakas, A., Wüthrich, M.V. (2024) Sensitivity-Based Measures of Discrimination in Insurance Pricing. Available at *SSRN*: https://ssrn.com/abstract=4897265

Contact details

- Email: lindholm@math.su.se
- Homepage: https://staff.math.su.se/lindholm/

Claims processing and costs under capacity constraints

Filip Lindskog, Stockholm University (presenter)

Mario Wüthrich, ETH Zürich

Abstract: Random delays between the occurrence of accident events and the corresponding reporting times of insurance claims is a standard feature of insurance data. The time lag between the reporting and the processing of a claim depends on whether the claim can be processed without delay as it arrives or whether it remains unprocessed for some time because of temporarily insufficient processing capacity that is shared between all incoming claims. We aim to explain and analyze the nature of processing delays and build-up of backlogs. We illustrate consequences for prediction and cost optimization, taking both delay-adjusted costs and fixed costs for claims settlement capacity into account.

Keywords: Claims processing, backlog, capacity constraints

References

1. Lindskog, F., Wüthrich, M. (2024). Claims processing and costs under capacity constraints Preprint. arXiv:2409.09091

Contact details

- Email: lindskog@math.su.se
- Homepage: https://staff.math.su.se/lindskog/

Meta-modelling paths of simple climate models using Neural Networks and Dirichlet polynomials: an application to DICE

Emmanuel GOBET, Centre de Mathématiques Appliquées (CMAP), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris,

Yushan LIU, Centre de Mathématiques Appliquées (CMAP), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris (presenter)

Gauthier Vermandel, Centre de Mathématiques Appliquées (CMAP), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris

Abstract: Our study focuses on climate models extensively employed in climate science and economic-climate research, which project temperature outcomes from carbon emission trajectories. Addressing the need for rapid evaluation in Integrated Assessment Models – critical tools for carbon emission mitigation policy analysis – we design a neural network (NN) meta-model as an efficient surrogate for mapping, in an infinite horizon setting, emission trajectories into temperature trajectories (usually modeled as coupled systems of differential equations). Our approach combines a projection on Generalized Dirichlet polynomials, whose coefficients are inputs of the NN and a suitable time change for handling infinite horizon: we prove that the quantity of interest is, under some assumptions, a smooth function of the inputs and therefore, is prone to accurate NN approximation. After a training with augmented Shared Socio-economic Pathways scenarios, the NN achieves high-fidelity approximations of the original climate model. Additionally, we establish theoretical accuracy guarantees for both the encoding and neural network approximation. Our numerical experiments demonstrate the framework's computational efficiency and accuracy. For full article, see [1].

Keywords: Climate model, Neural network, Meta-model, Dirichlet polynomials, Shared Socio-economic Pathways, Differential Equations

References

1. Gobet, Emmanuel, Liu, Yushan, & Vermandel, Gauthier. (2025, March). Meta-modelling paths of the DICE climate block using Neural Networks and Dirichlet polynomials. Working paper or preprint. Retrieved from https://hal.science/hal-04990321.

Contact details

- Email: yushan.liu@polytechnique.edu

Design of parametric insurance via machine learning and optimal combination with traditional insurance

Olivier Lopez, Ensae Institut Polytechnique de Paris (presenter)

Antoine Heranval, INRAE Avignon, Daniel Nkameni, Ensae Institut Polytechnique de Paris

Abstract: Parametric insurance is frequently mentioned as a method to reduce protection gaps, especially in the field of emerging risks. Compared to traditional insurance, the idea is to compute the compensation from a parameter (or index) that can be measured soon after the occurrence of the claim. The advantage is that payment is fast, and this might be appealing for the policyholder that needs to quickly receive funds to rebuild after an incident, rather than waiting for an expert to be sent on site. For the insurer, the costs related to claim management are considerably reduced. Moreover, the volatility of the risk is usually controlled since, by design, the parameter is quantity on which a significant amount of data has been collected, hence its distribution is well estimated statistically speaking. This comes of course with a price: by giving up on covering the risk itself and replacing it by a proxy, basis risk (defined here as the difference between the true loss experienced by the policyholder and the compensation) appears. This uncertainty regarding the fact that the parametric product will respond correctly to the needs of the customer explains a potential reluctance to subscribe to such solutions. In this work, we discuss the combination of parametric and traditional insurance. The idea is to identify situations (based on features observable just after the claim) where a fast compensation via the parametric product should be triggered, or if a traditional (and slower) compensation, based on the precise evaluation of the true loss, has to be prefered. The index of the parametric part of the product is designed via machine learning technique, with a specific loss function designed to reflect the preferences of the policyholders. We then use the model to identify the regions of the covariate space where the parametric product should be used. The viability of this "hybrid" product is studied via a model on insurance demand, and an illustration is provided in cyber insurance.

Keywords: Parametric insurance, neural networks, insurance demand, cyber risk.

References

1. Cesarini, L., Figueiredo, R., Monteleone, B., Martina, M.L.V. (2021) The potential of machine learning for weather index insurance. *Natural Hazards and Earth System Sciences*, **21(8)**, 2379–2405.
2. Clement, K.Y., Botzen, W.W., Brouwer, R., Aerts, J.C. (2018) A global review of the impact of basis risk on the functioning of and demand for index insurance. *International Journal of Disaster Risk Reduction*, **28**, 845–853.
3. Conradt, S., Finger, R., Spörri, M.. (2015). Flexible weather index-based insurance design. Climate *Risk Management*, **10**, 106–117.
4. Lopez, O., Nkameni, D. (2025). Combination of traditional and parametric insurance: calibration method based on the optimization of a criterion adapted to heavy tail losses. *Preprint,* https://hal.science/hal-04959706/

Contact details

- Email: olivier.lopez@ensae.fr
- Homepage: https://sites.google.com/view/sitepersonneldolivierlopez/home
- Social media: https://www.linkedin.com/in/olivier-lopez-a10b03128/

Mitigating Systemic Risk in Catastrophe Insurance: The Role of Human Judgment in Model Diversification

Despoina Makariou, Institute of Insurance Economics, University of St.Gallen

Agni Orfanoudaki, Said Business School, University of Oxford

Abstract: This study aims to investigate the systemic risk implications of widespread reliance on standardized vendor catastrophe models in the insurance and reinsurance industry. While these models guide critical decisions around capital allocation and risk transfer, their uniform use may lead to homogenized risk assessments and increased systemic vulnerability. Using empirical analysis of model outputs, actual loss data, and human adjustments on the model output, we aim to evaluate whether expert interventions can diversify risk evaluations and mitigate systemic risk. Our findings will contribute to the growing discourse on human-in-the-loop decision-making, examining the value of expert judgment in counterbalancing algorithmic biases.

Keywords: catastrophe risk models, systemic risk, model homogeneity, human-in-the-loop, expert judgment.

References

1. Heinrich, T., Sabuco, J., & Farmer, J. D. (2022). A simulation of the insurance industry: the problem of risk model homogeneity. Journal of Economic Interaction and Coordination, 17(2), 535-576.

Contact details

- Email: despoina.makariou@unisg.ch
- Homepage: myhomepage.com
- Repository: myrepository.com
- Social media: Linkedin, Twitter

Agentic AI applications in insurance

Gustavo Martinez, Mirai Solutions (presenter)

Roland Schmid, Mirai Solutions

Abstract: Insurers must navigate growing customer expectations, increasing regulatory obligations, labor-intensive processes, massive data volumes, and the challenges of modern risk management alongside new and evolving risks. In this context, agentic AI is set to revolutionize the insurance industry by streamlining complex business processes through groundbreaking automation. Unlike robotic process automation and traditional AI assistants, AI agents can independently analyze their environment, take initiative, apply memory and reasoning, break down intricate problems into manageable parts, assess multiple strategies, and evolve to meet user needs over time. By continuously learning and optimizing based on user interactions, these AI systems become increasingly tailored to specific business needs, promising unparalleled benefits in automating workflows and enhancing customer satisfaction. Building agentic applications requires assembling components from multiple domains —utilizing large language models, multi-modal data processing, advanced tool integration, security frameworks, and orchestration— and presents major transformation challenges, including data collection and integration, data analysis and model training, decision-making autonomy, and ongoing updates to support continuous learning and improvement. The deployment of agentic AI also carries significant risks and limitations, including data security concerns, computational complexity, knowledge constraints, external dependencies, and challenges related to decision-making autonomy. We introduce a strategic roadmap to guide insurers in building agentic AI applications, focusing on data strategy, choice of models and technology, building knowledge, and best practices.

Keywords: Agentic AI, machine learning, large language models, decision-making autonomy, data strategy.

References

1. Gotowska, A. (2024). What are AI agents?. *IBM Think*.
2. Bouchard, A., Smith-Bingham, A., Buvat, J., Nambiar, R., Kumar, T. (2025). The transformative impact of Gen AI and agentic AI. *Capgemini Research Institute*.
3. Korolov, M. (2024). AI agents will transform business processes and magnify risks. *CIO*.
4. Khiyara, A. (2024). Agent Swarms – an evolutionary leap in intelligent automation. *CIO*.

Contact details

- Email: info@mirai-solutions.com
- Homepage: mirai-solutions.ch
- Linkedin
- GitHub

Lost in Translation: Is Portfolio Analytics an Academic Exercise or a Strategic Asset?

Amaryllis Mouyiannou, Ph.D., Swiss Re (presenter)

Abstract: Despite advancements in data science and AI, portfolio analytics in Specialty Reinsurance often fails to translate into meaningful strategic decisions. Sophisticated models and dashboards promise better portfolio steering, yet adoption remains low due to organizational resistance, misalignment with underwriting workflows, and a lack of clear business integration.

This session explores why analytics often remain an academic exercise rather than a strategic asset and what it takes to bridge the gap. We will discuss how to translate data science-enabled projects into measurable and actionable business value, challenging assumptions about analytics adoption in reinsurance. Through real-world insights, we will spark a dialogue on designing solutions that are both methodologically advanced and impactful in practice.

The last mile of insurance analytics—the translation to business value—is also the first: without embedding business impact from the start, even the most powerful models risk being ignored.

Keywords: Portfolio Analytics, Reinsurance Strategy, Data Science Adoption, Business Impact

Contact details

- Email: amaryllis_mouyiannou@swissre.com
- Social media: https://ch.linkedin.com/in/amaryllis-mouyiannou-ab7a6250

Fully Transparent Machine Learning: Exact Factor Table Representation of GBMs

Lucas Muzynoski, Avenue Analytics (presenter)

Abstract: The insurance industry has been slow to adopt machine learning techniques due to the "black box" nature of many models, which challenges the transparency and compliance required by regulators. While SHAP values offer post-hoc interpretability for individual predictions, they don't provide the full model transparency that regulators demand. I present a novel solution that transforms Gradient Boosting Machines (GBMs) into fully transparent, regulator-ready factor tables that mirror the format of traditional Generalized Linear Models (GLMs).

The methodology shows that any GBM can be exactly represented as a set of factor tables. To make them more interpretable, three key innovations are introduced to LightGBM: controlled tree depth constraints, L0 penalties for feature interactions, and multi-objective tuning that balances accuracy with interpretability. This exact representation means every prediction can be traced through clearly defined factor tables - achieving true transparency rather than just post-hoc explanations of individual predictions.

The resulting solution is a large step towards automating pricing model development, saving weeks of model development time while producing more accurate models. Insurance carriers can now confidently deploy sophisticated machine learning models with mathematical guarantees of transparency and interpretability that fully satisfy regulatory requirements.

Keywords: Insurance pricing, gradient boosting, model transparency, factor tables, interpretability

References

1. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, **29(5)**, 1189–1232.
2. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd Edition. Springer.
3. Wüthrich, M. V., & Merz, M. (2013). *Statistical Foundations of Actuarial Learning and Modeling.* Springer.
4. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems,* **30**, 4765–4774.

Contact details

- Email: info@avenueanalytics.com
- Homepage: Avenue Analytics
- Social media: LinkedIn Profile

Advancing Claims Document Processing with LLMs: From PoC to Production

Matej Otcenas (presenter) and Jakob Dambon, SwissRe

Abstract: Claims handling in commercial insurance involves document-heavy workflows, particularly in processing claims' First-Notice-of-Loss (FNOL) documents. Last year, we presented the results of a proof-of-concept (PoC) that leveraged Large Language Models (LLMs) for entity extraction from unstructured FNOLs. Since then, our model has evolved into a production system, significantly improving automation in claims processing.

Our latest implementation predominantly employs LLMs for entity extraction, replacing earlier regular expression approaches. Further, the transition to new LLM architectures (GPT 4o) has led to notable improvements in extraction rates and precision. Our ongoing work on policy ID matching has shown promising results. This remains a key challenge due to the complexity of policy structures, and we have conducted a dedicated PoC to explore different approaches. We will share our experiences from this effort, including successes and limitations.

With FNOLs now at the core of our focus, we continue refining our approach to balance automation and human-in-the-loop oversight. By combining cutting-edge LLM techniques with structured evaluation and reproducibility, we aim to drive efficiency while ensuring accuracy in claims handling. Our presentation will provide insights into the practical deployment of LLMs in real-world insurance workflows and discuss the next steps for further optimization.

Keywords: GenAI, entity extraction, policy matching

References

None.

Contact details

- Social media: LinkedIn
- Matej_Otcenas@swissre.com
- Jakob_Dambon@swissre.com

Advancing Non-Life Insurance Modelling with NetSimR

Yiannis Parizas, Actuary and Open-Source Developer

Abstract: Non-life insurers increasingly demand analytical solutions that are both rigorous and adaptable. NetSimR is an open-source R package that meets this need by integrating tools into a single framework to support actuarial modelling tasks related to pricing, reserving, and capital modelling. It offers functions to compute capped means, exposure curves, and increased limit factor curves for various severity distributions. Moreover, it includes functions to predict pure IBNR reserves when the reporting delay is modelled using either a gamma or lognormal distribution. Additionally, NetSimR provides three interactive Shiny tools: - a Claims Simulation Tool for modelling reinsurance structures via simulations over various severity distributions (including Pareto slicing); - a GLM Fitting Tool for flexible, user-friendly generalized linear model analysis of claims data; and - a Distribution Fitting Tool that enables practitioners to accurately fit claims frequency and severity distributions (including Pareto slicing).

Together, these tools enhance pricing assumptions and risk assessments. With over 100,000 downloads on CRAN, NetSimR demonstrates its value in delivering reproducible, data-driven insights for actuarial practice. This presentation will outline the package's core methodologies, showcase practical case studies, and discuss how these interactive tools add value to non-life insurance risk modelling—ultimately bridging the gap between academic research and everyday operational challenges in the insurance industry.

Keywords: NetSimR, non-life insurance, claims simulation, distribution fitting, GLM, actuarial modelling

References

1. Parizas, Y. (2023). Free for all. The Actuary. Available at: https://www.theactuary.com/2023/03/02/free-all
2. Parizas, Y. (2019). Escaping the triangle. The Actuary. Available at: https://www.theactuary.com/features/2019/06/2019/06/05/escaping-triangle

Contact details

- Email: yiannis.parizas@gmail.com
- Repository: https://cran.r-project.org/web/packages/NetSimR/index.html
- Social media: https://www.linkedin.com/in/yiannisparizas/

Analytical variable importance indices for Generalized Additive Models

Giovanni Rabitti, Heriot-Watt University

Abstract: Generalized Additive Models (GAMs) are a widely used statistical method.

In this talk, we aim to tackle the challenge of identifying the most influential variables in GAMs. We introduce a variance allocation approach based on the Shapley values. We derive a closed-form expression for these importance indices for smooth components, which allows for their computation on high-dimensional datasets and with any dependence structure.

Moreover, we demonstrate that, after one-hot encoding a categorical variable, the sum of the Shapley values across all levels yields the total Shapley value for the variable, which remains invariant to the choice of the reference category. Furthermore, we construct confidence intervals for such Shapley value, that keeps this invariance.

Finally, we prove that the p-value associated with a categorical variable is also invariant to the reference category, even though the Shapley values and p-values of each level change. We discuss the practical implication that when a variable's importance is negligible, it can be safely eliminated from the GAM, simplifying the model. Numerical simulations and applications illustrate our findings.

Keywords: Global sensitivity analysis, Variable importance and significance, Cooperative game theory.

References

1. Calvetti, L., Khorrami Chokami, A., Rabitti, G. On the significance and importance of categorical variables in generalized additive models. *Submitted*.
2. Khorrami Chokami, A., Rabitti, G. (2024). An exact game-theoretic variable importance index for generalized additive models. *Journal of Computational and Graphical Statistics* **33(4)**, 1276–1285.
3. Lo, A., Chernoff, H., Zheng, T., Lo, S. H. (2015). Why significant variables aren't automatically good predictors. *Proceedings of the National Academy of Sciences of the United States of America* **112(45)**, 13892–13897.

Contact details

- Email: g.rabitti@hw.ac.uk

Price Leakage in Demand Models

Michael Ramati, Earnix (presenter)

Luba Orlovsky, Earnix

Abstract: Target leakage is a known notion in machine learning, where models include inputs that are available only when the target is, thus effects of other inputs are suboptimally estimated, leading to highly biased models. Demand models, where the price effect is estimated, pose an interesting variant to this notion, where the price (rather than target) leaks, leading to highly biased price effect and predicted demand. As with target leakage, the resultant biases can be removed by carefully eliminating the leaked-to variables. However, whereas the task of identifying these leaked-to inputs is generally hard for target leakages, the specific settings of demand models make the identification of price leakages simple. Fortunately, this proposed identification method is not too specific: it is general enough to also identify target leakages in the typical settings of demand models.

We introduce the notion of price leakage, compare it to the well known notion of target leakage, and present them both in the framework of causal graphs. We also propose a simple method to identify these leakages in the typical settings of demand models, discuss the potentially negative ramification of eliminating such leaked-to variables, and suggest a simple cure to avoid it.

Keywords: Target leakage, Demand models, Price effect, Causal analysis

References

1. Kapoor S, Narayanan A. (2023). *Leakage and the reproducibility crisis in machine-learning-based science*. Patterns.
2. Pearl, J. and Glymour, M. and Jewell, N.P. (2016). *Causal Inference in Statistics: A Primer*. Wiley.
3. Hastie, T., Tibshirani, R., Friedman, J. (2013). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

Contact details

- Email: michaelr@earnix.com
- Homepage: earnix.com
- Social media: Linkedin, Twitter

Granular mortality modeling with temperature and epidemic shocks: a three-state regime-switching approach

Jens Robben, Research Centre for Longevity Risk (RCLR), University of Amsterdam, The Netherlands (presenter)

Karim Barigou, Institute of Statistics, Biostatistics and Actuarial Science (ISBA), Louvain Institute of Data Analysis and Modeling (LIDAM), UCLouvain, Louvain-la-Neuve, Belgium

Torsten Kleinow, Research Centre for Longevity Risk (RCLR), University of Amsterdam, The Netherlands

Abstract: This paper develops a granular regime-switching framework to model mortality deviations from seasonal baseline trends driven by temperature and epidemic shocks. The framework features three states: (1) a baseline state that captures observed seasonal mortality patterns, (2) an environmental shock state for heat waves, and (3) a respiratory shock state that addresses mortality deviations caused by strong outbreaks of respiratory diseases due to influenza and COVID-19. Transition probabilities between states are modeled using covariate-dependent multinomial logit functions. These functions incorporate, among others, lagged temperature and influenza incidence rates as predictors, allowing dynamic adjustments to evolving shocks. Calibrated on weekly mortality data across 21 French regions and six age groups, the regime-switching framework accounts for spatial and demographic heterogeneity. Under various projection scenarios for temperature and influenza, we quantify uncertainty in mortality forecasts through prediction intervals constructed using an extensive bootstrap approach. These projections can guide healthcare providers and hospitals in managing risks and planning resources for potential future shocks.

Keywords: granular mortality modeling, regime-switching, environmental shocks, respiratory shocks

References

1. Robben, J., Barigou, K., Kleinow, T. (2025). Granular mortality modeling with temperature and epidemic shocks: a three-state regime-switching approach. *arXiv preprint arXiv:2405.18020*

Contact details

- Email: j.robben@uva.nl
- Homepage: jensrobben.github.io
- Repository: github.com/jensrobben/GranMoMoRS
- Social media: www.linkedin.com/in/jens-robben-4a075b1aa/

Strategic Asset Allocation for Insurance Product Development: a Machine Learning approach

- Francesca Nuzzo, PhD, Generali Italia (presenter)

- Leonardo Ruggieri, Generali Italia (presenter)

Abstract: This study explores the application of machine learning (ML) techniques for strategic asset allocation (SAA) in the insurance sector, with a specific focus on index replication methods. The study proposes an innovative framework that leverages the synergy between deep learning models and generative AI to develop a highly personalized approach to facilitate a move from broad risk categories to highly individualized client risk profiles.

The deep learning methodology balances tracking error minimization with real-world constraints such as transaction costs, turnover limits, and asset concentration rules. Notably, a key development of this project is the integration of "Paul", a proprietary agent-based generative AI solution, that provides macroeconomic insights, which are incorporated into the ML-driven asset allocation model to enhance portfolio management by dynamically adjusting investment decisions based on changing market conditions and learning from historical performance to optimize allocation strategies. This synergy between ML and LLMs offers a novel approach to blending qualitative and quantitative insights in portfolio construction.

The findings demonstrate the feasibility of using ML-based index replication in managing insurance-linked investment products, particularly Unit-Linked funds. The proposed model not only improves the efficiency of investment strategies but also enables the creation of customized insurance products tailored to specific client needs. Future developments will focus on expanding the application of these techniques to a broader range of asset classes and financial instruments. This research highlights the strategic role of artificial intelligence in modern insurance and asset management, positioning ML-driven solutions as a key driver for innovation in the industry.

Keywords: SAA, Index Replication, Deep Learning, Generative AI

References

1. Zhang, C., Liang, S., Lyu, F., & Fang, L. (2020). Stock-Index Tracking Optimization Using Auto-Encoders. Frontiers in Physics, **8**, 388.
2. Hong, Y., Kim, Y., Kim, J., & Choi, Y. (2022). Index Tracking via Learning to Predict Market Sensitivities. arXiv preprint arXiv:2209.00780
3. Xing, F. (2025). Designing heterogeneous LLM agents for financial sentiment analysis. *ACM Transactions on Management Information Systems* **16/25**, 99-122.

Contact details

- Email: francesca.nuzzo@generali.com, leonardo.ruggieri@generali.com
- Social media: https://www.linkedin.com/in/maria-francesca-nuzzo-ph-d-b339b270/, https://www.linkedin.com/in/leonardoruggieri/

Non-parametric insurance loss modelling using variable-knot splines

Dimitrina S. Dimitrova, Vladimir K. Kaishev and Emilio L. Sáenz Guillén (presenter)

Faculty of Actuarial Science and Insurance, Bayes Business School, City St George's, University of London

Abstract: We introduce a novel free-knot spline method for the simultaneous estimation of the probability density function (pdf) and cumulative distribution function (cdf) of a random variable based on a given data sample. The proposed method, named DDFS, leverages the property that both pdf and cdf spline models are defined on a common set of knots, and their coefficients are connected. The method combines sequential estimation of the knots using minimum distance, with maximum likelihood estimation of the coefficients, and it is readily applicable to both univariate and bivariate cases. Furthermore, we show that the underlying spline representation of the pdf (cdf) can be viewed as a mixture of linear combinations of Dirichlet random variables, enabling the establishment of strong consistency and asymptotic normality of the DDFS estimates. Extensive numerical analysis demonstrates that DDFS constitutes a highly competitive alternative to well-established nonparametric density estimation techniques.

In recent years, the actuarial literature has seen a proliferation of parametric models being proposed for the modelling of insurance loss data (see, e.g., [2], for a review), presenting practitioners with an often cumbersome model selection dilemma. In addition, model misspecification remains a major threat, potentially leading to biased risk assessments as well as invalid inference and prediction. In this regard, DDFS provides a robust alternative, effectively capturing heavy-tailed behaviour without the need for complex distributional assumptions.

Keywords: nonparametric density estimation, variable-knot splines, boundary bias, heavy-tailed distributions, risk measures

References

1. Marambakuyana, W.A., Shongwe, S.C. (2024). Composite and mixture distributions for heavy-tailed data — An application to insurance claims. *Mathematics* **12**(2), 335.

Contact details

- Email: Emilio.Saenz-Guillen@citystgeorges.ac.uk
- Homepage: https://emilioluissaenzguillen.github.io/
- Repository: https://github.com/emilioluissaenzguillen
- Social media: https://www.linkedin.com/in/emilioluissaenzguillen/

Auto Insurance Fraud Detection: Machine Learning and Deep Learning Applications

Meryem Yankol-Schalck, IPAG Business School (presenter)

Abstract: Insurance fraud poses a significant challenge to effective efforts to detect and prevent fraud. This is due to a number of factors, including data imbalance, evolving fraud patterns, and the significant number of false negatives. This paper focuses on the application of recent developments in deep learning algorithms for this purpose. A comparative analysis of both machine learning and deep learning algorithms was performed to identify the most efficient results using two different datasets with distinct properties. This study uses seven different machine learning algorithms to predict the likelihood of a claim being fraudulent: logistic regression (LR), random forest (RF), XGBoost, AdaBoost classifier (ADA), extra trees classifier (ET), light gradient boosting machine (LGBM) and decision tree (DT).Then, the study extends the analysis by incorporating Convolutional Neural Networks (CNNs). Three convolutional neural network architectures with 11, 17 and 20 layers were used to improve the accuracy of fraud detection. The aim was to improve the overall performance and reliability of fraud detection by using two different datasets. Based on key metrics, our results show that fraud detection is more accurate in the real database because the large dataset size ensures robust model training and generalisation. The results also show that different architectures of the deep learning model would perform differently on each dataset due to variations in data characteristics. However, our results do not confirm that deep learning algorithms outperform traditional ML algorithms, which challenges the common assumption of CNN superiority and suggests that the choice of algorithm should be context dependent, taking into account the characteristics of the dataset and the specific requirements of the problem. Taken together, these results highlight the need for nuanced fraud detection approaches that consider dataset size, imbalance and model selection, rather than relying on generalised assumptions about algorithmic performance. The implementation of SMOTE in CCNs methods yielded a modest improvement in fraud detection rate. CNN models did not provide a more robust solution for fraud detection, particularly when dealing with imbalanced datasets and striving to reduce false negatives in real-world settings.

Keywords: Fraud detection; Automobile insurance; CNN and Machine Learning

References

1. Yankol-Schalck, M. (2025). Auto Insurance Fraud Detection: Leveraging Cost Sensitive and Insensitive Algorithms for Comprehensive Analysis. Insurance: Mathematics and Economics.Volume 122, Pages 44-60.
2. Banulescu-Radu, D., & Yankol-Schalck, M. (2023). Practical guideline to efficiently detect insurance fraud in the era of machine learning: A household insurance case. Journal of Risk and Insurance
3. Yankol-Schalck, M. (2022). The value of cross-data set analysis for automobile insurance fraud detection. Research in International Business and Finance, 63, 101769.
4. Schalck, C., & Yankol-Schalck, M. (2021). Predicting French SME failures: new evidence from machine learning techniques. Applied Economics, 53(51), 5948-5963.

Contact details

- Email: meryem.schalck@ipag.fr
- Homepage: https://www.ipag.edu/meryem-schalck?tab=onglet1
- Social media: https://www.linkedin.com/in/meryem-yankol-schalck-phd-558b0539/ # New Document

Elevating Trust in High-Stakes Decisions Using Glass-Box Models and Robust Feature Selection

Matthias Linder, Judith C. Schneider, Brandon Schwab (presenter)

Institute for Risk and Insurance, Leibniz University Hannover, Hannover, Germany

Abstract: In high-stakes domains such as finance and insurance, machine learning models must balance predictive accuracy with interpretability and robustness. Traditional black-box approaches, while powerful, often lack transparency and stability, raising concerns about trust and regulatory compliance (Reinhardt, 2023). This paper presents a novel framework that combines Neural Additive Models (NAMs) (Agarwal et al., 2021) with a structured feature selection approach to enhance both interpretability and robustness. We introduce a principled method for selecting stable, high-impact features by leveraging statistical significance testing and bootstrap aggregation to ensure consistency under data perturbations. Our approach is evaluated using both a widely studied public insurance dataset and a proprietary dataset from a European insurer, demonstrating its practical applicability in real-world pricing scenarios. Results demonstrate that NAMs, when coupled with principled feature selection, not only match or exceed the predictive performance of black-box models but also provide transparent, stable justifications for AI-driven decisions. The study underscores the regulatory and managerial advantages of intrinsically interpretable models, advocating for a shift towards explainable AI solutions in high-stakes environments.

Keywords: Explainable AI, Neural Additive Models, Model Interpretability, Deep Learning

References

1. Reinhardt, K. (2023). Trust and trustworthiness in AI ethics. *AI and Ethics, 3(3), 735-744*.
2. Agarwal, R., Melnick, L., Frosst, N., Zhang, X., Lengerich, B., Caruana, R., & Hinton, G. E. (2021). Neural additive models: Interpretable machine learning with neural nets. *Advances in Neural Information Processing Systems, 34, 4699-4711.*

Contact details

- Email: brandon.schwab@insurance.uni-hannover.de
- Social media: www.linkedin.com/in/brandon-schwab-abb0791ab

The Credibility Transformer

Ronald Richman, InsureAI and University of the Witwatersrand, South Africa (presenter)

Salvatore Scognamiglio, University of Naples Parthenope, Italy

Mario V. Wüthrich, RiskLab, Department of Mathematics, ETH, Zurich, Switzerland

Abstract: Inspired by the large success of Transformers in Large Language Models, these architectures are increasingly applied to tabular data. This is achieved by embedding tabular data into low-dimensional Euclidean spaces resulting in similar structures as time-series data. We introduce a novel credibility mechanism to this Transformer architecture. This credibility mechanism is based on a special token that should be seen as an encoder that consists of a credibility weighted average of prior information and observation based information.

We demonstrate that this novel credibility mechanism is very beneficial to stabilize training, and our Credibility Transformer leads to predictive models that are superior to state-of-the-art Transformer models.

Through comprehensive experimentation on a French motor third party liability claims dataset, we show that our approach significantly outperforms classical feed-forward neural networks and existing Transformer architectures for tabular data. We further explore explainability aspects of the fitted models, demonstrating how the credibility mechanism adapts to different risk characteristics in the portfolio.

Keywords: Non-life Pricing, Transformer, Credibility, Tabular data, Feature-engineering, Entity embedding

References

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I. (2017). Attention is all you need. *arXiv:1706.03762v5*
2. Bühlmann, H., Straub, E. (1970). Glaubwürdigkeit für Schadensätze. *Bull Swiss Assoc Actuaries* **1970**, 111-131.
3. Gorishniy, Y., Rubachev, I., Khrulkov, V., Babenko, A. (2021). Revisiting deep learning models for tabular data. *Advances in Neural Information Processing Systems* **34**, 18932-18943.
4. Shleifer, S., Weston, J., Ott, M. (2021). NormFormer: Improved Transformer pretraining with extra normalization. *arXiv:2110.09456*
5. Wüthrich, M.V., Merz, M. (2023). *Statistical foundations of actuarial learning and its applications*. Springer Actuarial.

Contact details

- Email: ron@insureai.co
- Repository: https://github.com/RonRichman/CredibilityTransformer/tree/main

Copula Models of Correlation in Insurer Loss Reserves

Mark Shoun, Ledger Investing, Inc

Abstract: Property and casualty insurers in the US are required to report their carried loss reserves by line of business and accident year on an annual basis. Year-over-year changes in these carried reserves are not only correlated across development within accident years, but also across accident years, lines of business, and carriers within a given calendar year. These correlations have significant implications on the expected volatility of insurers' calendar-year underwriting results, and on the potential tail risk associated with loss portfolio transfer transactions.

Here we use Bayesian copula models fitted to a large corpus of US statutory property and casualty insurer filings to quantify the degree of correlation between changes in loss reserves and the aforementioned covariates. We discuss what the posterior distributions derived from model parameters imply about how insurers set and update loss reserves in practice. Finally, we illustrate the application of the model to estimate tail risk for a hypothetical portfolio of loss portfolio transfer transactions.

Keywords: Loss reserving; correlation.

References

Contact details

- Email: mark@ledgerinvesting.com
- Social media: linkedin.com/in/mark-shoun

Enhancing Contract Wordings Analysis with Generative AI: A Timeline of Efficiency and Accuracy

Pratyush Singh, Data Scientist, Swiss Re (presenter)

Charilaos Tsarouchas, Lead Data Scientist, Swiss Re

Abstract: The paper discusses a novel application of generative AI in the analysis of reinsurance contracts, particularly Life and Health Treaties. This innovative approach aims to revolutionize contract management and analysis by leveraging advanced analytics. The project involves parsing contracts from digital formats like images or PDFs into structured formats like MS Excel and JSON while preserving layout information and masking sensitive information in the contracts. The application utilizes advanced machine learning models and large language models to extract, summarize, and compare clauses from different contracts, enabling experts to track the evolution of terms over time. Retrieval-Augmented Generation (RAG) is used to extract small and hidden but essential information and integrate it into the final structured output. The solution is scalable, handling tens of thousands of documents efficiently and cost-effectively, utilizing GPT only when necessary. The implementation streamlines the contract analysis process, enhances accuracy and efficiency, and provides valuable insights for business decision-making. The structured digital twin of the contract enables automated compliance checks, integration with ERP or CRM systems, and contract lifecycle management. Additionally, the validation cycle of our predictions is outlined, completing the Human-in-the-loop (HITL) process. The output is displayed on a dashboard for quick insights and access to 25 years of data at fingertips. This approach provides an efficient method for generating insights from 25 years of digital contracts, thus accelerating the work process of lawyers and contract experts by at least 50% and up to 80%. It offers a robust solution for managing complex documents, ultimately driving better business outcomes and fostering a deeper understanding of contractual terms.

Keywords: Contract Analysis, Generative Artificial Intelligence, Unstructured Data, Contract Wordings

References

1. Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, Douwe Kiela (2021). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.
2. Adam Roegiest and Radha Chitta (2024). Answering Questions in Stages: Prompt Chaining for Contract QA.
3. Alexandre Yukio Ichida and Felipe Meneguzzi (2021). Detecting Logical Relation In Contract Clauses.
4. Nelson F. Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni and Percy Liang (2023). Lost in the Middle: How Language Models Use Long Contexts.

Contact details

- Email: pratyush_singh@swissre.com
- Homepage: Pratyush Singh
- Publications: Google Scholar
- Linkedin: https://www.linkedin.com/in/spratyush02/

LLMs for Claims Processing: A Fully Local and Compliant Solution

Dr Malgorzata Smietanka, UCL (presenter)

Abstract: This presentation explores the potential of Large Language Models (LLMs) in the context of insurance claims processing, with a particular focus on a fully local, regulation-compliant implementation.

The session begins with a brief clarification of LLM capabilities and limitations, their relevance to the insurance domain, and an overview of applicable regulatory obligations. Then, the session features a live demo of an operational LLM-powered platform that automates key stages of the claims workflow—namely, data extraction, exclusion identification, and document categorisation—applied to insurance claims.

The demonstrated solution operates entirely within the insurer's secure infrastructure. It supports full compliance with GDPR requirements, including data minimisation, removing PII, purpose limitation, and auditability. The solution is also aligned with the EU AI Act's regulatory framework for high-risk AI systems, meeting criteria such as human oversight, decision explainability, and ongoing monitoring.

This talk is particularly relevant for data scientists, claims professionals, and compliance teams seeking to understand how AI can be deployed responsibly and effectively within regulated insurance environments.

Keywords: Large Language Models, Claims Processing, GDPR, EU AI Act, Insurance AI, Document Automation, Compliance, Data Privacy

References

1. European Commission (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
2. European Union (2016). General Data Protection Regulation (GDPR).
3. Koshiyama et al. (2021). Towards Algorithm Auditing: A Survey on Managing AI Risks in Insurance. *Journal of Risk and Financial Management*.

Contact details

- Email: ucabmw5@ucl.ac.uk

Statistical Learning of Trade Credit Insurance Network Data with Applications to Ratemaking and Reserving

Tsz Chai Fung, Georgia State University (presenter)

Woongchae Yoo, Georgia State University

Spark C. Tseung, University of Toronto

Abstract: Trade credit insurance (TCI) is a specialized line of property and casualty insurance, protecting businesses against financial losses due to buyer's insolvency. Predictive modeling for TCI claims poses formidable challenges due to the data's complexity, yet remains underexplored in the literature. Leveraging six years of detailed TCI data from an Asian TCI insurer, we develop a bivariate, network-augmented Generalized Linear Mixed Model (GLMM) to jointly model claim probability and reporting time gaps. Our model integrates extended-order degree centrality and random effects at the business and policy levels, adjusted for data incompleteness, to capture claim histories, reporting time gaps, and network relationships specific to TCI data. To implement a feasible workaround for the high-dimensional integrations required by individual random effects, we propose a scalable Stochastic Expectation-Maximization (SEM) algorithm. Data analysis using this TCI dataset demonstrates that our model significantly outperforms benchmark models in both model fit and predictive accuracy, highlighting the effectiveness of our approach for improved ratemaking and reserving in TCI.

Keywords: Property and Casualty (P&C) Insurance; Random Effects; Ratemaking and Reserving; Second-Order Degree Centrality; Social Network

References

1. Yoo, W., Tseung, S.C., Fung, T.C. (2025). Statistical Learning of Trade Credit Insurance Network Data with Applications to Ratemaking and Reserving. Working paper.

Contact details

- Email: tfung@gsu.edu
- Homepage: https://robinson.gsu.edu/profile/tsz-chai-fung/
- Social media: Linkedin

Dynamic hierarchical graph neural networks for spatiotemporal prediction of flood-related claims

George Tzougas, Department of Actuarial Mathematics and Statistics Heriot-Watt University, Edinburgh, UK (presenter)

Jing Zou, Technische Universität Dresden; Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany

Ostap Okhrin, Technische Universität Dresden; Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany

Abstract: The aim of this paper is to develop a dynamic hierarchical Graph Neural Network (GNN) framework for spatiotemporal regression, see, e.g., Ma t al. (2023), to predict flood-related insurance claims. Our model utilizes a global graph where nodes—representing postal codes—contain local graph structures that capture policy-level information while addressing imbalances in property distributions over time. Incorporating geographical covariates, building and content characteristics, and vulnerability levels, we also integrate the excessive rainfall index and wind speed index at the postal code level to enhance predictive performance.

For interpretability, we leverage the contextual embedding-based GNN from Vallelado t al. (2024) to visualize pairwise feature interactions by including an interaction network layer at the policy level. To address the high skewness of zeros in the response variable, we also incorporate a Zero-Inflated Negative Binomial (ZINB) module into our GNN framework, drawing inspiration from methods used in highly sparse single-cell RNA-sequencing analysis, see Risso et al.(2018) and Gan t al. (2022). Furthermore, we implement a feature selection process using subnetworks to optimize the ZINB parameters for mean and dispersion.

We apply our GNN-ZINB framework to an empirical study using data from Athens, Greece, spanning 257 postal code areas from 2013 to 2022, demonstrating its potential for improving flood insurance design.

Keywords: Apple, Lemon, Kiwi

References

1. Ma, M., Xie, P., Teng, F., Wang, B., Ji, S., Zhang, J., Li, T. (2023). HiSTGNN: Hierarchical spatio-temporal graph neural network for weather forecasting. Information Sciences 648, 119580.
2. Villaizán-Vallelado, M., Salvatori, M., Carro, B., Sanchez-Esguevillas, A. J. (2024). Graph neural network contextual embedding for deep learning on tabular data. Neural Networks 173, 106180.
3. Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S., Vert, J. P. (2018). A general and flexible method for signal extraction from single-cell RNA-seq data. Nature Communications 9(1), 284.
4. Gan, Y., Huang, X., Zou, G., Zhou, S., Guan, J. (2022). Deep structural clustering for single-cell RNA-seq data jointly through autoencoder and graph neural network. Briefings in Bioinformatics 23(2), bbac018.

Contact details

- Email: george.tzougas@hw.ac.uk
- Homepage: https://researchportal.hw.ac.uk/en/persons/george-tzougas
- Social media: https://www.linkedin.com/in/george-tzougas-648711294/

An Augmented Variable Dirichlet Process Mixture model for the analysis of dependent lifetimes

Francesco Ungolo, University of New South Wales (presenter)

Patrick J. Laub, University of New South Wales

Abstract: The analysis of insurance and annuity products issued on multiple lives requires the use of statistical models which account for lifetime dependence. This paper presents a Dirichlet process mixture-based approach that allows to model dependent lifetimes within a group, such as married couples, accounting for individual as well as group-specific covariates. The model is analyzed in a fully Bayesian setting and illustrated to jointly model the lifetime of male–female couples in a portfolio of joint and last survivor annuities of a Canadian life insurer. The inferential approach allows to account for right censoring and left truncation, which are common features of data in survival analysis. The model shows improved in-sample and out-of-sample performance compared to traditional approaches assuming independent lifetimes and offers additional insights into the determinants of the dependence between lifetimes and their impact on joint and last survivor annuity prices.

Keywords: Dependent lifetimes, survival analysis, Dirichlet process, Bayesian analysis, life insurance, MCMC, mixture models

References

1. Ungolo, F., and Laub P.J. (2025). *An Augmented Variable Dirichlet Process Mixture Model for the Analysis of Dependent Lifetimes*. *ASTIN Bulletin 55* **(1): 50–75**.

Contact details

- Email: f.ungolo@unsw.edu.au
- Homepage: https://scholar.google.com/citations?user=Ty_f1dgAAAAJ&hl=it
- Repository: https://github.com/ungolof/AVDPM
- Social media: f_ungolo (Twitter)

Feature and quantile selection for the actuarial climate index: everything, everywhere, all at once

Mathias Valla, Chaire DIALog, Institut Louis Bachelier, Paris, France & Aix-Marseille University, Marseille, France (present)

Jose Garrido, Concordia University, Montreal, Canada (present)

Abstract: The actuarial climate index (ACI) has traditionally been computed as the average of standardized anomalies across several climatic variables, a method that uniformly weights features and lacks a statistically robust mechanism for feature and quantile selection.

In this presentation, we challenge these limitations by introducing an advanced statistical framework that leverages Time-Penalised Trees (TpT) to dynamically select and weight the features based on their predictive power for extreme events and climate-related mortality, while also rigorously assessing the interdependence among the climatic variables to refine the quantile selection process.

By focusing on key risk areas and integrating a more adaptable, data-driven methodology, this approach enhances the interpretability and sensitivity of the index, thereby offering insurers and policymakers a more tailored tool for assessing and managing evolving climate risks. Preliminary results indicate significant improvements in the index's ability to capture localized risk variations and provide a nuanced understanding of climate impacts on insurance claims, paving the way for further research and broader applications in actuarial science.

Keywords: Actuarial Climate Index, Feature Selection, Time-Penalised Trees (TpT), Extreme Compound Events, Climate Risk, Mortality, Quantile Selection

References

1. ACI Actuaries Climate Index. Development and design. American Academy of Actuaries, Canadian Institute of Actuaries, Casualty Actuarial Society, Society of Actuaries, 2016. URL : https://actuariesclimateindex.org/wp-content/uploads/2019/05/ACI.DevDes.2.20.pdf.
2. Steve Jackson and The American Academy of Actuaries. Actuaries Climate Risk Index, Preliminary findings. 2020
3. Jose Garrido, Xavier Milhaud, Anani Olympio. The definition of a French actuarial climate index; one more step towards a European index. 2023.
4. Nan Zhou, Jose Luis Vilar-Zanon, Jose Garrido, Antonio Jose Heras-Martinez. Measuring climate change from an actuarial perspective: a survey of insurance applications. Global Policy, 2024, 15 (S7), pp.34-46. 10.13140/RG.2.2.28395.96806.
5. Mathias Valla. Time-penalised trees (TpT): a new tree-based data mining algorithm for time-varying covariates. Annals of Mathematics and Artificial Intelligence, 2024.

Contact details

- Email: mathias.valla@gmail.com
- Homepage:
- Repository:
- Social media:

From claim counts to interarrival times using a small neural framework

Anne van der Scheer, Perunum Actuarieel Advies

Abstract: Claim counting plays a central role in actuarial modelling, with distributions such as Poisson, Negative Binomial, and their zero-inflated variants commonly used in Generalised Linear Models (GLMs). Recently, there has been a growing interest in integrating traditional models with neural network approaches.

In the context of assumption-free machine learning, it is natural to incorporate the precise occurence dates of claims, using this additional data to model claim interarrival times rather than just aggregated counts. To this end, we propose a small neural framework designed to estimate the cumulative distribution of continuous time-to-event data. This framework enables the direct derivation of probability density and hazard functions, with likelihood maximisation that accounts for left-, right-, and interval censoring. Furthermore, the model can handle multiple competing risks, such as different claim types (e.g. total loss and own fault).

We demonstrate the practical application of the framework through a synthetic example. Despite variations in underlying processes, the neural network is capable of accurately reproducing claim interarrival times and counts for different sub-populations.

In conclusion, the proposed framework makes a promising contribution to the accurate and insightful modelling of claim-generating processes. By providing a more granular connection to policy terms and conditions, it has the potential to enhance pricing. Additionally, this framework offers broader applicability as a versatile tool for time-to-event analysis.

Keywords: claim counts, claim interarrival times, time-to-event analysis, competing risks, neural network

References

1. Chilinski, P., & Silva, R. (2020, August). Neural likelihoods via cumulative distribution functions. In *Conference on Uncertainty in Artificial Intelligence* (pp. 420-429). PMLR.
2. Katzman, J. L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., & Kluger, Y. (2018). DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC medical research methodology, 18,* 1-12.
3. Lee, C., Zame, W., Yoon, J., & Van Der Schaar, M. (2018, April). Deephit: A deep learning approach to survival analysis with competing risks. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).

Contact details

- Email: a.vanderscheer@perunum.nl
- Homepage: https://perunum.nl
- Repository: https://github.com/perunum/claim-interarrival-times
- Social media: https://www.linkedin.com/in/anne-van-der-scheer-698a462

Machine learning in an expectation-maximisation framework for nowcasting

Paul Wilsens, KU Leuven (presenter)

Katrien Antonio, KU Leuven, University of Amsterdam

Gerda Claeskens, KU Leuven

Abstract: Information is often only partially observable. In decision making, this may cause under- or overestimation of underlying risk. Leveraging the available information to model the complete information is called nowcasting within the literature. In practical nowcasting applications, partial information is often caused by reporting delays. In this paper, we propose an expectation-maximisation framework that uses machine learning techniques to model both the occurrence as well as the reporting process of events. We allow for the inclusion of information specific to the occurrence and reporting periods as well as information related to the entity for which events occurred. Additionally, we demonstrate how deep learning techniques can be adapted for use in a nowcasting application. With simulation experiments, we show that we can effectively model both the occurrence and reporting of events when dealing with high-dimensional covariate information. In the presence of non-linear effects, we show that our methodology outperforms existing expectation-maximisation frameworks that rely on generalised linear models.

Keywords: nowcasting, machine learning, EM algorithm

References

1. Crevecoeur, J., Antonio, K., Desmedt, S., and Masquelein, A. (2023). Bridging the gap between pricing and reserving with an occurrence and development model for non-life insurance claims. ASTIN Bulletin: The Journal of the IAA, 53(2):185–212.
2. Delong, L., Lindholm, M., and W¨uthrich, M. V. (2021). Gamma mixture density networks and their application to modelling insurance claim amounts. Insurance: Mathematics and Economics, 101:240–261.
3. Lawless, J. (1994). Adjustments for reporting delays and the prediction of occurred but not reported events. Canadian Journal of Statistics, 22(1):15–31.

Contact details

- Email: paul.wilsens@kuleuven.be

Tests for Auto-Calibration

Mario Wüthrich, ETH Zurich (presenter)

Abstract: Auto-calibration is an important property of regression functions in actuarial applications such as insurance pricing. Basically, auto-calibration ensures that there is no systematic cross-subsidy between the different price cohorts in a pricing scheme.

Comparably little is known about statistical testing of auto-calibration. Denuit et al. (2014) recently published a test with an asymptotic distribution that is not fully explicit and its evaluation needs non-parametric Monte Carlo sampling.

In a simpler set-up, we present three test statistics with fully known and interpretable asymptotic distributions.

Keywords: Auto-calibration, regression modeling, isotonic regression, statistical modeling, insurance pricing, machine learning

References

1. Denuit, M., Huyghe, J., Trufin, J., Verdebout, T. (2024). Testing for auto-calibration with Lorenz and concentration curves. *Insurance: Mathematics and Economics* **117**, 130-139.
2. Wüthrich, M.V. (2024). Auto-calibration tests for discrete finite regression functions. *arXiv*:2408.05993.
3. Wüthrich, M.V., Merz, M. (2023). *Statistical Foundations of Actuarial Learning and its Applications*. Springer Actuarial. https://link.springer.com/book/10.1007/978-3-031-12409-9

Contact details

- Email: mario.wuethrich@math.ethz.ch
- Homepage: https://people.math.ethz.ch/~wueth/
- Social media: https://www.linkedin.com/in/mario-w%C3%BCthrich-aa97aa303/

LSTM-Based Coherent Mortality Forecasting for Developing Countries

Jose Garrido, Concordia University

Yuxiang Shang, Xi'an Jiaotong–Liverpool University

Ran Xu, Xi'an Jiaotong–Liverpool University (presenter)

Abstract: This paper studies a long short-term memory(LSTM)-based coherent mortality forecasting method for developing countries or regions. Many of such developing countries have experienced a rapid mortality decline over the past few decades. However, their recent mortality development trend is not necessarily driven by the same factors as their long-term behavior. Hence, we propose a time-varying mortality forecasting model based on the life expectancy and lifespan disparity gap between these developing countries and a selected benchmark group. Here, the mortality improvement trend for developing countries is expected to converge gradually to that of the benchmark group during the projection phase. More specifically, we use a unified deep neural network model with LSTM architecture to project the life expectancy and lifespan disparity difference, which further controls the rotation of the time-varying weight parameters in the model. This approach is applied to three developing countries and three developing regions. The empirical results show that this LSTM-based coherent forecasting method outperforms classical methods, especially for the long-term projections of mortality rates in developing countries.

Keywords: Coherent mortality forecasting, LSTM, Developing countries, Life expectancy, Lifespan disparity

References

1. Bohk-Ewald, Christina, Marcus Ebeling, and Roland Rau. 2017. Lifespan disparity as an additional indicator for evaluating mortality forecasts. Demography 54: 1559–77.
2. Li, Nan, and Ronald D. Lee. 2005. Coherent mortality forecasts for a group of populations: An extension of the Lee–Carter method. Demography 42: 575–94.
3. Marino, Mario, Susanna Levantesi, and Andrea Nigri. 2023. A neural approach to improve the Lee–Carter mortality density forecasts. North American Actuarial Journal 27: 148–165.

Contact details

- Email: ran.xu@xjtlu.edu.cn
- Homepage: https://scholar.xjtlu.edu.cn/en/persons/RanXu
- Repository:
- Social media:

How does granularity affect motor insurance claim predictions in a telematics setting?

Juan Sebastian Yanez, University of Barcelona (presenter)

Montserrat Guillen, University of Barcelona

Abstract: Telematics data have tremendously impacted the development of fairer and more precise motor insurance products, such as Pay-How-You-Drive (PHYD) insurance policies. This is because telematics data's accurate and granular nature allows insurers to dynamically estimate accident risk levels based on a client's driving behaviour and risk exposure. However, the collection of this information significantly deviates from traditional motor insurance policies based on non-telematics covariates, which are not observed dynamically as clients drive. This paper aims to bridge the gap between classical non-telematics models and their more modern counterparts by studying various levels of granularity in telematics data sets (e.g., data collected weekly, monthly, or annually). Therefore, it provides insight into the trade-off between more granular and dynamic telematics structures, such as weekly models, and less granular structures that are more akin to classical insurance models. This study benefits from having access to telematics data from a major Spanish insurer, allowing us to investigate driving behaviors and accident patterns among 17,405 policyholders over two years (2017-2018). The dataset includes demographic and telematics variables, providing a comprehensive basis for understanding accident frequency dynamics.

Keywords: telematics; road accident; rate-making; granularity; Pay-How-You-Drive.

References

Contact details

- Email: yanez.juan_sebastian@ub.edu
- Homepage: NA
- Repository: NA
- Social media: NA

A Sub Neural Network Approach for Forecasting Climate-Related Claim Costs in Property Insurance

Mr. Yubo Rasmussen, Heriot-Watt University (presenter)

Dr. Alex Jose, Heriot-Watt University

Dr. George Tzougas, Heriot-Watt University

Dr. Fraser Daly, Heriot-Watt University

Abstract: Given the increasing frequency and severity of climate-driven extreme weather events, it is essential to understand the underlying factors contributing to climate-related claim costs within property insurance portfolios. This study is motivated by the characteristics of a property insurance portfolio in Greece, including various property-insurance-related risk characteristics along with geographical information and regional susceptibility to extreme weather events, notably flooding. Specifically, we propose integrating mixed Pareto regression models within a neural network framework that allows for feature-dependent mean and dispersion parameters, effectively modeling the claim costs. We assess the predictive accuracy of these neural-network-based models against traditional regression-based approaches to determine the neural network models' effectiveness in capturing complex, non-linear interactions inherent in the climate-related dataset.

Keywords: Climate Change, Property and Contents Insurance, Neural Networks, Pareto

References

Contact details

- Email: yr2001@hw.ac.uk
- Social media: Linkedin

Fine-grained Mortality Forecasting with Deep Learning

Huiling Zheng, University College London (presenter)

Hai Wang, University College London

Rui Zhu, Bayes Business School

Jing-Hao Xue, University College London

Abstract: Fine-grained mortality forecasting has gained momentum in actuarial research due to its ability to capture localised, short-term fluctuations in death rates. This paper introduces MortFCNet, a deep-learning method that predicts weekly death rates using region-specific weather inputs. Unlike traditional Serfling-based methods and gradient-boosting models that rely on predefined fixed Fourier terms and manual feature engineering, MortFCNet automatically learns patterns from raw time-series data without needing explicitly defined Fourier terms or manual feature engineering. Extensive experiments across over 200 NUTS-3 regions in France, Italy and Switzerland demonstrate that MortFCNet consistently outperforms both a standard Serfling-type baseline and XGBoost in terms of predictive accuracy. Our ablation studies further confirm its ability to uncover complex relationships in the data without feature engineering. Moreover, this work underscores a new perspective of exploring deep learning for advancing fine-grained mortality forecasting.

Keywords: Fine-grained, Mortality Forecasting, XGBoost, Deep Learning, Multiple Populations.

References

1. Robben, J., Antonio, K., & Kleinow, T. (2024). The short-term association between environmental variables and mortality: Evidence from europe.
2. Richman, R., & Wüthrich, M. V. (2021). A neural network extension of the lee–carter model to multiple populations. Annals of Actuarial Science, 15(2), 346–366.
3. Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting us mortality. Journal of the American Statistical Association, 87(419), 659–671.

Contact details

- Email: huiling.zheng.16@ucl.ac.uk

The Influence of Climate Change on Insurance Sustainability: Evidence from Spanish Agricultural Insurance

Nan Zhou, Complutense University of Madrid (presenter)

José L. Vilar-Zanón, Complutense University of Madrid

Abstract: As extreme weather events become more frequent, climate change-related risk is emerging as a major concern for insurers due to its impact on insured losses and its potential threat to business sustainability. Building on the North American Actuarial Climate Index™ (ACI), our research develops and applies a nested framework of actuarial climate indices for the Iberian Peninsula -particularly Spain - that includes the Iberian Actuarial Climate Index (IACI), the Spanish Actuarial Climate Index (SACI), and the Provincial Spanish Actuarial Climate Index (pSACI). Through a comprehensive analysis of hailstorm impacts in Spain's wine-grape sector, using claims data from 1990 to 2022, we use advanced statistical methods - including regression, quantile regression, generalized linear mixed models (GLMMs), and linear quantile mixed models (LQMMs)—to quantify how incremental increases in these climate indices correlate with higher claim frequencies and more severe losses. Notably, extreme losses (e.g., at the 99th percentile) escalate disproportionately, underscoring the importance of tail-risk-sensitive approaches. Our findings highlight significant regional variability, particularly at the provincial level, and underscore the need for climate-responsive underwriting strategies and Solvency II-compliant risk measures. By offering a province-level monetization of climate change effects on insurance premiums and solvency capital requirements (SCR), this research provides actionable insights for product design, asset-liability management, and regulatory compliance. Overall, these results advocate for the transition from conventional mean-based risk models to adaptive, tail-focused methodologies, positioning scalable climate indices as crucial tools for 21st-century risk governance in the insurance industry.

Keywords: Climate Change, Insurance Sustainability, Agriculture Insurance, Actuarial Climate Index, Generalized Linear Mixed Model, Linear Quantile Mixed Model

References

1. Zhou, N., Vilar-Zanón, J. L., Garrido, J., & Heras Martínez, A.-J. (2023). On the definition of an actuarial climate index for the Iberian peninsula. *Anales Del Instituto De Actuarios Españoles*, (29), 37–59. doi:10.26360/2023_3.
2. Zhou, N. & Vilar-Zanón, J. L. (2024). Impact Assessment of Climate Change on Hailstorm Risk in Spanish Wine Grape Crop Insurance: Insights from Linear and Quantile Regressions. *Risks*, (12). doi:10.3390/risks12020020.
3. Zhou, N. & Vilar-Zanón, J. L. (2024). Climate Change and Crop Insurance: Geographical Heterogeneity in Hailstorm Risk for Wine Grapes in Spain. Under revision. Available at: https://hdl.handle.net/20.500.14352/107232

Contact details

- Email: zhounan@ucm.es

Index of presenters

Amaryllis Mouyiannou, 41
Amin Karbassi, 34
Anne VanderScheer, 60
Aurelien Couloumy, 16

Brandon Schwab, 51

Carlos Arocha, 9
Charlotte Jamotton, 33
Chris Halliwell, 30
Claudio Giancaterino, 25
Conor Goold, 27

Daniel Jakobi, 32
Despoina Makariou, 39

Emanuele Fabbiani, 20
Emilio Saenz Guillen, 49

Filip Lindskog, 36
Francesco Ungolo, 58
Freek Holvoet, 31
Fung Tsz, 56

George Tzougas, 57
Giovanni Rabitti, 45
Guangyuan Gao, 21
Gustavo Martinez, 40
Gwen Chan, 13

Hirbod Assa, 10
Huiling Zheng, 66

Jens. Robben, 47
Jose. Garrido, 22
Juan deOyarbide, 17
Juan. Yanez, 64

Karol Gawlowski, 23

Leonardo Ruggieri, 48
Lucas Muzynoski, 42

Malgorzata Smietanka, 55
Manuel Caccone, 11
Marie Cote, 15
Mario Wuthrich, 62
Mark Shoun, 53
Markus Gesmann, 24
Matej Otcenas, 43
Mathias Lindholm, 35
Mathias Valla, 59
Meryem Schalck, 50
Michael Ramati, 46
Mick Cooney, 14

Nan Zhou, 67

Olivier Lopez, 38

Paul Wilsens, 61
Pierre Goffard, 26
Pratyush Singh, 54

Quentin Guibert, 28

Ran Xu, 63
Rasmussen Yubo, 65
Robert Carruthers, 12

Salvatore Scognamiglio, 52
Samuel Gyamerah, 29

Tian Dong, 18

Yiannis Parizas, 44
Yushan Liu, 37
Yuval Dror, 19