Insurance Data Science Conference
15 - 17 June 2022

Programme and Book of Abstracts

Università Cattolica del Sacro Cuore, Milano

Scientific committee:

- Katrien Antonio (KU Leuven and University of Amsterdam)
- Christophe Dutang (Université Paris Dauphine)
- Arthur Charpentier (Université du Québec à Montréal)
- Gian Paolo Clemente (Università Cattolica del Sacro Cuore, Milano)
- Markus Gesmann (Insurance Capital Markets Research, London)
- Ioannis Kyriakou (City, University of London)
- Davide de March (Markel, London)
- Silvana Pesenti (University of Toronto)
- Rosalba Radice (City, University of London)
- Ronald Richman (SA Taxi, South Africa)
- Nino Savelli (Università Cattolica del Sacro Cuore, Milano)
- Jürg Schelldorfer (Swiss Re, Zurich)
- Giorgio Alfredo Spedicato (Leithà, Milan)
- Andreas Tsanakas (City, University of London)
- Mario Wüthrich (ETH Zurich)
- Diego Zappa (Università Cattolica del Sacro Cuore, Milano)
- Rui Zhu (City, University of London)

Organizers:

- Gian Paolo Clemente, Università Cattolica del Sacro Cuore, Milan
- Markus Gesmann, Insurance Capital Markets Research
- Nino Savelli, Università Cattolica del Sacro Cuore, Milan
- Andreas Tsanakas, City, University of London
- Diego Zappa, Università Cattolica del Sacro Cuore, Milan
- Francesco Della Corte, Università Cattolica del Sacro Cuore, Milan
- Gabriele Pittarello, Università degli Studi di Roma - La Sapienza

Venue, Università Cattolica del Sacro Cuore

The congress will take place at Università Cattolica del Sacro Cuore (Largo Gemelli 1, Milan) in the centre of Milan.

The university has a rich history, which you can read about here and here.

A brief description about what to see in Milan is available online and more information about its main monuments and tourist attractions can be found here.

Getting there

- From Malpensa airport, we suggest to take the train: Malpensa - Milano Cadorna
- From the Central Station, we suggest to take the metro M2 (green line) from the Central Station to Sant'Ambrogio.
- From Linate airport, it is possible to reach the University via bus number 73 (from Linate airport to Piazza Velasca) and then 20 minutes walk. Alternatively a shuttle connects Linate Airport to the Central Station.

More information about how to reach the campus can be found on the University web site.

Synthetic Program and Social Events

Insurance
Data
Science

Thursday, 16 June Morning

9.00-9.30	*Opening Session:* E. Beccalli, M. Gesmann, N. Savelli	
Room	G110 Gemelli	
9.30-10.30	*Parallel 1*	*Parallel 2*
Room	G110 Gemelli	G111 Salvadori
Chair	C. Blier-Wong	D. Zappa
Speaker 1	**M. Wuthrich**	**H.B. Lim**
Speaker 2	**R. M. Verschuren**	**A. Jose**
Speaker 3	**B. DC Campo**	**A. Streich**
10.30-11.00	Coffee Break	
11.00-12.00	*Plenary Session*	
Room	G110 Gemelli	
Chair	G.P. Clemente	
Invited Speaker	**J. Trufin**	
12.00-12.30	*Lightening 1*	*Lightening 2*
Room	G110 Gemelli	G111 Salvadori
Chair	G. Bormetti	F. Della Corte
Speaker 1	**J. Schelldorfer**	**H.Y.J. Yung**
Speaker 2	**S. Ng**	**A.I. Mugwe**
Speaker 3	**J. Dambon**	**A.U. Montero**
Speaker 4	**S. Flaig**	**R. Van Der Zwaan**
12.30-14.00	Lunch	

online speakers
speakers in person

Thursday, 16 June Afternoon

14.00-15.00	*Parallel 3*	*Parallel 4*	*Organized by Leithà*
Room	G110 Gemelli	G111 Salvadori	G.118 Benedetto XV
Chair	M. Pigeon	L. Delong	A. Tirri
Speaker 1	**M. Bladt**	**J. Robben**	**G. Rianna**
Speaker 2	**F. Ungolo**	**G. Pittarello**	**A. Castellarin**
Speaker 3	**S. Schnurch**	**J. Ko**	**A. Petruccelli**
15.00-16.00	*Plenary Session*		
Room	G110 Gemelli		
Chair	D. Zappa		
Invited Speaker	**M. Senn & P. Hogan**		
16.00-16.30	Coffee Break		
16.30-17.00	*Lightening 3*	*Lightening 4*	
Room	G110 Gemelli	G111 Salvadori	
Chair	O. Lopez	M. Gesmann	
Speaker 1	**J. Ponnet**	**A. Riva**	
Speaker 2	**C. Giancaterino**	**A. Zatsepin**	
Speaker 3	**G. Rabitti**	**U. Korn**	
Speaker 4	**A. Badescu**	**V. Sriram**	
17.10-18.10	*Parallel 5*	*Parallel 6*	
Room	G110 Gemelli	G111 Salvadori	
Chair	M. Gesmann	G. Bormetti	
Speaker 1	**C. Blier-Wong**	**Z. Li**	
Speaker 2	**D. Biancalana**	**Y. Havrylenko**	
Speaker 3	**M. Shoun**	**R. Pusz**	

Friday, 17 June Morning/Afternoon

9.30-10.30	*Parallel 7*	*Parallel 8*
Room	G110 Gemelli	G111 Salvadori
Chair	M. Wuthrich	R. Richman
Speaker 1	**H. Zakrisson**	**B. Wong, Y. Li**
Speaker 2	**J. Schelldofer**	**R.Metulini**
Speaker 3	**E. Menvouta**	**O. Lopez**
10.30-11.15	Coffee Break	
11.15-12.15	*Plenary Session*	
Room	G110 Gemelli	
Chair	N. Savelli	
Invited Speaker	**F. Parente**	
12.30-14.00	Lunch	
14.00-14.30	*Lightening 5*	*Lightening 6*
Room	G110 Gemelli	G111 Salvadori
Chair	M. Ludovsky	F. Baione
Speaker 1	**W.F. Chong**	**G. Stupfler**
Speaker 2	**D. Giorgi**	**S. Sangari**
Speaker 3	**S. Kessy**	**S. Levantesi**
Speaker 4	**M. Vhudzijena**	**O. Laverny**
14.40-15.40	*Parallel 9*	*Parallel 10*
Room	G110 Gemelli	G111 Salvadori
Chair	W.F. Chong	R. Schmid
Speaker 1	**M. Marino**	**A. Dræge**
Speaker 2	**S. Scognamiglio**	**K. Bett**
Speaker 3	**M. Ludkovski**	**V. Arannil**
15.40-16.00	*Closing Remarks:* A. Tsanakas, Members of Organizing Committee	

The Conference will offer two social events:

- Welcome Event will be held on Wednesday afternoon (5-7 p.m.) at Cour d'Honneur of Catholic University (Milan, Largo Gemelli 1)

- Gala Dinner will be held on Thursday evening (from 8-8.30 p.m.) at Palazzo Visconti di Modrone (Milan, Via Cino del Duca 8).
  The location of the Gala Dinner can be reached from the university by metro (red line from Cadorna to San Babila) or by walk (25-30 minutes approximatively).

Timetable (CEST)

Thursday 16 June, 2022

9.00-9.30 Opening: Room G.110 Gemelli

Elena Beccalli – Dean of the faculty of Banking, Finance and Insurance
Markus Gesmann – Member of the Scientific committee of IDS
Nino Savelli – Member of the Scientific committee of IDS and head of the organizing committee of IDS in Catholic University

9.30-10.30 Regular Sessions

Regular Session 1 – Room G.110 Gemelli	**Regular Session 2 – Room G.111 Salvadori**
Hybrid Mode – Speakers in-person	**Speakers partially on-line partially in-person**
Chair: **Christopher Blier-Wong**	Chair: **Diego Zappa**
M. Wüthrich, Eth Zurich, *Deep Composite Regression Model*	**H.B. Lim**, The University of Iowa, *Mortality Forecasting with Neural Tangent Kernel Regression* (with N.D. Shyamalkumar, S. Tao) on-line
R. M. Verschuren, University of Amsterdam, *Frequency-Severity Experience Rating based on Latent Markovian Risk Profiles*	**A. Jose**, Heriot-Watt University, *Predictive modelling for admission rates related to respiratory diseases in the US* (with A. Macdonald, G. Tzougas, G. Streftaris) on-line
B. DC Campo, KU Leuven, *A data-driven approach to construct a hierarchical structure* (with K. Antonio, J. Robben)	**A. Streich**, Analytics, Life & Health Europe at PartnerRe, *A Stochastic Compartmental Model for Pandemic Risk Assessment* (with F. Gomez, P. Hogan) in-person

10.30-11.00 Coffee Break

11.00-12.00 Plenary Session – Room Gemelli

Chair: **Gian Paolo Clemente**

Invited Speaker: Julien Trufin, Université libre de Bruxelles, *Non-life insurance pricing: boosting trees and diagnostic tools to compare competing models.*

12.00-12.30 Lightning Sessions

Lightning Session 1 – G.110 Gemelli	**Lightning Session 2 – Room G.111 Salvadori**
Hybrid Mode – Speakers in-person	**Speakers partially on-line partially in-person**
Chair: **Giacomo Bormetti**	Chair: **Francesco Della Corte**
J. Schelldorfer, Swiss Re, *A discussion on the similarities and differences between insurance risk modelling and standard machine learning techniques*	**H. Y. J. Yung**, UNSW Sydney, *Modelling the mortality for China's oldest-old* (with K. Hanewald, A. Villegas) - on-line
S. Ng, Vantage Risk, *Applying Simple String Matching (NLP) in Casualty and Specialty Reinsurance Pricing and Risk Management*	**A. I. Mugwe**, Strathmore University, *Enhancing Food Security in Africa with a Predictive Early Warning System on Extreme Weather Phenomena* (with J, Gachanja, B. Muriithi, J. Olukuru, A. Wairegi, I. Rutenberg) – on-line

J. Dambon, Swiss Re, *Modeling Container Shipping Delay with Random Effects: A Comparison of Methods*	**A. U. Montero**, Université de Lausanne, *Cost of Dying in Switzerland: Taking a Glimpse of Medical Expenses in the Last Year of Life* (with J. Wagner) on-line
S. Flaig, University of Oldenburg, *Scenario generation for market risk models using generative neural networks* (with G. Junike)	**R. van der Zwaan**, MavenBlue, *Need for Speed - GPU acceleration for insurance* (with J. van Bruggen, M. Smith) – in person

12.30-14.00 Lunch

14.00-15.00 Regular and Organized Sessions

Regular Session 3 G.110 Gemelli	**Regular Session 4 G.111 Salvadori**	**Organized Session 1 by Leithà Room G.118 Benedetto XV**
Hybrid Mode – Speakers partially online partially in-person	**Hybrid Mode – Speakers in-person**	**Hybrid Mode – Speakers in-person**
Chair: **Mathieu Pigeon**	Chair: **Lukasz Delong**	Chair **Antonio Tirri**
M. Bladt, University of Lausanne, *Matrix regression: models, algorithms, and applications (*with H. Albrecher, M. Bladt, J. Yslas) – in person	**J. Robben**, KU Leuven, *A hierarchical reserving model for reported non-life insurance claims* (with J. Creveceour, K. Antonio)	**G. Rianna**, Fondazione CMCC Centro Euromediterraneo sui Cambiamenti Climatici, *European Extreme Events Climate Index (E3CI)* (with A. Tirri, F. Repola, F. Lo Conti, G. Barbato, P. Mercogliano, G.A. Spedicato)
F. Ungolo, Technische Universität München, *Affine_mortality: R tools for estimation, comparison and projection of affine mortality models* (with M. Sherris, L. P. D. M. Garces, Y. Zhou) – in person	**G. Pittarello**, Università degli Studi – La Sapienza, *Bayesian Neural Networks applied to individual Chain-Ladder reserving* (with G.P. Clemente, D. Zappa)	**A. Castellarin**, Università di Bologna, Geomorphic flood hazard mapping: from floodplain delineation to flood-hazard characterization (with A. Magnini, M. Lombardi, A. Bujari, P. Mattivi, M. Patella, G. Bitelli, F. Lo Conti, A. Tirri)
S. Schnürch, Fraunhofer Institute for Industrial Mathematics, *Accounting for COVID-19-Type Shocks in Mortality Modeling: A Comparative Study* (with T. Kleinow, A. Wagner) on-line	**J. Ko**, SAS Institute, *Claims reserving, simulation engines* (with B. Fannin)	**A. Petruccelli**, Leithà, Terraferma: an interactive tool for insurance seismic risk awareness in Italy (with A. Tirri, L. Ferraresi)

15.00-16.00 Plenary Session: Room G.110 Gemelli

Chair: **Diego Zappa**

Invited Speaker: Markus Senn, Head of Analytics, Life & Health Europe at Partner Re **& Patrick Hogan**, Senior Data Scientist at Partner Re, *Some like it Bayesian: The allure, obstacles, and rewards*

16.00-16.30 Coffee Break

16.30-17.00 Lightning Sessions

Lightning Session 3 – Room G.110 Gemelli	**Lightning Session 4 – Room G.111 Salvadori**
Hybrid Mode – Speakers in-person	**Speakers on-line**
Chair: **Olivier Lopez**	Chair: **Markus Gesmann**
J. Ponnet, KU Leuven, *Estimation of the enhanced concordance probability in linearithmic time* (with J.Raymaekers, R. Vanoirbeek, T. Verdonck)	**A. Riva**, Università degli Studi di Roma – La Sapienza, *Strategy optimization in a dynamical financial analysis environment through evolutionary reinforcement learning*
C. Giancaterino, Catholic University of Milan, *Machine Learning Interpretability in Lapse Prediction for Non-Life Insurance Premium*	**A. Zatsepin**, VSK insurance company, *Reserves, tariff rates, portfolio management. All in One: Machine Learning + stochastic loss reserving* (with A. Kvitchenño)
G. Rabitti, Heriot-Watt University, *Bottom-up construction of rating system using sensitivity measures* (with A. Vallarino, A.K. Chokami)	**U. Korn**, Ledger Investing, *A New Approach to Forecasting Insurance Loss Ratios*
A. Badescu, University of Toronto, *On the Logit-weighted Reduced Mixture of Experts models with insurance applications* (with S. T. Chai Fung, S. Lin)	**V. Sriram**, Guy Carpenter, *AI Systems for Insurance Data Prep* (with J, Fan, N. Liu)

17.10-18.10 Regular Sessions

Regular Session 5 **Room G.110 Gemelli**	**Regular Session 6** **Room G.111 Salvadori**
Hybrid Mode – Speakers in-person	**Speakers on-line**
Chair: **Markus Gesmann**	Chair: **Giacomo Bormetti**
C. Blier-Wong, Université Laval, *Insurance ratemaking with images* (with H. Cossette, L. Lamontagne, E. Marceu)	**Z. Li**, *A general framework for modelling claim count data in general insurance based on the local mixed Poisson net* (with G. Tzougas)
D. Biancalana, Università degli Studi di Roma – La Sapienza, *Health insurance claims prediction with GAMLSS* (with F. Baione)	**Y. Havrylenko**, Technical University of Munich, *Algorithmic detection of interacting variables for generalized linear models via neural networks* (with J. Heger)
M. Shoun, Ledger Investing, *Domain-Specific Languages for Reserve Modeling*	**R. Pusz**, *Warsaw School of Economics*, Pure premium calculation for flood risk based on spatial information using R

Friday, 17 June, 2022

9.30-10.30 Regular Sessions

Regular Session 7 – Room G.110 Gemelli	**Regular Session 8 – Room G.111 Salvadori**
Hybrid Mode – Speakers in-person	**Hybrid Mode – Speakers partially in-person and partially on-line**
Chair: **Mario Wuthrich**	Chair: **Ronald Richman**
H. Zakrisson, Stockholm University, *A Collective Reserving Model With Claim Openness* (with M. Lindholm)	**B. Wong & Y. Li,** University of New South Wales, *Stochastic Ensemble Loss Reserving* (with B. Avanzi, A. Xian) – in-person
J. Schelldorfer, Swiss Re, *LocalGLMnet: A Deep Learning Architecture for Actuaries*	**R. Metulini**, University of Salerno, *Forecasting flood risk exposure using mobile phone traffic flows' data,* (with M. Carpita) – in-person
E. J. Menvouta, KU Leuven, *Comparing machine learning models for micro-level reserving* (with R. Vanoirbeek, T. Verdonck)	**O. Lopez**, Sorbonne University and Detralytics, *Identification of the network structure to evaluate the impact on cyber attacks on an insurance portfolio* (with C. Hillairet, L. d'Oultremont, B. Spoorenberg, M. Thomas) – in-person

10.30-11.15 Coffee Break

11.15-12.15 Plenary Session: Room G.110 Gemelli

Chair: **Nino Savelli**

Invited Speaker: Fausto Parente, Executive Director of the European Insurance and Occupational Pensions Authority (*EIOPA*), *AI, data and insurance: Protecting policyholders*

12.15-14.00 Lunch

14.00-14.30 Lightning Sessions

Lightning Session 5 – Room G.110 Gemelli	**Lightning Session 6 – Room G.111 Salvadori**
Hybrid Mode – Speakers partially in-person and partially on-line	**Speakers on-line**
Chair: **Michael Ludkovski**	Chair: **Fabio Baione**
W. F. Chong, Heriot-Watt University, *Pseudo-Model-Free Hedging for Variable Annuities via Deep Reinforcement Learning*, (with H. Cui, Y. Li) – in-person	**G. Stupfler**, ENSAI & CREST, Extreme conditional risk estimation in heavy-tailed heteroscedastic regression models (with S. Girard, A. Usseglio-Carleve)
D. Giorgi, CNRS, Sorbonne Université, *IBMPopSim: a package for the efficient simulation of individual-based population models* (with S. Kaakaï, V. Lemaire) – in-person	**S. Sangari**, Kennesaw State University, *Under-reporting correction in Cyber Incidents* (with E. Dallal)
S. R. Kessy, University of New South Wales, *Combination of Mortality Rate Forecasts From Multiple Starting Points* (with M. Sherris, A. Villegas, J. Ziveyi) – online	**S. Levantesi**, Sapienza University of Rome, *Multi-country clustering-based forecasting of healthy life expectancy* (with A. Negri, G. Piscopo)
M. Vhudzijena, UNSW Sydney, *Mortality Heterogeneity and Clustering using Joint Body Mass Index and Self-Reported Health Trajectories* (with M. Sherris, A. Villegas, J. Ziveyi) - online	**O. Laverny**, University of Lyon & SCOR SE, *Estimation of high dimensional gamma convolutions through random projections*

14.40-15.40 Regular Sessions

Regular Session 9 **Room G.110 Gemelli**	**Regular Session 10** **Room G.111 Salvadori**
Hybrid Mode – Speakers in-person	**Speakers on-line**
Chair: **Wing Fung Chong**	Chair: **Roland Schmid**
M. Marino, Sapienza University of Rome, *Transfer learning for boosting mortality table* (with G.A. Spedicato)	**A. Dræge**, Frende Forsikring, *Using cosine similarity for recommending insurance products* (with H. Midtgarden Golid, F. Dorn)
S. Scognamiglio, University of Naples "Parthenope", *Calibrating the Lee-Carter and the Poisson Lee-Carter models via Neural Networks*	**K. Bett**, Strathmore University, *Weather index-based Crop Insurance using Machine Learning*, J. Olukuru
M. Ludkovski, University of California Santa Barbara, *Joint modeling of State-level mortality in US* (with D. Padilla)	**V. Arannil**, Amazon Web Services, *Applying computer vision for high precision 360 degree car damage assessment* (with A. Roy)

15.40-16.00 Closing Remarks: Room G.110 Gemelli

Andreas Tsanakas– Member of the Scientific committee of IDS
Gian Paolo Clemente, Francesco Della Corte, Silvia Facchinetti, Gabriele Pittarello, Nino Savelli, Diego Zappa - Local Organizing committee of IDS

Contents

1 Regular Session 1

Thursday, 9.30-10.30
Room G110 Gemelli
Chair: *Christopher Blier-Wong*
Mode: *Hybrid - speakers in person*

1.1 Deep Composite Regression Model

Mario Wuthrich, ETH Zurich

Abstract: A main difficulty in actuarial claim size modeling is that there is no simple off-the-shelf distribution that simultaneously provides a good distributional model for the main body and the tail of the data. In particular, covariates may have different effects for small and for large claim sizes. To cope with this problem, we discuss a deep composite regression model whose splicing point is given in terms of a quantile of the conditional claim size distribution rather than a constant. To facilitate M-estimation in such models, we consider and characterize the class of strictly consistent scoring functions for the triplet consisting of the quantile, as well as the lower and upper expected shortfall beyond that quantile. In a second step, this elicitability result is applied to fit deep neural network regression models.

Keywords: Deep composite regression, elicitability, splicing model.

References

1. Fissler, T., Merz, M., Wuthrich, M.V. (2021). Deep quantile and deep composite model regression. arXiv:2112.03075

Contact information

- Email: mario.wuethrich@math.ethz.ch
- Homepage: https://people.math.ethz.ch/ wmario/

1.2 Frequency-Severity Experience Rating based on Latent Markovian Risk Profiles

Robert Matthijs Verschuren, University of Amsterdam

Abstract: Bonus-Malus Systems are widely employed as a commercial form of experience rating within non-life insurance. However, these systems traditionally consider a customer's number of claims irrespective of their sizes, even though these components are dependent in practice. We therefore propose a novel joint experience rating approach based on latent Markovian risk profiles in this paper to allow for a positive or negative individual frequency-severity dependence. The latent profiles evolve over time in a Hidden Markov Model to capture updates in a customer's claims experience, making claim counts and sizes conditionally independent. We show that the resulting risk premia lead to a dynamic, claims experience-weighted mixture of standard credibility premia, where the profile assignment probabilities include a customer's observed claims experience *a posteriori*. We additionally allow a customer's risk characteristics to affect these credibility premia, and hence their posterior weights, and efficiently estimate all (prior) parameters through an empirical Bayes version of the Baum- Welch algorithm. The proposed joint experience rating approach is applied to a Dutch automobile insurance portfolio and identifies customer risk profiles with distinctive claiming behavior. These profiles, in turn, enable us to better distinguish between customer risks.

Keywords: Experience rating, frequency-severity dependence, dynamic latent risk profiles, Hidden Markov Model, Expectation-Maximization, automobile insurance.

Contact information

- Email: r.m.verschuren@uva.nl
- Homepage: www.uva.nl/profile/r.m.verschuren
- LinkedIn: (https://www.linkedin.com/in/robertverschuren)

1.3 A data-driven approach to construct a hierarchical structure

Bavo DC Campo, KU Leuven (presenter)
Katrien Antonio, KU Leuven
Jens Robben, KU Leuven

Abstract: In a workers' compensation product, we need to account for hierarchically structured risk factors with a large number of categories such as the industry- and branch-code. The industry code commonly identifies the first hierarchical level and the branch code the second level, nested within the first level. Within statistics, we use random effect models to handle these type of risk factors and within machine learning, we commonly incorporate these using an encoding scheme [1] or embeddings [2]. One of the drawbacks of these approaches is that they regard the hierarchical structure as fixed and result in a numerical representation for each separate category. We present a data-driven hierarchical clustering method that effectively groups categories with a similar risk profile at different hierarchical levels. We outline both a top-down (starts at the industry-level) and bottom-up (starts at the branch-level) approach to construct a data-driven hierarchical structure. Both approaches apply the following procedure at each hierarchical level. First, we engineer features that convey information on the risk profile of each category. Hereto, we use estimated random effects obtained with a generalized linear mixed model for damage rates on the one hand and claim frequencies on the other hand. Next, we also use embeddings [3, 4] to extract textual information. These features are used as input in a clustering algorithm (e.g. hierarchical cluster analysis [5]), which groups similar categories together.

Keywords: Nested classification, Supervised learning, Unsupervised learning, Clustering

References:

1. Micci-Barreca, D. (2001). A preprocessing scheme for high-cardinality categorical attributes in classification and prediction problems. *SIGKDD explorations* **3(1)**, 27-32

2. Kang, W.C., Cheng, D.Z., Yao, T., Yi, X., Chen, T., Hong, L., Chi, E.H. (2020). Learning to embed categorical features without embedding tables for recommendation. *arXiv*

3. Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv*

4. Cer, D., Yang, Y., Kong, S., Hua, N., Limtiaco, N., John, R.S., Constant, N., Guajardo-Cespedes, M., Yuan, S., Tar, C., Sung, Y.H., Strope, B., Kurzweil, R. (2018). Universal sentence encoder. *arXiv*

5. Hartigan, J.A. (1975). *Clustering Algorithms.* Wiley.

Contact details:

- Homepage: https://bavodc.github.io/

- Repository: https://github.com/BavoDC

- Social media: https://be.linkedin.com/in/bavo-de-cock-campo-584087b6

2 Regular Session 2

Thursday, 9.30-10.30
Room G111 Salvadori
Chair: *Diego Zappa*
Mode: *Hybrid - speakers partially on-line, partially in-person*

2.1 Mortality Forecasting with Neural Tangent Kernel Regression

Hong Beng (Ben) Lim (presenter), The University of Iowa
N. D. Shyamalkumar, The University of Iowa
Siyang Tao, Ball State University

Abstract: Best practices concerning neural network architectures tend to be application-specific, which is a significant hurdle for neural network applications in actuarial science. For many network architectures, kernel regression using Neural Tangent Kernels (NTKs) is equivalent to training neural networks with layers of infinite width [1]. NTK regression is significantly faster than training its finite-width counterpart, and empirically it has also consistently delivered comparable performance. Importantly, NTK regression sidesteps the need to specify the number of nodes in each layer while simplifying the training process. This talk explores the pros and cons of using NTK regression in lieu of finite-width neural network architectures for mortality forecasting [2], an area of active interest in actuarial science.

Keywords:Recurrent neural network, deep learning, kernel regression

References:

1. Jacot, A., Gabriel, F., and Hongler, C. (2018). Neural tangent kernel: Convergence and generalization in neural networks. *Advances in Neural Information Processing Systems (NIPS), 2018*:8571-8580.

2. Richman, R. and Wuuml;thrich, M.V. (2019). Lee and Carter go Machine Learning: Recurrent Neural Networks. Tutorial, SSRN. URL https://ssrn.com/abstract=3441030.

Contact details:

- Email: hongbeng-lim@uiowa.edu

2.2 Predictive modelling for admission rates related to respiratory diseases in the US

Alex Jose, Heriot-Watt University (presenter)
Angus S Macdonald, Heriot-Watt University
George Tzougas, Heriot-Watt University
George Streftaris, Heriot-Watt University

Abstract: In this work we investigate rates of admission to hospital (or other health facility) due to respiratory diseases in a US working population, and their dependence on a number of demographic and health insurance related factors. In earlier research, a number of generalized-linear-type predictive models have been fitted to the number of admissions and to analogous events in critical illness data, under both frequentist and Bayesian approaches, see, for instance, Ozkok-Dodd et al. (2015) and [1]. Here, we follow the Combined Actuarial Neural Network (CANN) approach of [2] and we assume that admission numbers can be modeled by embedding both the Poisson and Negative Binomial count regression models. For the latter, see [5]. The aim is to explore the gains in predictive power obtained by taking into account missing interactions of non-multiplicative type. When nesting the count regression models into networks, we consider two approaches: (i) we pre-calibrate the linear component of the models, as suggested by [2] and (ii) we calibrate both the linear and non-linear components simultaneously. The results from our numerical illustration indicate that CANN approach (ii) provides an improved predictive performance compared to the count regression models, plain vanilla neural network models and CANN approach (i). Finally, following the setup of [3], we used nagging predictors averaging over random calibrations of the neural network-based models to provide more accurate and stable predictions from a single run.

Keywords: predictive modelling, neural network, actuarial, morbidity, CANN

References

1. Arik, A., Dodd, E., Cairns, A., Streftaris, G. (2021). Socioeconomic disparities in cancer incidence and mortality in England and the impact of age-at-diagnosis on cancer mortality. Plos one, 16(7), e0253854.

2. Wüthrich, M.V., Merz, M. (2019). Yes, we CANN!. ASTIN Bulletin: The Journal of the IAA, 49(1), 1-3.

3. Richman, R., Wüthrich, M. V. (2020). Nagging predictors. Risks, 8(3), 83.

4. Schelldorfer, J., Wuthrich, M. V. (2019). Nesting classical actuarial models into neural networks. Available at SSRN 3320525.

5. Tzougas, G. and Li, Z. (2022). Neural Network Embedding of the Mixed Poisson Regression Model for Claim Counts.

Contact details

- Email: aj61@hw.ac.uk

2.3 A Stochastic Compartmental Model for Pandemic Risk Assessment

Andreas Streich, Analytics, Life Health Europe at PartnerRe (presenter)
Florian Gomez, Head of Global Methodology, Life & Health at PartnerRe
Patrick Hogan, Analytics, Life & Health Europe at PartnerRe

Abstract: Pandemics are a main driver of losses for both life and non-life (re-)insurance. Until COVID-19, the risk assessment mostly relied on the few historic examples, most importantly the Spanish flu. The COVID-19 ongoing COVID-19 pandemic has shown that other pathogens (namely the corona viruses) need to be taken into account. Also, modern societies and governments drastically respond to novel threats and implement measures such as social distancing, mask mandates, and lockdowns. Besides the properties of the pathogen itself, these reactions and medical progress (namely the development of vaccines) are key influences on the course of a pandemic, and thus also the resulting outcomes.

To account for the new insights from the current COVID-19 pandemic, we have built a stochastic, discrete, age-structured compartmental model as basis for the study of pandemic risk. The squire package [1] was used as starting point, and the model heavily relies on the package odin [2], which allows to reproducibly simulate stochastic state-space models. We calibrated the model on the recent pandemics and found that the characteristics of societies' reaction are quite stable over time. This in turn led us to introduce several parameter sets to model different levels of reaction speed and stringency between groups of countries. For other important factors such as pathogen parameters and medical progress, parameter distributions have been defined in close collaboration with medical experts.

The framework allows us to sample a large number of potential pathogens. For each of them, we repeatedly run the stochastic model to gather the diverse courses of the pandemic. Together, this provides many potential outcomes, which we use as a basis to derive both summary and tail statistics to assess pandemic risk. Additionally, through the stochastic nature of our approach, also complex deals can be accurately priced.

Keywords: Pandemic Risk, compartmental modelling, stochastic simulation

References:

1. https://github.com/mrc-ide/squire

2. FitzJohn, R. G., Knock, E. S., Whittles, L. K., Perez-Guzman, P. N., Bhatia, S., Guntoro, F., Watson, O. J., Whittaker, C., Ferguson, N. M., Cori, A., Baguelin, M., & Lees, J. A. (2021). Reproducible parallel inference and simulation of stochastic state space models using odin, dust, and mcstate. Wellcome Open Research, 5, 288. https://doi.org/10.12688/wellcomeopenres.16466.2

Contact details:

- Email: andreas.streich@partnerre.com

3 Plenary Session

Thursday, 11.00-12.00
Room G110 Gemelli
Chair: *Gian Paolo Clemente*
Mode: *Hybrid - speaker in person*
Invited Speaker: *Julien Trufin*

3.1 Non-life insurance pricing: boosting trees and diagnostic tools to compare competing models.

E Julien Trufin

Abstract: Thanks to its outstanding performances, boosting has rapidly gained wide acceptance among actuaries. To speed up calculations, boosting is often applied to gradients of the loss function, not to responses (hence the name gradient boosting). When the model is trained by minimizing Poisson deviance, this amounts to apply the least-squares principle to raw residuals. This exposes gradient boosting to the same problems that lead to replace least-squares with Poisson Generalized LinearModels (GLM) to analyze low counts (typically, the number of reported claims at policy level in lines). In a first time, this talk shows that boosting can be conducted directly on the response under Tweedie loss function and log-link, by adapting the weights at each step. Modern data science tools are effective to produce predictions that strongly correlate with responses. Model comparison can therefore be based on the strength of dependence between responses and their predictions. Positive expectation dependence turns out to be attractive in that respect. In a second time, the present talk proposes an effective testing procedure for this dependence concept and applies it to compare two models. Empirical illustrations using insurance loss data demonstrate the relevance of the approach for model selection in supervised learning. The most positively expectation dependent predictor can then be autocalibrated to obtain its balance-corrected version that appears to be optimal with respect to Bregman, or forecast dominance. Under autocalibration, it is shown that Lorenz curve and concentration curve coincide and that the integral of the concentration curve is equivalent to Gini coefficient.

4 Lightning Session 1

Thursday, 12.00-12.30
Room G110 Gemelli
Chair: *Giacomo Bormetti*
Mode: *Hybrid - speakers in person*

4.1 A discussion on the similarities and differences between insurance risk modelling and standard machine learning techniques.

Juerg Schelldorfer, Swiss Re (presenter)

Abstract: In industry, it is regularly encountered that there is some confusion about the similarities and differences between insurance risk modelling and standard machine learning techniques. In academy, the papers of B. Efron (2020), G. Shmueli (2010) and L. Breiman (2001) are addressing the topic. We have developed a simple overview table which helps to eliminate most of these ambiguities. In this talk, we discuss this table while covering quantitative and qualitative aspects. The talk will consist only of a single slide for the whole talk. The goal of the presentation is to help actuaries, data scientists, management and insurance specialists to better understand the conceptual commonalities and differences between the two disciplines. The table shall also serve to decide which approach is appropriate for a given modelling challenge.

The table looks as in figure 4.1:

	Insurance Risk Modeling	Standard Machine Learning
Foundation	Distribution and uncertainty	Point estimate and algorithm
Mathematical foundation	Statistical model	Numerical optimization
Modelling target	Probabilistic forecast	Point forecast
Statistical distributions	Non-Gaussian (asymmetric, skewed)	Gaussian (symmetric)
Signal-to-noise (SNR) ratio	Small	High
Mathematical model selection «criteria»	• Predictability (in-sample) • Stability and robustness (long-term) • Smoothness • Parsimony • Interpretability / explainability • -	• Predicatability (out-of-sample) • Stability and robustness (short-term) • - • Anti-parsimony • Black-box • Computability
Non-mathematical model selection «criteria»	• Causality / truth between predictors and predictant • Inclusion of expert knowledge • Human adjustability of models	• Correlation, train/test paradigm • - • -
Non-technical considerations	• Regulatory framework • Political and social aspects	• Ethics and fairnesss • Accountability and transparency
Professional associations	• Professional standards	• Ethical codes of conduct

This work has been done as part of the "Data Science" working group of the Swiss Association of Actuaries (SAA). The group publishes material that discuss the use of machine learning techniques for actuarial applications, see the group's website www.actuarialdatascience.org for further information.

Keywords: Risk Modelling, Machine Learning, prediction, estimation, Data Science

References:

1. B. Efron (2020) Prediction, Estimation, and Attribution. *Journal of the American Statistical Association* **115:539**, 636-655

2. G. Shmueli (2010) To explain or to Predict? *Statistical Science* **25/3**, 289-310

3. L. Breimann (2001) Statistical Modeling: The Two Cultures. *Statistical Science* **16/3**, 199-215

Contact details

- Email: Juerg_Schelldorfer@swissre.com
- Homepage: www.actuarialdatascience.org
- Repository: https://github.com/JSchelldorfer/ActuarialDataScience

4.2 Applying Simple String Matching (NLP) in Casualty and Specialty Reinsurance Pricing and Risk Management

Shirley Ng (presenter), Vantage Risk

Disclaimer: The views and opinions expressed in this article are those of the authors and do not necessarily the postings, strategies or opinions of Vantage Risk.

Abstract: In reinsurance aggregation roll-up time, significant manpower goes into matching the entity of the underlying insureds for each cedent. Each submission is unique and they do not always follow the same structure. String Matching is a billion dollar question in Reinsurance Aggregation. There is one insurtech start up selling a SaaS product based on string matching energy rigs. Numerous vendors sell their string matching products, with each of them them claiming to have a secret sauce. Knowing and understanding your data is key to string matching. Since each company knows its data best, I would like to go through some simple NLP techniques that you can apply to achieve the same result, if not better, without paying a premium. Google have one of the best address matching algorithms. I will not go into this area. The focus is on lines of business outside Property Catastrophe and Terrorism, such as Marine, Energy, Transport, Liability and Cyber insurance. The NLP algorithm is simple yet powerful. It can improve work efficiency and enable data enrichment. You can apply the same techniques at the point of submission for pricing. The algorithms are widely used in the UK personal lines market. When an insurer get a quote for car / home insurance in the UK, the aggregator websites pre-populate information once they know your car registration number / home address. By the end of this presentation, you will know how to build your own string matching algorithm.

Keywords: Natural Language Processing, Reinsurance, Marine, Energy, Cyber, Liability, String Matching, Distance Matrix, Clustering

References:

1. Kevin P. Murphy (2012). *Machine Learning: A Probabilistic Perspective, Chapter 25 Clustering*The MIT Press

2. scikit-learn developers (2007-2021).

3. Mark P.J. van der Loo (2014). *The stringdist Package for Approximate String Matching* The R Journal Vol. 6/1, June 2014

Contact details:

- Email: shirley.ng21@imperial.ac.uk

4.3 Modeling Container Shipping Delay with Random Effects: A Comparison of Methods

Jakob Dambon, Swiss Re, Zurich (presenter)

Abstract: Over the last years, the global supply chain experienced several disruptions, which lead to delays of container shipments. From an insurance's and logistics' perspective, it is vital to recognize key factors of delay as they are associated with demurrage fees and further disruptions in the supply chain. Here, the interpretability of such delay models is essential to gain insights and to prevent future delays. In an applied use case, we analyze the shipment data of an international operating retailer. We compare a wide range of generalized linear fixed and mixed effect models as well as more difficult to interpret boosting-based approaches with respect to two properties. First, we are interested in the key drivers of delay and the over all model interpretability. Second, we compare the predictive accuracy of all models in cross validation. We showcase the advantages of generalized linear mixed effect models (GLMM) over similar fixed effect models in terms of interpretability and predictive accuracy. While the GLMM does not reach the same predictive performance as boosting approaches, it allows us to draw more precise conclusions about the main driving factors of container delays.

Keywords: Generalized Linear Mixed Effect Models (GLMM), Supply Chain

References:

1. Chen, T., Guestin, C. (2016). XGBoost: A Scalable Tree Boosting System *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 781-794 DOI

2. Jiang, J. (2007). *Linear and Generalized Linear Mixed Models and Their Applications* Springer.

3. Sigrist, F. (2021). Gaussian Process Boosting. *Preprint* Arxiv

4. Viellechner, A., Spinler, S. (2020). Novel Data Analytics Meets Conventional Container Shipping: Predicting Delays by Comparing Various Machine Learning Algorithms *Proceedings of the 53rd Hawaii International Conference on System Sciences — 2020* 1278-1287.

Contact details:

- Email: jakob_dambon@swissre.com

- Linkedin: https://www.linkedin.com/in/jakob-dambon/)

-Twitter: https://twitter.com/jakobdambon

4.4 Scenario generation for market risk models using generative neural networks

Solveig Flaig, Deutsche Rückversicherung AG, University of Oldenburg, Germany (presenter)
Gero Junike, University of Oldenburg, Germany

Abstract: Insurance companies calculating market risk under Solvency 2 with an internal model are obliged to generate financial market scenarios. Traditionally, this is solved by using a Monte-Carlo simulation with financial mathematical models in an economic scenario generator (ESG). We present here Generative Adversarial Networks (GANs) as an alternative solution. For optimization of the GAN infrastructure, hyperparameter optimization and validation, we develop new performance measures and provide a consistent, assumption-free framework for the evaluation of the scenario generation. The main properties of these evaluation measures, mainly based on nearest-neighbour distances, are presented.

Finally, we compare the results of a GAN-based ESG to the classical ESG approach. For this comparison, we use the benchmark portfolios created by EIOPA for the market and credit risk comparison study (MCRCS). Those portfolios are available for asset- and liability side and represent average exposure profiles of insurance companies in Europe.

The comparison shows that the GAN-based ESG leads to similar results than the current approaches used in regulatory approved internal market risk models in Europe. In comparison to the currently used ESGs, they have the advantage that they are assumption-free, are able to model more complex dependencies, have no need of a time-consuming calibration process and can include new risk factors easily.

Keywords: Generative Adversarial Networks, Economic Scenario Generators, nearest neighbour distance, market risk modelling, Solvency 2

References:

1. Ian Goodfellow, J. Pouget-Abadie, and M. Mirza et al (2014). Generative adversarial nets. Advances in neural information processing systems.

2. Magnus Wiese, Robert Knobloch, Ralf Korn and Peter Kretschmer (2020). Quant gans: Deep generation of financial time series. Quantitative Finance, 20(9):1419–1440.

3. Rao Fu, Jie Chen, Shutian Zeng, Yiping Zhuang, and Agus Sudjianto (2019).

Time series simulation by conditional generative adversarial net. arXiv preprint arXiv:1904.11419.

4. Mark F Schilling (1986). Multivariate two-sample tests based on nearest neighbors. Journal of the American Statistical Association, 81(395):799–806.

Contact details:

- Email: solveig.flaig@deutscherueck.de

5 Lightning Session 2

Thursday, 12.00-12.30
Room G111 Salvadori
Chair: *Francesco Della Corte*
Mode: *Hybrid - speakers partially online, partially in person*

5.1 Modelling the mortality for China's oldest-old

Ho Yan Joey Yung, School of Risk Actuarial Studies, UNSW Sydney; Australian Research Council (ARC) Centre of Excellence in Population Ageing Research (CEPAR) (presenter)
Katja Hanewald, School of Risk Actuarial Studies, UNSW Sydney; Australian Research Council (ARC) Centre of Excellence in Population Ageing Research (CEPAR)
Andrés Villegas, School of Risk Actuarial Studies, UNSW Sydney; Australian Research Council (ARC) Centre of Excellence in Population Ageing Research (CEPAR)

Abstract: We develop a new modelling framework for the mortality of the oldest-old, the population aged 80 and over. We propose a multifactor model that combines a classic parametric oldest-old mortality model, such as the Kannisto model, with the survival tree. To improve the model accuracy, we implement the survival tree in the ensemble bagging technique. We formulate the framework to apply to left-truncated and right-censored data. To illustrate the use of the model, we apply the model to individual-level data for individuals aged 80 to 115 from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) from 1998 to 2018. Model comparisons show that our proposed model outperforms other candidate models in fitting and prediction based on the oldest-old sample from the CLHLS. We find that in the CLHLS sample, the rate of increase in the force of mortality decelerates with age at around age 105. In addition, we analyse the impact of different covariates on the oldest-old mortality rate. We report and visualise the new estimated force of mortality for a range of subgroups based on different covariates.

Keywords: multifactor mortality model, oldest old, survival tree

Contact details:

- Email: hyjoeyyung@gmail.com

5.2 Enhancing Food Security in Africa with a Predictive Early Warning System on Extreme Weather Phenomena

Alvin Igobwa, Strathmore University(presenter)
Jeremy Gachanja, Strathmore University; Betsy Muriithi, Strathmore University
John Olukuru, Strathmore University
Angeline Wairegi, Centre for Intellectual Property and Information Technology Law (CIPIT)
Isaac Rutenberg, Centre for Intellectual Property and Information Technology Law (CIPIT);

Abstract: Climate change is predicted to exacerbate Africa's, already, precarious food security. Climate models, by accurately forecasting future weather events, can be a critical tool in developing countermeasures to reduce crop loss [1,2], decrease adverse effects on animal husbandry and fishing, and even help insurance companies determine risk for agricultural insurance policies – a measure of risk reduction in the agricultural sector that is gaining prominence. In this paper, we investigate the efficacy of various open-source climate change models and weather datasets in predicting drought and flood weather patterns in northern and western Kenya and discuss practical applications of these tools in the country's agricultural insurance sector. We identified two models that may be used to predict flood and drought events in these regions. The combination of Artificial Neural Networks (ANNs) and weather station data was the most effective in predicting future drought occurrences in Turkana and Wajir with accuracies ranging from 78% to 90%. In the case of flood forecasting, Isolation Forests models using weather station data had the best overall performance. The above models and datasets may form the basis of a more objective and accurate underwriting process for agricultural index-based insurance that are powered by predictive climate change models with the aim of democratizing the risk information. This should assist in the objectivity of the risk calculation process and reducing the ethical problems that arise in human bias in terms of determining the cost of coverage.

Keywords: Climate Change, Food Security, Climate Modelling, Kenya, Drought, Floods, Agricultural Insurance, Insurance Based Index

References:

1. T. D.S.G., T. C., O. H. and H. B., "Adaptation to climate change and variability: farmer responses to intraseasonal precipitation trends in South Africa.," Clim Chang, vol. 83, no. 3, pp. 301 - 322, 2007.

2. C. Songok, E. Kipkorir and E. Mugalavai, "Integration of indigenous knowledge systems into climate change adaptation and enhancing food security in Nandi and Keiyo districts, Kenya.," in Experiences of climate change adaptation in Africa., Hamburg, Springer, 2011, pp. 69 - 95

Contact details:

- Email: amugwe@strathmore.edu

5.3 Cost of Dying in Switzerland: Taking a Glimpse of Medical Expenses in the Last Year of Life

Andrey Ugarte Montero (presenter)
Joel Wagner

Abstract: Medical expenses tend to increase as individuals grow older. Specifically, in case of illness, the last year before death may represent a particularly costly period as the need for medical attention intensifies. By using a novel private dataset containing over a million records of medical claims filed by individuals during the last year of their life to their health insurer, our research seeks to shed light on the medical costs before dying in Switzerland. With the help of machine learning algorithms, we identify and quantify the main effects that drive a person's medical expenses in this critical period. Our findings provide a better understanding of the costs linked to hospitalization before death, the role played by the age, and the differences in costs based on geographical regions, among other factors.

Keywords: Medical expenses, health insurance, end-of-life costs

Contact details

- Email: andrey.ugarte@unil.ch
- Linkedin: Andrey Ugarte Montero

5.4 Need for Speed - GPU acceleration for insurance

Ruben van der Zwaan, MavenBlue (presenter)
Joost van Bruggen, MavenBlue
Marcel Smith, MavenBlue

Abstract: Speed is not a 'nice-to-have', it is a currency that can be traded freely to solve more complex problems or adapt a new workflow. For example, we see in practice that obtaining results online vs batch shortens ORSA cycles to weeks in processes that used to span months. GPU acceleration has the potential to speed up typical workloads in insurance settings by a huge [1] factor by exploiting parallelism. Tasks in insurance are exceptionally suited to parallel processing on the GPU as many workloads are parallel, but don't always fit a neural network implementation. For example, take stochastic projections in an ESG and per policy processing. Although GPUs are not new and are extensively used in neural networks there is a steep learning curve in enabling these speedups. Reimplementing existing code in C with all the idiosyncrasies of C and CUDA is no mean feat. However, with some practical knowledge and new tools it becomes much easier to take this step. In this talk we show how speed can lead to better processes in practice, the underlying fundamentals of achieving massive speedup, but also how this can be practically achieved. We do this with practical examples.

Keywords: ESG, GPU, Parallel computing

Contact details

- Email: ruben@mavenblue.com
- Homepage: mavenblue.com

[1] We observe 100x speedups in ESG and balance projections, and sometimes much higher. Obviously, it's directly connected to the underlying problem and hardware.

6 Regular Session 3

Thursday, 14.00-15.00
Room G110 Gemelli
Chair: *Mathieu Pigeon*
Mode: *Hybrid - speakers in person*

6.1 Matrix regression: models, algorithms, and applications

Martin Bladt, University of Lausanne (presenter)
Hansjoerg Albrecher, University of Lausanne
Mogens Bladt, University of Copenhagen
Jorge Yslas, University of Liverpool

Abstract: The task of modeling claim severities is addressed when data is not consistent with the classical regression assumptions. This framework is common in several lines of business within insurance and reinsurance, where catastrophic losses or heterogeneous sub-populations result in data difficult to model. Their correct analysis is required for pricing insurance products, and more generally, for risk management. In this talk, we propose to use regression models based on phase-type distributions. We first investigate regressing on the underlying inhomogeneous Markov intensity, generalizing the proportional hazards model. We subsequently show how variants of these models can be successfully applied to both loss and mortality modeling. The latter has the multi-state setting as a natural framework for describing the aging process. Finally, we introduce a model where covariates act on the initial probabilities of the underlying chain, which play the role of expert weights. The basic properties of such models are computed in terms of matrix functionals, and denseness properties are derived, demonstrating their flexibility. Effective estimation strategies are proposed, mainly based on the EM algorithm and weighted multinomial logistic regression models, and we provide illustrations using real-world datasets.

Keywords: matrix distributions, regression, motor insurance, mortality modeling

References:

1. Bladt, M. (2022). Phase-type distributions for claim severity regression modeling. *ASTIN Bulletin: The journal of the IAA*, 1-32.

2. Bladt, M., Yslas, J. (2022+). Phase-type mixture-of-experts regression for loss severities. *Preprint, University of Lausanne*. arXiv:2111.00581.

3. Albrecher, H., Bladt, M., Bladt, M. Yslas, J. (2022+). Mortality modeling and regression with matrix distributions. *Preprint, University of Lausanne*. arXiv:2011.03219.

Contact details:

- Email: martin.bladt@unil.ch
- Homepage: martinbladt.com
- Repository: github.com/martinbladt/matrixdist_1.0

6.2 affine_mortality: R tools for estimation, comparison and projection of affine mortality models

Francesco Ungolo, Lehrstuhl für Finanzmathematik, Technische Universität München (presenter)
Michael Sherris, School of Risk & Actuarial Studies, Australian Research Council, Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales
Len P. D. M. Garces, School of Risk & Actuarial Studies, Australian Research Council, Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales
Yuxin Zhou, School of Risk & Actuarial Studies, Australian Research Council, Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales

Abstract: Affine mortality models, developed in continuous time, are well suited to longevity risk applications including pricing and capital management. A major advantage of this mortality modelling approach is the availability of closed-form cohort survival curves, consistent with the assumed time dynamics of mortality rates. The tools developed in affine_mortality allow to efficiently estimate the parameters of several models, including the canonical Blackburn-Sherris, the Arbitrage-Free Nelson-Siegel and Cox-Ingersoll-Ross by means of maximum likelihood with the univariate Kalman filter. This method turns out to be faster and more robust compared to traditional estimation methods which heavily use large matrix multiplication and inversion. The application of these tools to the estimation of mortality models using age-cohort mortality data from five different countries showed that these models fit historical mortality data well, provide good out-of-sample performance, and are robust with respect to the set of age-cohort data used for parameter estimation. The tools in affine_mortality allow to produce quantitative measures and graphical checks to assess the goodness of fit of the models and to compare them. In addition, the projected cohort survival curves can be readily obtained, as useful for modelling longevity risk where both mortality models and financial models are required, as in the case of longevity-linked securities.

Keywords: Affine mortality, Longevity Risk, R, State-space models, Kalman-Filter

References:

1. Ungolo, F., Sherris, M., Zhou Y. (2022). Estimation, Comparison and Projection of Multi-factor Age-Cohort Affine Mortality Models, Submitted.

2. Ungolo, F., Sherris, M., Garces L.P.D.M., Zhou Y. (2022). affine_mortality: R tools for estimation, analysis and projection of affine mortality models.

Contact details:

- Email: francescoungolo@gmail.com

- Homepage: https://www.groups.ma.tum.de/mathfinance/personen/doktoranden-und-postdocs/dr-francesco-ungolo/

- Repository: https://github.com/ungolof/affine_mortality

- Twitter: https://twitter.com/f_ungolo

6.3 Accounting for COVID-19-Type Shocks in Mortality Modeling: A Comparative Study

Simon Schnürch, Department of Financial Mathematics, Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, Germany (presenter)
Torsten Kleinow, Department of Actuarial Mathematics and Statistics and the Maxwell Institute for Mathematical Sciences, School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK
Andreas Wagner, Faculty of Management Science and Engineering, Karlsruhe University of Applied Sciences, Karlsruhe, Germany

Abstract: Mortality shocks such as the one induced by the COVID-19 pandemic have substantial impact on mortality models. We describe how to deal with them in the period effect of the Lee-Carter model (Lee and Carter, 1992). The main idea is to not rely on the usual normal distribution assumption as it is not always justified. We consider a mixture distribution model based on the peaks-over-threshold method, a jump model and a regime switching model and introduce a modified calibration procedure to account for the fact that varying amounts of data are necessary for calibrating different parts of these models. We perform an extensive empirical study for nine European countries, comparing the models with respect to their parameters, goodness of fit and forecasting performance. Moreover, we define five exemplary scenarios regarding the future development of pandemic-related mortality. As a result of our evaluations, we recommend the peaks-over-threshold approach for applications with a possibility of extreme mortality events.

Keywords: COVID-19, Lee-Carter model, mortality forecasting, mortality modeling, mortality shocks

References

1. Chen, H., Cox, S. H. (2009). Modeling Mortality With Jumps: Applications to Mortality Securitization. *Journal of Risk and Insurance* **76/3**, 727-751.

2. Chen, H., Cummins, J. D. (2010). Longevity bond premiums: The extreme value approach and risk cubic pricing. *Insurance: Mathematics and Economics* **46/1**, 150-161.

3. Lee, R. D., Carter, L. R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* **87/419**, 659-671.

4. Milidonis, A., Lin, Y., Cox, S. H. (2011).Mortality Regimes and Pricing. *North American Actuarial Journal* **15/2**, 266-289.

5. Schnürch, S., Kleinow, T., Wagner, A. (2021). Accounting for COVID-19-Type Shocks in Mortality Modeling: A Comparative Study. Available at SSRN: https://ssrn.com/abstract=3979826.

Contact details

- Email: simon.schnuerch@itwm.fraunhofer.de; t.kleinow@hw.ac.uk; andreas.wagner@h-ka.de

7 Regular Session 4

Thursday, 14.00-15.00
Room G111 Salvadori
Chair: *Lukasz Delong*
Mode: *Hybrid - speakers in person*

7.1 A hierarchical reserving model for reported non-life insurance claims

Jonas Creveceour, Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat), Data Science Institute, Hasselt University
Jens Robben, Leuven Research Center on Insurance and Financial Risk Analysis (LRisk), KU Leuven, Belgium (presenter)
Katrien Antonio, Leuven Research Center on Insurance and Financial Risk Analysis (LRisk), KU Leuven, Belgium

Abstract: Traditional non-life reserving models largely neglect the vast amount of information collected over the lifetime of a claim. This information includes covariates describing the policy, claim cause as well as the detailed history collected during a claim's development over time. We present the hierarchical reserving model as a modular framework for integrating a claim's history and claim-specific covariates into the development process. Hierarchical reserving models decompose the joint likelihood of the development process over time. Moreover, they are tailored to the portfolio at hand by adding a layer to the model for each of the events registered during the development of a claim (e.g. settlement, payment). Layers are modelled with statistical learning (e.g. generalized linear models) or machine learning methods (e.g. gradient boosting machines) and use claim-specific covariates. As a result of its flexibility, this framework incorporates many existing reserving models, ranging from aggregate models designed for run-off triangles to individual models using claim-specific covariates.

We illustrate our proposed methodology using our dedicated ¡tt¿'R'¡/tt¿ package called ¡tt¿'hirem'¡/tt¿. We show how to calibrate the hierarchical reserving models and how to simulate the future development of claims. Layer-specific outcome variables can currently be modelled with generalized linear models (GLMs) or gradient boosting models (GBMs), but an extension to other machine learning methods (e.g. neural networks) is possible and part of future work. The proposed methodology is illustrated on a real insurance data set and on a large number of generated portfolios from a simulation engine representing several realistic development scenarios, which is also part of the R package. We compare our method's performance with that of the chain ladder method under the different development scenarios.

Keywords: individual claims reserving, generalized linear models, gradient boosting models, simulation machine, R package

References

1. Crevecoeur, J., Robben, J., Antonio, K. (2021). A hierarchical reserving model for reported non-life insurance claims. Available at arXiv: https://arxiv.org/abs/1910.12692.

2. Jonas Crevecoeur and Jens Robben. hirem: Hierarchical Reserving Models, 2021. URL https://github.com/jonascrevecoeur/hirem. R package version 1.0.

3. Gabrielli, A., V Wuuml;thrich, M. (2018). An individual claims history simulation machine. Risks, 6(2), 29.

4. Avanzi, B., Taylor, G., Wang, M., Wong, B. (2021). SynthETIC: an individual insurance claim simulator with feature control. Insurance: Mathematics and Economics, 100, 296-308.

5. Gabrielli, A. (2021). An individual claims reserving model for reported claims. European Actuarial Journal, 11(2), 541-577.

Contact details

- Email: jens.robben@kuleuven.be

- Homepage: https://www.kuleuven.be/wieiswie/en/person/00131219

7.2 Bayesian Neural Networks applied to individual Chain-Ladder reserving

Gabriele Pittarello (presenter),Sapienza, University of Rome
Gian Paolo Clemente, Catholic university of the Sacred Heart
Diego Zappa, Catholic university of the Sacred Heart

Abstract: In standard claims reserving, practitioners aggregate individual data into triangles to model the underlying process with some well known reserving technique. In individual claims reserving, data are disentangled and treated as supervised learning problem. The peculiar form of a reserving data set forces the practitioner to deal with truncated time-series. A possible solution available in the literature is to select a one-period supervised learning model that process each development year separately. One possible approach is to base the individual predictions on multilayer perceptrons. On the other hand, while stochastic claims reserving assess both process uncertainty and model uncertainty, multilayer perceptrons do not provide an estimate of the variability which is necessary for capital requirement purposes. The novel of this work is to provide the first bayesian neural networks implementation of an individual one-period model based on the individual Chain-Ladder that allows to assess the individual reserves.

Keywords: individual claims reserving, bayesian neural networks, machine learning, non-life insurance

References

1. Wuthrich, Mario V.. "Neural networks applied to chain-ladder reserving." European Actuarial Journal 8 (2018): 407-436.

2. Kuo, Kevin. "Individual Claims Forecasting with Bayesian Mixture Density Networks." ArXiv abs/2003.02453 (2020)

3. Gabrielli, Andrea. "An individual claims reserving model for reported claims." European Actuarial Journal (2021)

Contact details

- Email: gabriele.pittarello@uniroma1.it - Repository: https://github.com/gpitt71 - Social media: Linkedin

7.3 Comparison of Simulation Engines for Claims Data

Brian Fannin, Casualty Actuarial Society
Jordan Ko, SAS Institute (presenter)

Abstract: The Institute and Faculty of Actuaries working party on Machine Learning in Reserving aims to advance the use of advanced analytics methods within non-life claims reserving and the co-authors have been active members of this working party. One major stumbling block in this exercise is the availability of granular claims data. While the ultimate goal is to apply machine learning methods on real claims data, simulated data may be beneficial in the intermediate steps for several reason. First, researchers may create specific trends that the machine learning algorithms should capture. Second, training the algorithm on a simulated data set may allow one to separate subjective effects such as human bias or inconsistency. Indeed, several papers have addressed the need of simulation engines [1,3,4,5] and introduce additional features relevant for the industry [2]. Several simulation engines have been developed by various actuaries, working independently. In order to contrast the underlying assumptions and facilitate comparisons of the output, the IFoA working party has developed a Shiny app which implements four different simulation engines. Users may engage with the specific assumptions for the individual engines while also being able to recognize some of the common features. Further topics to be investigated and general advice in the usage of the simulated data will be proposed.

Keywords: Claims reserving, simulation engines

References

1. Avanzi, B., Taylor, G., Wang, M., Wong, B. (2021). SynthETIC: an individual insurance claim simulator with feature control. *Insurance: Mathematics and Economics* **100**, 296–308.

2. Avanzi, B., Taylor, G., Wang, M.(2021). SPLICE: A Synthetic Paid Loss and Incurred Cost Experience Simulator. arXiv:2109.04058

3. Bear, R., Shang, K., You, H. (2020). cascsim: Casualty Actuarial Society Individual Claim Simulator. R package version 0.4. https://CRAN.R-project.org/package=cascsim

4.Fannin, B. (2022). imaginator: Simulate General Insurance Policies and Losses. R package version 1.0.0. https://github.com/casact/imaginator

5. Gabrielli, A., Wuthrich, M. (2018) Individual Claims History Simulation Machine. *Risks* **6**, 29.

Contact details

- Email: jordan.ko@sas.com

- Repository: https://github.com/jordanko/

- Social media: https://www.linkedin.com/in/jordanko/

8 Organized Session 1 by Leithà

Thursday, 14.00-15.00
Room G118 Benedetto XV
Chair: *Antonio Tirri*
Mode: *Hybrid - speakers in person*

8.1 European Extreme Events Climate Index (E3CI)

Guido Rianna, Fondazione CMCC Centro Euromediterraneo sui Cambiamenti Climatici (presenter)
Antonio Tirri, Leithà Unipol Group
Francesco Repola, Fondazione CMCC Centro Euromediterraneo sui Cambiamenti Climatici
Francesco Lo Conti, Leithà Unipol Group
Giuliana Barbato, Fondazione CMCC Centro Euromediterraneo sui Cambiamenti Climatici
Paola Mercogliano, Fondazione CMCC Centro Euromediterraneo sui Cambiamenti Climatici
Giorgio Alfredo Spedicato, Leithà Unipol Group

Abstract: The European Extreme Events Climate Index (E3CI) is a 'in progress experiment 'to define a synthetic index aimed at providing information about the areas affected by different types of weather-induced hazards and the severity of such events. E3CI permits equipping Europe with an index similar to actuariesclimateindex.org, developed over North America. Currently, in E3CI, five main hazards are included: cold and heat stresses, droughts,heavy precipitations, intense winds. For each hazard, an indicator has been identified after a deep desk review; it is then calculated on a monthly basis in terms of standardized anomaly respect to a reference period (1981-2010). The components are computed exploiting weather forcing from the fifth generation of reanalysis [1] made available by European Center Medium Weather Forecast (ECMWF) in Copernicus Climate Change Service (C3S) with a delay of only 5 days, global coverage,horizontal resolution of 0.25° and hourly temporal resolution. E3CI data are made freely available at Country level through the dashboard hosted on the https://www.ifabfoundation.org/it/e3ci/ for visualization and the https://e3ci.dataclime.com/, developed by CMCC Foundation, for downloading. In next months, several improvements are expected to be implemented: suitable indicators for forest fires and hail events will complement the Index; furthermore, while the current version of E3CI can supply information for back-analysis of past criticalities, atmospheric inputs provided by climate projections up to 2100 under different greenhouse gases concentration scenarios will permit the computation of E3CI over future time horizons supporting planning processes for several purposes. E3CI, promoted by Foundation Big Data and Artificial Intelligence for Human Development (IFAB), is developed by Fondazione CMCC Euro-Mediterranean Center on Climate Change (CMCC) and Leithà Unipol Group.

Keywords: weather-induced hazards, reanalysis, back-analysis, climate change

References

1. Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C,

2. Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R,

3. Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti

4. G, Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut JN (2020) The ERA5 global reanalysis.

5. Q. J. R. Meteorol. Soc 146(730):1999–2049. https://doi.org/10.1002/qj.3803

Contact details

- Email: guido.rianna@cmcc.it

- Homepage: https://www.cmcc.it/it/people/rianna-guido

- Social media: https://www.linkedin.com/in/guido-rianna-85abb7200/

8.2 Geomorphic flood hazard mapping: from floodplain delineation to flood-hazard characterization

Andrea Magnini, Univ. Bologna
Michele Lombardi, Univ. Bologna
Armir Bujari, Univ. Bologna
Pietro Mattivi, Univ. Bologna
Marco Patella, Univ. Bologna
Gabriele Bitelli, Univ. Bologna
Francesco Lo Conti, Leithà, Unipol Group
Antonio Tirri, Leithà, Unipol Group
Attilio Castellarin, Univ. Bologna

Abstract: Several contributions in the literature show the application of geomorphic indices retrieved from the analysis of Digital Elevation Models (DEMs) for flood hazard mapping. DEM-based techniques are generally trained on reference inundation or flood-hazard maps and are nowadays well established and simpler alternatives to resource-intensive hydrodynamic models for flood hazard mapping. Our study highlights and addresses some limitations of the conventional application of such techniques, which are mostly targeting floodplain delineation, contributing to advancing our understanding of how to fully exploit their potential for computationally efficient and geographically consistent characterization of flood hazards across large geographical regions. We focus on three important aspects: (a) the accuracy, availability, and information content of input information (i.e., DEMs and reference flood-hazard maps); (b) how significantly we can optimize the efficiency of the computational pipeline and the integration of various software libraries available in the literature when such techniques are applied on large and very large DEMs; (c) how to best profit from the outcome of geomorphic flood hazard assessment. Our results (a) show the remarkable role played by input information; (b) exemplify the huge potential offered by computational pipeline optimization; (c) suggest that geomorphic flood hazard maps using continuous indices (e.g. inundation p-value; raw geomorphic index; etc.) should always be preferred to binary flood-hazard map obtained by thresholding the continuous indices themselves (i.e. differentiating between likely and unlikely floodable pixels).

8.3 Terraferma: an interactive tool for insurance seismic risk awareness in Italy

Antonio Petruccelli, Leithà Srl – Unipol Group
Antonio Tirri, Leithà Srl – Unipol Group
Luigi Ferraresi, Leithà Srl – Unipol Group

Abstract: Due to its geographical location and its tectonic conformation, Italy is one of the most seismic countries in the world. The inestimable urban and architectural heritage makes the peninsula one of the countries mostly subjected to ingent economic losses. However, this relevant exposure is not always accompanied by an adequate and commensurate awareness of the seismic risk. Fully developed by Leithà, AI-based solutions factory of the Unipol Group with special expertises on Natural Events, for Incontra Assicurazioni, Terraferma is a simple and extremely innovative graphic tool that responds to the abovementioned need. The tool allows the users to have access to a complete set of informations about seismic hazard, updated in real time, to which Italy is currently exposed. In particular, it is possible to visualize, over a 3D map, the latest earthquakes recorded by the seismic network of the Istituto Nazionale di Geofisica e Vulcanologia (INGV), together with other informations, such as the maximum intensity at location and the correspondent seismic risk classification, and a simulation of the seismic waves propagating. Thanks to the special selection masks, users can also specify reference addresses, time intervals, magnitude ranges or reference radii within which display earthquakes. Terraferma results then in a unique tool for providing a 360° knowledge around seismic risk for special needs, such as when purchasing an earthquake insurance policy on property goods.

Keywords: earthquakes, seismic risk, data visualization, user experience

References:

www.incontra-terraferma.it leitha.eu www.incontraassicurazioni.it

Contact details: - Email 1: antonio.petruccelli@leitha.eu
- Email 2: antonio.tirri@leitha.eu
- Email 3: luigi.Ferraresi@leitha.eu

9 Plenary Session

Thursday, 15.00-16.00
Room G110 Gemelli
Chair: *Diego Zappa*
Mode: *Hybrid - speakers in person*
Invited Speakers: *Markus Senn, Patrik Hogan*

9.1 Some like it Bayesian: The allure, obstacles, and rewards

Markus Senn
Patrick Hogan

Abstract: The past decade has witnessed the emergence of new, powerful tools for Bayesian analysis, improving computational efficiency and usability. These tools have made the methods more accessible to modelers, and ever since, early adopters have advertised Bayesian modelling for its tremendous flexibility and rich information on estimated parameters. Despite this promise and improved accessibility, the insurance industry seems to have adopted Bayesian practices only gradually. We will give a subjective account on the allure, obstacles, and rewards of adopting Bayesian methods at PartnerRe, illustrating our journey with practical case studies.

10 Lightning session 3

Thursday, 16.30-17.00
Room G110 Gemelli
Chair: *Olivier Lopez*
Mode: *Hybrid - speakers in person*

10.1 Estimation of the enhanced concordance probability in linearithmic time

Jolien Ponnet, Katholieke Universiteit Leuven (presenter)
Jakob Raymaekers, Universiteit Maastricht
Robin Vanoirbeek, Allianz Belgium
Tim Verdonck, Universiteit Antwerpen

Abstract: An important property of the pricing models used by insurance companies is their capability of discriminating high and low risks from each other. Therefore, we introduced a performance measure based on the concordance probability, also known as the c-index. It is the probability that a randomly selected comparable pair of observations with their predictions is a concordant pair. Hence, the closer to 1, the better the predictive model can discriminate low from high risks. This measure is adapted to the specific needs of the frequency and severity models [1]. For example, we can require that observed payments need to differ enough from each other, before they are considered as being comparable. There is namely no need to distinguish payments that are nearly identical. This adaptation is called the enhanced concordance probability. The intuitive implementation of the (enhanced) concordance probability requires a quadratic computation time. Especially for large data sets, this run time can become problematic, which is why we first suggested two approximations: the marginal approximation and the k-means approximation [2]. However, these are only approximations and especially when comparing the performance of models, we want to avoid a bias. Therefore we propose a new algorithm that exactly computes the (enhanced) concordance probability in linearithmic time, which is proven and empirically confirmed. The good performance of this algorithm is illustrated on large insurance data sets.

Keywords: C-index, efficient algorithm, performance measure, mergesort

References:

1. Ponnet, J., Van Oirbeek, R., Verdonck, T. (2021). Concordance probability for insurance pricing models. *Risks* **9(10)**, 178.

2. Van Oirbeek, R., Ponnet, J., Verdonck, T. (2021). Computational efficient approximations of the concordance probability in a big data setting. *Under Review*

Contact details

- Email: jolien.ponnet@kuleuven.be

- Social media: https://www.linkedin.com/in/jolienponnet

10.2 Machine Learning Interpretability in Lapse Prediction for Non-Life Insurance Premium

Claudio Giorgio Giancaterino, Catholic University of Milan (presenter)

Abstract: An issue for every Company is the high number of churning customers. Company's customers lapse or quit insurance policies every day. Some of them for unpredictable reasons but others for the competitive environment. So, many customers leave insurance policies for reasons that a Company can prevent with the help of the churn prediction. In this work, from the collection of portfolio contracts, will be predicted the probability of lapses, with the both use of traditional and modern machine learning, also facing issues linked with imbalanced classes. The growing of open source programming communities combined with the advances in computing power have given the opportunity to develop and use models with more predictive power than traditional ones. Anyway, ensemble algorithms and neural networks are complex models referred as "black-boxes" than linear regression, decision trees and logistic regression referred as interpretable models. For this reason, it's becoming more common to combine the predictive power with a toolbox to understand results from machine learning.

Keywords: Churn Prediction, Imbalanced Classification, Ensemble Algorithms, Logistic Regression, Neural Networks, Machine Learning Interpretability,Global Model-Agnostic Methods, Local Model-Agnostic Methods, Permutation Feature Importance, Partial Dependence Plot, Individual Conditional Expectation, SHAP, LIME.

References

1. C. Lorentzen, M. Mayer, (2020).*Peeking into the Black Box: An Actuarial Case Study for Interpretable Machine Learning*, SSRN.

2. L. Baeder, P. Brinkmann, E. Xu (2021).*Interpretable Machine Learning for Insurance*, Society of Actuaries.

3. S. Masís (2021). *Interpretable Machine Learning with Python*, Packt.

Contact details

- Email: c.giancaterino@gmail.com

- Repository: https://github.com/claudio1975/Machine_Learning_Interpretability_in_Lapse$_{P}rediction$

- Social media: https://www.linkedin.com/in/claudiods/

10.3 Bottom-up construction of rating system using sensitivity measures

Giovanni Rabitti, Heriot-Watt University (presenter)
Arianna Vallarino, University of Turin
Amir Khorrami Chokami, University of Turin

Abstract: The ratemaking process is a key task in the insurance activity. In this work, we consider the the construction of a rating system for insurance pricing based on the most important risk factors. To this aim, we firstly compute the individual premium with the quantile regression models of Baione and Biancalana (2019, 2021). Then, we apply global sensitivity measures, in particular the Shapley effects (Owen, 2014; Plischke et al., 2021), to find which are the risk factors that explain together as much variability as possible of the insurance premiums. We construct risk classes aggregating the policyholders according to the selected risk factors. Finally, we introduce a new measure of intra-class heterogeneity of the rating system. This measure allows us to compare rating systems based on different combinations of the rating factors. Using the well-known Australian car insurance dataset, we show that the rating system based on the selected risk factors minimizes this measure.

Keywords: Quantile Regression, Shapley Effects, Intra-class hetereogeneity

References

1. Baione, F., Biancalana, D. (2019). An individual risk model for premium calculation based on quantile: a comparison between generalized linear models and quantile regression. *North American Actuarial Journal* **23/4**, 573-590.

2. Baione, F., Biancalana, D. (2021). An application of parametric quantile regression to extend the two-stage quantile regression for ratemaking. *Scandinavian Actuarial Journal* **2021/2**, 156-170.

3. Owen, A. B. (2014). Sobol'indices and Shapley value. *SIAM/ASA Journal on Uncertainty Quantification* **2/1**, 245-251.

4. Plischke, E., Rabitti, G., Borgonovo, E. (2021). Computing Shapley effects for sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification* **9/4**, 1411-1437.

Contact details

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

10.4 On the Logit-weighted Reduced Mixture of Experts models with insurance applications

Andrei Badescu, University of Toronto - presenter Samson Tsz Chai Fung, Georgia State University Sheldon Lin, University of Toronto

abstract: The logit-weighted reduced mixture of experts model (LRMoE) is a flexible yet analytically tractable non-linear regression model. In this talk, we present an extension of the Expectation-Conditional-Maximization (ECM) algorithm that efficiently fits the LRMoE to random censored and random truncated regression data. Using real automobile insurance data sets, the usefulness and importance of the proposed algorithm are demonstrated through an actuarial application.

keywords: Em algorithm, Mixture of Experts, CLaim Reserving..

1. Fung, T.C., Badescu, A. and Lin, X.S. (2021). "A new class of severity regression models with an application to IBNR prediction," North American Actuarial Journal, 25(2), 206-231.

2. Tseung, S.C, Badescu, A., Fung, T.C. and Lin, X.S. (2021). "LRMoE.jl: a software package for flexible actuarial loss modelling using mixture of experts regression model," Annals of Actuarial Science, 15(2), 419-440.

3. Fung, T.C., Badescu, A. and Lin, X.S. (2019). "A class of mixture of experts models for general insurance: theoretical developments," Insurance: Mathematics and Economics, 89, 111-127.

Contact details

- Email: andrei.badescu@utoronto.ca - Repository: - Social media:

11 Lightning session 4

Thursday, 16.30-17.00
Room G111 Salvadori
Chair: *Markus Gesmann*
Mode: *Hybrid - speakers online*

11.1 Strategy optimization in a dynamical financial analysis environment through evolutionary reinforcement learning

Andrea Riva, Universita degli Studi di Roma 'LA SAPIENZA '

Abstract: This thesis develops a reinforcement learning framework to solve insurance control problems. A Dynamic Financial Analysis model is formulated to represent the environment in which a non-life insurance company operates. Based on the modelled environment, a multi-objective stochastic control problem is formalized by defining the company's control variables and target quantities to optimize. To avoid a modelling bottleneck induced by analytic techniques, two computational methods, neural networks and symbolic regression, have been adopted to approximate candidate strategies. Depending on the approximation method, strategies are represented by a specific set of parameters. Therefore, the search for optimal strategies boils down to the search for an optimal configuration of such parameters. To this end, an evolutionary inspired search algorithm has been adopted and compared to a Uniform Monte Carlo Search. Numerical results show that the proposed framework can find optimal strategies regardless of the underlying insurance model complexity or number of control variables.

Keywords: Dynamic Financial Analysis, Stochastic Control, Reinforcement Learning, Evolutionary Optimization.

References

1. Barto A. G. and Sutton R. S. (2018). Reinforcement Learning: An Introduction. *The MIT Press*.

2. Brunton S., Proctor J. and Kutz J. (2015). Discovering governing equations from data: Sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, vol. 113, pp. 3932-3937.

3. Eling M., Parnitzke T. and Schmeiser H. (2006) Management Strategies and Dynamic Financial Analysis. *University of St.Gallen*, 10.2139/ssrn.943788.

4. Gadmer A., Kaufmann R. and Klett R. (2001). Introduction to Dynamic Financial Analysis. *ASTIN Bulletin*, vol. 31, no. 1, pp. 213-249.

5. Taylor G. (2008). A Simple Model of Insurance Market Dynamics. *North American Actuarial Journal*, vol. 12, no. 10.1080.

Contact details

- Email: andrea.riva@uniroma1.it - Repository: https://github.com/AndreaRiva1991

- Social media: https://www.linkedin.com/in/andrea-riva-81069854/

11.2 Reserves, tariff rates, portfolio management. All in One: Machine Learning + stochastic loss reserving.

Andrey Zatsepin, VSK insurance company (presenter) ¡br¿Alexander Kvitchenño, Yandex

Abstract: Typically, the reserves estimating, the setting of tariff rates and portfolio management do not lay within one single model. We propose a comprehensive model that allows you to obtain:

- estimates of the IBNRs;
- reserves allocation for each policy period;
- estimates of ultimate losses on a policy year basis, distributed by months of events and, as a result;
- future cash flows and net risk premiums for each policy;
- a comparison of expected and actual results with estimates of probability of such difference according to the model and;
- possibility to identify problematic portfolio segments for making management decisions.

Normally, 1-year policies are considered within tariffication process. For the purposes of consideration, we propose to divide 1-year policies into 13 parts referring to each calendar month of the policy duration. Such an approach provides a better understanding of the seasonality by calendar month as well as within whole period of the policy duration. The proposed comprehensive model is based on regression models of machine learning, such as gradient boosting (LightGBM) and special neural network architecture (TabNet), and stochastic reserving methods.

Keywords: Loss reserving, tariff development, machine learning in insurance, deep learning in insurance

References

1. Clark, David R., "LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach," CAS Forum, Fall 2003

2. Computational Actuarial Science With R by Arthur Charpentier (editor), 2015, Boca Raton, FL: CRC Press

3. Chen, Tianqi Guestrin, Carlos. (2016). XGBoost: A Scalable Tree Boosting System.

Contact details

-E-mail: andrej-zacepin@yandex.ru

11.3 A New Approach to Forecasting Insurance Loss Ratios

Uri Korn, Lead Actuary, Ledger Investing (presenter)

Abstract: The traditional procedure to forecasting insurance loss ratios relies heavily on data adjustments in the form of loss trends and estimated rate changes. The forecast is then selected as an average of these adjusted loss ratios. Loss trends are difficult to calculate and have extreme leverage on the final projections. Rate changes are time consuming to estimate and only tell a fraction of the story, as re-underwriting, policy wording changes, and unintended mix changes have significant impacts as well. Even if these adjustments can be considered accurate, simply selecting an average of these adjusted historical values does not take into account any observed changes that may have already manifested in the experience. Finally, these factors ignore the single most important consideration for a program's expected profitability: the past success or lack of in the ability to have actually achieved that profitability. This talk will focus on the methodology developed at Ledger Investing to address these issues for the purposes of evaluating and securitizing reinsurance portfolios. Instead of attempting to adjust past years, the insurance process itself along with its changes are modeled using a mean reverting state space model. This method is able to adapt to any shifts in the program's history while leveraging industry data. It also allows for a direct analysis of the program's unadjusted historical performance.

Keywords: Reinsurance Pricing, Forecasting, State Space Models

References

1. Korn, Uri. (2021). *A simple method for modeling changes over time.* Variance Journal.

Contact details

- Email: uri@ledgerinvesting.com

11.4 AI Systems for Insurance Data Prep

Varun Sriram, Guy Carpenter (Presenter)
Fan, Guy Carpenter (Presenter)
Ni Liu, Guy Carpenter (Presenter)

Abstract: Data often comes in messy and unorganized from multiple different systems, such that the first job in data science is resolving and cleaning this dataset. In commercial insurance, this is amplified with datasets being recorded in disparate ways across various systems. AI is increasingly relevant in insurance, and this paper explores its relevance in insurance data preparation: specifically, how to leverage NLP, ML algos and AI software systems to automate company resolution, string enumeration matching, and ultimately other data prep operations on insurance policy data. First, de-duplicating and augmenting complete information on insured corporations (for instance, augmenting revenue and industry code, or determining Goog, Google Inc, and Alphabet refer to the same general entity) are of critical importance for insurance risk analysis. However, insurers often receive only company names and addresses. We have developed a corporation resolution framework leveraging python, Databricks Spark, and labeling web interfaces, to fuzzy match corporations in incoming insurance policy data with a source of truth dataset. This framework is flexible and robust to handle company name variations, missing address information, and different user preferences. It achieves SOTA performance or better compared to several vendor models in terms of match rate, precision, and speed. Next, data prep involves adhering to universal standards. In insurance, open data standards exist to standardize policy listings, and loss modeling inputs. ML software systems provide a practical solution for mapping to these data standards and mapping to standard string enumerations. We have developed an E2E web UI and backend system for ingesting data in standard schemas, running validations, and correcting validation failures. The correction of validation failures is assisted by ML models, which provide suggestions on possible fixes. Specifically, we developed ML models to map column names to standard names in insurance exposures schema, and ML models to map insurance lines to a standard classification. The models leverage NLP, feature engineering, and stacked ensembles, and are continuously updated in an MLOps framework, where Human-In-the-Loop feedback on the web UI adds to the training dataset. Future work can extend on these AI systems to further automate data prep in new insurance lines, data standards, and entities. For example, the company resolution framework can be generalized into broader entity resolution, to de-duplicate and augment other insured entities or objects. Similarly, more ML models can be created to facilitate correcting other validation issues.

Keywords: Entity Resolution, MLOps, NLP

References

1. Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. (2017). Bag of Tricks for Efficient Text Classification. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 427–431.

2. Vassilis Christophides, Vasilis Efthymiou, Themis Palpanas, George Papadakis, and Kostas Stefanidis. (2020). An Overview of End-to-End Entity

Resolution for Big Data. *ACM Comput. Surv.* 53, 6, Article 127.

Contact details

- Email(s): varunsriram93@hotmail.com, zfan2209@gmail.com, ni.liu.nina@gmail.com

12 Regular session 5

Thursday, 17.10-18.10
Room G110 Gemelli
Chair: *Markus Gesmann*
Mode: *Hybrid - speakers in person*

12.0.1 Insurance ratemaking with images

Christopher Blier-Wong, Université Laval (presenter)
Hélène Cossette, Université Laval Luc Lamontagne, Université Laval
Etienne Marceau, Université Laval

Abstract: We explore using images as a novel data source for the ratemaking of homeowners insurance contracts. Consider a customer who provides his address to the quoting system of an insurance company. On an optional basis, a computer system can then obtain the image of the house from an online resource like Google Street View and extract the required pricing information from the image. Such a process is beneficial for the customer since less information is required during the quoting process, thus improving the shopping experience. This process is also beneficial for insurance companies since the information provided may be more reliable. In this talk, we present such an analysis and extract a dataset of images showing the facades of houses from Google Street View. Street View data has proved useful in various academic fields like urban planning, transportation and socio-economic studies [1], and we investigate its use in actuarial science. This talk explores two methods of leveraging facade images of potential customers. First, we extract the rating characteristics from the images like building value and building age using a convolutional neural network. Second, we use the images as input to a pricing model. To do so, we extract intermediate representations from the image using a convolutional neural network, following the framework prescribed in [2] and use the representations as inputs to a predictive model. Throughout, we discuss technical and social issues of using such data in a rating model.

Keywords: Ratemaking, images, convolutional neural networks, representation learning

References

1. Biljecki and Ito (2021). Street view imagery in urban analytics and GIS: A review. Landscape and Urban Planning

2. Blier-Wong, Baillargeon, Cossette, Lamontagne and Marceau (2021+). Rethinking representations in P&C actuarial science with deep neural networks.

Contact details

- Email: christopher.blier-wong.1@ulaval.ca

12.1 Health insurance claims prediction with GAMLSS

Fabio Baione, Sapienza University of Rome
Davide Biancalana, Sapienza University of Rome (presenter)

Abstract: Regression models represent a powerful tool to fit point estimate (e.g. mean, median, quantile) of the dependent variable, conditional on the values of the independent variables. Anyway, in some cases the restriction on the underlying distribution assumed is not suitable to "best" fit the loss distribution which is one of the main goal of risk assessment and especially in insurance pricing. To this aim Generalized additive models for location, scale and, shape (GAMLSS) are a flexible class of regression models for analyzing data allowing for a relevant extension of distributions assumed for the response variable even in the presence of truncated or censored data. Given their flexibility and thriftiness they can represent a valid tool for solving actuarial problems even in the presence of big data. We apply GAMLSS to price an health insurance which provides coverage for several categories of medical expenditure such as surgery, dental care or hospitalization. It is worth noting, that the coverage exhibits high repeatability, hence the number of claims for each policyholder per year can be much greater than one, with high probability. Moreover, depending on the claims category, we observe low or high frequencies and highly skewed claim size distributions. Finally, the presence of policy limitations such as deductibles and policy limits implies truncated or censored distributions. For these reasons, in order to summarize this large amount of data and information, GAMLSS represents a statistical parsimonious tool that allows the prediction of the one-year health care expenditures at an individual level for several categories of medical expenditure. A comparison with traditional regression models is carried in the application.

Keywords: Health insurance, Regression models, Gamlss

References

1. I. Duncan, M. Loginov M. Ludkovski (2016). Testing Alternative Regression Frameworks for Predictive Modeling of Health Care Costs, *North American Actuarial Journal*, **20:1**, 65-87

2. E.W. Frees PhD, FSA, J. Gao PhD M.A. Rosenberg PhD, FSA (2011). Predicting the Frequency and Amount of Health Care Expenditures, *North American Actuarial Journal*, **15:3**, 377-392.

3. Mc Cullagh, P. A Nelder, J.A. . Generalized Linear Models, Taylor Francis.

4. R. Rigby, M. Stasinopoulos (2011). Generalized Linear Model for Location, Scale and Shape. *Journal of the Royal Statistical Society.* **54(3)**: 507–554

Contact details

- Email(s): fabio.baione@uniroma1.it, davide.biancalana@uniroma1.it

12.2 Domain-Specific Languages for Reserve Modeling

MarkMark Shoun, Ledger Investing (presenter)

Abstract: Modern Bayesian tools such as Stan allow insurance practitioners to fit a dizzying variety of models to loss triangles. Stan's power and efficiency allow it to sample from very complex models, but its facilities for model specification can make it difficult to extend and modify elaborate models. Several other projects (such as 'brms', 'rstanarm', and 'rethinking') aim to offer streamlined interfaces for specifying models in Stan, but each of these have important limitations for loss reserve modeling. In this talk, we present a short but complete implementation of a domain-specific language for loss reserve modeling. Features of the language include semantically-aware indexing of parameter groups; imputation of missing data; relative references to previous development lags and accident years; hierarchical parameters; and flexible specifications of likelihoods and model priors. The language itself is implemented in Python, with a mixed strategy of transpilation to Stan for model fitting and interpretation for data transformation and model prediction.

Keywords: Loss Reserving, Domain-Specific Languages, Stan, Python, Bayesian Methods

References

1. Stan Development Team. (2022). Stan Modeling Language Users Guide and Reference Manual, Version 2.29. https://mc-stan.org

2. Bürkner P (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, **80(1)**, 1–28.

3. Goodrich B, Gabry J, Ali I, Brilleman S (2020). rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.1, https://mc-stan.org/rstanarm.

4. Richard McElreath (2021). rethinking: Statistical Rethinking book package. R package version 2.13.2.

Contact details

- Email: mark@ledgerinvesting.com

- LinkedIn: https://www.linkedin.com/in/mark-shoun/

13 Regular session 6

Thursday, 17.10-18.10
Room G111 Salvadori
Chair: *Giacomo Bormetti*
Mode: *Hybrid - speakers online*

13.1 A general framework for modelling claim count data in general insurance based on the local mixed Poisson net

George Tzougas, Maxwell Institute for Mathematical Sciences and Department of Actuarial Mathematics and Statistics, Heriot-Watt University,
Ziyi Li (presenter)

Abstract: Over the last few decades, mixed Poisson (MP) regression models have been widely used for studying count regression type problems in many scientific fields such as sociology, econometrics, manufacturing, engineering, agriculture, biology, biometrics, genetics, medicine, sports, marketing and insurance. However, due to the ongoing emergence of a range of new technologies and data analytics tools, the highly data driven general insurance sector has become to pay increasing attention to deep learning models which often outperform classical regression models which are too restrictive to learn complex features from data. On the other hand, mitigating interpretability issues is a crucial step for enabling actuaries to safely incorporate deep learning into their predictive modelling techniques. In this study, inspired by the work of Richman and Wuthrich (2021) and extending on our recent work by Tzougas and Li (2022), we introduce the local mixed Poisson net, which we call LocalMPnet, for modelling claim count data using an interpretable deep learning architecture. The proposed modelling framework provides sufficient flexibility for (i) modelling different levels of overdispersion, which is a direct consequence of unobserved heterogeneity due to systematic effects in the data, since as it is well known the tails of mixed Poisson distributions are similar to the tails of their mixing distributions, see Perline (1998) and (ii) taking advantage of the philosophy of representation learning while preserving the linear structure of MP regression models by using a skip connection which provides a linear modeling part around which the network model is built. For expository purposes we use choose the Negative Binomial (NB) deviance loss and we fit the LocalNBnet on a portfolio of motor insurance data. The results from indicate that the LocalNBnet provides better predictive performance than the NB regression and the plain-vanilla neural network.

Keywords: Deep learning, neural networks, mixed Poisson regression model, Negative Binomial, variable selection, explainable deep learning, attention layer, claim count data, Shapley values, SHapley Additive exPlanations (SHAP), integrated gradients

References

1. Perline, R. (1998). Mixed Poisson distributions tail equivalent to their mixing distributions. *Stat. Comput.***38**, 229-233.

2. Richman, R., and Wuthrich, M. V. (2021). LocalGLMnet: interpretable deep learning for tabular data. *Available at SSRN 3892015.*

3. Tzougas, G., and Li, Z. (2022). Neural Network Embedding of the Mixed Poisson Regression Model for Claim Counts. *(in submission).*

Contact details

- Email: George.Tzougas@hw.ac.uk

13.2 Algorithmic detection of interacting variables for generalized linear models via neural networks

Yevhen Havrylenko, Technical University of Munich (presenter),
Julia Heger, University of Augsburg

Abstract: To predict insurance claim frequencies, insurance companies often use generalized linear models (GLMs) due to the interpretability of these models, legacy systems and regulation. To continually improve GLMs, pricing actuaries manually adjust these GLMs via sophisticated choice of interactions of variables. This process is time-consuming, depends much on expert judgement and relies on visual performance indicators. To overcome the above challenges, we develop an algorithmic procedure for detecting the next-best interaction that is missing in a benchmark GLM. Our approach uses Neural Networks, does not rely on visual evaluation of interactions and does not require re-training the benchmark GLM. In our talk, we describe three main parts of the proposed interaction-detection approach, show its performance on an open-source data set, and briefly comment on its performance on a proprietory data set with over 10 million observations.

Keywords: generalized linear model, neural network, interaction, interpretability, insurance claim frequency

References

1. Henckaerts, R., Cote, M.P., Antonio, K., Verbelen, R. (2021). Boosting insights in insurance tariff plans with tree-based machine learning methods. *North American Actuarial Journal* **1/2020**, 120-122.

2. Wuethrich, M.V. (2020). Bias regularization in neural network models for general insurance pricing. *European Actuarial Journal* **1/2020**, 120-122.

Contact details

- Email: yevhen.havrylenko@tum.de
- Linkedin: https://www.linkedin.com/in/yevhenhavrylenko/
- Researchgate: https://www.researchgate.net/profile/Yevhen-Havrylenko

13.3 Pure premium calculation for flood risk based on spatial information using R

Robert Pusz, Warsaw School of Economics

Abstract: With the introduction of the spatial information technology, inspection if a given building is within flood zone became possible. However, in order to find out the level of water in case of flood, additional information is necessary, especially the topography. This information can be obtained from the point cloud coming from the Airborn Laser Scanning (ALS). Such scanning is possible with Light Detection And Ranging (LIDAR) scanner. Two products of ALS point cloud are of particular interest: the numerical terrain model and buildings. Providing additional information from topographic object database enables the use of this knowledge in order to forecast the level of flooding of a given object. Thanks to this, it is possible to deduce pure insurnce premium for the flood risk more precisely. For this purpose, the Digital Terrain Model (DTM) was first generated from ALS point cloud. Joining that information with flood zone, the Numerical Water Mirror Model (NMZW) has been generated. Finaly these two raster models allow to generate the Numerical Water Depth Model and the flooding level for each particular building. All calculations, including spatial analyzes were done with the use of R packages: lidR, sf, raster, tmap, mapview anglr and mesh2.

Keywords: Flood, Premium, DMT

References

1. Weżyk P.(editor), (2015). *Podrecznik dla uczestników szkoleń z wykorzystania produktów LiDAR.*

2. *Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32007L0060.

14 Regular session 7

Friday, 9.30-10.30
Room G110 Gemelli
Chair: *Mario Wuthrich*
Mode: *Hybrid - speakers in person*

14.1 A Collective Reserving Model With Claim Openness

Henning Zakrisson, Stockholm University (presenter)
Mathias Lindholm, Stockholm University

Abstract: The present paper introduces a simple aggregated reserving model based on claim count and payment dynamics, which allows for claim closings and reopenings. The modelling starts off from individual Poisson process claim dynamics in discrete time, keeping track of accident year, reporting year and payment delay. This modelling approach is closely related to the one underpinning the so-called double chain-ladder model, and it allows for producing separate reported but not settled and incurred but not reported reserves, but leverages data availability in a nested model for claim openness, allowing for a more flexible and data-driven approach. The introduced model allows for use of flexible parametrisations in terms of, for example, generalised linear models (GLM) whose parameters can be estimated based on aggregated data using quasi-likelihood theory. Moreover, it is possible to obtain interpretable and explicit moment calculations, as well as having consistency of normalised reserves when the number of contracts tend to infinity. Further, by having access to simple analytic expressions for moments, it is computationally cheap to bootstrap the mean squared error of prediction for reserves. The performance of the model is illustrated using a flexible but simple and interpretable GLM parametrisation evaluated on non-trivial simulated claims data. This numerical illustration indicates a clear improvement compared with models not taking claim closings and re-openings into account. The results are also seen to be of comparable quality with machine learning models for aggregated data not taking claim openness into account.

Keywords: Reserving, GLM, RBNS, IBNR, Simulated claims, Data science

References

1. Wahl, F., Lindholm, M. and Verrall, R. (2019) The collective reserving model. *Insurance: Mathematics and Economics*, **87**, 34–50.

2. Gabrielli, A. and Wütrich, M.V. (2018) An individual claims history simulation machine. *Risks* **6(2)**, 29.

3. Lindholm, M., Verrall, R., Wahl, F. and Zakrisson, H. (2020) Machine learning, regression models, and prediction of claims reserves. *Casualty Actuarial Society E-Forum* **Summer 2020**, 1–47.

Contact details

- Email: zakrisson@math.su.se
- Linkedin: linkedin.com/in/henningzakrisson/

14.2 LocalGLMnet: A Deep Learning Architecture for Actuaries

Dr. J. Schelldorfer, Swiss Re (presenter)

Abstract: The purpose of this talk is to discuss the LocalGLMnet architecture (introduced by R. Richman and M. Wü;thrich (2021)), which is tailored to the needs of actuaries. The LocalGLMnet is a flexible network architecture for tabular data that allows for variable selection, the study of interactions, gives nice interpretations and allows to rank variable importance. The LocalGLMnet architecture is inspired by the structure and properties of generalized linear models (GLMs). It preserves the linear structure of GLMs, but it makes the coefficients of the linear predictors feature dependent. The LocalGLMnet architecture is similar to attention layers. Attention layers are a recently introduced new way of building powerful networks by extracting more important feature components from embeddings by giving more weight (attention) to them. We exemplify the LocalGLMnet on a publicly available accident insurance data set. We show the nice properties and and visuals of the model and we have identified variables that can be dropped from the model. This case study has been done as part of the "Data Science" working group of the Swiss Association of Actuaries (SAA). The group publishes tutorials that discuss the use of machine learning techniques for actuarial applications. The tutorials are self-explanatory and its code and data is publicly available on the website www.actuarialdatascience.org.

Keywords: LocalGLMnet, neural network, deep learning, variable selection, interactions, explainable artificial intelligence, XAI, generalized linear model, GLM, tabular data, variable importance

References

1. Schelldorfer, J., Wuuml;thrich, M. V. (2021) LocalGLMnet: A Deep Learning Architecture for Actuaries, *SSRN Manuscript*, **ID 3900350**, SSRN

2. Richman, R., Wuuml;thrich, M. V. (2021) LocalGLMnet: interpretable deep learning for tabular data, *SSRN Manuscript*, **ID 3892015**,SSRN

3. Richman, R., Wüthrich, M. V. (2021) LASSO Regularization within the LocalGLMnet Architecture, *SSRN Manuscript*, **ID 3927187**, SSRN

Contact details

- Email: Juerg_Schelldorfer@swissre.com

- Homepage: www.actuarialdatascience.org

- Repository: https://github.com/JSchelldorfer/ActuarialDataScience

14.3 Comparing machine learning models for micro-level reserving

Emmanuel Jordy Menvouta, Katholieke Universiteit Leuven (presenter)
Robin Vanoirbeek, Allianz Belgium
Tim Verdonck, Universiteit Antwerpen

Abstract: A central part of an insurance company is the management of its future cash flows and solvency capital. To this end, insurers have to set aside reserves to cover outstanding claims liabilities. Macro-level reserving models such as the chain-ladder, focus on aggregated data organized in a so-called run-off triangle and produce reserves for the whole portfolio. However, aggregating data results in a loss of information which may be useful to accurately estimate reserves. Micro-level reserving models use claim-specific data to produce reserves for individual claims. Machine learning models have been proposed for the task of micro-level reserving [1], [2], [3], [4], [5]. These machine learning models can detect non-linear associations and process large amount of structured or unstructured data and hence are expected to perform better than classical macro-level reserving models. However, such machine learning micro-level models have not been compared to each other leading to a fragmented literature and no indication of their benefits or when to use one method rather than another. In this presentation, we propose a comparative study of these machine learning models and investigate their properties on both simulated and real individual claims data from an insurance company. Different performance measures are used to investigate the trade-offs in the different model architectures.

Keywords: Individual claims reserving, Machine learning, RBNS, Time series

References

1.Van Oirbeek, R., Menvouta, E.J., Ponnet, J., Verdonck, T. (2022). mCube : Multinomial Micro-level reserving Model.

2.Bettonville, C., d'Oultremont, L., Denuit, M. , Trufin, J. , Van Oirbeek, R., (2020).Matrix calculation for ultimate and 1-year risk in the Semi-Markov individual loss reserving model. *Scandinavian Actuarial Journal*

3.Gabrielli, A. An individual claims reserving model for reported claims. (2021) *European Actuarial Journal*. **11**, 541-577.

4.Delong, L., Lindholm, M., Wuthrich, M. V. (2021). Collective reserving using individual claims data. *Scandinavian Actuarial Journal*

5. Crevecoeur, J., Robben, J., Antonio, K., (2021). A hierarchical reserving model for reported non-life insurance claims.

Contact details

- Email: emmanueljordy.menvoutankpwele@kuleuven.be

- Social media: https://www.linkedin.com/in/emmanuelmenvouta

15 Regular session 8

Friday, 9.30-10.30
Room G111 Salvadori
Chair: *Ronald Richman*
Mode: *Hybrid - speakers partially in person and partially online*

15.1 Stochastic Ensemble Loss Reserving

Benjamin Avanzi, University of Melbourne
Yanfeng Li, University of New South Wales (presenter)
Bernard Wong, University of New South Wales
Alan Xian, Macquarie University

Abstract: Loss reserving generally focuses on identifying a single model that can generate superior predictive performance. However, different loss reserving models specialise in capturing different aspects of loss data. This is recognised in practice in the sense that results from different models are often considered, and sometimes combined. For instance, actuaries may take a weighted average of the prediction outcomes from various loss reserving models, often based on subjective assessments. However, such methods are typically based on ad-hoc rules. In this paper, we propose a framework to objectively combine (i.e. ensemble) multiple stochastic loss reserving models such that the strengths offered by different models can be utilised effectively. A notable innovation of our framework is that it is tailored for the features inherent to reserving data. These include, for instance, accident, development, calendar, and claim maturity effects. Importantly, the relative importance and scarcity of data across accident periods renders the problem distinct from the traditional ensembling techniques in statistical learning. Our ensemble reserving framework is illustrated with a complex synthetic dataset. In the results, the optimised ensemble outperforms both (i) traditional model selection strategies, and (ii) an equally weighted ensemble. In particular, the improvement occurs not only with central estimates but also relevant quantiles, such as the 75th percentile of reserves which is of key interest to both insurers and regulators.

Keywords: Aggregate Loss reserving, Ensemble learning, Linear Pool, Distributional forecast

References

1. Taylor, G. (2000). *Loss reserving: an actuarial perspective.* Kluwer Academic Publishers.

2. England, P.D., Verrall, R.J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, 443–544.

3. Shapland, M.R. (2022). *Using the hayne mle models: A practitioner's guide.* CAS Monograph

4. Gneiting, T., Balabdaoui, F., Raftery, A.E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 243–268.

5. Gneiting, T., Ranjan, R. (2013). Combining predictive distributions. *Electronic Journal of Statistics* **7**, 1747–1782.

Contact details

- Email 1: b.avanzi@unimelb.edu.au
- Email 2: yanfeng.li@student.unsw.edu.au
- Email 3: bernard.wong@unsw.edu.au
- Email 4: alan.xian@mq.edu.au

15.2 Forecasting flood risk exposure using mobile phone traffic flows' data

Rodolfo Metulini, University of Salerno (presenter)
Maurizio Carpita, University of Brescia

Abstract: Flooding risk exposures maps traditionally assume amount of presences (either people or vehicles) constant over time, although crowding, expecially in metropolitan areas, is a highly dynamic process in the real world. Real-time monitoring (and forecasting) of mobility is thus a relevant aspect for metropolitan areas subjected to flooding risk. To obtain a dynamic measure for the risk of exposure we make use of the Information and communication technologies (ICT) source represented by mobile phone network data, that have already been used with the aim of producing dynamic information on people's presences and flood risk maps for areas with hydrogeological criticality (Balistrocchi et al., 2020; Metulini and Carpita, 2021). Specifically, in this work we use mobile phone origin-destination (OD) signals on traffic flows by Telecom Italia Mobile (TIM) users among different census areas (ACE of ISTAT, the Italian National Statistical Institute) of the Mandolossa (a critical zone with flood episodes in the north-west of the city of Brescia), recorded at hourly basis from September 2020 to August 2021, with the aim of forecasting the exposure risk and thus to make outbreaks' early detection and warning to who is transiting through that area, in case the density is higher than a pre-specified threshold. Preliminar results based on an Harmonic Dynamic Regression (HDR) (Hyndman and Athanasopoulos, 2021) to model the complex seasonality of the data and on a statistical algorithm for threshold dectection inspired by that of Farrington at al. (1996) show a promising performance.

Keywords: Flood risk; Origin-destination data; high frequencies time series; mobile phone data; traffic flows.

Acknowledgement MoSoRe(at)UniBS (Infrastrutture e servizi per la Mobilità Sostenibile e Resiliente) Project of Lombardy Region, Italy (CallHub ID 1180965; bit.ly/2Xh2Nfr)

References

1. Balistrocchi, M., Metulini, R., Carpita, M., Ranzi, R. (2020). Dynamic maps of human exposure to floods based on mobile phone data. *Natural Hazards and Earth System Sciences* **20(12)**, 3485-3500.

2. Farrington, C.P., Andrews, N.J., Beale, A.D., Catchpole, M.A. (1996). A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **159(3)**, 547-563.

3. Hyndman, R. J., Athanasopoulos, G. (2018). Forecasting: principles and practice. *OTexts*.

4. Metulini, R., Carpita, M. (2021). A spatio-temporal indicator for city users based on mobile phone signals and administrative data. *Social Indicators Research,* **156(2)**, 761-781.

Contact details

- Email: rmetulini@unisa.it

- Homepage: https://docenti.unisa.it/044956/home

15.3 Identification of the network structure to evaluate the impact on cyber attacks on an insurance portfolio

Olivier Lopez, Sorbonne University and Detralytics (presenter)
Caroline Hillairet, Ensae Paris
Louise d'Oultremont, Detralytics
Brieuc Spoorenberg, Detralytics
Maud Thomas, Sorbonne University

Abstract: Cyber risk is a direct corollary of the digitalization of the economy, and of the increase of data science related tools. Cyber insurance is a way to protect industries against the consequences of a cyber attack. But the considerable increase of malicious acts in this field, linked to a professionalization of the criminal ecosystem, is an important limit to the development of a resilient market. The difficulty to properly quantify the risk may slow down the expansion of cyber insurance, especially when the perspective of systemic event - accumulation of claims striking simultaneously a large number of policyholder, like the Wannacry or NotPetya episodes in 2017 - is not unrealistic. In this work, we develop a general framework adapted to the construction of accumulation scenario. These tools may be used to evaluate the robustness of a portfolio when it faces a contagious cyber event. The model takes into account the connectivity between different group of policyholder, since the topology of the network may tend to slow down or accelerate the propagation of such an event. A particular attention is devoted to the way the network structure can be determined from a combination of portfolio data and macro-level statistics. The core of the procedure relies on a Poisson lognormal model, estimated through a variational EM algorithm.

Keywords: Cyber insurance, network, contagion, variational EM

References

1. D'Oultremont, L., Lopez, O., Spoorenberg, B. (2021). Modeling accumulation scenarios in cyber risk. *Detra Note*

2. Hillairet, C., Lopez, O., d'Oultremont, L., Spoorenberg, B. (2021). Cyber contagion: impact of the network structure on the losses of an insurance portfolio. *Preprint*

Contact details

- Email: olivier.lopez@sorbonne-universite.fr

- Homepage: https://sites.google.com/view/sitepersonneldolivierlopez/home

- Social media: https://fr.linkedin.com/in/olivier-lopez-a10b03128

16 Plenary Session

Friday, 11.00-12.30
Room G110 Gemelli
Chair: *Nino Savelli*
Mode: *Hybrid - speakers in person* **Invited Speaker:** *Fausto Parente*, Executive Director of the European Insurance and Occupational Pensions Authority (EIOPA), AI, data and insurance: Protecting policyholders

17 Lightning Session 5

Friday, 14.00-14.30
Room G110 Gemelli
Chair: *Michael Ludkovski*
Mode: *Hybrid - speakers partially in person and partially online*

17.1 Pseudo-Model-Free Hedging for Variable Annuities via Deep Reinforcement Learning

Wing Fung Chong, Heriot-Watt University (presenter)
Haoen Cui, Georgia Institute of Technology
Yuxuan Li, University of Illinois at Urbana-Champaign

Abstract: This talk proposes a two-phase deep reinforcement learning (RL) approach to address model miscalibration for hedging variable annuity contracts with GMMB rider in the Black-Scholes financial and constant force of mortality actuarial market environment. In the training phase, an infant RL agent interacts with a pre-designed training environment, collects sequential anchor-hedging reward signals, and gradually learns how to hedge the contracts. As expected, after a sufficient number of training steps, the trained RL agent hedges, in the training environment, equally well as the correct Delta while outperforms misspecified Deltas. In the online learning phase, the trained RL agent interacts with the market environment in real time, collects single terminal reward signals, and self-revises its hedging strategy. The hedging performance of the further trained RL agent is demonstrated via an illustrative example. First, by comparing their sample means of terminal profit and loss from simulated scenarios, in most future trajectories, within a reasonable amount of time, the further trained RL agent is able to exceed the hedging performance by the correct Delta from the market environment and the incorrect Delta from the training environment. Second, through a more delicate hypothesis testing analysis, similar conclusions can be drawn in a fair amount of future trajectories. Finally, snapshots of empirical density functions, among the future trajectories, of the sample means of terminal profit and loss from simulated scenarios by each hedging strategy, shed light on the high probability that, the further trained RL agent is indeed able to self-revise the hedging strategy from a very sub-optimal one to a nearly optimal one close to the correct Delta from the market environment.

Keywords: Two-phase deep reinforcement learning, Variable annuities hedging, Training phase, Sequential anchor-hedging reward signals, Online learning phase, Single terminal reward signals, Hedging strategy self-revision

Contact details

- Email: alfred.chong@hw.ac.uk - Homepage: https://sites.google.com/view/wingfungalfredchong/

17.2 Combination of Mortality Rate Forecasts From Multiple Starting Points

Salvatory R. Kessy, University of New South Wales (presenter)
Michael Sherris, University of New South Wales
Andrés M. Villegas, University of New South Wales
Jonathan Ziveyi, University of New South Wales

Abstract: An important aspect of estimating many commonly used mortality models is selecting the time used. This is especially important in the light of changing mortality trends. For trend extrapolation, this requires determining the length of the base period. Many mortality models proposed in the literature assume a random walk with drift for trends to forecast mortality rates. For this assumption, mortality rate forecasts are highly sensitive to the calibration period, especially with changing mortality trends and structural changes in mortality patterns. One approach to this problem is to average forecasts from the same model across multiple starting points to account for numerous structural changes and the impact of the fitting period. We use this approach with Generalized Age-Period-Cohort (GAPC) mortality models to obtain mortality forecasts by averaging the same model across multiple calibration periods. We propose and assess three different choices of the combination weights. In the first scheme, we generate the forecast combination by using equal weights to average the individual forecasts; the second weighting method assigns heavier weights to predictions that use more recent data; and in the third approach, we determine the forecast combination by using cross-validation weights to average the individual forecasts. To illustrate the approach, we focus on three GAPC models: the Lee-Carter model, Age-Period-Cohort (APC) model, and the Plat model. We apply our method to Australian female data from the Human Mortality Database because the Lee-Carter model trend is non-linear and shows structural changes. In the out-of-sample forecast analysis, we find that combining forecasts from multiple fitting periods produces a lower mean squared error of mortality rate forecasts than the original mortality models with a fixed fitting period. This supports the model-fitting strategy of averaging mortality rate forecasts from multiple starting points for improving out-of-sample forecast performance. The results also show that the impact of selecting the historical period for fitting extrapolative mortality models can be greater than the choice of the mortality models that we considered.

Keywords: Structural breaks and forecasting, Average forecasts across models and windows, uncertainty, age-period-cohort model, mortality forecasting

References

1. Kessy, Salvatory R., Michael Sherris, Andrés M. Villegas, and Jonathan Ziveyi. 2021. "Mortality Forecasting Using Stacked Regression Ensembles." Scandinavian Actuarial Journal, November, 1–36. https://doi.org/10.1080/03461238.2021.1999316.

2. Booth, Heather, John Maindonald, and Len Smith. 2002. "Applying Lee-Carter Under Conditions of Variable Mortality Decline." Population Studies 56 (3): 325–36

3. Pesaran, M Hashem, and Allan Timmermann. 2007. “Selection of Estimation Window in the Presence of Breaks.” Journal of Econometrics, 28.

Contact details

- Email: s.kessy@unsw.edu.au - LinkedIn: https://www.linkedin.com/in/salvatory-kessy-a9b7a615b/ - Repository: https://github.com/kessysalvatory/CoMoMo

17.3 IBMPopSim: a package for the efficient simulation of individual-based population models

Daphné Giorgi, Laboratoire de Probabilité, Statistique et Modélisation, CNRS, Sorbonne Université (presenter)

Sarah Kaakaï, Laboratoire Manceau de Mathématiques and Risk and Insurance Institute, Le Mans Université

Vincent Lemaire, Laboratoire de Probabilité, Statistique et Modélisation, Sorbonne Université

Abstract: The IBMPopSim package ¡https://daphnegiorgi.github.io/IBMPopSim/¿ aims at simulating the random evolution of heterogeneous populations, called stochastic Individual Based Models (IBMs). The package allows users to simulate population evolution in which individuals are characterized by their age and some characteristics, and where the population is modified by different types of events including births/arrivals, death/exit events, or changes of characteristics. The frequency at which an event can occur to an individual can depend on his age and characteristics, but also on the other individuals' characteristics (interactions).

Such models have a wide range of applications. For instance, IBMs can be used for simulating the evolution of an heterogeneous insurance portfolio with selection or for validating mortality forecasts. In this presentation, we propose an illustration of such applications.

IBMPopSim overcomes the limitations of time consuming IBMs simulations. This is done by implementing new efficient algorithms based on thinning methods, which are compiled using the Rcpp library, whle providing a user friendly interface. We also provide tools for analyzing outputs, such a age-pyramids or life tables.

Keywords: Individual Based Models, heterogeneous populations, actuarial science, insurance portfolio simulation.

References

1. Giorgi, D., Kaakaï, S., Lemaire, V. (2020). IBMPopSim: Individual Based Model Population Simulation. ¡https://daphnegiorgi.github.io/IBMPopSim/¿

Contact details

- Email: giorgi@lpsm.paris - Homepage: https://www.lpsm.paris/pageperso/giorgi/

17.4 Mortality Heterogeneity and Clustering using Joint Body Mass Index and Self–Reported Health Trajectories

Michelle Vhudzijena, School of Risk and Actuarial Studies, UNSW Business School, University of New South Wales, Sydney, Australia (presenter)

Michael Sherris, School of Risk and Actuarial Studies, UNSW Business School, University of New South Wales, Sydney, Australia

Andrés Villegas, School of Risk and Actuarial Studies, UNSW Business School, University of New South Wales, Sydney, Australia

Jonathan Ziveyi, School of Risk and Actuarial Studies, UNSW Business School, University of New South Wales, Sydney, Australia

Abstract: Mortality models that capture heterogeneity are becoming increasingly important in pricing longevity risk products. This is pertinent given the vast amounts of individual level data that provide additional information on health status, socio-economic circumstances and financial wellbeing over time. There is not much work in the actuarial literature on the objective determination of mortality risk profiles using individual level data. This work leverages clustering techniques developed in machine learning to create homogeneous groups for older adults in the US Health and Retirement Study ($N = 9,815$) where the data are segmented into distinct mortality risk profiles. We apply an unsupervised k-means 3D algorithm to jointly cluster body mass index (BMI) and self-reported health trajectories. We identify three clusters: normal, stable BMI and declining very good health (A); normal, stable BMI and declining fair health (B); high, increasing BMI and declining good health (C). One–Way Anova tests show that the clusters are unique across different socio–economic characteristics and pairwise tests show that the differences between clusters are statistically significant ($p < 0.0001$) even after adjusting for multiple testing. We find very strong evidence ($p < 0.0001$) of an association between mortality and cluster allocation. Pairwise, we have very strong evidence of differences in mortality amongst clusters ($p < 0.0001$). The mortality odds are $0.303(0.274, 0.334)$ times lower in Cluster A than in Cluster B. They are $0.461(0.412, 0.516)$ times lower in Cluster A than in Cluster C, and $1.525(1.391, 1.671)$ times higher in Cluster B than Cluster C. As such, the estimated predicted probabilities of death are consistently highest in the normal, stable BMI and declining fair health cluster at different ages for both males and females. Consequently, the different annuity prices for the clusters demonstrate that accounting for heterogeneity can lead to fairer pricing and potentially greater demand for annuity products. The methods applied can also be used for long–term care annuities and experience studies.

Keywords: Mortality heterogeneity, longitudinal data, k–means clustering 3D, body mass index trajectories, self reported health trajectories, underwritten annuities.

References

1. Villegas, Andrés M., and Steven Haberman. 2014. "On the Modeling and Forecasting of Socioeconomic Mortality Differentials: An Application to Deprivation and Mortality in England." North American Actuarial Journal 18 (1): 168–93. https://doi.org/10.1080/ 10920277.2013.866034.

2. Zajacova, Anna, Snehalata Huzurbazar, Mark Greenwood, and Huong Nguyen. 2015. "Long-Term BMI Trajectories and Health in Older Adults: Hierarchical Clustering of Functional Curves." Journal of Aging and Health 27 (8): 1443–61. https://doi.org/10.1177/0898264315584329.

3. Genolini, Christophe, Xavier Alacoque, Mariane Sentenac, and Catherine Arnaud. 2015. "Kml and kml3d: R packages to cluster longitudinal data." Journal of Statistical Software 65 (4): 1–34. https://doi.org/10.18637/jss.v065.i04.

4. Meyricke, Ramona, and Michael Sherris. 2013. "The determinants of mortality heterogeneity and implications for pricing annuities." Insurance: Mathematics and Economics 53 (2): 379–87. https://doi.org/10.1016/j.insmatheco.2013.06.002.

5. Xu, Mengyi, Michael Sherris, and Ramona Meyricke. 2019. "Systematic Mortality Improvement Trends and Mortality Heterogeneity: Insights from Individual–Level HRS Data." North American Actuarial Journal 23 (2): 197–219. https://doi.org/10.1080/10920277.2018. 1513369.

Contact details

- Email: m.vhudzijena@student.unsw.edu.au - Homepage: https://www.unsw.edu.au/staff/michelle-vhudzijena

18 Lightning Session 6

Friday, 14.00-14.30
Room G111 Salvadori
Chair: *Fabio Baione*
Mode: *Hybrid - speakers online*

18.1 Extreme conditional risk estimation in heavy-tailed heteroscedastic regression models

Stéphane Girard, Inria Grenoble Rhône-Alpes
Gilles Stupfler, ENSAI CREST (presenter)
Antoine Usseglio-Carleve, Avignon Université

Abstract: We build a general theory for the estimation of extreme conditional risk measures in heteroscedastic regression models with heavy-tailed noise. Our approach is supported by general results of independent interest on residual-based extreme value estimators in heavy-tailed regression models, and is intended to cope with covariates having a large but fixed dimension. We demonstrate how our results can be applied to the estimation of extreme conditional quantiles and expectiles in a wide class of important examples, among which linear models, single-index models as well as ARMA and GARCH time series models. Our estimators are showcased on a numerical simulation study and on real sets of actuarial and financial data.

Keywords: Expectiles, Extreme value analysis, Heavy tails, Heteroscedasticity, Quantiles, Regression models, Residual-based estimators, Tail empirical process

References

1. Girard, S., Stupfler, G., Usseglio-Carleve, A. (2021). Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models, *Annals of Statistics* **49(6)**, 3358-3382.

Contact details

- Email: gilles.stupfler@ensai.fr - Homepage: https://ensai.fr/en/equipe/stupfler-gilles/ - LinkedIn: linkedin.com/in/gilles-stupfler-5713651ab - Twitter: @GillesStupfler

18.2 Under-reporting correction in Cyber Incidents

Seema Sangari, School of Data Science and Analytics, Kennesaw State University (Presenter)
Dr. Eric Dallal, Verisk Extreme Events Solutions, Cyber Solutions

Abstract: Under-reporting is a well documented problem in cyber incidents, a consequence of the potential reputational risk and of consequent financial impacts. A significant proportion of incidents are never disclosed to the public unless they involve a breach of protected data. Commonly, a proportion-based approach is applied to resolve the problem of under-reporting. Such an approach consists of comparing the incident rate for a–typically smaller–data set with complete reporting to the implied incident rate of the (incomplete) large data set and finding the appropriate correction factor. We employ a similar approach here, using cyber insurance claims and corresponding policy data (the complete data set) to correct a data set of publicly reported incidents. The purpose of this research is to quantify under-reporting with respect to multiple dimensions: company revenue, industry, and incident categorization. Notably, we have found that there is a substantial difference in under-reporting, a factor of 100, as a function of these variables. The resulting under-reporting model provides the needed correction for incident frequencies derived from data sets of publicly reported incidents. This research provides insight to the level of cyber risk hidden from the public and allows researchers to develop cyber risk models without access to closely guarded cyber insurance claims data, thus opening the field of cyber risk modeling more widely to academia and other researchers with less access to such information.

Keywords: Cyber Insurance, Cyber Risk, Under-reporting ¡!–

References

1. Hazell, L. and Shakir, S. A. (2006). Under-reporting of adverse drug reactions: A systematic review. *Drug Safety*, 29(5):385–396.

2. Rosinska, M., Pantazis, N., Janiec, J., Pharris, A., Amato-Gauci, A. J., Quinten, C., Schmid, D., Sasse, A., van Beckhoven, D., Varleva, T., Blazic, T. N., Hadjihannas, L., Koliou, M., Maly, M., Cowan, S., Rutel, K., Liitsola, K., Salminen, M., Cazein, F., Pillonel, J., Lot, F., Gunsenheimer-Bartmeyer, B., Nikolopoulos, G., Paraskeva, D., Dudas, M., Briem, H., Sigmundsdottir, G., Igoe, D., O'Donnell, K., O'Flanagan, D., Suligoi, B., Konova, ˇS., Erne, S., ˇCaplinskien, I., Schmit, A. F. J. C., Melillo, J. M., Melillo, T., de Coul, E. O., van Sighem, A., Blystad, H., Rosinska, M., Aldir, I., Martins, H. C., Mardarescu, M., Truska, P., Klavs, I., Diaz, A., Axelsson, M., and Delpech, V. (2018). Potential adjustment methodology for missing data and reporting delay in the HIV surveillance system, European Union/European Economic Area, 2015. *Eurosurveillance*, 23(23).

Contact details

- E-mail: SSangar1@students.kennesaw.edu, EDallal@verisk.com - Social media: https://www.linkedin.com/in/seemasangari/

18.3 Multi-country clustering-based forecasting of healthy life expectancy

Susanna Levantesi, Sapienza University of Rome (presenter)
Andrea Nigri, Bocconi University, Milan
Gabriella Piscopo, University Federico II of Naples

Abstract: Healthy Life Expectancy (HLE) is an indicator that measures the number of years individuals are expected to live free of disease or disability. Its forecasting is essential to ensure the sustainability of pension systems, plan the provision of health care to increasingly elderly populations, appropriately price Long Term Care insurance products. In this paper, we propose a methodology in order to simultaneously forecast HLE for groups of countries, investigating similarities in the HLE pattern. We use a functional data clustering to the multivariate time series of HLE of different countries to catch similarities among them. Thus, we carry out the HLE simultaneous forecasting of the populations within each cluster through a multivariate random walk. We estimate our model using HLE at birth data, provided by the Global Burden of Disease Study for about 30 countries worldwide.

Keywords: Healthy Life Expectancy, Forecasting, Functional data clustering.

References

1. Levantesi, S., Nigri, A., Piscopo, G. (2022) Clustering-based simultaneous forecasting of life expectancy time series through Long-Short Term Memory Neural Networks. *International Journal of Approximate Reasoning*, 140, 282-297

2. Permanyer, I., Spijker, J., Blanes, A. (2022) On the measurement of healthy lifespan inequality. *Population Health Metrics*,20,1

3. Permanyer, I., Trias-Llimós, S., Spijker, J. (2021) Best-practice healthy life expectancy vs. life expectancy: Catching up or lagging behind? *Pnas* 118(46)

4. Robine J.M. (2021) Ageing populations: We are living longer lives, but are we healthier? *United Nation Population Division* UN DESA/POP/2021/TP/NO.2

5. Wang, A., Abbas, K.M., Abbasifard, M. et al. (2020) Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: A comprehensive demographic analysis for the Global Burden of Disease Study. *Lancet* 396, 1160-1203
.

Contact details

- Email: susanna.levantesi@uniroma1.it

18.4 Estimation of high dimensional gamma convolutions through random projections

Oskar Laverny, University of Lyon SCOR SE

Abstract: The multivariate generalized gamma convolutions is a class of distributions defined by a convolutional semi-parametric structure. Their flexible dependence structures, the marginal possibilities and their useful convolutional expression makes them appealing to the practitioner. However, fitting such distributions when the dimension gets high is a challenge. We propose stochastic estimations procedures based on (shifted) cumulants approximation of a Laguerre integrated square error, evaluated on random projections of the dataset. Through the analysis of our loss via tools from Grassmannian cubatures, sparse optimization on measures and Wasserstein gradient flows, we show the convergence of the stochastic gradient descent to a proper estimator of the high dimensional distribution. We propose several examples on both low and high-dimensional settings.

Keywords: Generalized gamma convolutions, estimation, high-dimension, sparse measure, gradient flows,

References

1. Thorin, O. (1977) On the Infinite Divisibility of the Pareto Distribution. *Scandinavian Actuarial Journal*.

2. Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities.* Springer New York.

3. Chizat, L. (2021) Sparse optimization on measures with over-parameterized gradient descent. *Mathematical Programming.*

Contact details

- Email: oskar.laverny@univ-lyon1.com - Homepage: actuarial.science - Github: github.com/lrnv

19 Regular session 9

Friday, 14.40-15.40
Room G110 Gemelli
Chair: *Wing Fung Chong*
Mode: *Hybrid - speakers in person*

19.1 Transfer learning for boosting mortality table

Mario Marino, Sapienza University of Rome (presenter)
Giorgio Alfredo Spedicato, Data Science Manager and Actuary @ Unipol Group

Abstract: Mortality modeling is crucial in life and social security insurance, as mortality rates are a key driver of both insurance products rates and the evaluation of associated liabilities. Considering the biometric heterogeneity characterizing the policyholders, the construction of mortality tables based on the mortality experience analysis represents the best estimate of the portfolio mortality. Such an assessment involves statistical procedures for processing observed mortality data and collecting the policyholder mortality rates for various risk factors combinations, as the age at policy issue, the policy duration, the policyholder gender, the smoker status, the socio-economic status and so on. In addition, the so-constructed mortality tables may also constitute the starting point for forecasting future mortality. However, the described approach can be challenging for individual companies whose portfolio mortality experience is limited. Alterative mortality assumptions may be there used, based either on industry mortality tables or at country level, that however may imply some basis risk. Therefore, a robust framework to "update" mortality tables to insurer's portfolio mortality experience, taking into account the policyholders individual characteristics, may significantly help. To this end, our work concerns the so-called tansfer learning in order to derive the insurer's portfolio mortality starting from an a priori mortality experience. In particular, starting from reference mortality estimates we employ Gradient Boosted Trees and Deep Neural Networks models to boost such a priori experience into portfolio mortality, taking into account the available subject-level covariates. To test our proposal we perform an empirical application considering the SOA dataset, which contains industry mortality table representing the reference mortality as well as individual-level covariates. Finally, we draw conclusions about our proposal in terms of predictive performances and practical suitability for a practitioner actuary.

Keywords: Portfolio mortality modeling, transfer learning, Gradient Boosting Trees, Deep Neural Networks.

References

1. Deprez, P., Shevchenko, P.V., Wütrich, M. (2017). Machine learning techniques for mortality modeling. *European Actuarial Journal***7**, 337-352.

2. Perla, F., Richman, R., Scognamiglio, S., Wütrich, M. (2021). Time-series forecasting of mortality rates using deep learning. *Scandinavian Actuarial Journal* **7**,572-598.

3. Society of Actuaries (2019). 2009-2015 Individual Life Insurance Mortality Experience Report. in *Society of Actuaries:2009-2015 Individual Life Insurance Mortality Experience Report.*

Contact details

- Email: m.marino@uniroma1.it

- Homepage: https://web.uniroma1.it/memotef/users/marino-mario

19.2 Calibrating the Lee-Carter and the Poisson Lee-Carter models via Neural Networks

Salvatore Scognamiglio, University of Naples "Parthenope" (Italy)

Abstract: This paper introduces a neural network approach for fitting the Lee-Carter [1] and the Poisson Lee-Carter model [2] on multiple populations. We develop some neural networks that replicate the structure of the individual LC models and allow their joint fitting by simultaneously analysing the mortality data of all the considered populations. The neural network architecture is specifically designed to calibrate each individual model using all available information instead of using a population-specific subset of data as in the traditional estimation schemes. A large set of numerical experiments performed on all the countries of the Human Mortality Database (HMD)[3] shows the effectiveness of our approach. In particular, the resulting parameter estimates appear smooth and less sensitive to the random fluctuations often present in the mortality rates' data, especially for low-population countries. In addition, the forecasting performance results significantly improved as well.

Keywords: Mortality modelling, Multi-population mortality modelling, neural networks, Lee–Carter model, Human Mortality Database.

References

1. Lee, R.D., Carter, L.R. (1992). Modelling and forecasting us mortality. *Journal of the American Statistical Association* **87/419**, 659-671.

2. Brouhns, N., Denuit, M., Vermunt, J.K. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics* **31/3**, 373-393.

3. Human Mortality Database. (2020). University of California, Berkeley (USA), and Max Planck Institute for Demo-graphic Research (Germany). Available at https://www.mortality.org/Public/CitationGuidelines.php.

Contact details

- Email: salvatore.scognamiglio@uniparthenope.it

- Social media: https://it.linkedin.com/in/salvatore-scognamiglio-328383131

19.3 Joint modeling of State-level mortality in US

Mike Ludkovski, University of California Santa Barbara (presenter)
Doris Padilla, UC Santa Barbara

Abstract: We study age-period mortality models for the United States Mortality Database that presents a high quality dataset for the 50 U.S. states. Our first goal is to design joint multi-population models that can fuse information across states and crystallize disparities among their mortality surfaces. Our second goal is to use external socio-economic covariates about state-level incomes, demographics and economies to obtain insights on the relationship of age-specific mortality and such covariates. We apply the framework of multi-output Gaussian processes (Huynh and Ludkovski, 2021) to study methods for grouping states and efficiently infer the global correlation structure. Our analysis is done in R and is extensively illustrated with interactive maps.

Keywords: longevity modeling, Gaussian processes, multi-population mortality

References

1. USMD: United States Mortality DataBase. University of California, Berkeley (USA). Available at usa.mortality.org

2. Huynh, N., Ludkovski, M. (2021) Multi-Output Gaussian Processes for Multi-Population Longevity Modeling, Annals of Actuarial Science, 15(2), 318-345.

Contact details

- Email: ludkovski@pstat.ucsb.edu
- Homepage: ludkovski.faculty.pstat.ucsb.edu
- Social media: @MLudkovski (twitter)

20 Regular session 10

Friday, 14.40-15.40
Room G111 Salvadori
Chair: *Roland Schmid*
Mode: *Hybrid - speakers online*

20.1 Using cosine similarity for recommending insurance products

Anders Drage, Frende Forsikring (Presenter)
Heidi Midtgarden Golid, Frederic Dorn, Frende Forsikring

Abstract: A key factor for insurance companies to make profit is to sell new insurance polices. When selling new polices, it is a huge advantage to have knowledge about the likely needs of a given customer, so that the marketing or sales process can be targeted. We have developed an approach that uses our internal customer portfolio to predict the likely next product an existing customer would need. The insurance product portfolios for customers similar to a target customer is found by using cosine similarity, and are used to predict the next likely product a customer would need. This method allows us both to have individual product recommendations in all communication with the customers, as well as identifying all customers that are likely to buy a given product. We will present the method and give a demonstration on how it is implemented in Frende. The work has further been used in marketing campains, and we will show some results.

Keywords: Cosine similarity, product recommendation, marketing

Contact details

- Email: anders.drage@frende.no
- Social media: Linkedin

20.2 Weather index-based Crop Insurance using Machine Learning

Kelvin Bett, Strathmore University (presenter)
Dr. John Olukuru, Strathmore University

Abstract:

Background: Unfavorable weather conditions and natural disasters pose a significant risk to the predictability of income for farmers. This exacerbates the risk on food security since it limits the ability of a farmer to invest in the next harvest. Weather index-based crop insurance is a practical solution in enhancing the resilience of farmers to weather related shocks thereby enhancing food security. The financial instrument has however proven costly to scale to small-holder farmers (who account for c.75% of produce in Kenya [3] owing to high yield measurement costs. In comparison with traditional insurance, weather index-based insurance is less expensive to administer, which can lead to more affordable contracts and faster payments to farmers, who often need the funds for timely planting in the subsequent season [1]. This research aims to enhance food security by reducing the costs associated with crop insurance, thereby increasing the uptake of the product. This is achieved using Artificial Intelligence (AI) algorithms.

Methods: In situ weather stations and remote sensing satellites are used to get weather related information including, rainfall, temperature and soil moisture. The information from the different sources is used to create a weather index. Machine learning is used to predict the rainfall and soil moisture of a specific region based on historical rainfall and soil moisture data going back forty years. The generated prediction is used to determine if the growing period for a farmer will result in a trigger. The trigger is based off, of historical data where the soil moisture is below an established trigger. Claim payments are based on the realization of an objectively measured weather variable (e.g., rainfall) that is correlated with production losses.

Conclusion: The prediction accuracy of the machine learning model plays a significant part in the insurance instrument. A high accuracy in the model will ensure that a farmer gets a payout for expected low-yields based on weather inputs. Crop insurance is recognized to be a basic instrument for maintaining stability in farm income, through promoting technology, encouraging investment, and increasing credit flow in the agricultural sector. It contributes to self-reliance among farmers, since in cases of crop loss they can claim compensation as a matter of right. Thus, crop insurance cushions the shock of crop loss by assuring farmers protection against natural hazards beyond their control [2]. Weather pattern crucially affects dynamics of soil moisture, which in turn drives vegetation dynamics. Weather index-based insurance is a product designed to help farmers cope with the risk of uncertain weather patterns. This research done in partnership with Twiga EU.

Keywords: Crop Insurance, Machine Learning, Data Analytics, Weather, Weather index-based, Food Security

References

1. Sibiko, K., Veettil , P., Qaim , M. (2018). Small farmers' preferences for weather index insurance: insights from Kenya. Agriculture Food Security.

2. Nagentran, Rajendran. (2017). Economic Analysis of Crop Insurance A Critical Review. Shanlax International Journal of Economics.

3. Salami, A., Kamara, A., Brixiova, Z. (2010). Smallholder Agriculture in East Africa: Trends, Constraints and Opportunities. Retrieved 8 February 2022, from https://www.afdb.org/sites/default/files/documents/publications/working_105_pdf_d.pdf

Contact details

- Email(s): kbett@strathmore.edu, jolukuru@strathmore.edu

- Homepage: http://www.ilabafrica.ac.ke/, https://strathmore.edu/

20.3 Applying computer vision for high precision 360 degree car damage assessment

Vinayak Arannil, Amazon Web Services (presenter)
Atanu Roy, Amazon Web Services

Abstract: In this paper, we demonstrate an effective multi-stage solution to perform car damage assessment using multiple state-of-the-art computer vision models, so that insurance companies can prevent claim leakage and improve claim resolution time by automating part of their claim process. This can also be used during customer onboarding, if an insurer requires identification of all existing damages on the car exterior while creating the policy. Unlike other available approaches, our solution can perform 360 degree damage assessment using multiple camera views and accurately detect and assess severity for a wide range of damages, like hairline scratches, dots, minor dents right up to heavy dents and tears. Additionally, we also detect the part associated with each damage which is very important for claim amount settlement, as the cost varies based on the part involved and the extent of damage. On a proprietary real life dataset, our solution achieved a precision of 92% and a recall of 94%.

Keywords: car damage assessment, vehicle insurance, instance segmentation, image classification, vision transformer, maskrcnn

References

1. T. -Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 936-944, doi: 10.1109/CVPR.2017.106.

2. Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick, "Mask RCNN", IEEE International Conference on Computer Vision (ICCV), 2017.

3. S. Xie, R. Girshick, P. Dollár, Z. Tu and K. He, "Aggregated Residual Transformations for Deep Neural Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 5987-5995, doi: 10.1109/CVPR.2017.634.

4. Dosovitskiy, Alexey et al. "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." ArXiv abs/2010.11929 (2021): n. pag.

Contact details

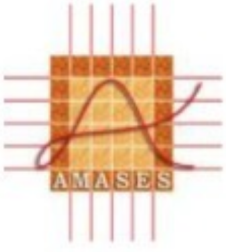
- Email: varannil@amazon.com

- Social media: Linkedin: https://www.linkedin.com/in/vinayak-arannil-69789768/

Gold sponsors

Silver sponsors

Partnerships



Swiss Re

Digital risk solutions and platform driven business

We embed digital technology in everything we do to simplify and accelerate processes and **optimize customer journeys**. This enables us to drive new tailored, accessible, and affordable re/insurance solutions to help **close the protection gap**. With the scalability of digital services via platforms and ecosystems we can protect more people and businesses from financial hardship faster.

- **Concrete use case** - What we mean by seamless, optimized customer journeys

  The satellite-based solution for natural catastrophes from Swiss Re and the satellite operator ICEYE does not require any damage photos from the insured. It records floods in high resolution even in complex terrain such as urban centres, where conventional flood modelling reaches its limits. Swiss Re gains access to data within 24 hours of an event and makes it available to partners on the CatNet® Premium platform. This enables them to record the damage more quickly and initiate payments. There are plans to extend the monitoring to other kinds of natural disasters, as well as to use it as an early warning system.

Tech partnerships for progress

With selected strategic **tech partners** for public clouds and data analytics we set the basis for the secure integration of our services, as well as the accessibility and scalability of our solutions. Throughout our value chain we take the opportunities to collaborate with specialized **innovation partners** for a win-win situation; be it with joint-ventures or as participant in industry ecosystems.

- **Concrete use case** – scaling via ecosystems with embedded digital insurance

  Best known example of our customer centric approach is the embedded affordable household contents and private liability insurance product, Hemsäker, iptiQ created for IKEA. With Hemsäker, furnishing a home can go hand-in-hand with buying home insurance. A fully digital journey means customers can get a quote and buy coverage in a few, easy steps.

Trust, Data governance, ethics

For a responsible digitalisation we foster rigorous standards and frameworks on digital governance, data protection and privacy, transparency, conduct and analytical model validation. Our digital transformation goes hand in hand with the fostering of data culture and innovation, the upskilling of employees in data literacy and agile working practices. **We are also actively addressing the ethical challenges of new technologies** and we participate actively in the evolving regulation e.g., EU on AI and ML. We want to ensure **fair treatment and inclusion** beyond regulation.

COMPANY PROFILE

Banca Mediolanum S.p.A. is the parent company of the Mediolanum Banking Group (present in 4 European countries) and of the Mediolanum Financial Conglomerate, which includes Italian and foreign insurance companies. Banca Mediolanum is one of the leading players in the banking and asset management market in Italy and Europe. Banca Mediolanum has been listed on the Italian Stock Exchange since 1996 and is part of the FTSE MIB.

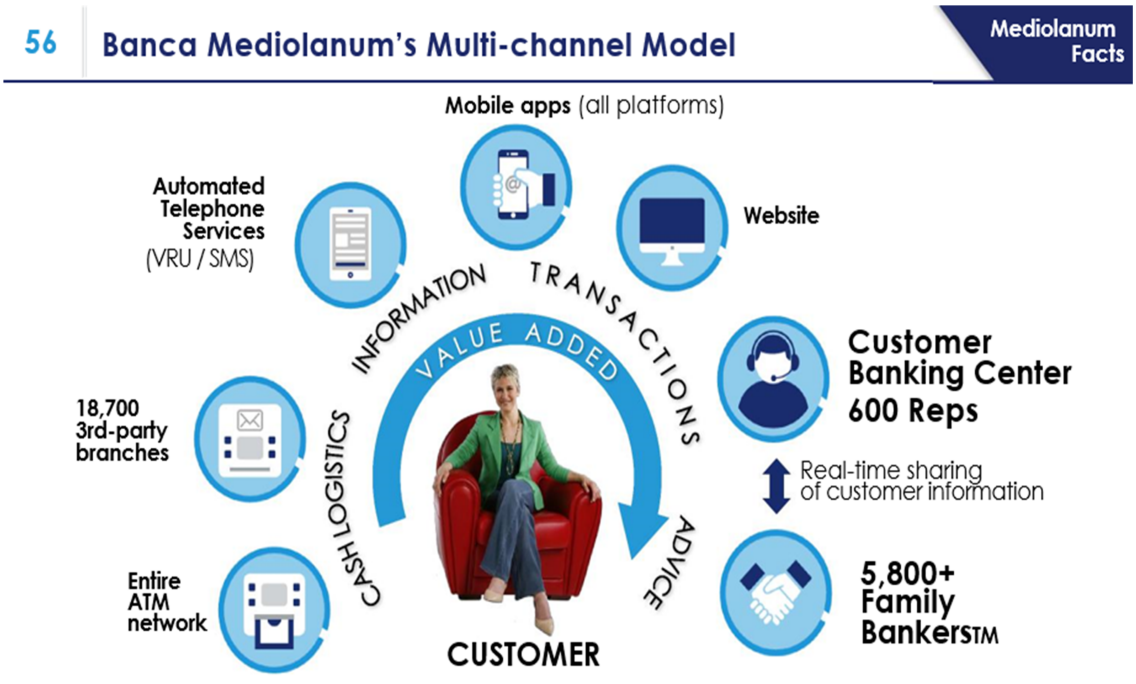
Its mission is to enhance the financial resources of Italian families placing the customer at the 'center' in providing solutions that best fit their needs, whether it be in the form of a mutual fund, an insurance policy, or a bank product.

Banca Mediolanum is an extremely innovative bank that offers customers the full range of products and services through its integrated multi-channel business model - when, where and how the customer chooses, thanks to a constant focus on innovation in terms of accessing and interacting with the Bank.

Banca Mediolanum has a sales network of 5,855 Family Bankers who are the financial advisors entrusted with the job meeting all the needs regarding investments, insurance, protection and retirement savings, assisting the customers with their most important financial decisions for themselves, their future and their families.

For the more sophisticated needs of high net worth individuals, the Wealth Management division offers personalized solutions such as capital protection, fiduciary services, inheritance and succession planning, art consultancy, tax & legal consultancy.

The Investment Banking division was formed to develop the business of Advisory and Investment Banking services with focus on SMEs.



Are we an Insurtech company? Sure, but we don’t brag about it!

We provide insurance and digital services but without being the white collars insurers: technology, creativity and innovation are the beating heart of our company.

Discover our Datalab

ADVANCED
We use technology and innovation applied to insurance but we are not the usual data lab

INNOVATIVE
We start from the needs of customers and create products and services that are easy to understand and practical to use, we are not the usual incubator

UNIQUE
Not the usual chatter: we like to speak clearly, simply and concretely.

We are pioneers in the insurance proposal in niche markets where it is not enough to use traditional models and techniques but it is necessary to **think outside the traditional schema**.

Do you have a disruptive idea that you think works but you've never had the courage to talk to someone about it?

Talk to us! We are always looking for people with **unconventional approaches** to apply to the European insurance market.

Contact 24hassistance to talk about your idea
christian.invernizzi@24hassistance.com

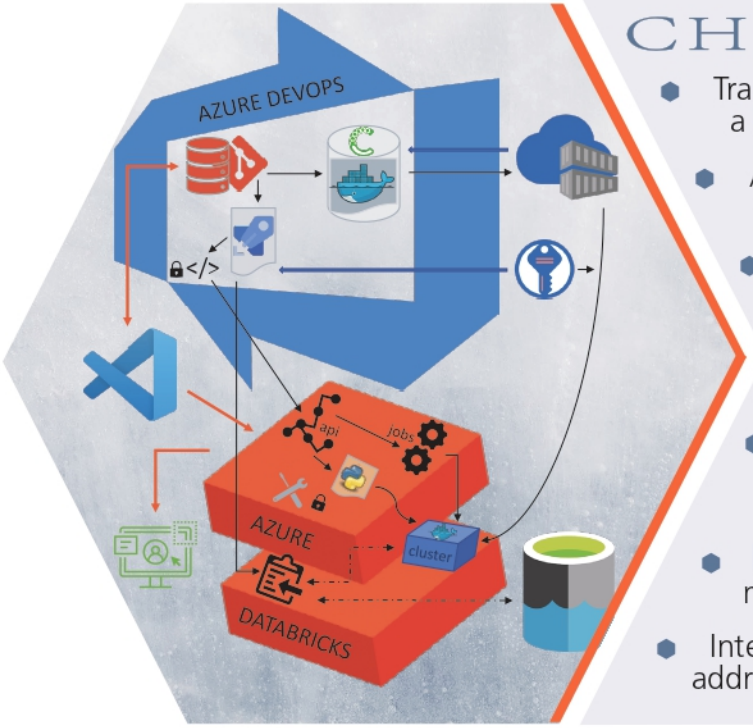
Azure Cloud Transition with Databricks & DevOps

Mirai Solutions

Scalable analytic application in the cloud using Azure DevOps & Databricks

We have successfully completed the redesign of a custom analytic retrocession engine, while laying the foundations for a broader analytics / data science platform, which will allow the team of a reinsurance company to enhance and scale their capabilities efficiently in an environment of enterprise-grade compliance and security.

Streamlined processing and efficient agile DevOps.



CHALLENGES

- Transition to a more or less blank cloud environment. From a core legacy application and data flows / model.
- Architect a performant, scalable next-gen solution, highly flexible and extendable, while assuring robustness.
- Simultaneous use for production as well as ad-hoc inquiries and exploratory what-if scenarios.

SOLUTION

- Custom Python / Spark application on Azure Databricks with Lakehouse architecture. MS Azure services like Data Factory for data flows and PowerBI for reports.
- Azure DevOps with repos & pipelines for agile development, CI/CD and "Infrastructure As Code" automation.
- Integrated platform features to manage/monitor resources, address permissions and security, backup, recovery, audit.

BENEFITS

- Decrease time to market, better insights:
  Faster and better assessment of various use cases and new potential business.
  Supports innovative ideas and complex modeling through enhanced flexibility and performance.
  Efficient resource handling and scalable big data capabilities with curated Spark.
- Cutting-edge technology:
  Services can be seemlessly upgraded to latest versions with the latest features.
- Dual mode with single source of truth:
  Azure solution enables scheduled execution of batch jobs and at the same time safe exploration of production data by actuaries via notebooks.

"Azure DevOps facilitates best practices of modern (Sec-)DevOps and CI/CD. From the start we set the project up with automated build pipelines, which we could later easily complement with release pipelines to ultimately have fully automated and governed deployments."

contact@mirai-solutions.com

"We are what we do" is not just a philosophy, it is our mission. Focusing on customers is what makes us different. It is why HDI Assicurazioni is truly "at your side every day".

We have a long history, but we look ahead. Our history is rooted in our country, Italy, where in 1881 a group of railway workers established "Società di Mutuo Soccorso", a mutual insurance association. Today, HDI Assicurazioni is the result of the path taken more than a century ago.

In 1997, HDI went another step further when it joined the Hannover-based Talanx Group (HDI Group offshoot), a major German insurance group of international standing. Talanx is the third largest insurance company in Germany, active in 150 countries around the world. As part of this network, HDI Assicurazioni has an Italian heart and a solid global vision.

In Italy, HDI relies on 500 insurance agents evenly distributed throughout the country and has expanded its footprint through partnerships with major players in the banking industry to complete its distribution network.

In 2021, with the acquisition of Amissima Assicurazioni SpA, HDI Assicurazioni moved to 11th place in the ranking of non-life insurers in Italy.

HDI Assicurazioni's strategic vision, in line with the Group's principles, is based on putting customers first. We want to be an insurance company that is always connected to our customers, capable of improving the quality of our offer and service. We strive to constantly improve the service we offer our policyholders by differentiating our offer and optimising our rates.

For HDI, being "at your side every day" means:

- developing tailor-made offers and services distributed on specific channels;
- upholding the principles of social responsibility and carefully considering the communities in which we live;
- building, over time, relationships based on trust, loyalty and mutual respect with people who believe in the importance of insurance as a useful tool to protect and improve their lifestyles.

List of Attendants (in-person and online)

Surname	Name	Affiliation
Abrignani Malchiorre	Danilo	Leithà
Andres	Hervé	Ecole nationale des ponts et chaussées
Arannil	Vinayak	Amazon Web Services
Badescu	Andrei	University of Toronto
Baione	Fabio	Sapienza University of Rome
Barbaglia	Luca	Mediolanum
Basler	Hannes	SCOR
Bauwelinckx	Yves-Cédric	Katholieke Universiteit Leuven
Beccalli	Elena	Università Cattolica del Sacro Cuore
Bett	Kelvin	Strathmore University
Biancalana	Davide	Sapienza University of Rome
Biessy	Guillaume	LinkPact - Sorbonne Université
Bladt	Martin	University of Lausanne
Blier-Wong	Christopher	Université Laval
Bormetti	Giacomo	University of Bologna
Borromini	Andrea	Vittoria Assicurazioni SpA
Boucher	Jean-Philippe	Université du Québec à Montréal (UQAM)
Bouse	Eilish	Grant Thornton
Breitig	Marco	Allianz Suisse
Brenda	Paola	EIOPA
Brufatto	Verena	Unipolsai Assicurazioni
Caccone	Manuel	Unipolsai Assicurazioni
Cairoli	Jacopo	Vittoria Assicurazioni SpA
Cannone	Eleonora	24h Assistance
Cardoso	Luca Antonio	Vittoria Assicurazioni SpA
Castellarin	Attilio	University of Bologna
Chan	Sophia	University of Toronto
Charpentier	Arthur	Université du Québec à Montréal (UQAM)
Cherkaoui Tangi	Yousra	Crest Ensae
Chiani	Claudio	Unipolsai Assicurazioni
Chong	Wing Fung	Heriot-Watt University
Clemente	Gian Paolo	Università Cattolica del Sacro Cuore
Cornioli	Elisa	Unipolsai Assicurazioni
Crippa	Mauro	Reinsurance Group of America (RGA)
Crupi	Giuseppe	De Angelis Savelli e associati
Dambon	Jakob	Swiss Re
D'Amico	Martina	HDI ASSICURAZIONI S.P.A.
De Cock Campo	Bavo	Katholieke Universiteit Leuven
De Luca	Ciro	Intesa San Paolo RBM Salute
De March	Davide	Markel International
Della Corte	Francesco	Università Cattolica del Sacro Cuore
Delong	Lukasz	SGH Warsaw School of Economics
Deslais	Thomas	Wakam
Donati	Paolo	Reale Mutua Assicurazioni
Doni	Angelo	ANIA

Continued on next page

Table 1 – *Continued from previous page*

Surname	Name	Affiliation
Dræge	Anders	Frende Forsikring
Elmer	Greg	Utica National Insurance Group
Facchinetti	Silvia	Università Cattolica del Sacro Cuore
Fan	Zijie	Guy Carpenter
Flaig	Solveig	University of Oldenburg, Germany
Frischknecht	Simon	Swiss Re
Gamberini	Daniela	Swiss Re
Gawlowski	Karol	AXA XL
Gesmann	Markus	Insurance Capital Markets Research
Ghiradi	Nicola	Leithà
Giancaterino	Claudio	Università Cattolica del Sacro Cuore
Giorgi	Daphné	Sorbonne Université
Glotzer	John Paul	Admiral Group
Gorge	Guillaume	CNAM (ext)
Grandis	Massimo	Mediolanum
Gross Lorenzi	Laura	Austral Re
Grzywaczyk	Witold	EY
Haringa	Martin	NN Group
Hartveit	Marit	Element Insurance AG
Haugen	Rune	Eika Forsikring AS
Havrylenko	Yevhen	Technical University of Munich
Heger	Julia	University of Augsburg
Henriksen	Sindre	Eika Forsikring AS
Heydari	Muhammad Javad	Insurance Research Center (IRC)
Hieber	Peter	Université de Lausanne
Hogan	Patrick	PartnerRe
Holvoet	Freek	Katholieke Universiteit Leuven
Hong	Bingyan	Faraday Underwriting
Igobwa Mugwe	Alvin	Strathmore University
Jose	Alex	Heriot-Watt University
Jules	Emilien	Swiss Re
Kang	Ethan	Accelerant Insurance
Karmazyn	Aleksandra	Aviva Services
Kedidi	Islem	university of sousse
Kessy	Salvatory	University of New South Wales
Khudeeda	Shekhzad	Bielefeld University
Klåpbakken	Øyvind	Eika Forsikring AS
Ko	Jordan	SAS Institute
Komanski	Piotr	Aviva
Korn	Uri	Ledger Investing
Krymina	Oxana	VSK insurance company
Kyriakou	Ioannis	Bayes Business School, City, University of London
Laverny	Oskar	University of Lyon - SCOR SE
Levantesi	Susanna	Sapienza University of Rome
Li	Ziyi	Heriot-Watt University
Li	Yanfeng	University of New South Wales

Continued on next page

Table 1 – *Continued from previous page*

Surname	Name	Affiliation
Lin	Hong Beng	The University of Iowa
Liu	Ni	Guy Carpenter
Lopez	Olivier	Sorbonne University and Detralytics
Ludkovski	Mike	University of California Santa Barbara
Macrì	Francesca	Mediolanum
Marchi	Lorenzo	Katholieke Universiteit Leuven
Marcus	Dean	Guy Carpenter
Marino	Mario	Sapienza University of Rome
Mashechkin	Alexey	Allianz Partners
Mazzone	Jacopo	Università Degli Studi di Firenze
Mazzoni	Luca	Università Degli Studi di Firenze
McCabe	Johnathan	Faraday
Menvouta	Emmanuel Jordy	Katholieke Universiteit Leuven
Merolla	Roberta	Zurich Insurance Plc
Meroni	Nicholas	24h Assistance
Metulini	Rodolfo	University of Salerno
Mitev	Nikolina	Generali Deutschland AG
Morini	Costanza	24h Assistance
Mtalai	Itre	Carleton University
Mura	Fabrizio	Unipolsai Assicurazioni
Nardi	Lorenzo	Mediolanum
Ng	Shirley	Vantage Risk
Ottini	Michele	Unipolsai Assicurazioni
Parente	Fausto	EIOPA
Passarello	Gloria	Leithà
Pesenti	Silvana	University of Toronto
Petruccelli	Antonio	Leithà
Pigeon	Mathieu	Université du Québec à Montréal (UQAM)
Pitidis	Alessio	24h Assistance
Pittarello	Gabriele	Sapienza University of Rome
Ponnet	Jolien	Katholieke Universiteit Leuven
Pusz	Robert	Warsaw School of Economics
Rabitti	Giovanni	Heriot-Watt University
Rasanga	Fiona	University of Edinburgh
Raynal	Etienne	Université Claude Bernard Lyon 1
Rianna	Guido	Centro Euromediterraneo sui Cambiamenti Climatici
Richman	Ronald	Old Mutual Insure and University of the Witwatersrand
Rinaldi	Riccardo	Reale Mutua Assicurazioni
Riva	Andrea	Sapienza University of Rome
Riva De Onestis	Alessandro	Intesa San Paolo RBM Salute
Robben	Jens	Katholieke Universiteit Leuven
Sangari	Seema	Kennesaw State University
Savelli	Nino	Università Cattolica del Sacro Cuore
Schauer	Philipp	Generali Deutschland AG
Schelldorfer	Juerg	Swiss Re
Schmid	Roland	Mirai Solutions

Continued on next page

Table 1 – *Continued from previous page*

Surname	**Name**	**Affiliation**
SchnuRch	Simon	Fraunhofer Institute for Industrial Mathematics ITWM
Scognamiglio	Salvatore	University of Naples "Parthenope"
Sebastianelli	Pasquale	Deloitte Consulting
Senn	Markus	PartnerRe
Serafini	Simone	University of Bologna
Sgrignoli	Paolo	Swiss Re
Shoun	Mark	Ledger Investing
Smallman	Luke	Admiral Group
Sologni	Eugenio	Leithà
Spedicato	Giorgio	Leithà
Spina	Marco	Studio Attuariale De Angelis - Savelli e Associati
Sriram	Varun	Guy Carpenter
Steinskog	Dag Johan	Norwegian Hull Club
Stocco	Giacomo	Università di Verona
Streftaris	George	Heriot-Watt University
Streich	Andreas	PartnerRe
Stupfler	Gilles	ENSAI - CREST
Szatkowski	Marcin	Warsaw School of Economics
Talone	Gaia Laura	Unipolsai Assicurazioni
Teleon	Maciej	Quantee
Tirri	Antonio	Leithà
Trufin	Julien	Université Libre de Bruxelles
Tsanakas	Andreas	Bayes Business School, City, University of London
Turcotte	Roxanne	Université du Québec à Montréal (UQAM)
Ugarte Montero	Andrey	Université de Lausanne
Ungolo	Francesco	Technische Universität München
Uratani	Tadashi	Hosei University, Tokyo
Urbano	Davide	Zurich Insurance Plc
Vagnoli	Matteo	Swiss Re
Van Der Zwaan	Ruben	MavenBlue
Vermassen	Olivier	Wakam
Verschueren	Eva	Katholieke Universiteit Leuven
Verschuren	Robert Matthijs	University of Amsterdam
Vhudzijena	Michelle	School of Risk and Actuarial Studies, UNSW Sydney
Vivas Occhipinti	Ana	Wakam
Vormberg	Alexandra	Generali Deutschland AG
Weiss	Adrien	Wakam
Wong	Bernard	University of New South Wales
Wuthrich	Mario	ETH Zurich
Yung	Ho Yan Joey	School of Risk and Actuarial Studies, UNSW Sydney
Zakrisson	Henning	Stockholm University
Zappa	Diego	Università Cattolica del Sacro Cuore
Zatonski	Tomasz	EY
Zatsepin	Andrey	VSK insurance company
Zhang	Yaojun	University of Leeds
Zhu	Rui	Bayes Business School, City, University of London