Claims modelling for climate risk

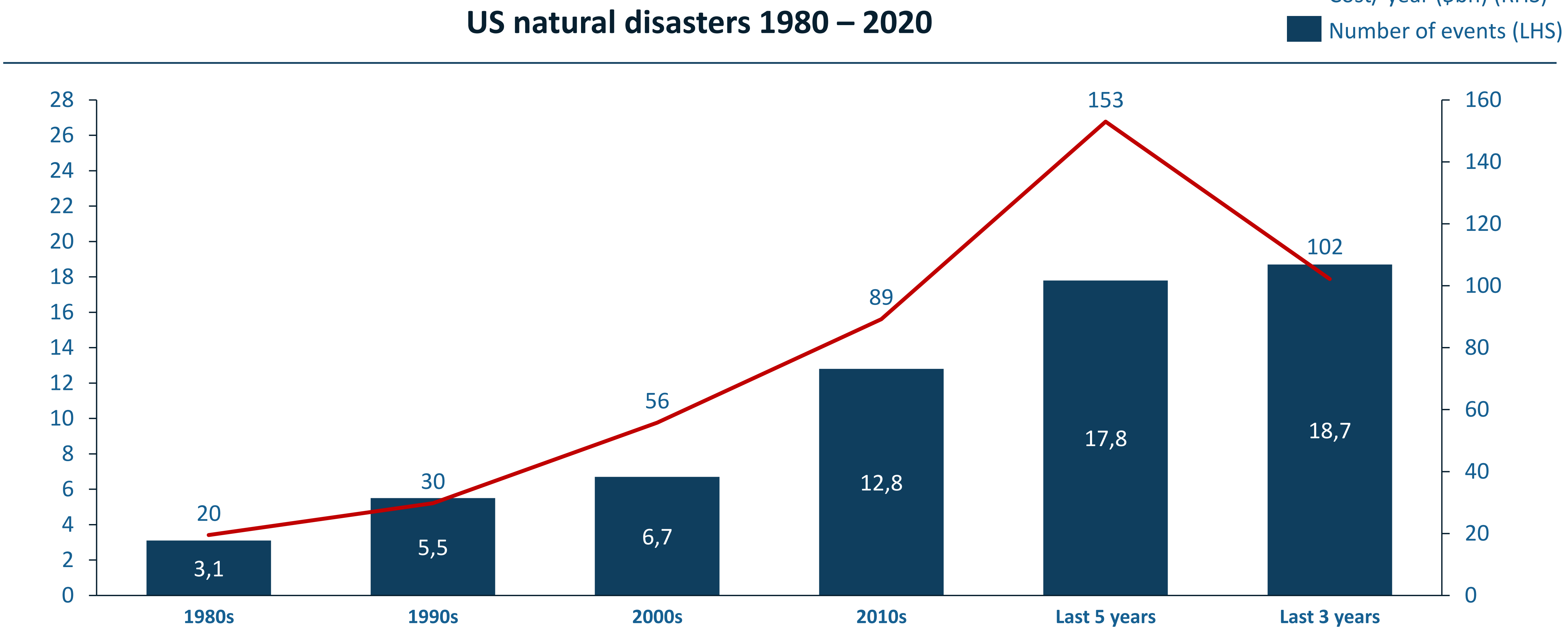
Ronald Richman

June 2024

Insurance Data Science

Background

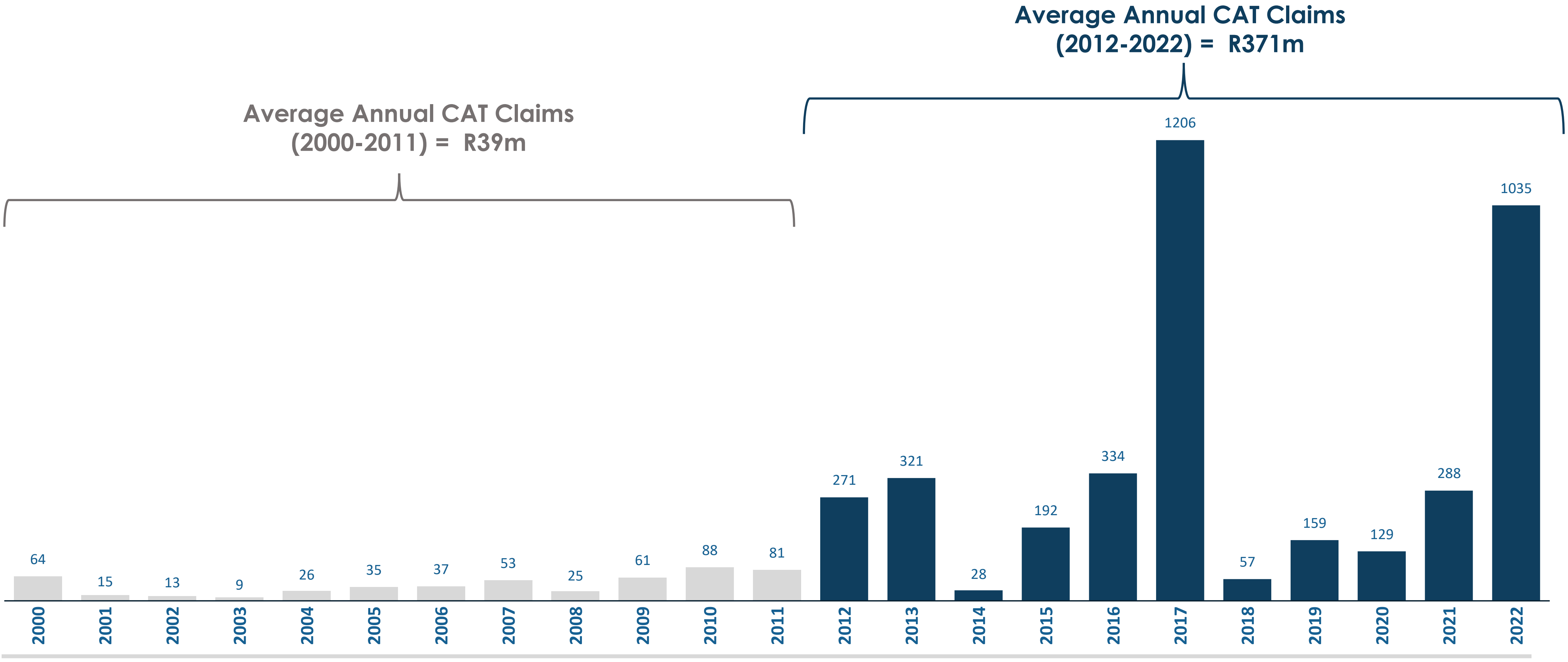
There has been an increase in both frequency and severity of natural disasters globally



Source: NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2022)

Severity Impact

Severity of weather-related claims has increased 10-fold over last decade



Source: Old Mutual Insure pricing data (inflation- and exposure-adjusted weather catastrophe claims) R'mil

Macro and Micro Modelling

Insurance
Data
Science

Macro view

- Pre-existing models of shocks to short-term insurance portfolio:
  - Earthquake
  - Hail
  - Wildfire
  - Flood
  - Windstorm
- Models calibrated to recent experience of these perils
- Run at a portfolio level
- <u>Can we modify these models to take climate change into account?</u>

Micro view

- Pricing data links individual policies in portfolio to claims data
- Can also acquire climate data looking at experience at granular level…
- … e.g. precipitation data in small areas for a long period
- <u>Can we link climate data to our traditional pricing to quantify effect of climate change?</u>

Micro - Short-term Weather Forecasting

- **Project aim**
  - Can we link climate data to our traditional pricing to quantify effect of climate change?
  - Incorporate highly granular precipitation data, curated by meteorologists, into traditional short-term pricing datasets.
  - Fit statistical models to observe predictive value of this addition.
  - Quantify the potential impact of using future predicted precipitation levels in rating processes
  - Quantify the impact of increased precipitation (driven by climate change and La Nina weather system) on insurance risk
- **Project with support from:**
  - University of the Witwatersrand (Prof. Rendani Mbhuva, Adam Balusik)
  - University of Pretoria (Prof. Willem Landman)
  - ETH Zürich (Prof. Dr. Mario V Wüthrich)
  - OMI Catastrophe & Climate Modelling (Caesar Balona)
- **Working paper in progress**

Micro - Short-term Weather Forecasting

- **Overview of steps taken**
  - Select one line of business
  - Geolocate LoB pricing file using external service provider
  - Obtained CHIRPS precipitation dataset
  - Created precipitation grid across SA at a 0.05’ longitude by 0.05’ latitude level of granularity (~25km^2)
  - Mapped geolocated pricing file to the precipitation grid
  - Fit Gradient Boosted Machines (GBMs) model to predict claims experience using factors used in the current pricing environment, with and without precipitation
  - Fit a Neural Net to disperse overall South African rainfall forecasts to a grid level
  - Refit models using forecasted rainfall
  - Analyzed model results on an actual and forecasted basis
    - Feature importance
    - Dependence plots
    - Predicted loss experience by yearly rainfall experience (actual and forecasted basis)

Geolocation – Data Engineering

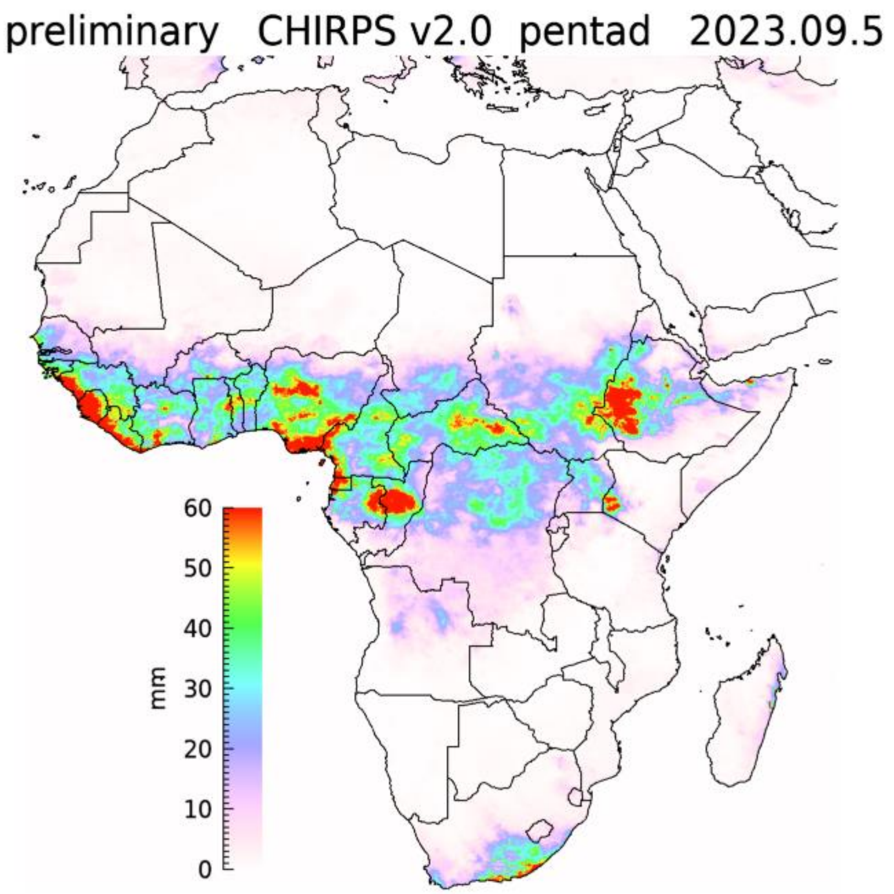
- **Data Considerations**
  - Geolocated LOB pricing file
    - ~ 13mil rows and many columns
  - CHIRPS precipitation dataset
    - ~ 19.5mil rows and 4 columns
  - Memory management and optimisation becomes very important
    - Python – Pandas
      - Batch processing
      - Memory efficient data storage
      - Minimum viable datatypes
      - Use vectorized operations where possible
      - Utilize GPU for modelling

Precipitation – CHIRPS Overview

Insurance
Data
Science

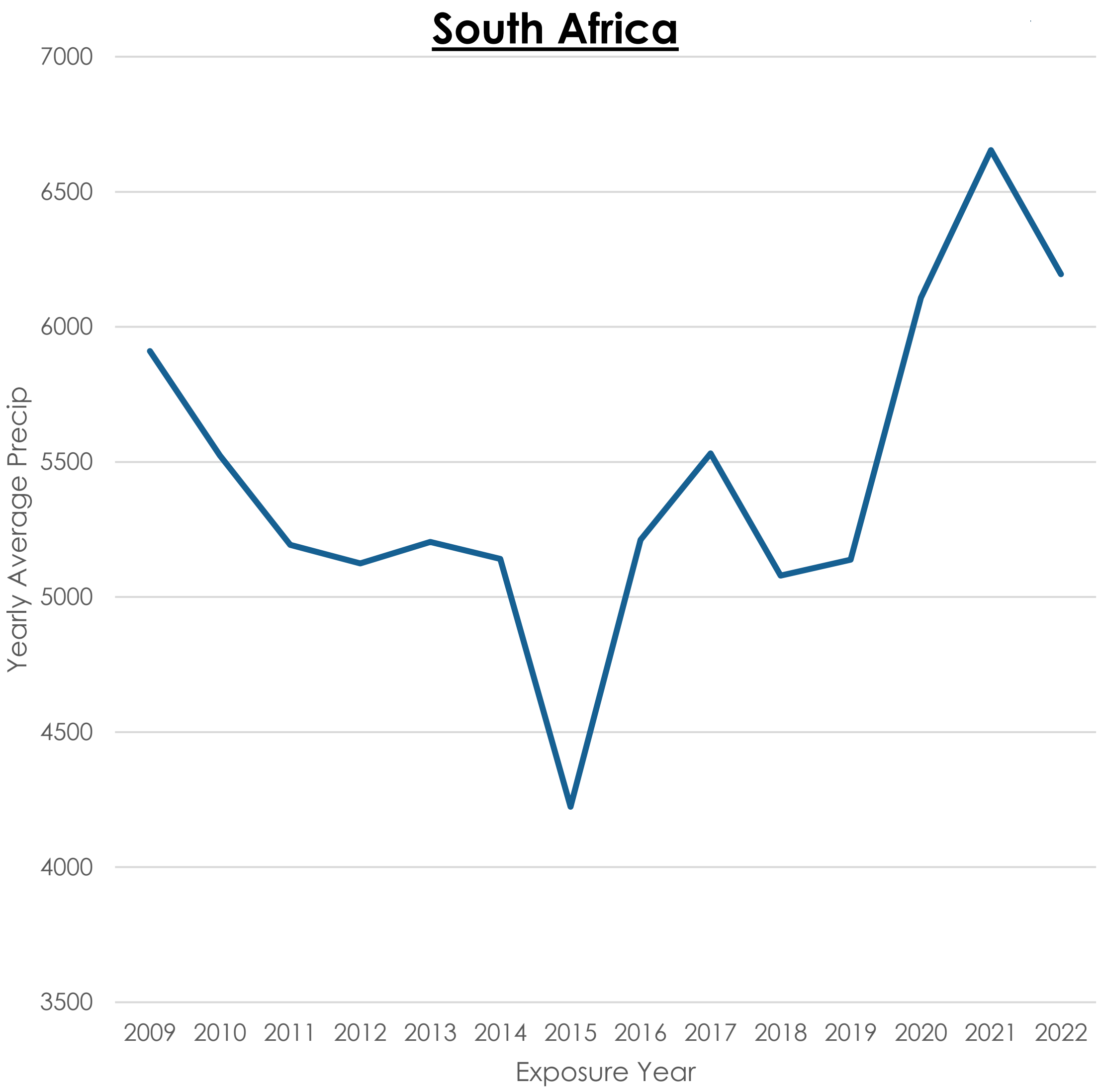
Climate
Hazards
Center
UC SANTA BARBARA

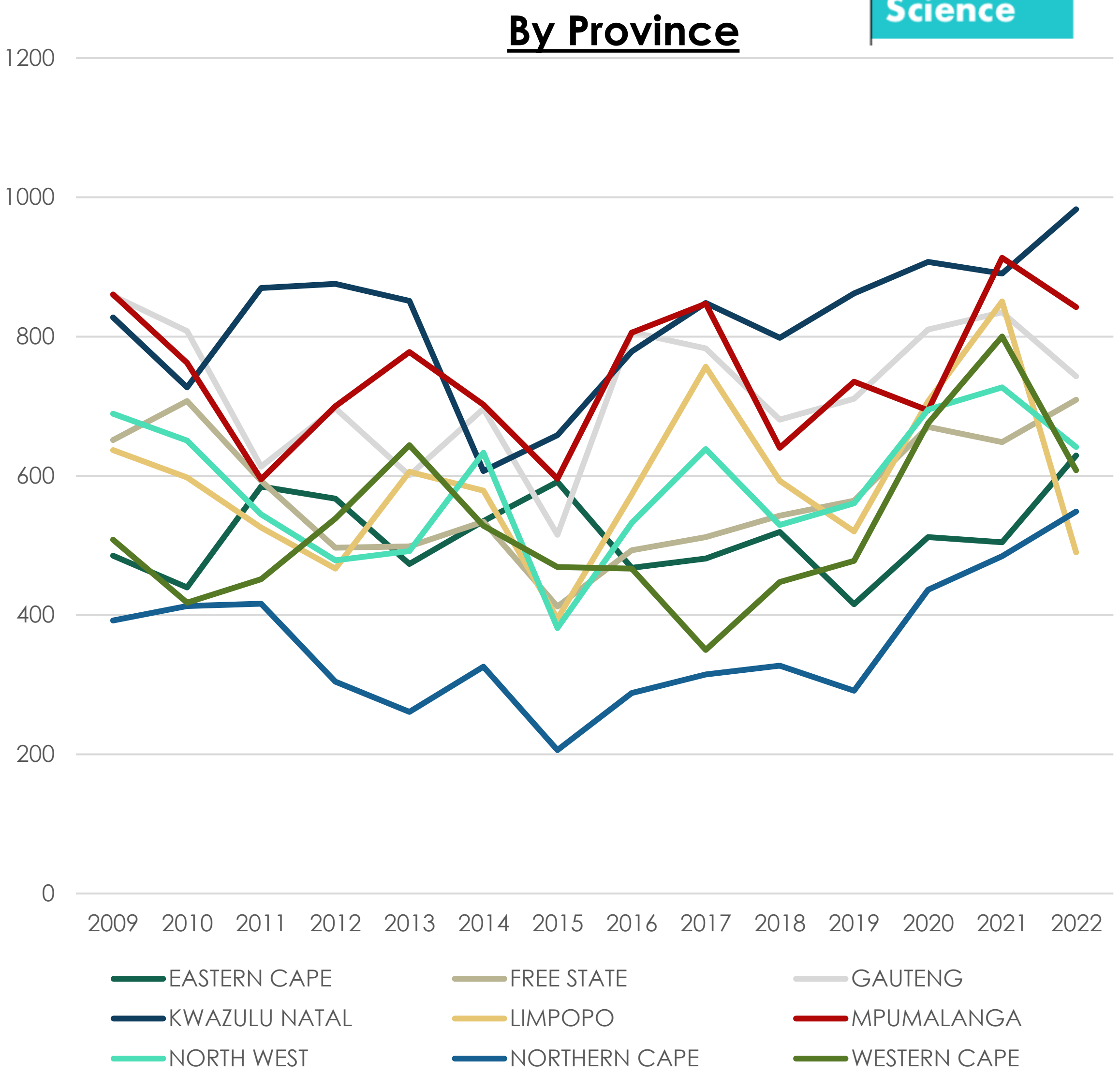
- **CHIRPS Dataset**
  - Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) is a 35+ year quasi-global rainfall data set.
  - Spanning 50°S-50°N (and all longitudes) and ranging from 1981 to near-present.
  - CHIRPS incorporates in-house climatology, 0.05° resolution satellite imagery, and in-situ station data to create gridded rainfall time series for trend analysis and seasonal drought monitoring.



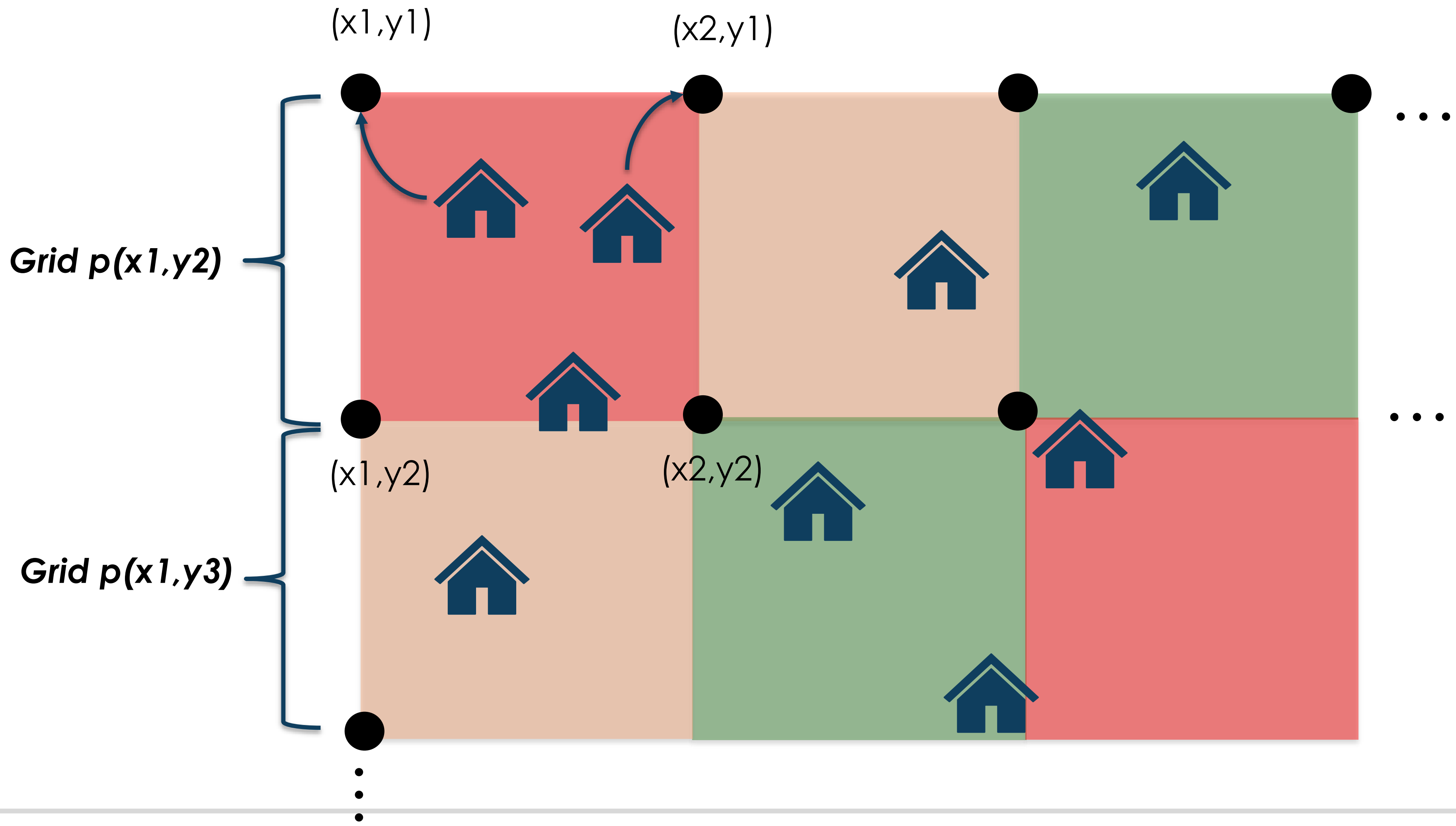
Precipitation – CHIRPS Visualisation

Insurance
Data
Science





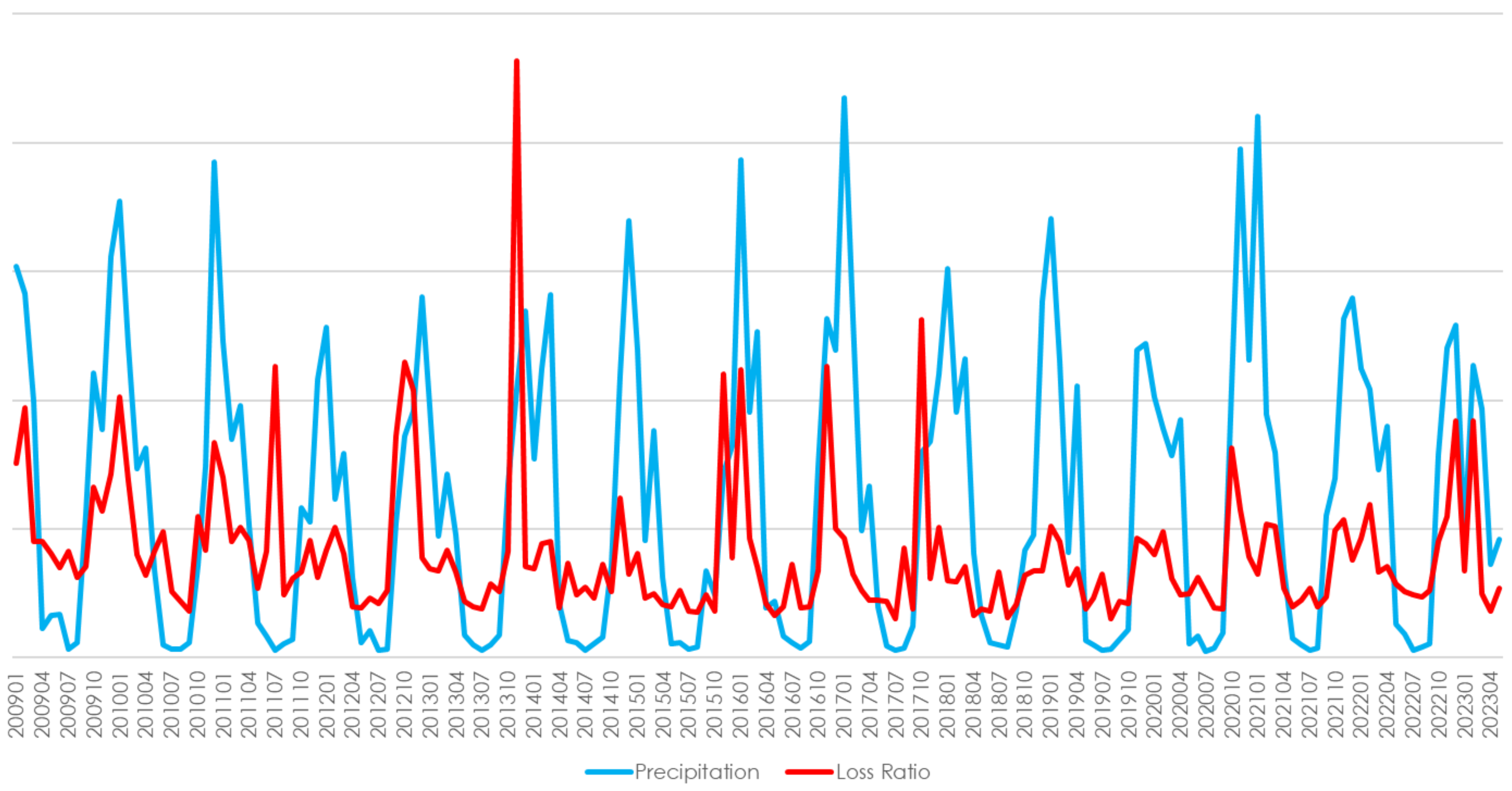
Linking Exposure to Precipitation – Join Logic



Linking Exposure to Precipitation - Visualisation

Insurance
Data
Science

Gauteng - Precipitation vs Loss Ratio



Modelling Implementation

Loss prediction given precipitation experience

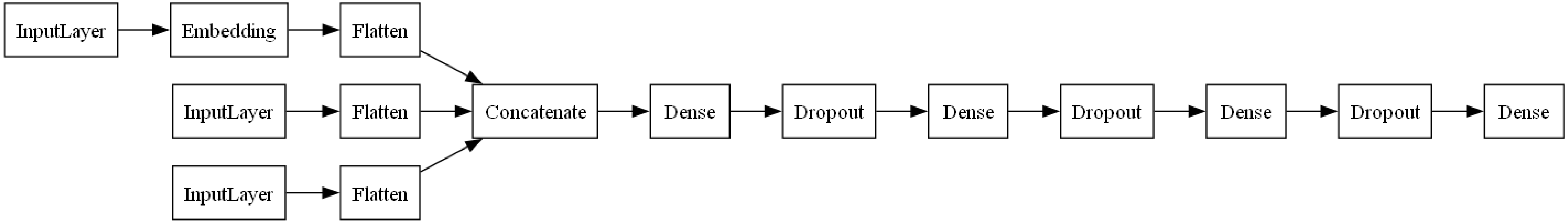
Frequency GBM	
Model	Gradient Boosted Machine
Form	Poisson Regression
Algorithm	LightGBM
Train/Test Split	Time-based
Loss function	Poisson Negative Log-Likelihood
Inputs	Traditional rating factors +- (Grid Precipitation)
Weight	Exposure
Output	Frequency
Validation score	Poisson Mean Deviance

Severity GBM	
Model	Gradient Boosted Machine
Form	Gamma Regression
Algorithm	LightGBM
Train/Test Split	Time-based
Evaluation metric	Gamma Negative Log-Loss Likelihood
Inputs	Traditional rating factors +- (Grid Precipitation)
Weight	Exposure
Output	Severity
Validation score	Gamma Mean Deviance

Modelling Implementation

Forecasting precipitation

Grid Dispersion NN	
Model	Neural Net
Form	Poisson Regression
Algorithm	Keras
Train/Test Split	Random
Loss function	Mean Squared Error
Inputs	Grid cell bounds, Overall precipitation prediction*, Calendar month
Output	Per grid cell precipitation
Validation score	MSE



Modelling Results

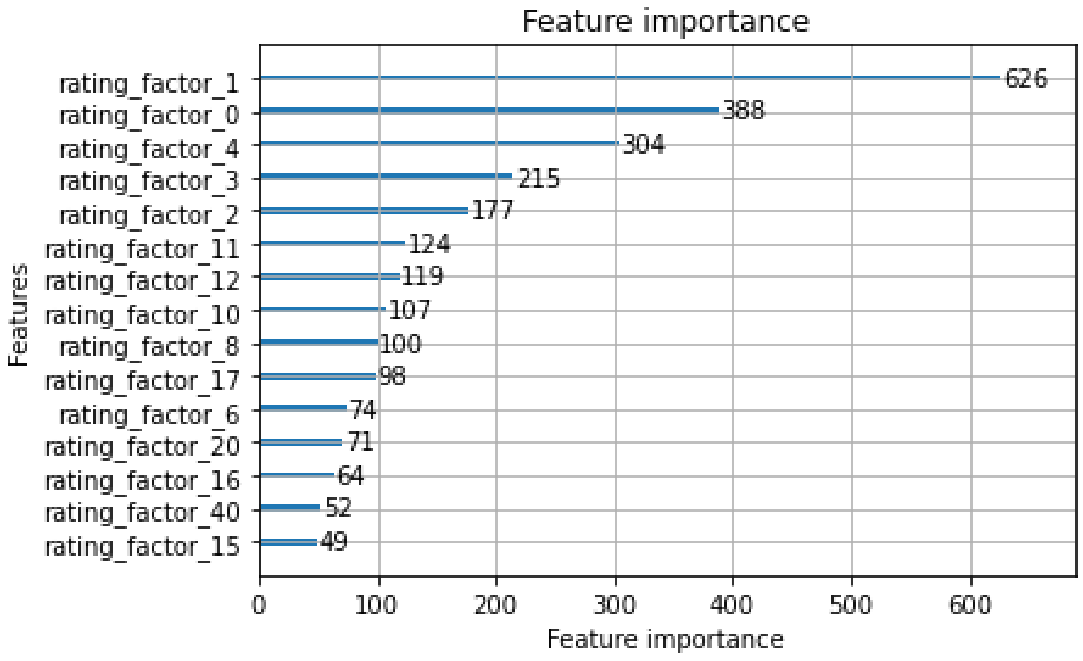
Out-of-sample validation scores

Model	Poisson/Gamma Deviance
Frequency GBM w/o precipitation	0.1687
Frequency GBM w/ actual precipitation	0.1679
Frequency GBM w/ forecasted precipitation	0.1683
Severity GBM w/o precipitation	1.7833
Severity GBM w/ actual precipitation	1.7465
Severity GBM w/ forecasted precipitation	1.7775

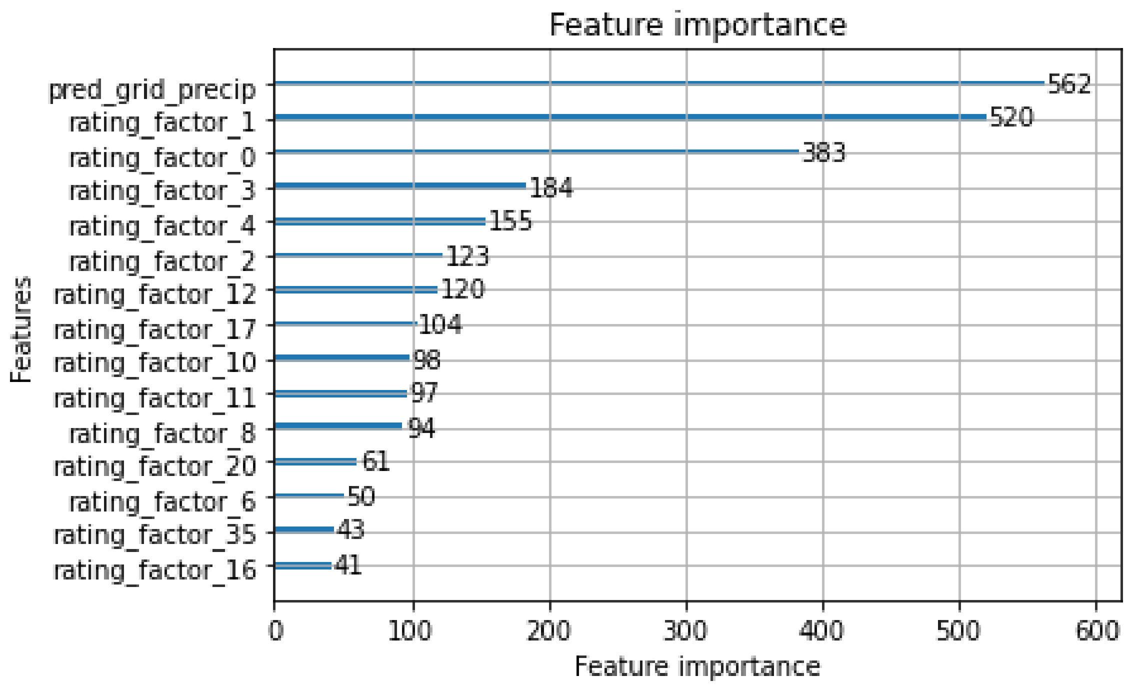
Modelling Results

Frequency GBM Implementation

Traditional Pricing Dataset



With Forecasted Precipitation Data

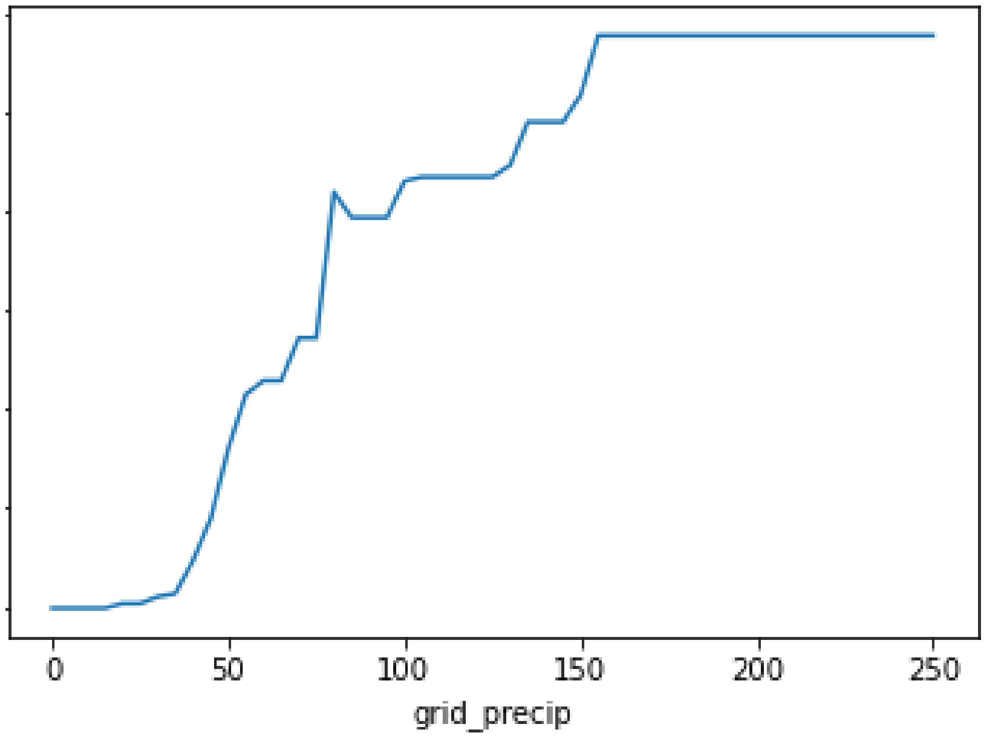


Modelling Results

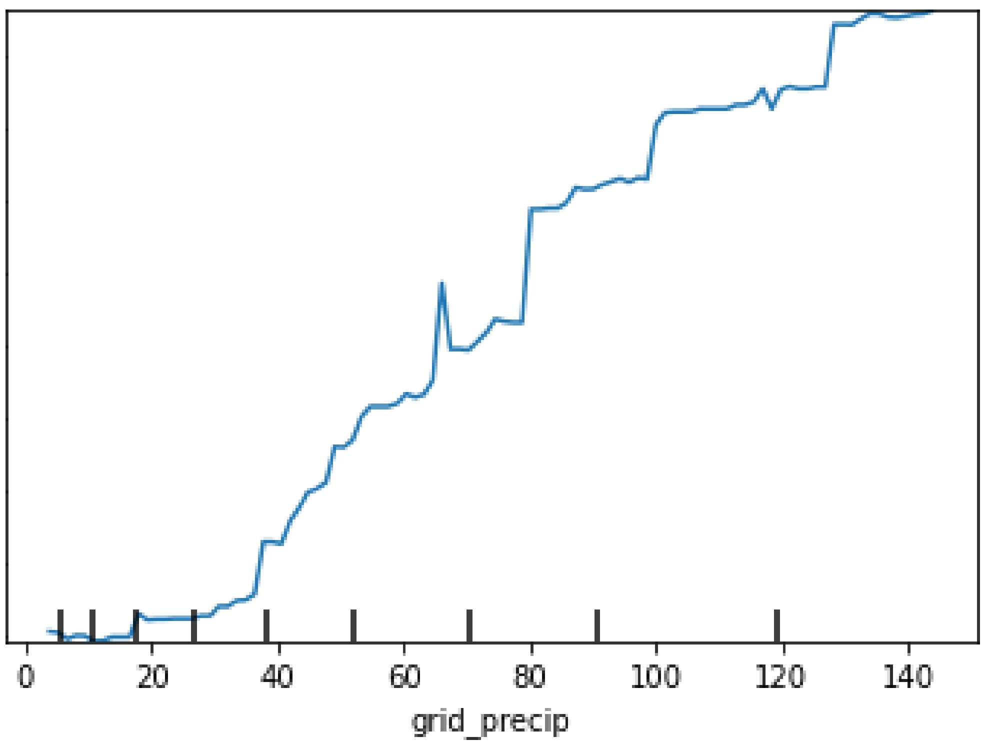
Insurance
Data
Science

Frequency GBM Implementation

Sample P/H Sensitivity (Base Risk Profile)



Partial Dependency Plot

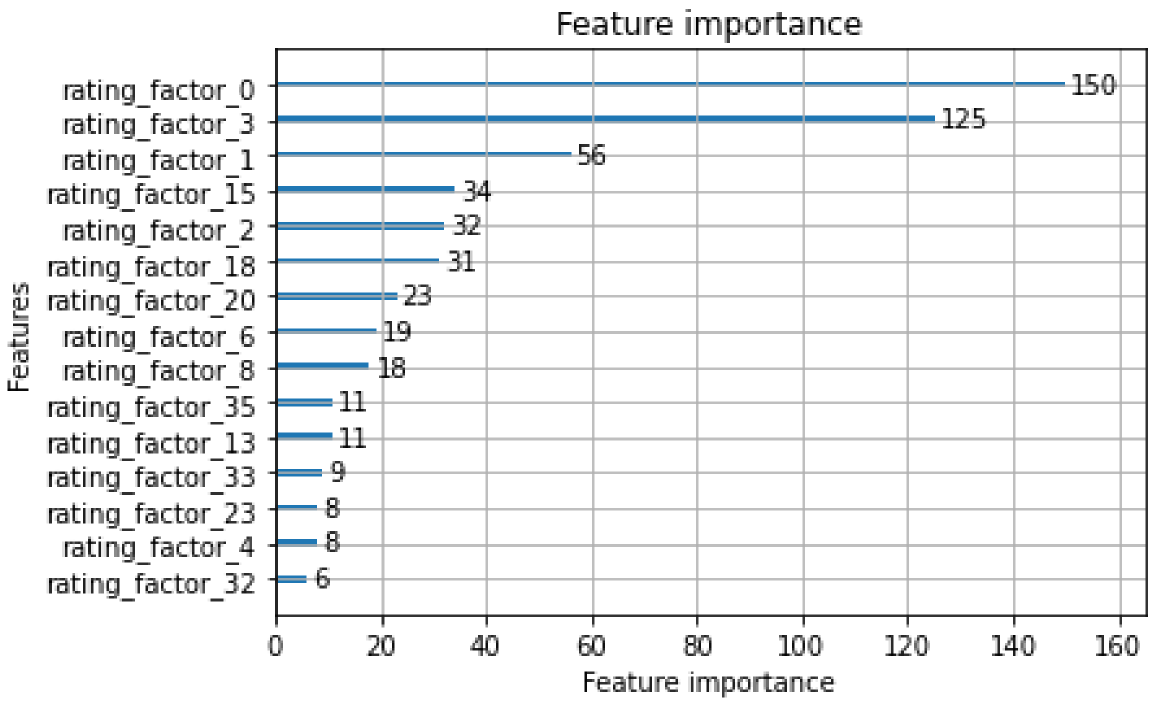


Modelling Results

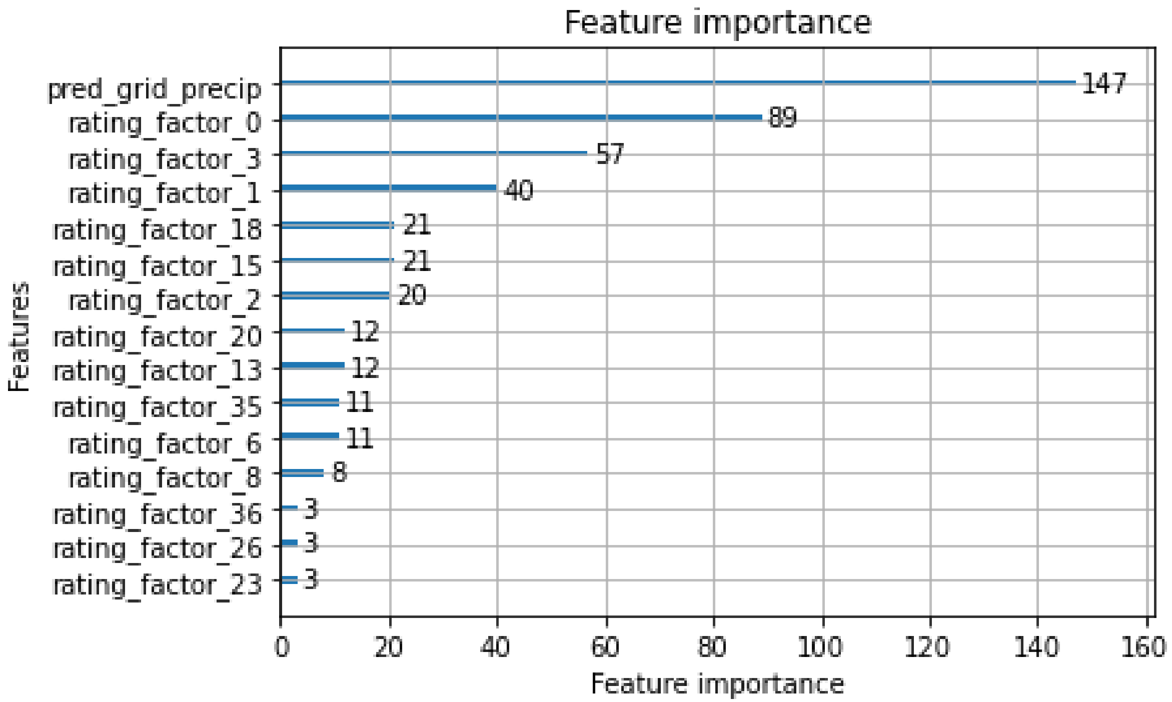
Insurance
Data
Science

Severity GBM Implementation

Traditional Pricing Dataset



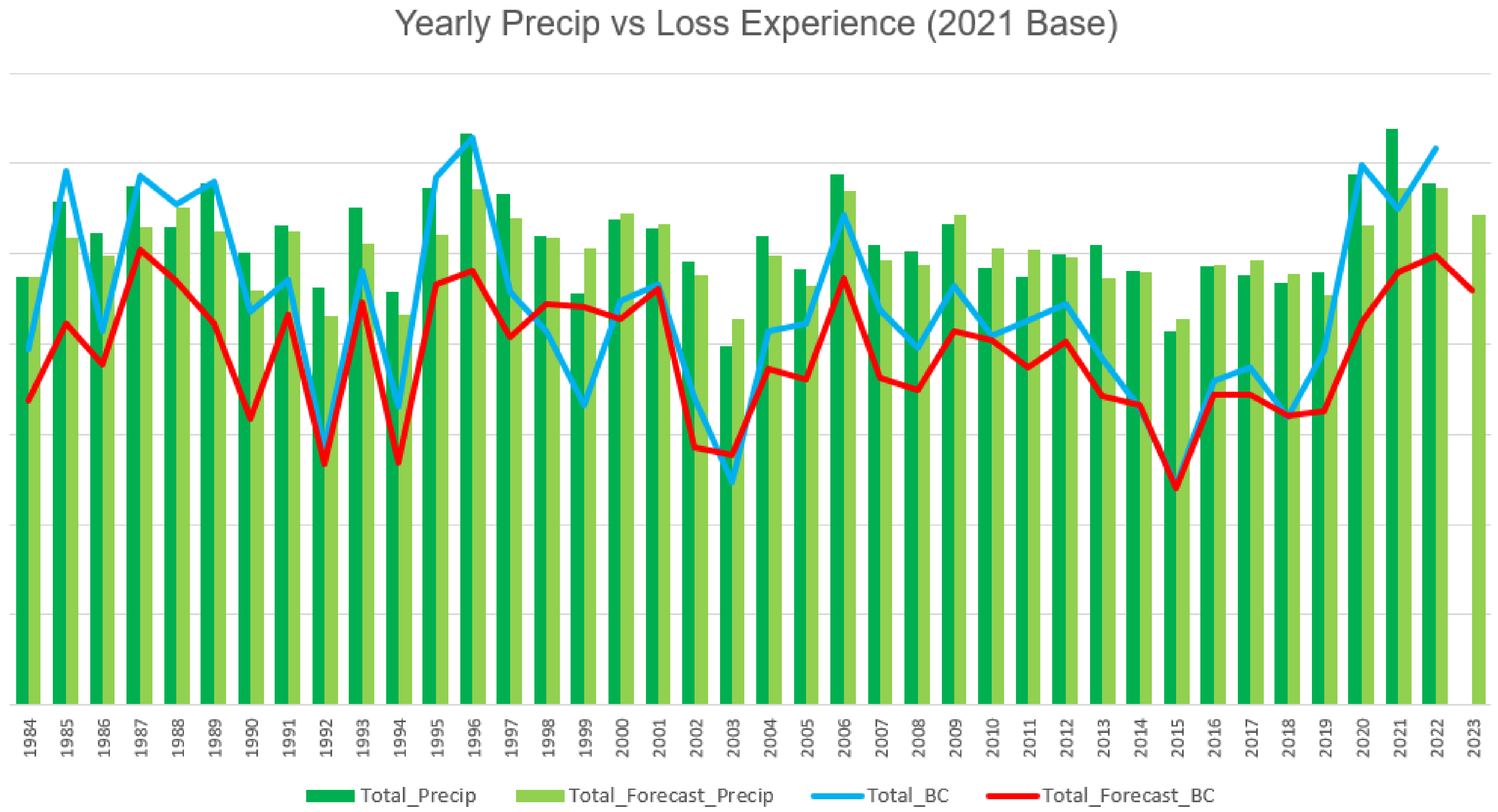
With Forecasted Precipitation Data



Modelling Results

Overall Book Sensitivity to Yearly Precipitation Experience (2021 Base)

Loss Experience



Precipitation induced volatility on loss experience was calculated to equal **61%** of the associated SAM measure of volatility.

Conclusion

- **Conclusions**
    - Shown that traditional short-term pricing datasets can be linked to open-source highly granular precipitation data
    - Demonstrated that precipitation data is a highly predictive factor when modelling insurance risk
    - Demonstrated the relationship between changes in actual precipitation and frequency and severity
    - Obtained precipitation forecasts that may be used for practical implementations (pricing/proactive risk management)
    - Demonstrated that precipitation forecasts provide similar predictive value
    - Obtained distribution of loss experience given differing years of precipitation experience for proactive risk management.