Institute and Faculty of Actuaries

# C4: Pricing Workshop – Overview and Practical Applications of Machine Learning Methods in Pricing

Bethan Faultless & Matthew Lambert

25 April 2019

# C4: Pricing Workshop

## – Overview and Practical Applications of Machine Learning Methods in Pricing

## Agenda

**Context of machine learning in pricing**

- **Decision trees**
- **Random forests**
- **Gradient boosting machines**

**Practical applications of ML techniques**

**Conclusions**

**Q&A**

**Objective:**

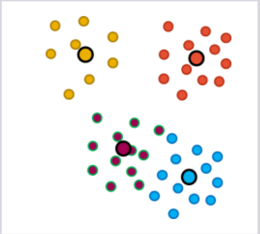
to give you a working knowledge of some machine learning methods that may be used to improve GLM results and/or offer valuable insights in their own right in the field of P&C insurance pricing

Institute and Faculty of Actuaries

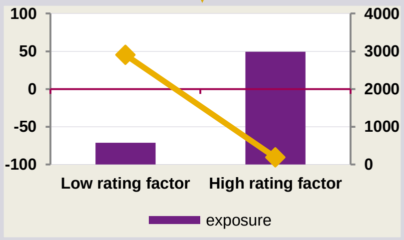
# Applications of machine learning in the insurance sector

Marketing and Distribution → Customer management → Underwriting and risk management → Pricing → Claims management → Asset management

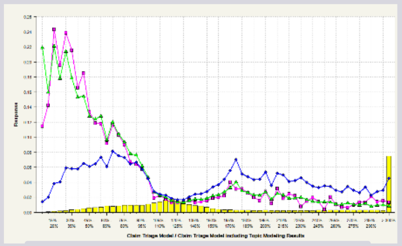
- Product development (UBI)
- Retention segmentation and program design
- Agent/Broker performance evaluation
- Fraud detection



P&C to Life cross-selling



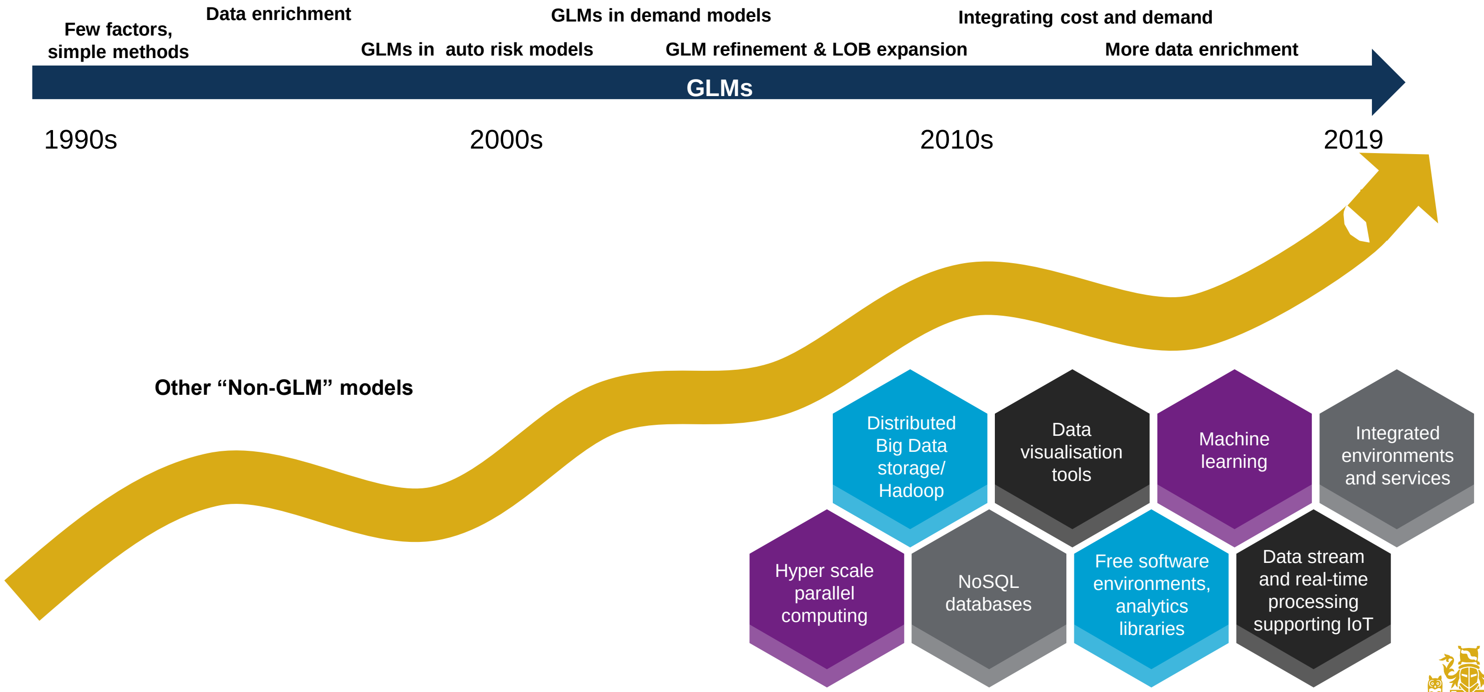

Web-scraping and feature design in commercial lines



Topic modeling and large loss modeling

Institute and Faculty of Actuaries

# This is not new….




Institute and Faculty of Actuaries

# What are these machine learning methods?

Ensembles

Classifications Trees

"Earth"

Regression Trees

Gradient Boosting Machines

K-nearest Neighbors

Elastic Net

Neural Networks

Naïve Bayes

Random Forests

K-Means Clustering

Principal Components Analysis

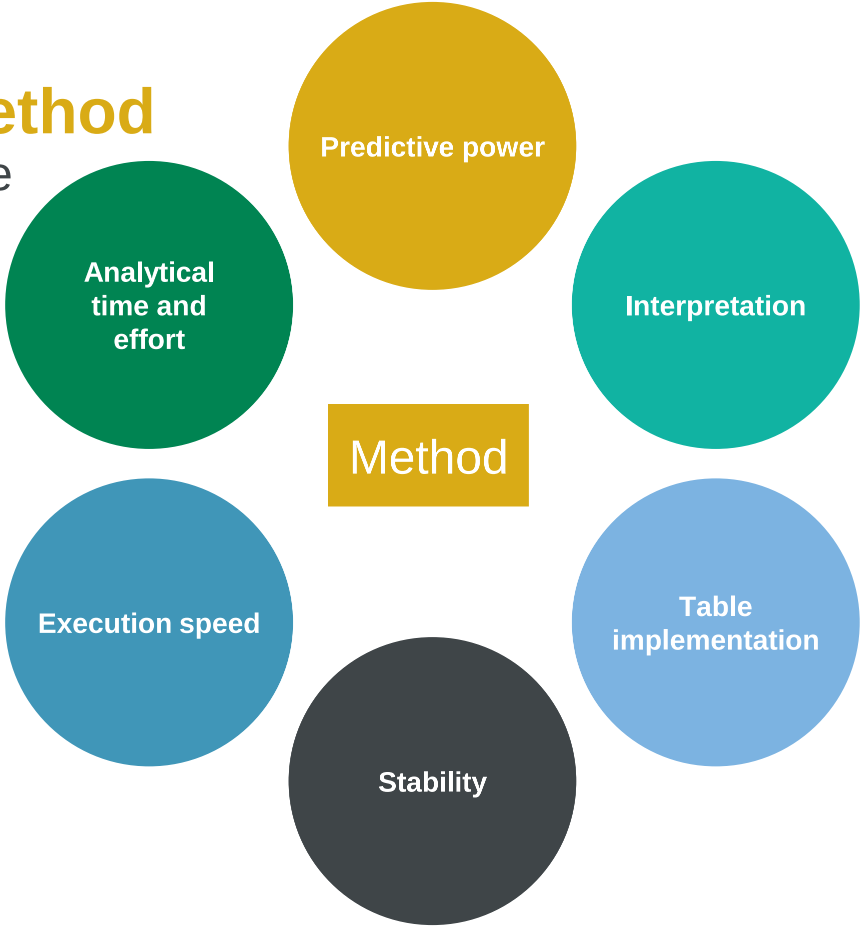
Lasso

Support Vector Machines

Ridge Regression

Institute and Faculty of Actuaries

# Choosing a method
## Dimensions of choice




Institute and Faculty of Actuaries


Institute and Faculty of Actuaries

# Focus on Trees

Ensembles

**Classifications Trees**

"Earth"

**Regression Trees**

Boosting

K-nearest Neighbours

Elastic Net

Neural Networks

Naïve Bayes

Random Forests

K-Means Clustering

Principal Components Analysis

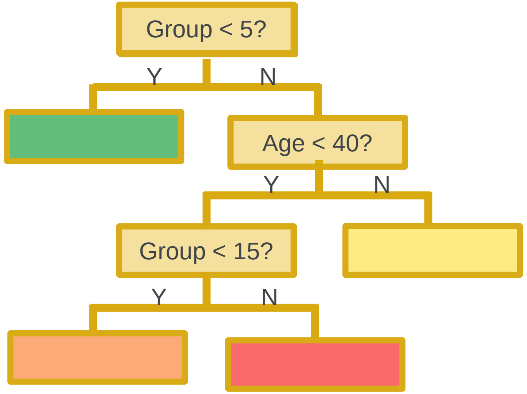
Lasso

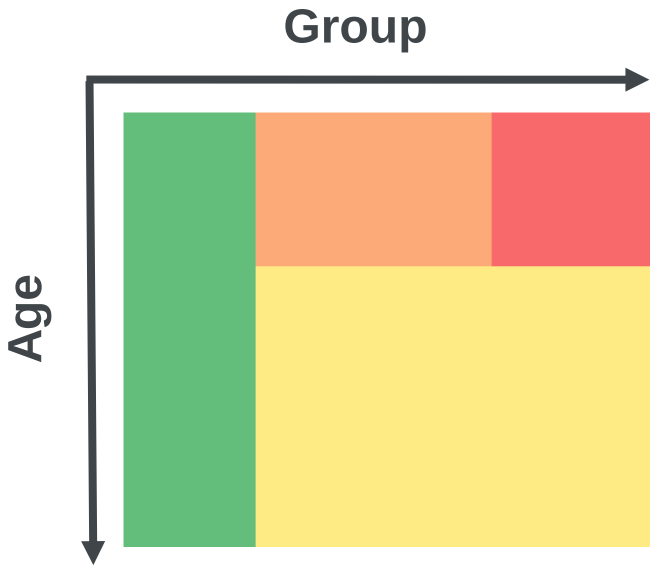
Support Vector Machines

Ridge Regression

Institute and Faculty of Actuaries

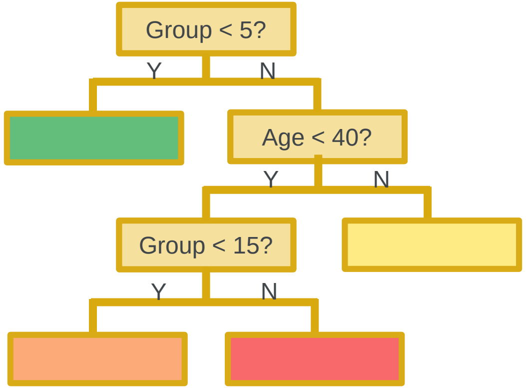
# Decision Trees

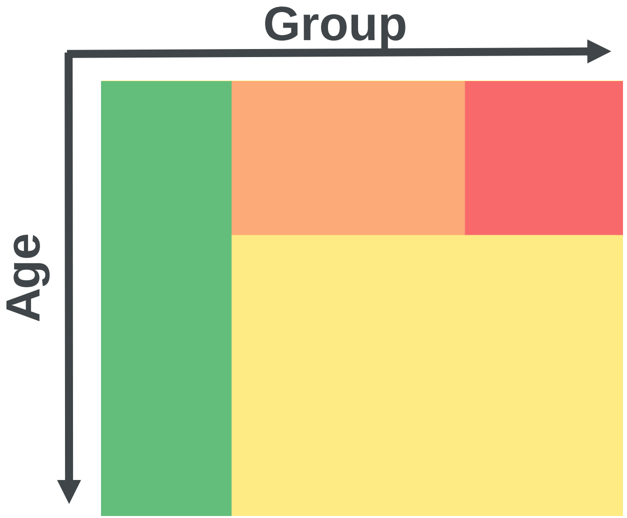



Institute and Faculty of Actuaries

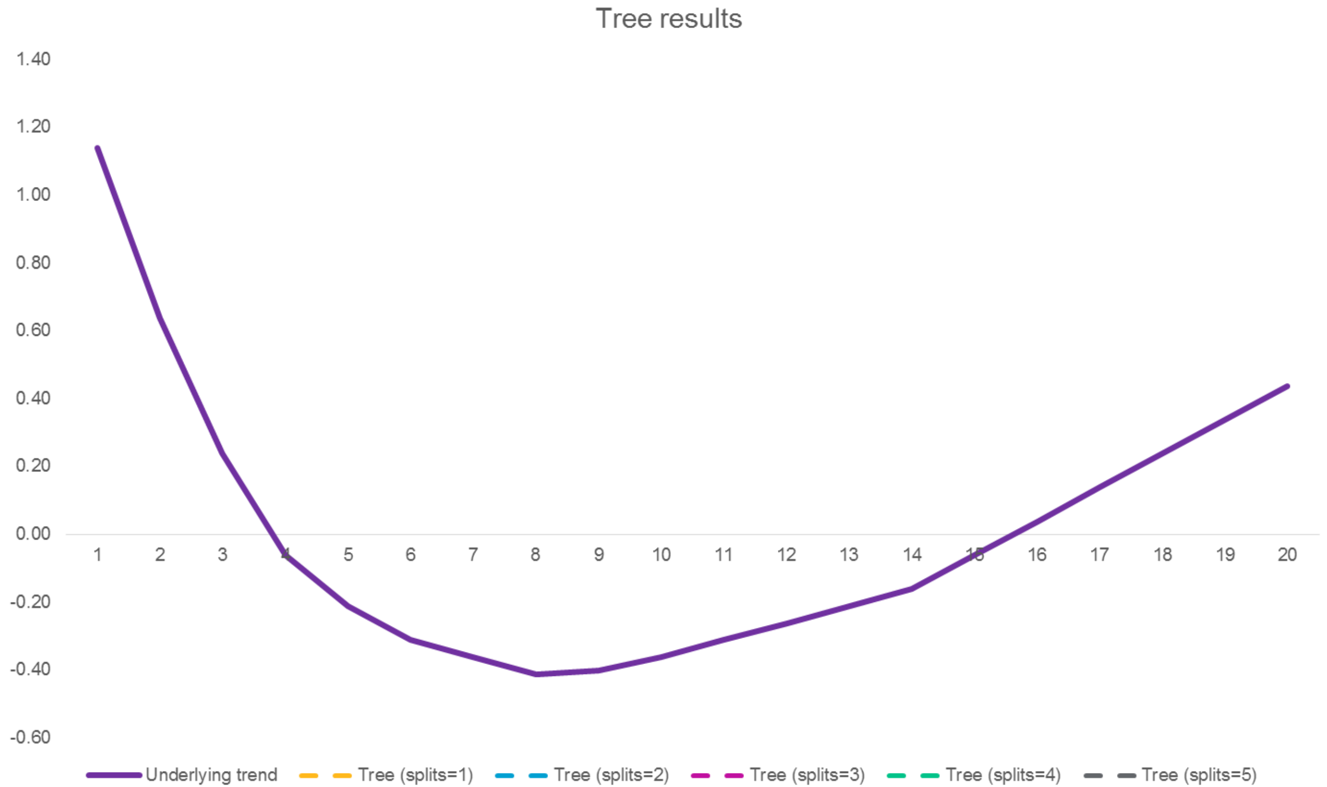
# Decision Trees



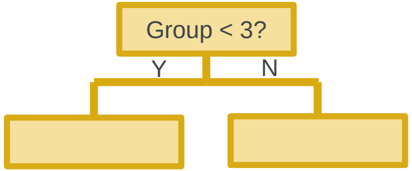



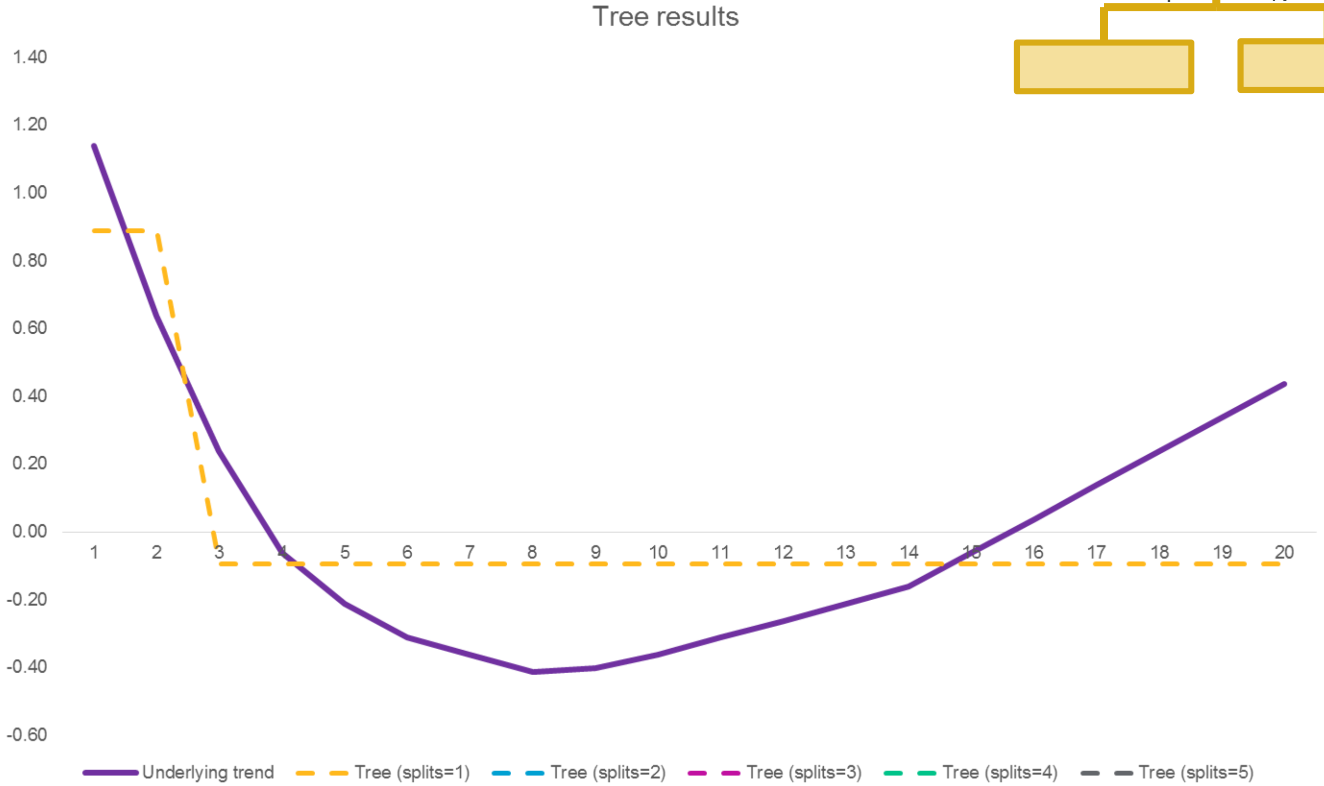

Institute and Faculty of Actuaries

# A simple Tree example

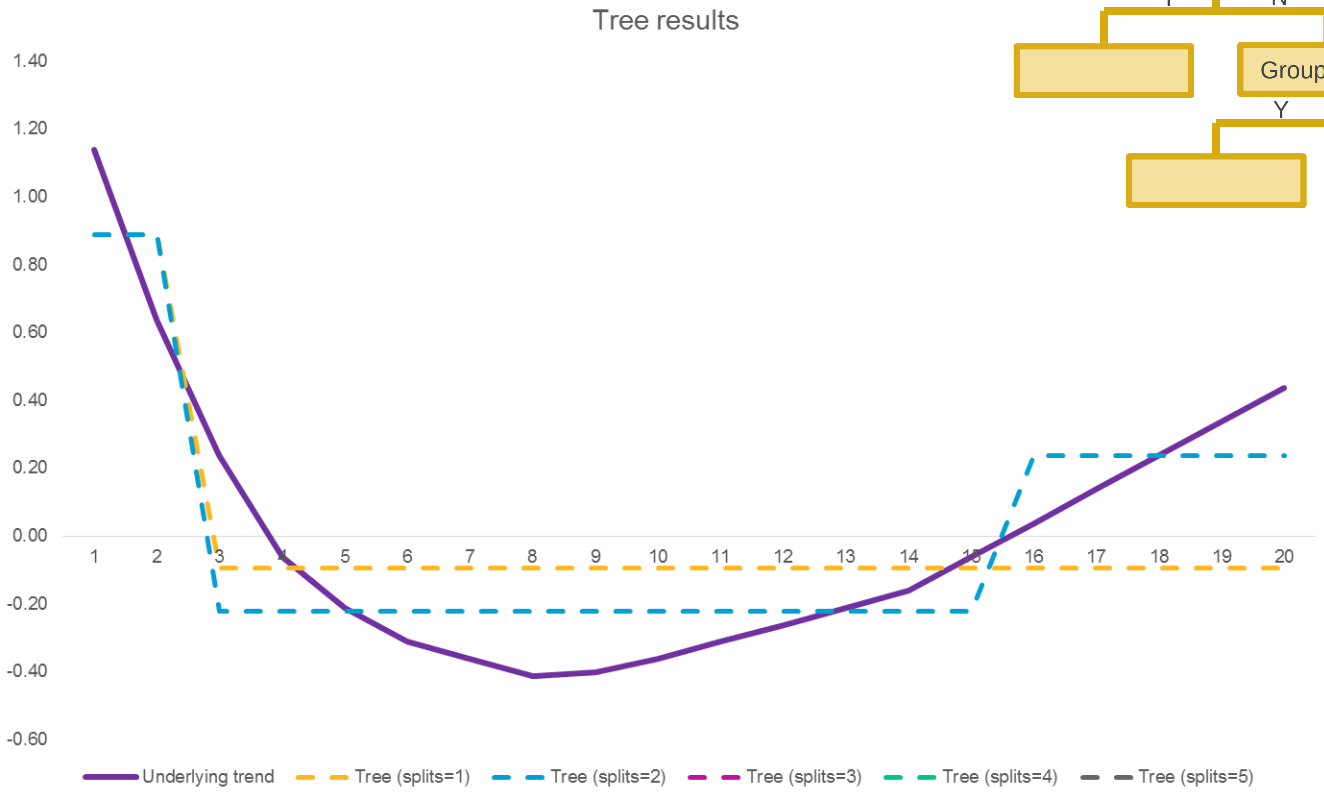


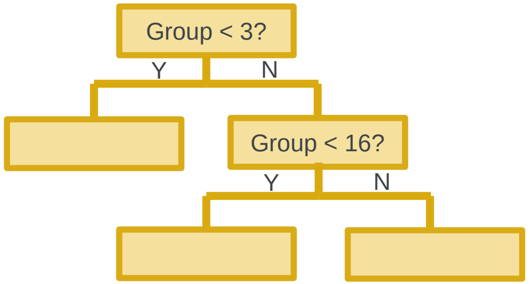
# A simple Tree example



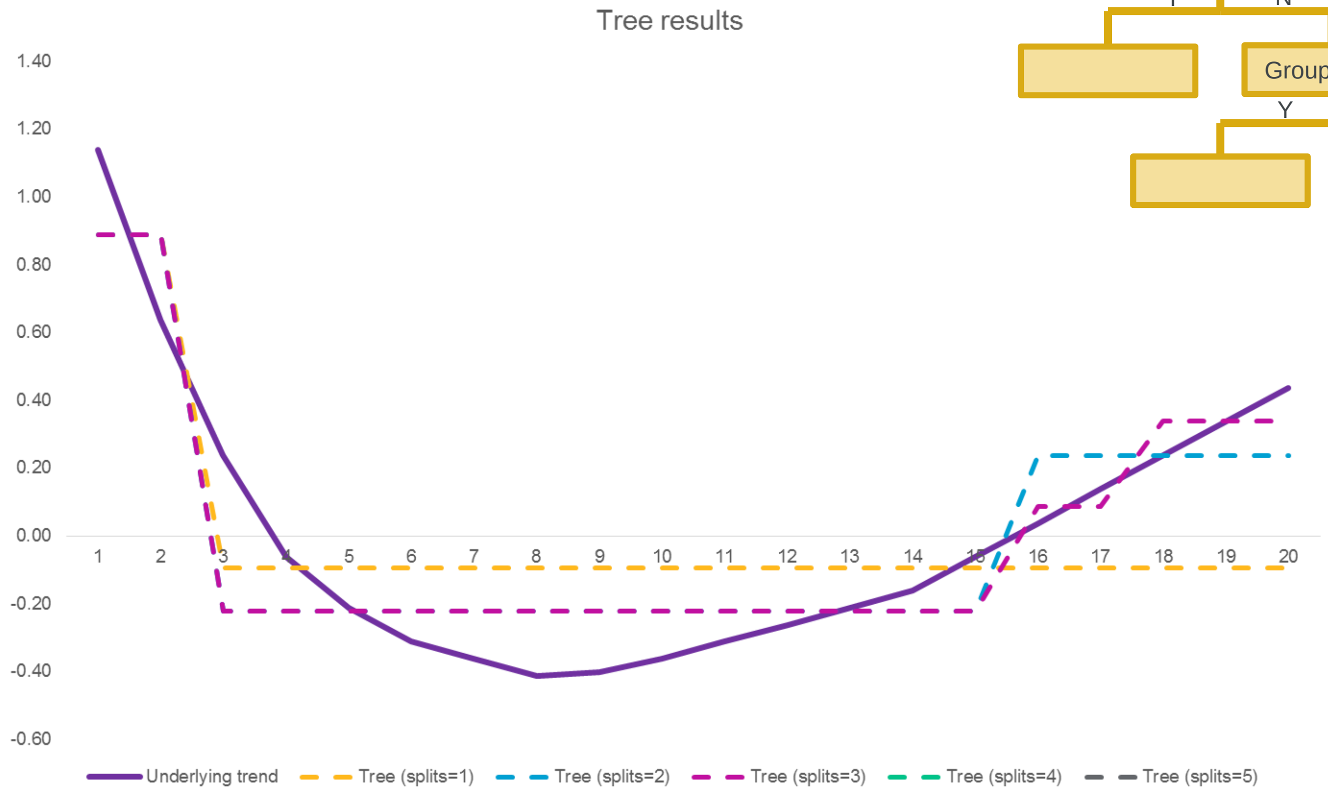



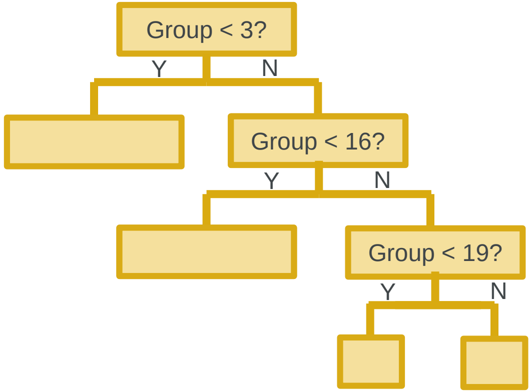
# A simple Tree example



# A simple Tree example

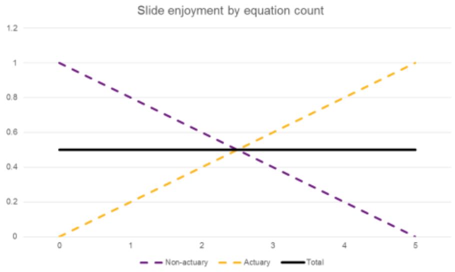


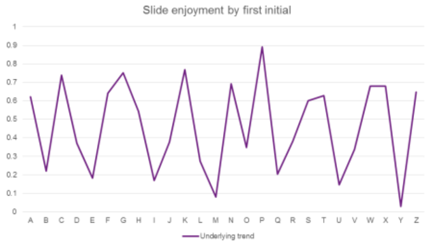

Institute and Faculty of Actuaries

# Shortcomings of using trees

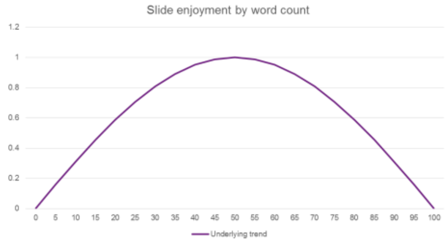
**They may miss interactions…**



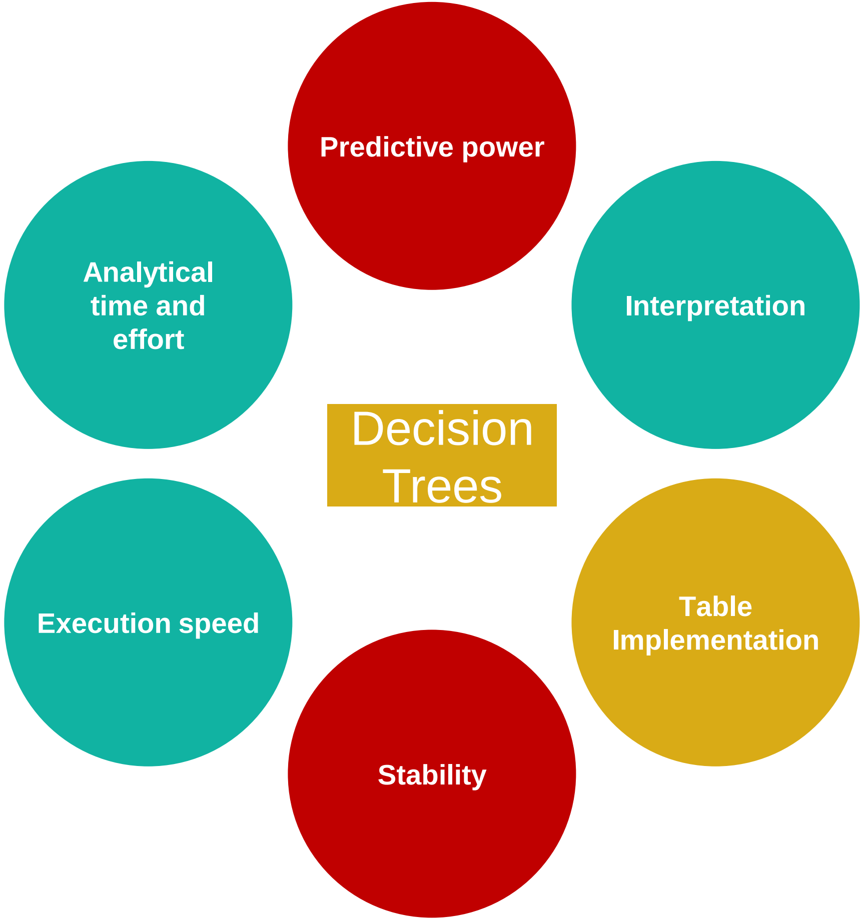

**… they may struggle with categorical variables….**




**…and they can be bad at turning points**




Institute and Faculty of Actuaries




Institute and Faculty of Actuaries

# Focus on Random Forests

| | | | | |
|---|---|---|---|---|
| Ensembles | Classifications Trees | "Earth" | Regression Trees | Gradient Boosting Machines |
| K-nearest Neighbours | Elastic Net | Neural Networks | Naïve Bayes | **Random Forests** |
| K-Means Clustering | Principal Components Analysis | Lasso | Support Vector Machines | Ridge Regression |

Institute and Faculty of Actuaries

# Random Forests

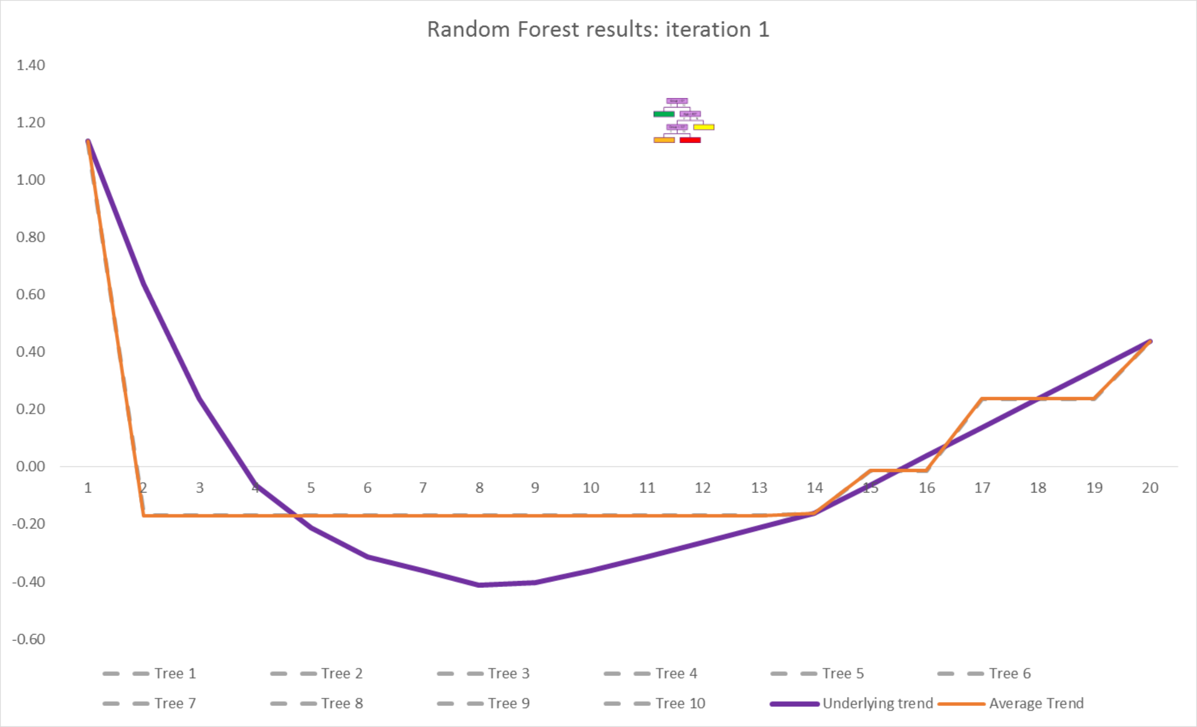
- **Tree 1:** **Prediction 1** = **Signal 1** + **Noise 1**
- **Tree 2:** **Prediction 2** = **Signal 2** + **Noise 2**
- **Tree 3:** **Prediction 3** = **Signal 3** + **Noise 3**
- …
- **Tree 1000:** **Prediction 1000** = **Signal 1000** + **Noise 1000**

- **Random Forest:**
- **Prediction = AVERAGE(Tree Predictions)**
- = **AVERAGE(Tree Signal)** + **AVERAGE(Tree Noise)**

- **Average Noise → 0** if the trees are independent
- Independence of trees achieved by fitting each tree to:
  - Random subset of data (bootstrap sample)
  - Random subset of factors
- **Average Signal → Underlying trend**, provided trees are complex enough to represent it
- This is **bagging** (**b**ootstrap **agg**regation) – fit **lots** of independent models and take an average

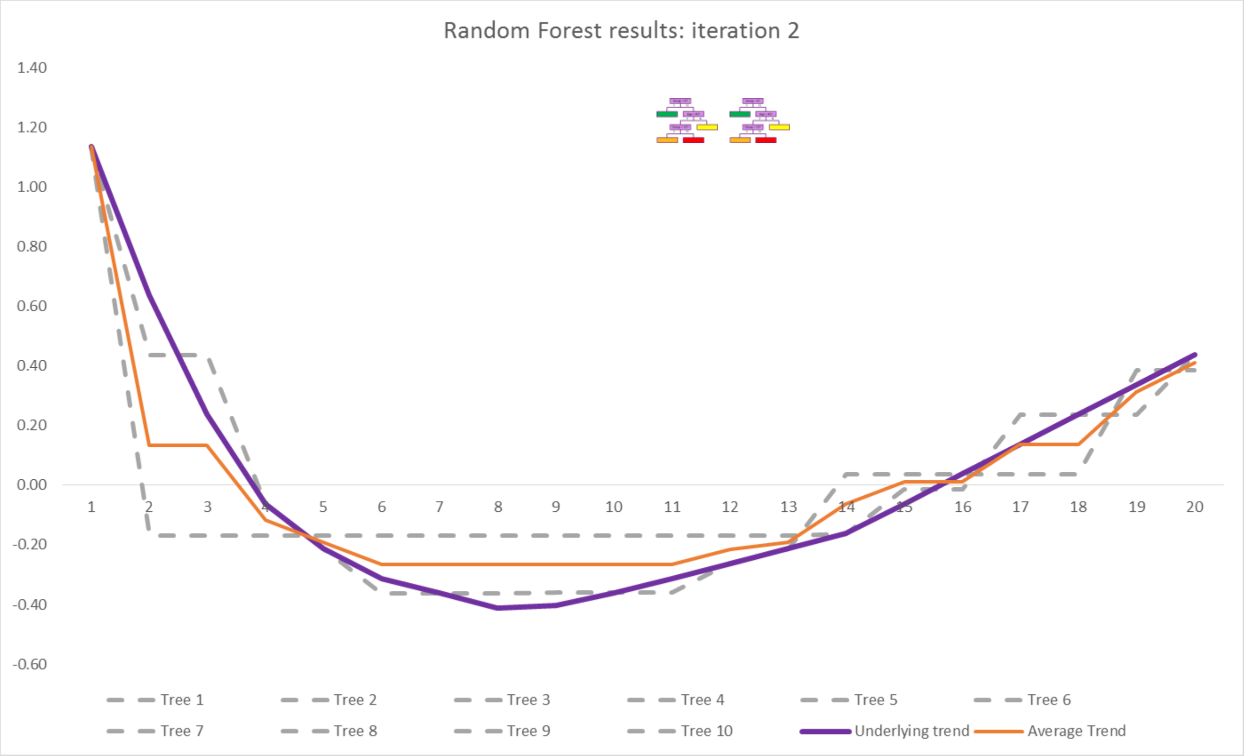
Institute and Faculty of Actuaries

# A simple Random Forest example



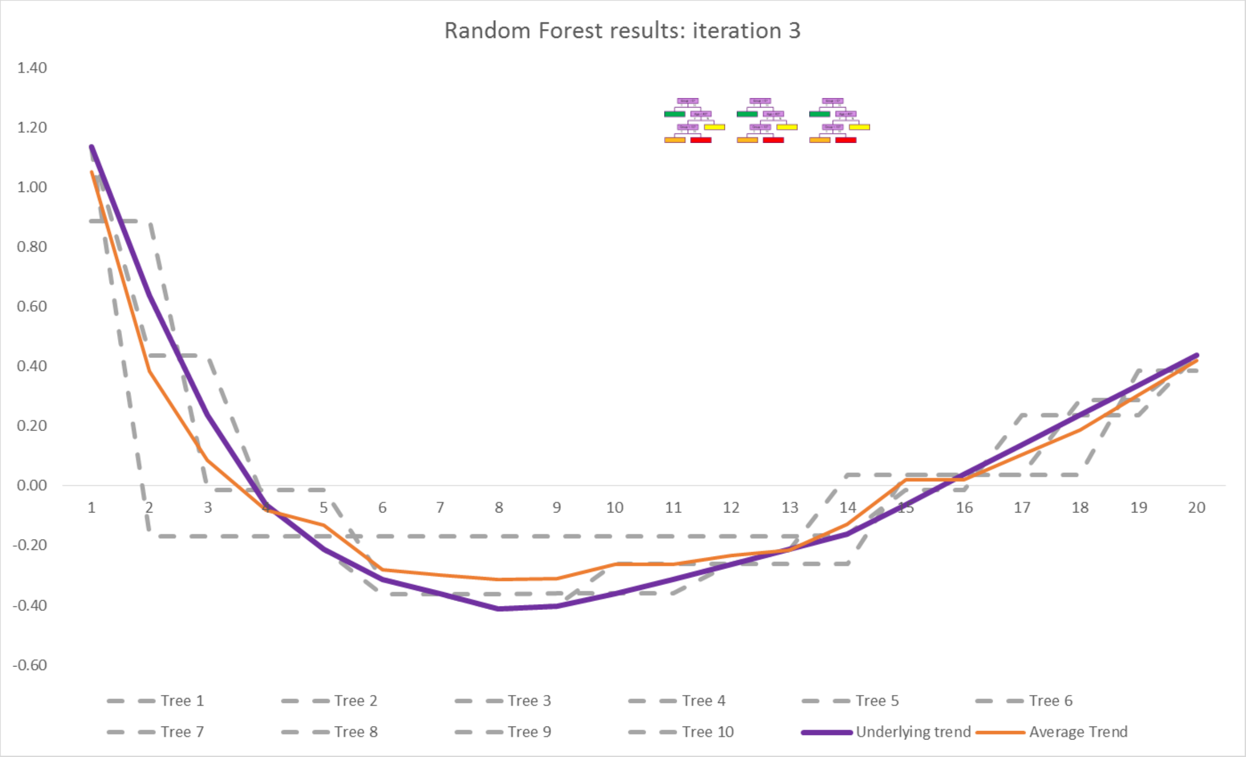

Institute and Faculty of Actuaries

# A simple Random Forest example



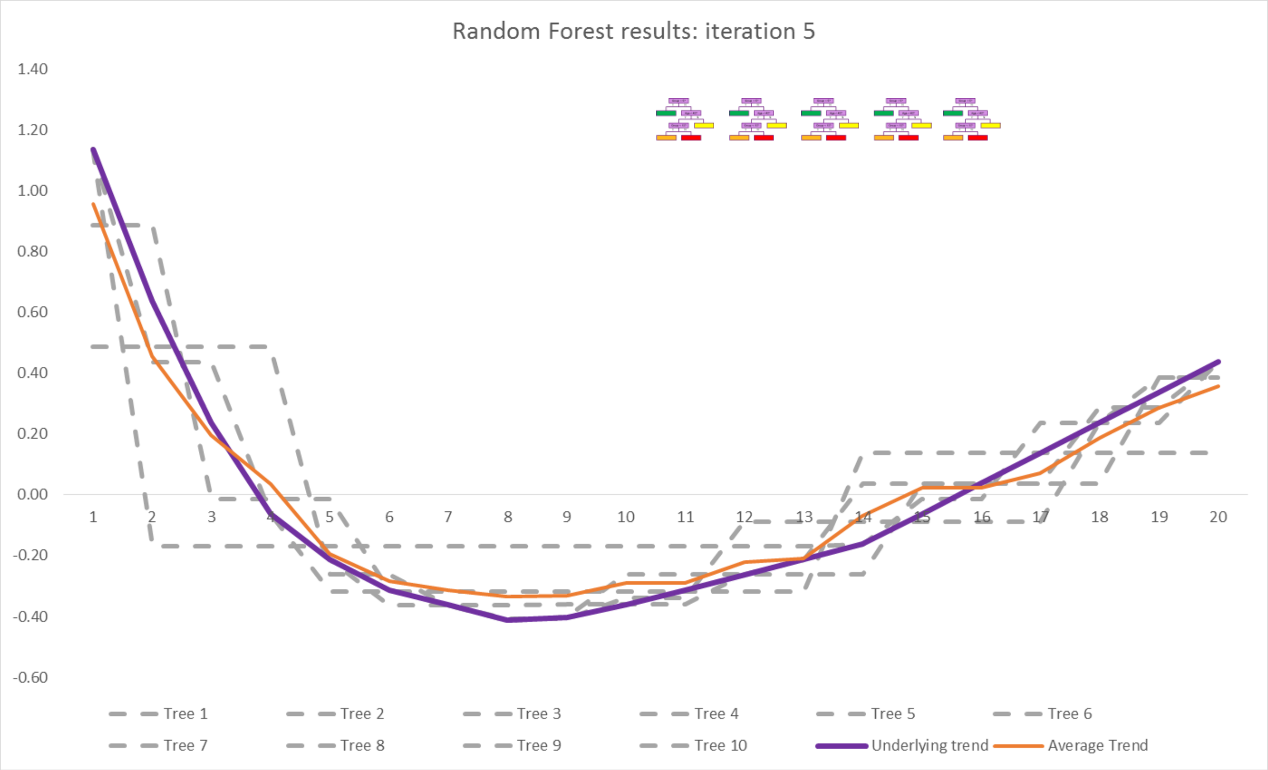

Institute and Faculty of Actuaries

# A simple Random Forest example



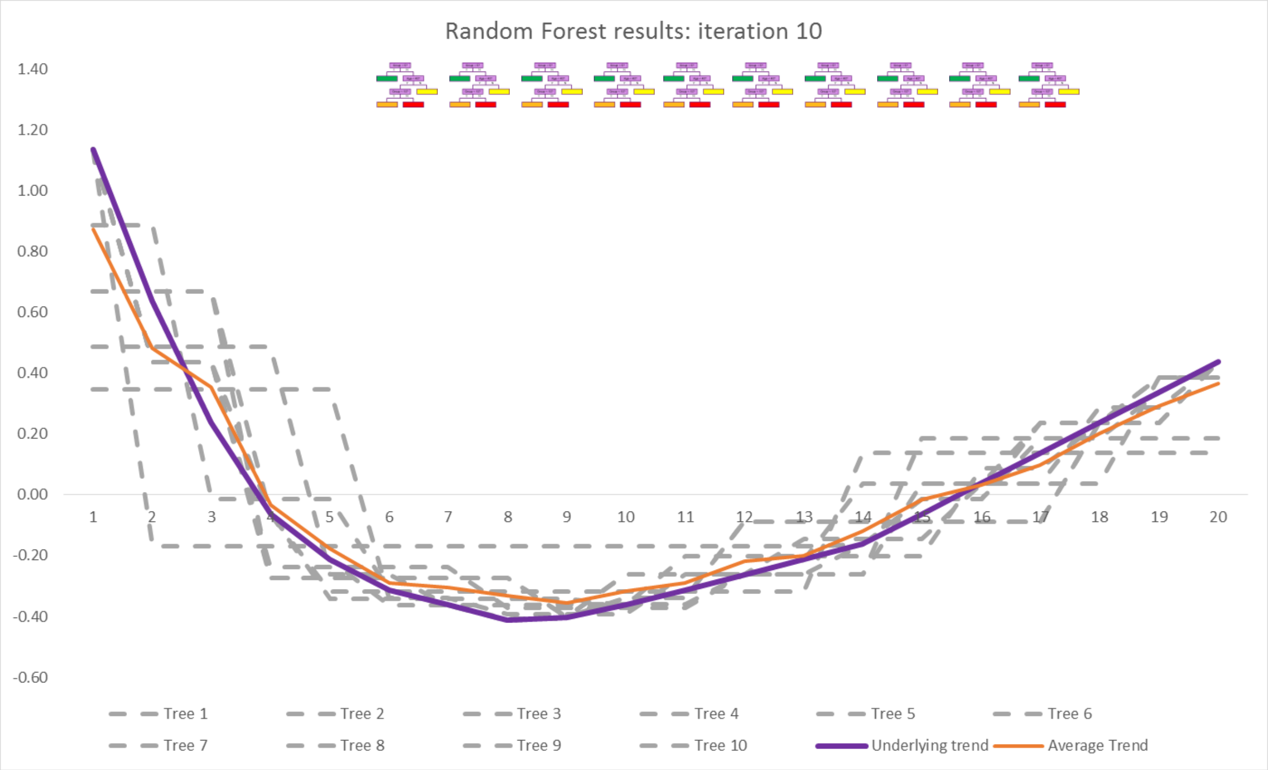

Institute and Faculty of Actuaries

# A simple Random Forest example

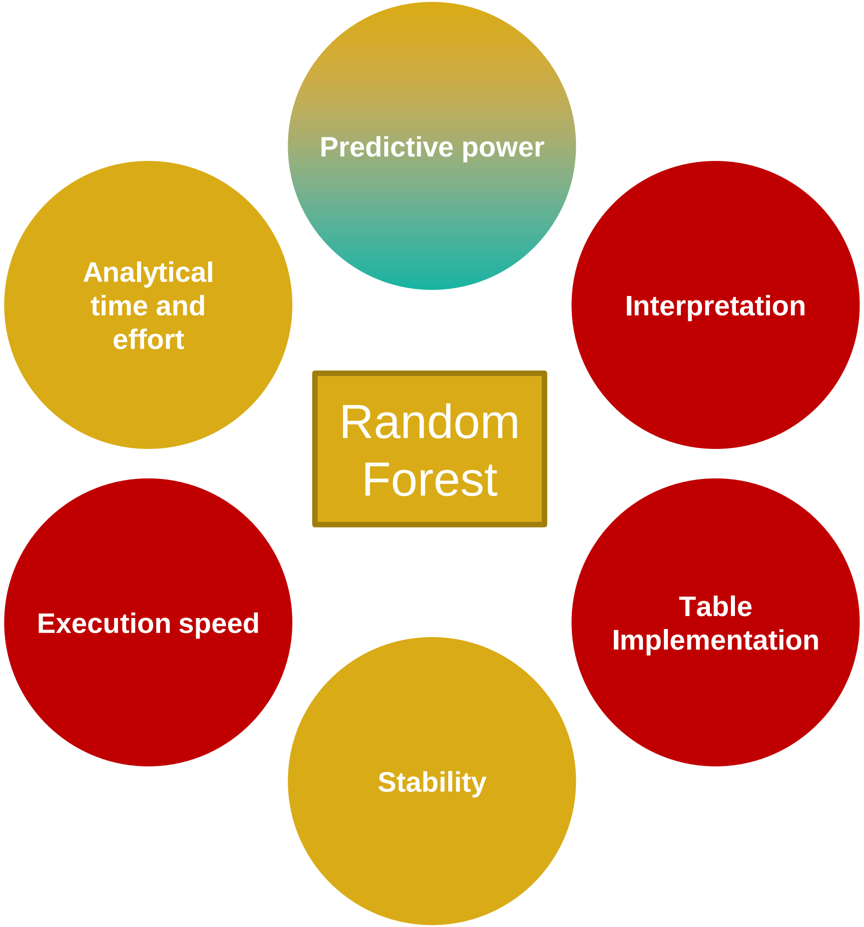



Institute and Faculty of Actuaries

# A simple Random Forest example




Institute and Faculty of Actuaries


Predictive power
Analytical time and effort
Interpretation
Random Forest
Execution speed
Table Implementation
Stability

# Focus on Gradient Boosting Machines

Ensembles

Classifications Trees

"Earth"

Regression Trees

**Gradient Boosting Machines**

K-nearest Neighbours

Elastic Net

Neural Networks

Naïve Bayes

Random Forests

K-Means Clustering

Principal Components Analysis

Lasso

Support Vector Machines

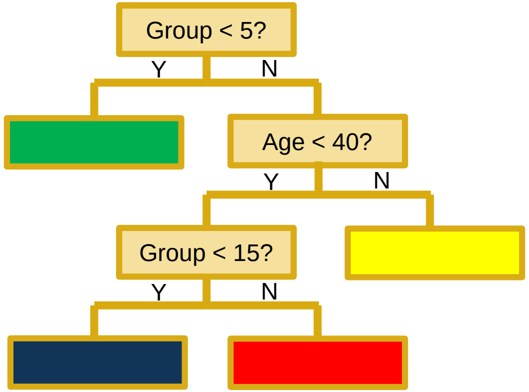
Ridge Regression

Institute and Faculty of Actuaries

# Gradient Boosted Machine or "GBM"

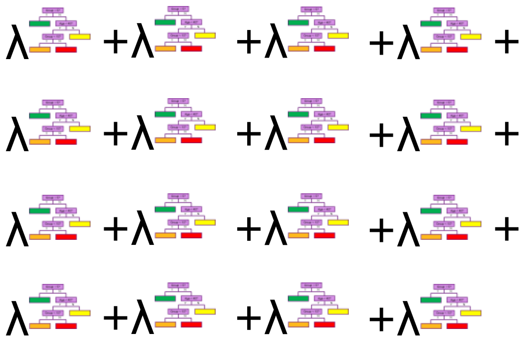
**A tree**

$f_i(x)$



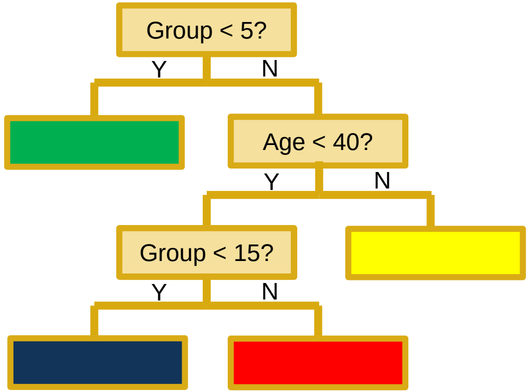

**A GBM**

$$f(x) = \lambda \sum_{n=1}^{N} f_n(x)$$



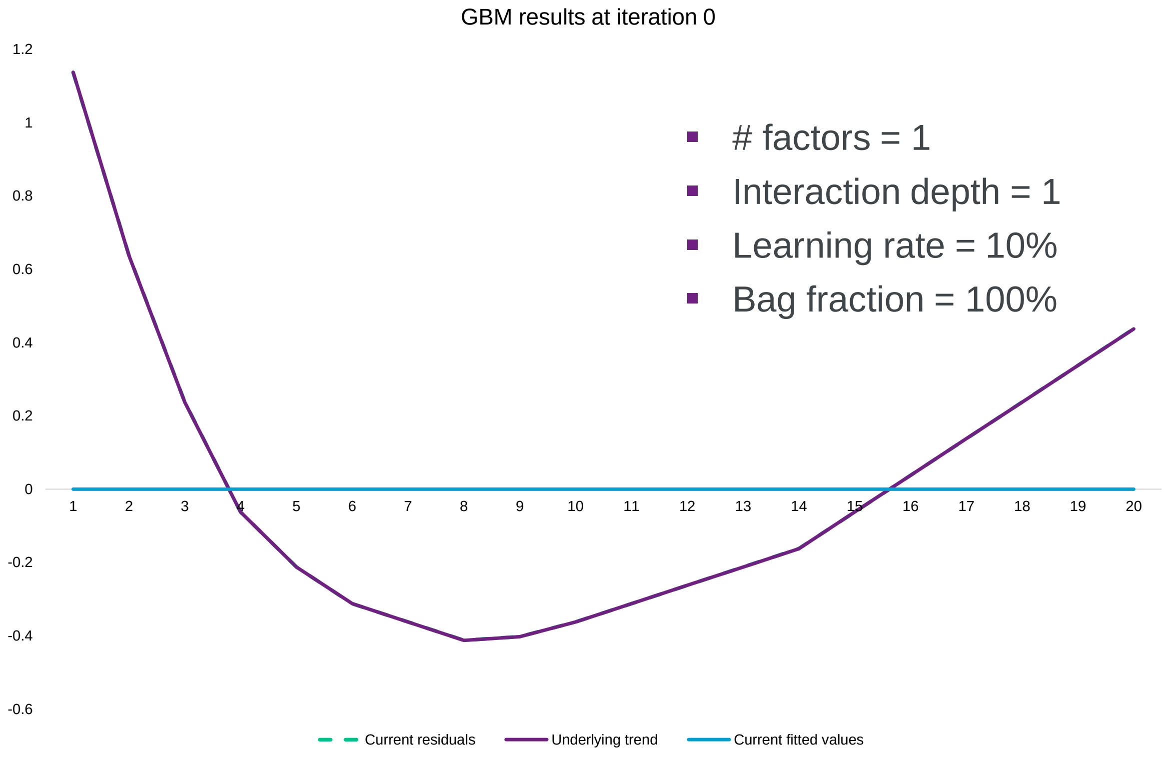
# Four main assumptions

- λ **Learning rate / "shrinkage"**
  - Amount by which the old model predictions are varied for the next model iteration
  - New model = Old + (Prediction x Learning rate)
- **Interaction depth**
  - Number of splits allowed on each tree (or the number of terminal nodes – 1)
- **N Number of trees** (iterations) allowed
- **Bag fraction**
  - Trees are fitted to a subset of the data (the bag fraction) on a randomized basis
  - Additional noise-reduction can be achieved by using a random subset of the available factors at each iteration




Institute and Faculty of Actuaries

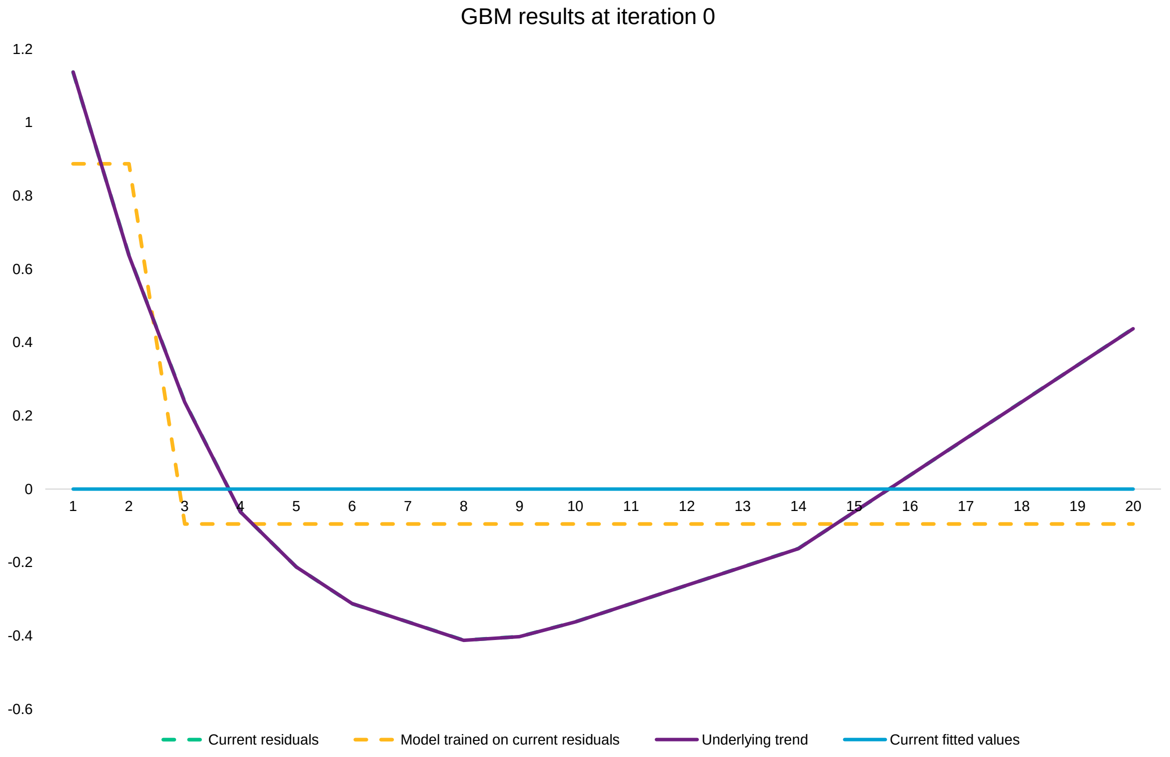
# A simple GBM example



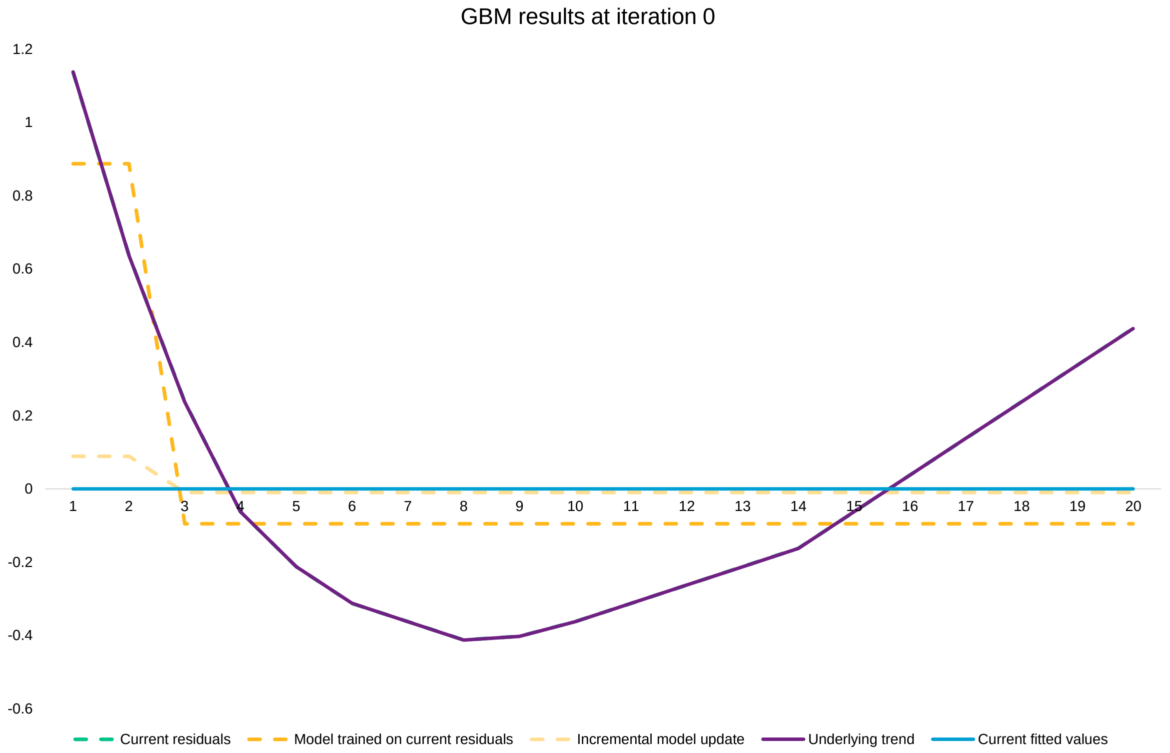

- # factors = 1
- Interaction depth = 1
- Learning rate = 10%
- Bag fraction = 100%

Institute and Faculty of Actuaries

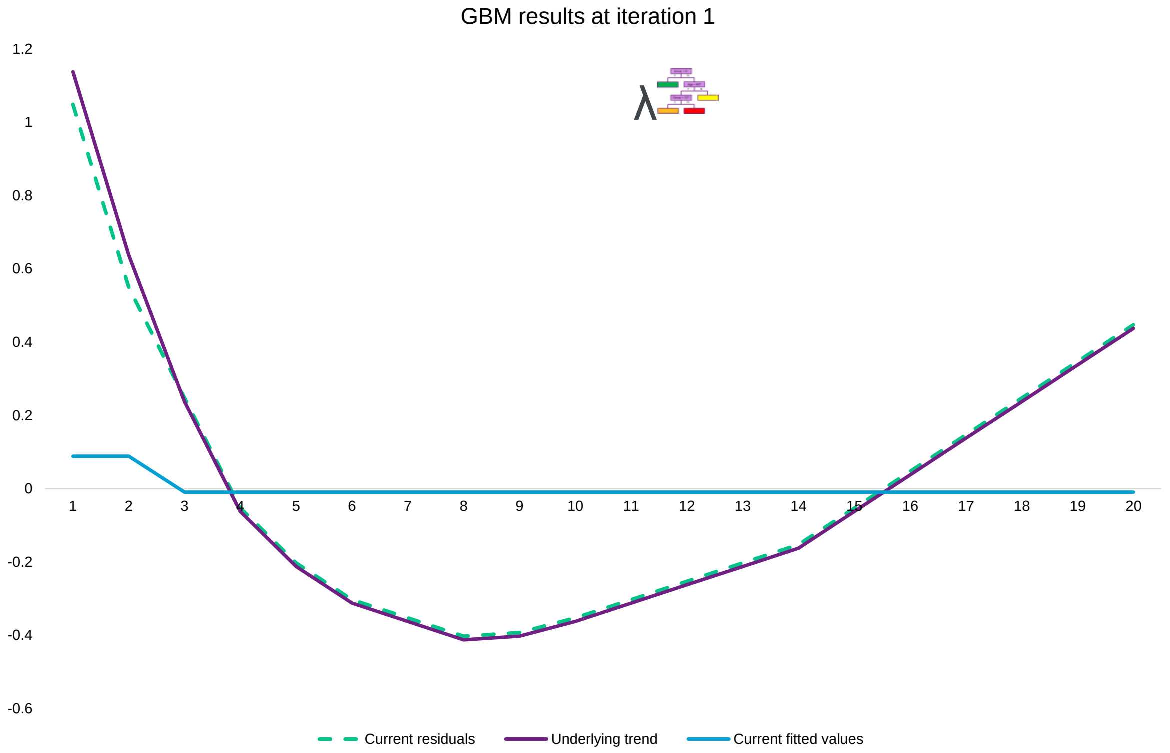
# A simple GBM example



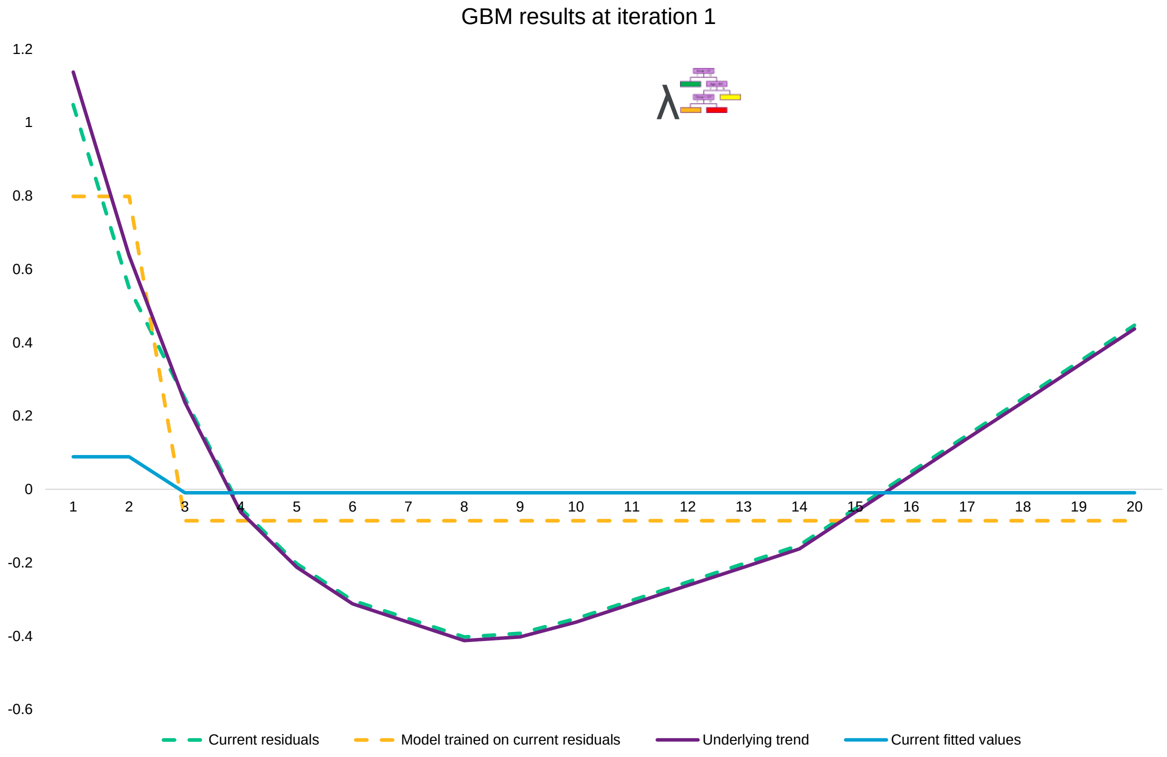
# A simple GBM example



# A simple GBM example

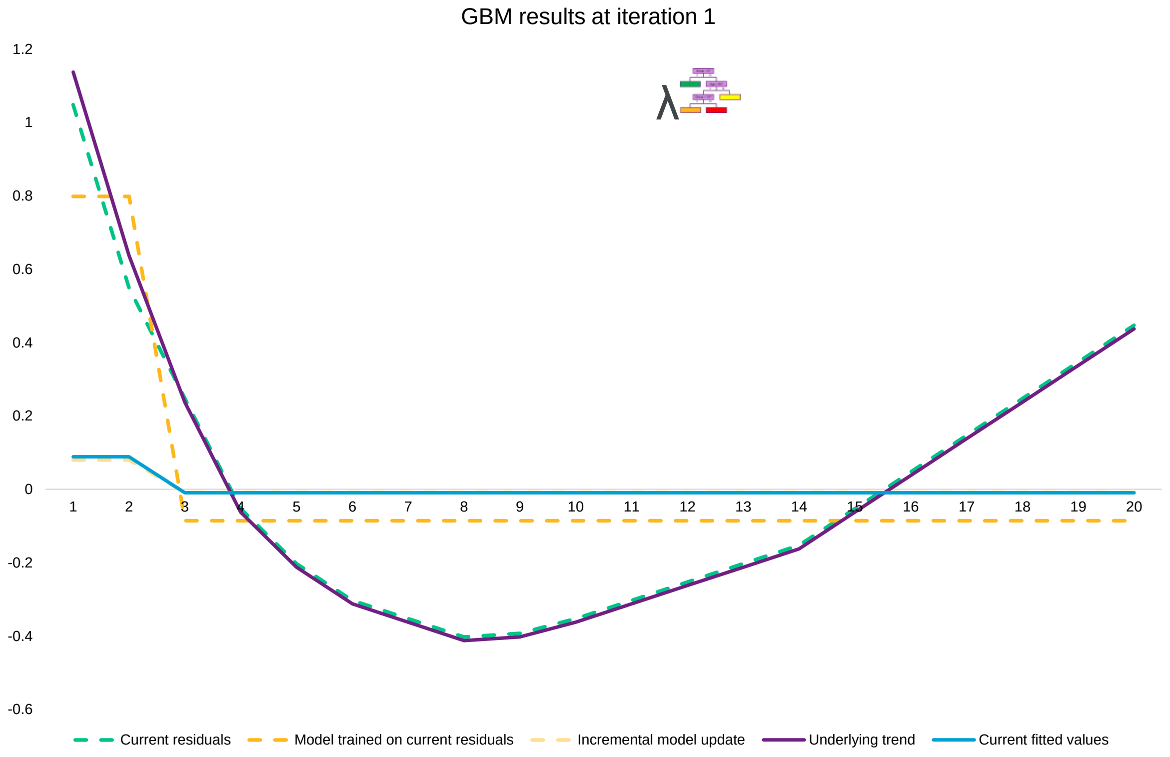


# A simple GBM example



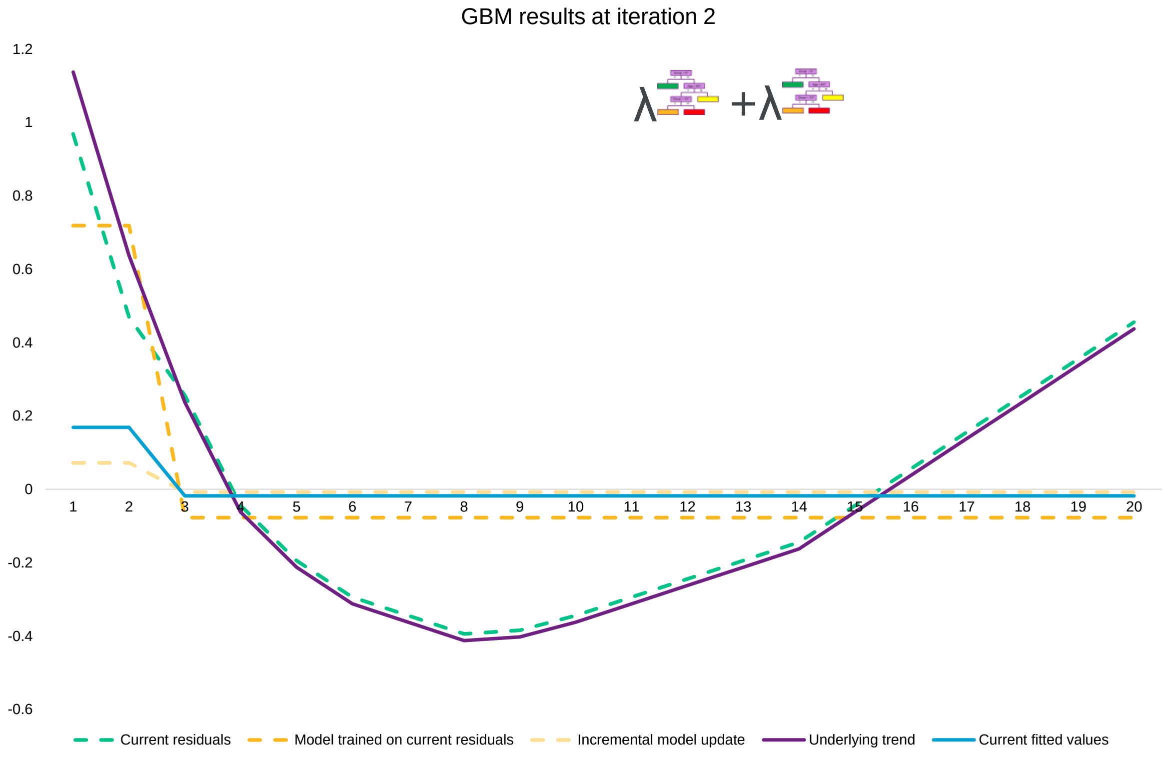

Institute and Faculty of Actuaries

# A simple GBM example

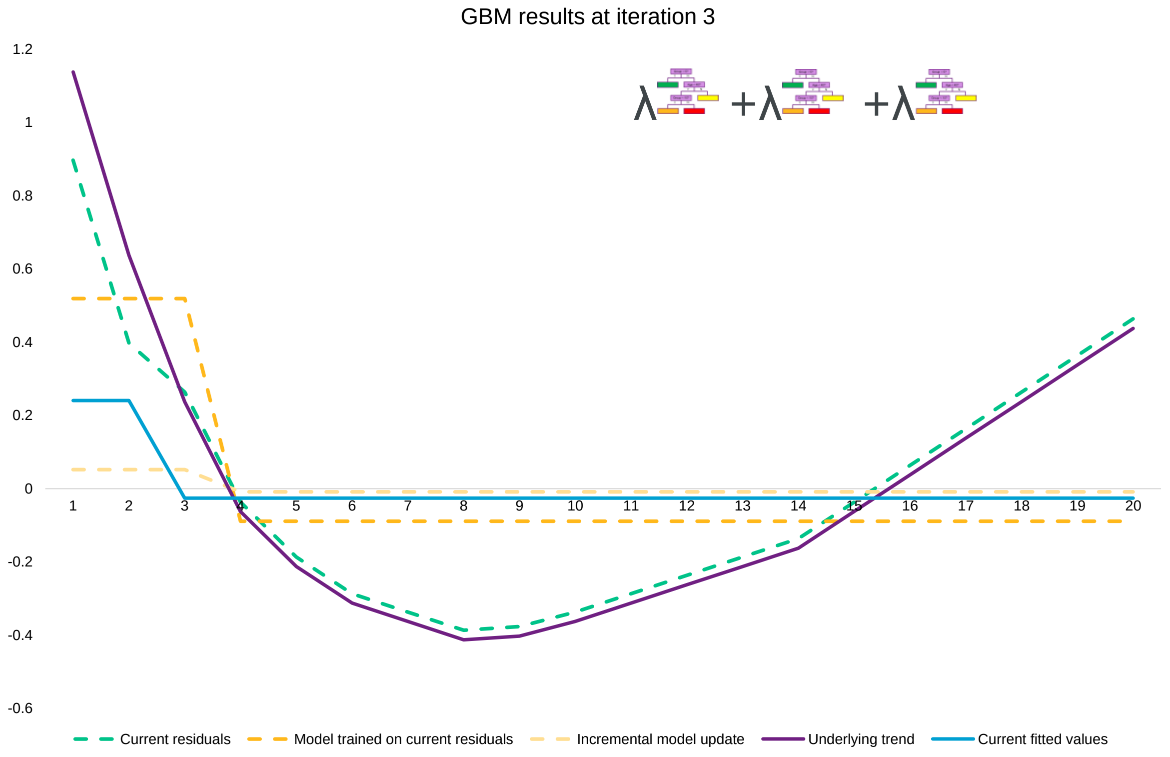



Institute and Faculty of Actuaries

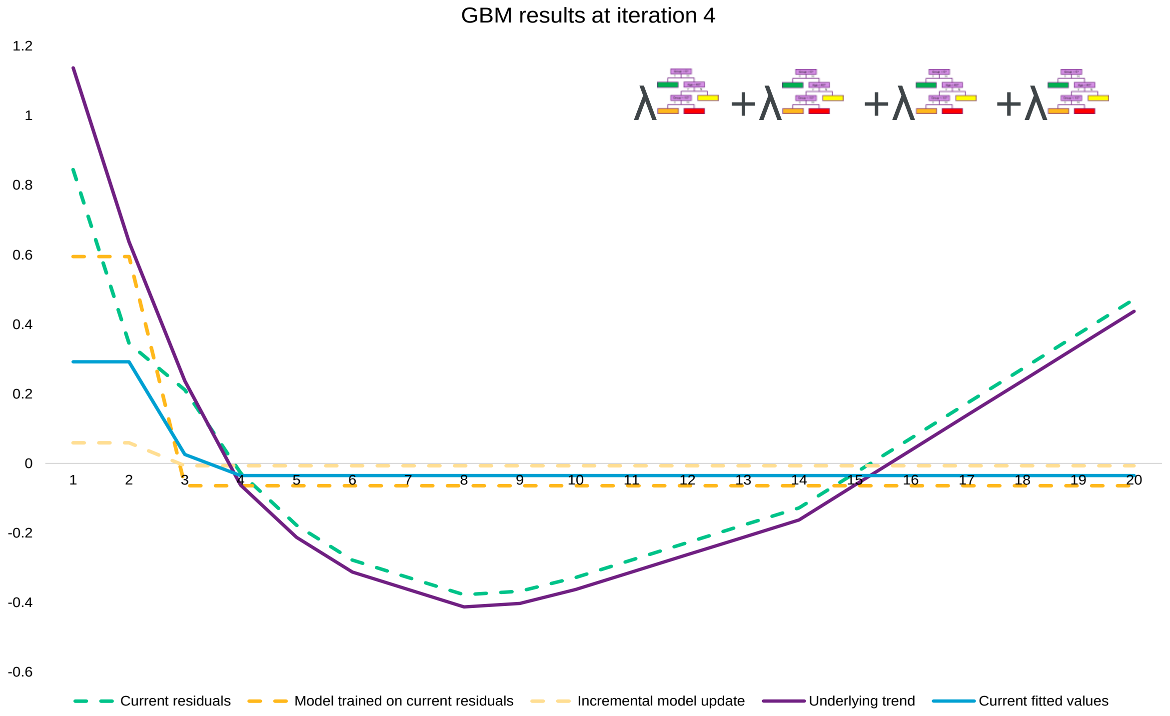
# A simple GBM example



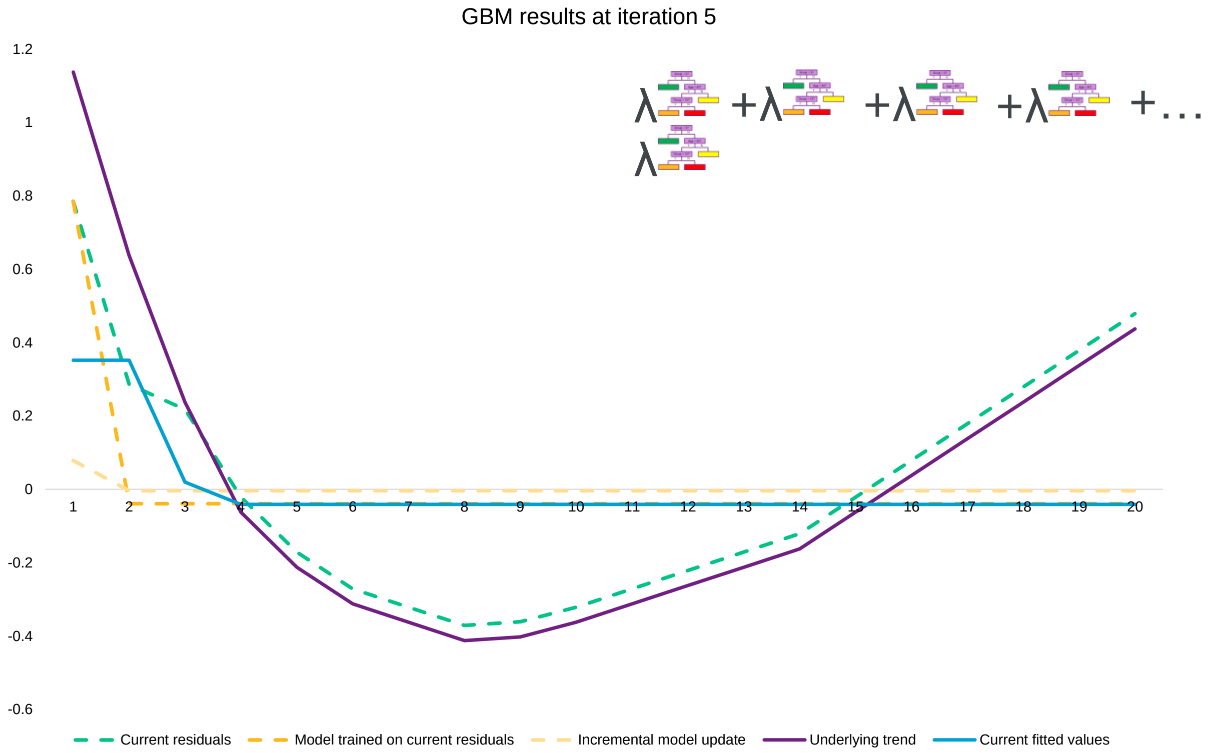
# A simple GBM example



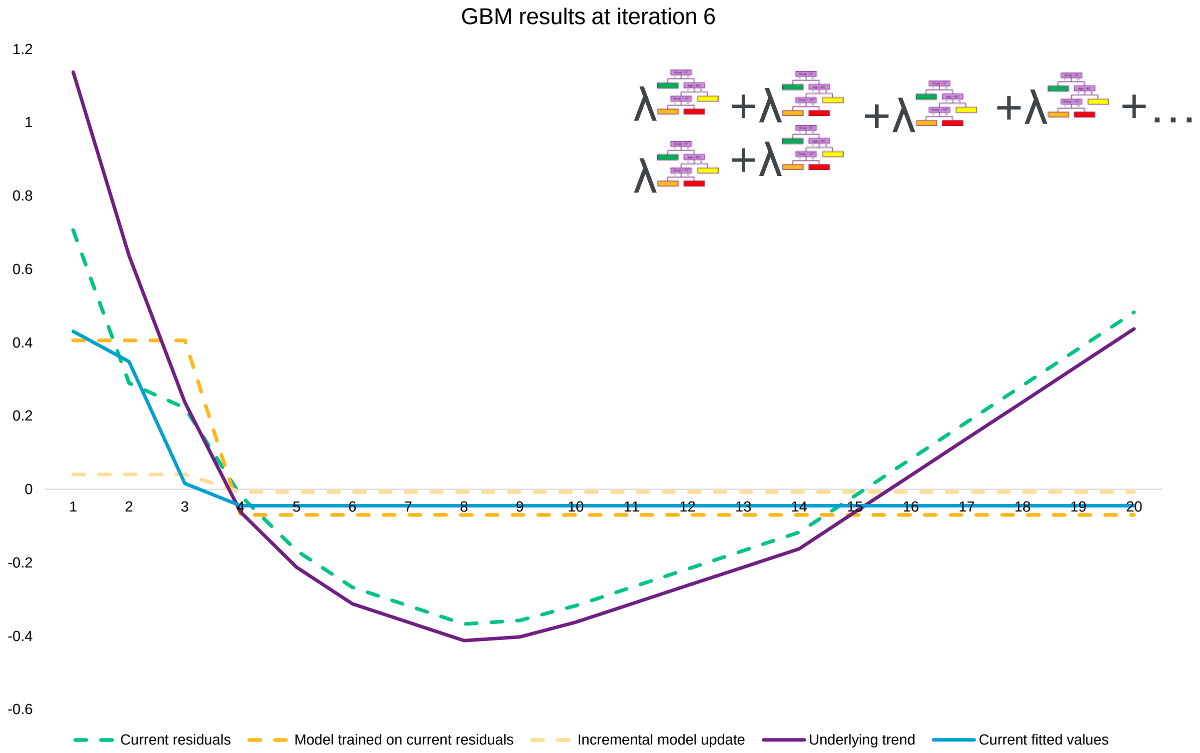
# A simple GBM example



# A simple GBM example

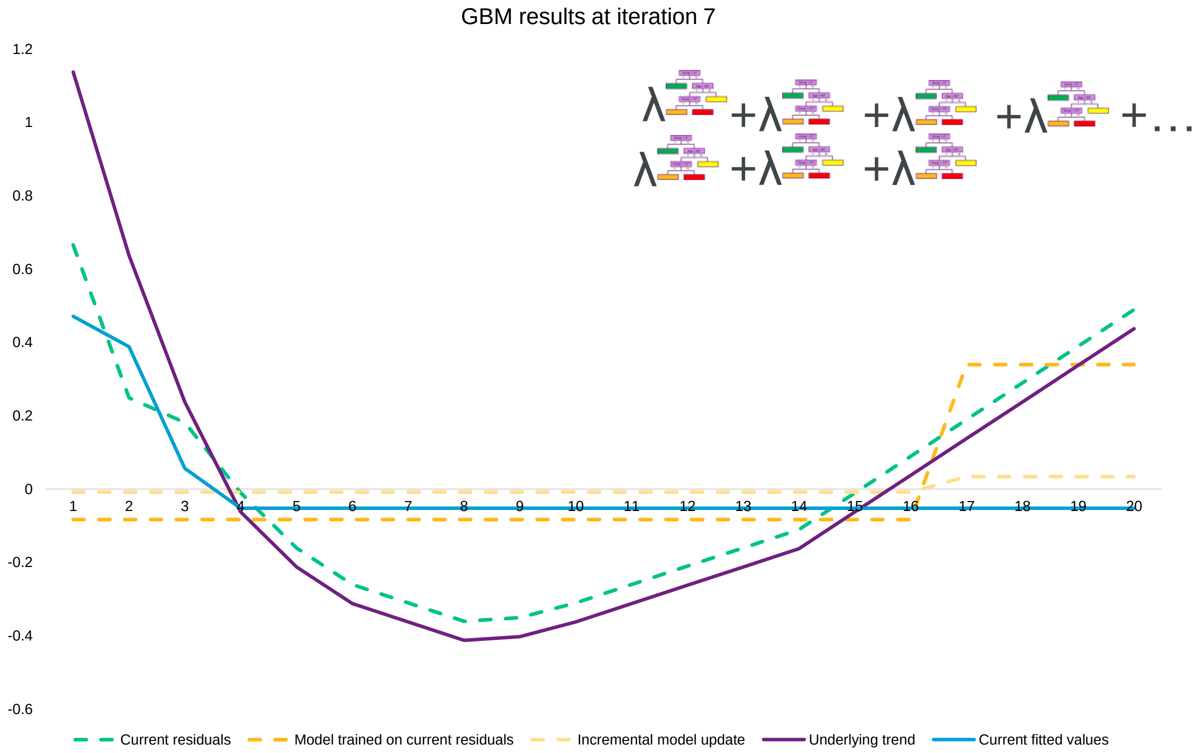


# A simple GBM example



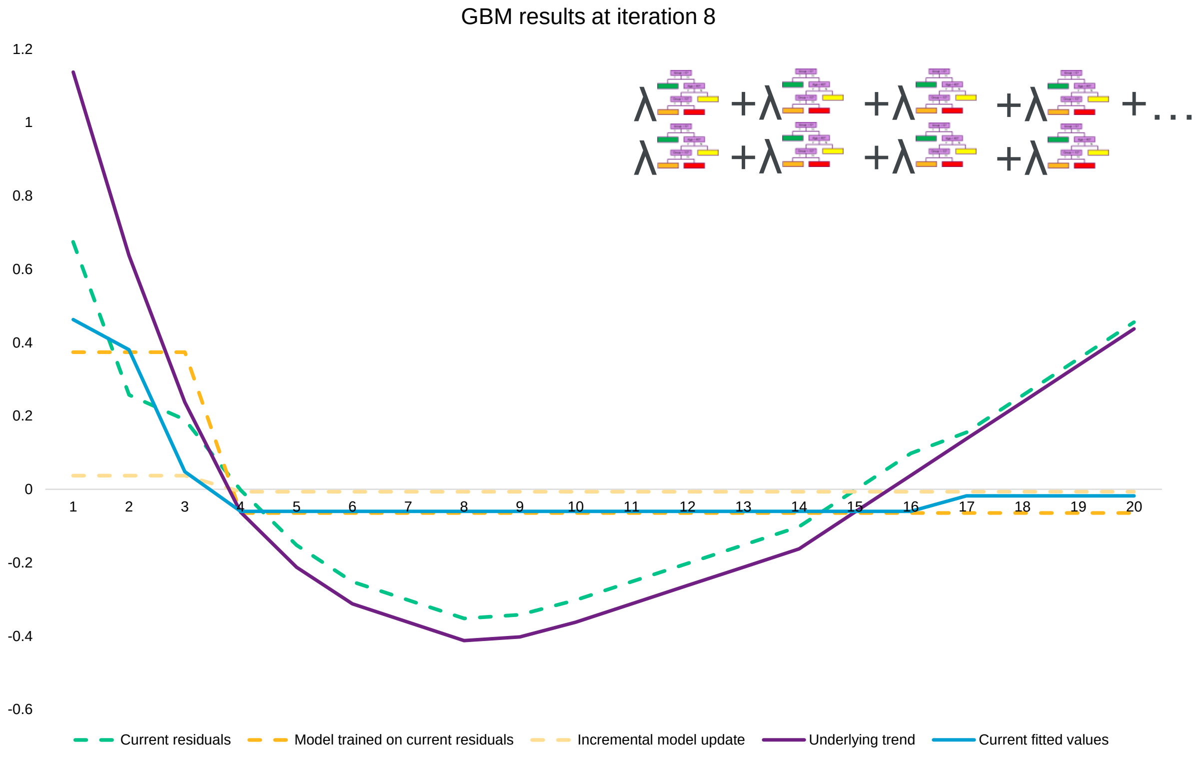

Institute and Faculty of Actuaries

# A simple GBM example



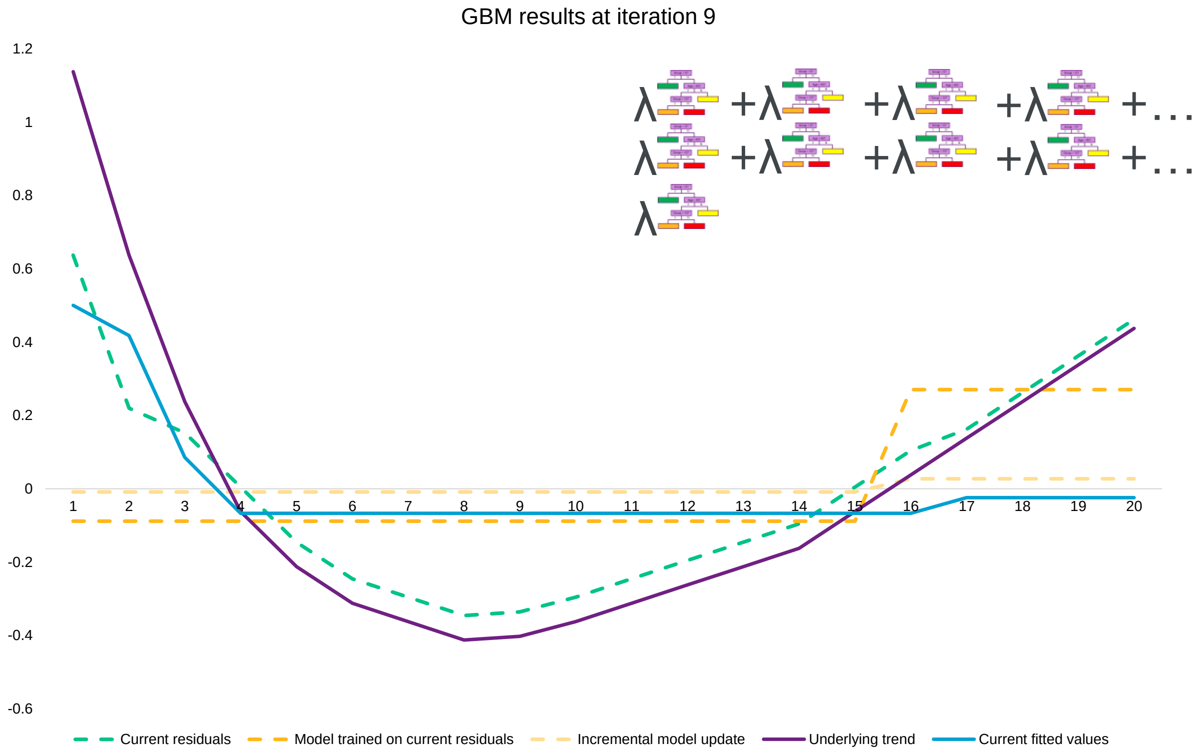

Institute and Faculty of Actuaries

# A simple GBM example

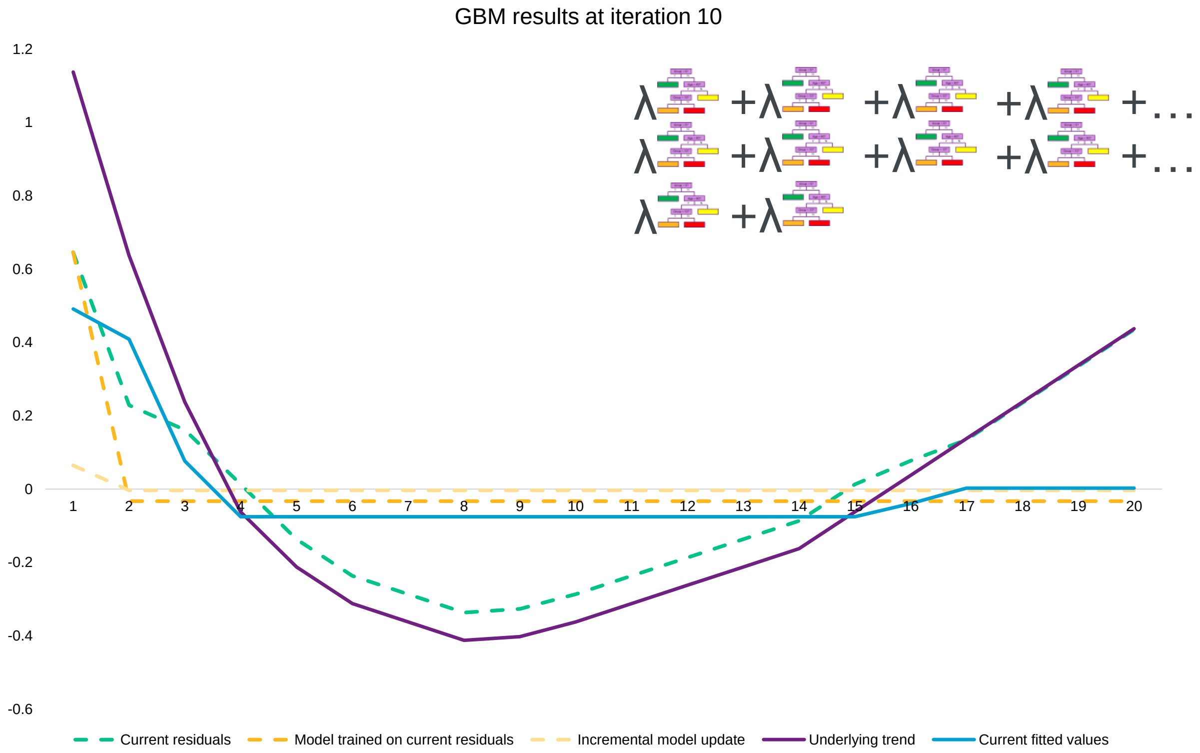



Institute and Faculty of Actuaries

# A simple GBM example

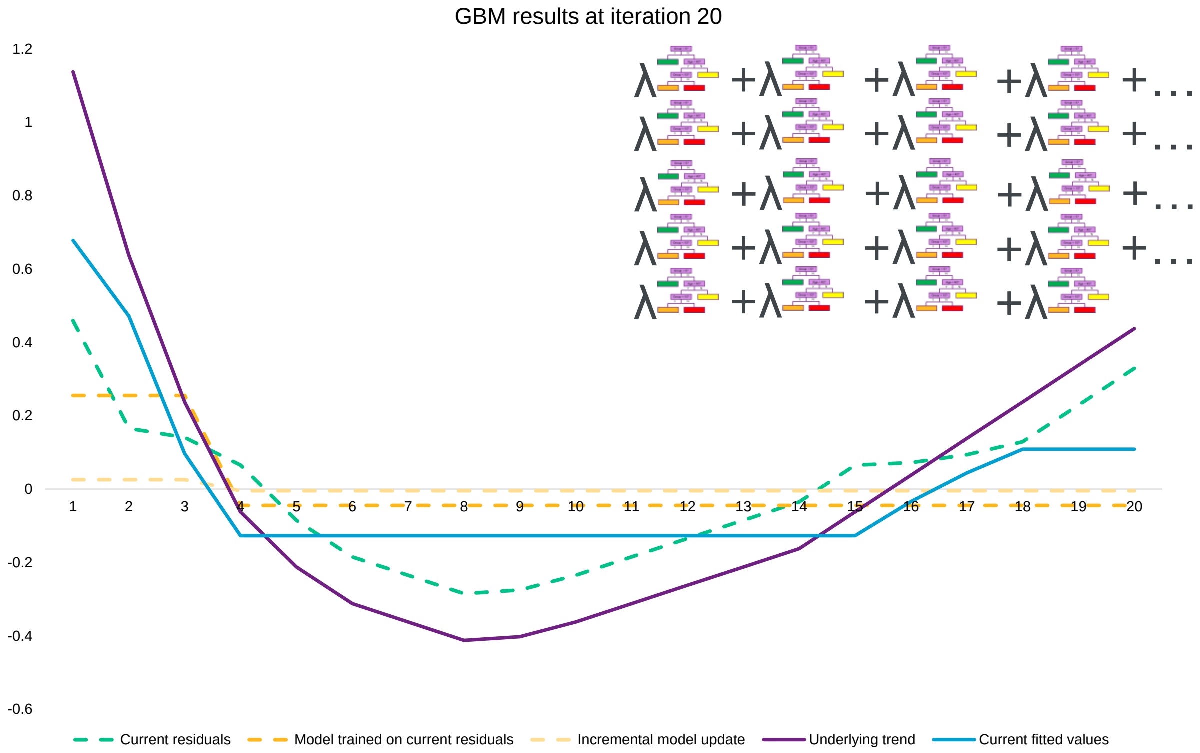


# A simple GBM example



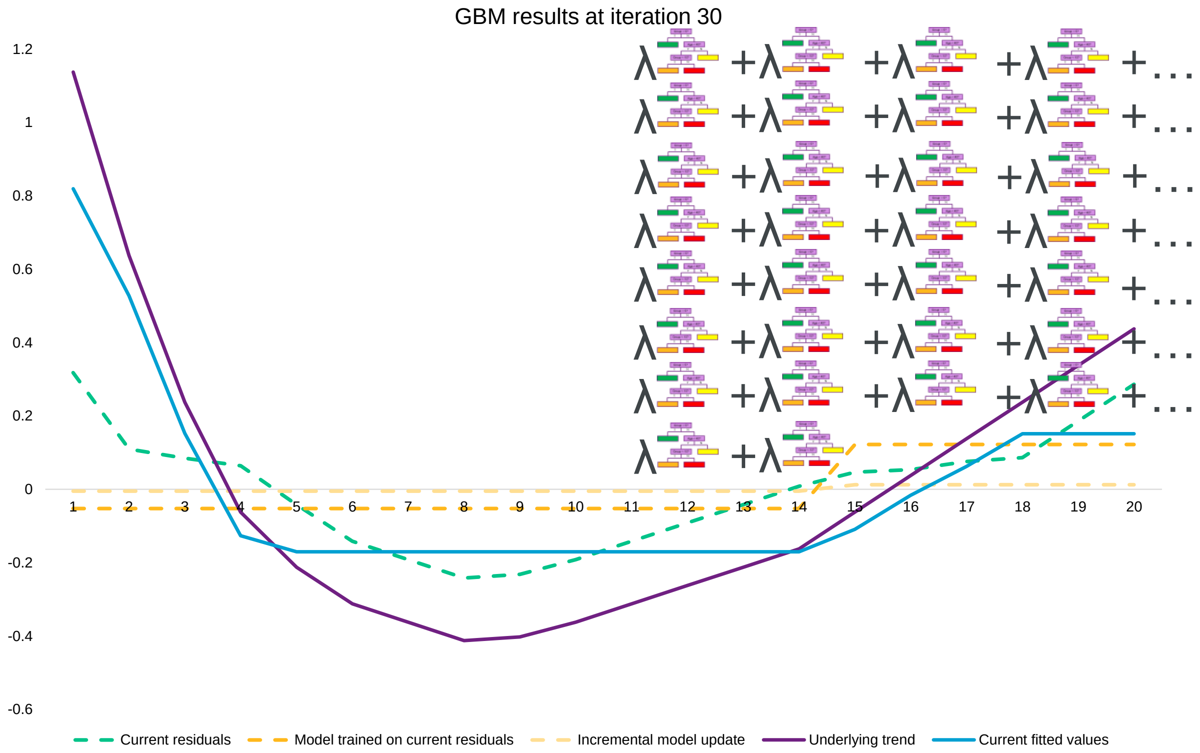

Institute and Faculty of Actuaries

# A simple GBM example



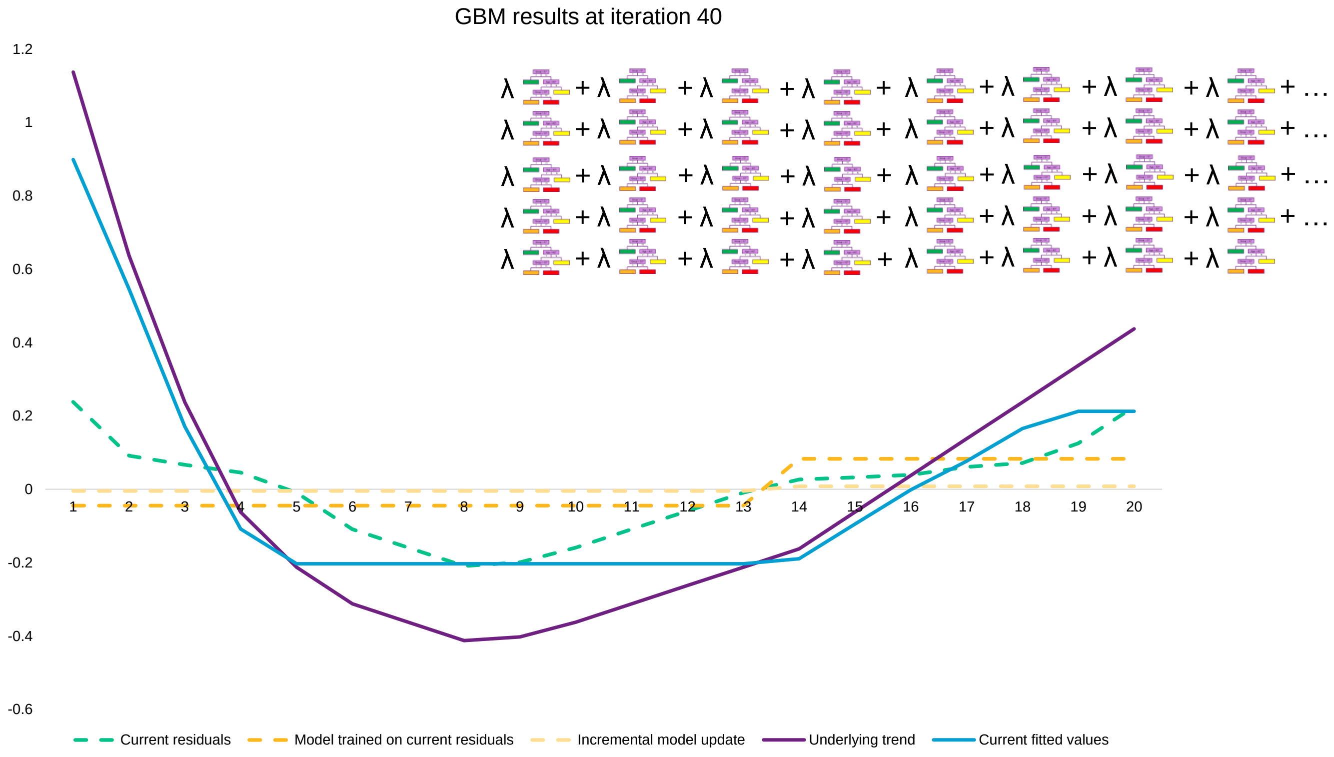

Institute and Faculty of Actuaries

# A simple GBM example



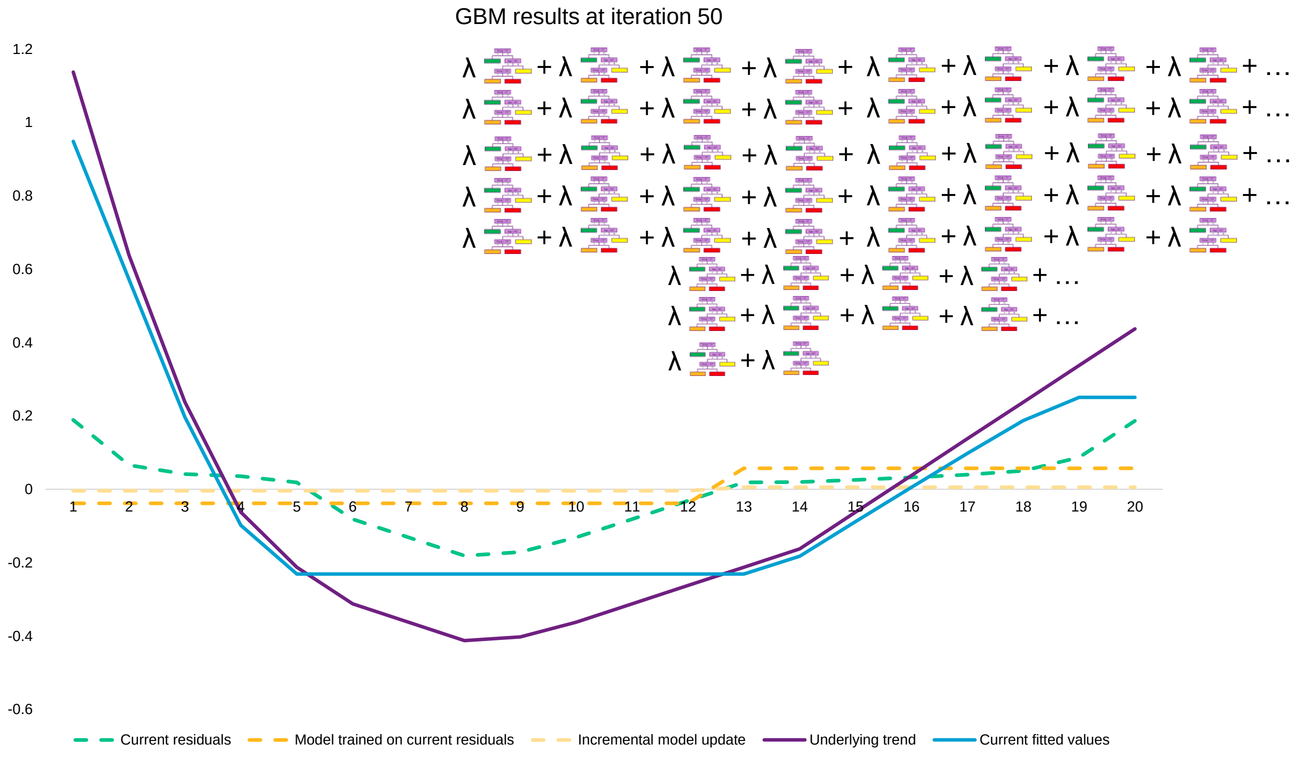

Institute and Faculty of Actuaries

# A simple GBM example



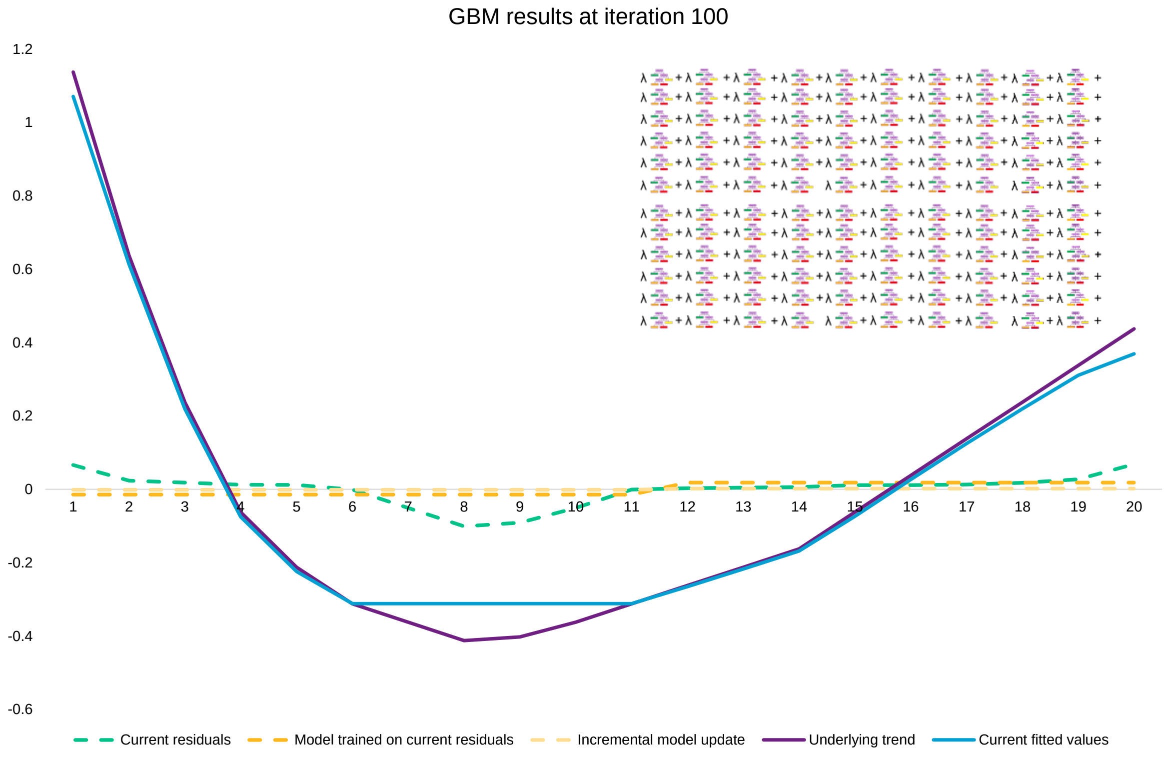

Institute and Faculty of Actuaries

# A simple GBM example

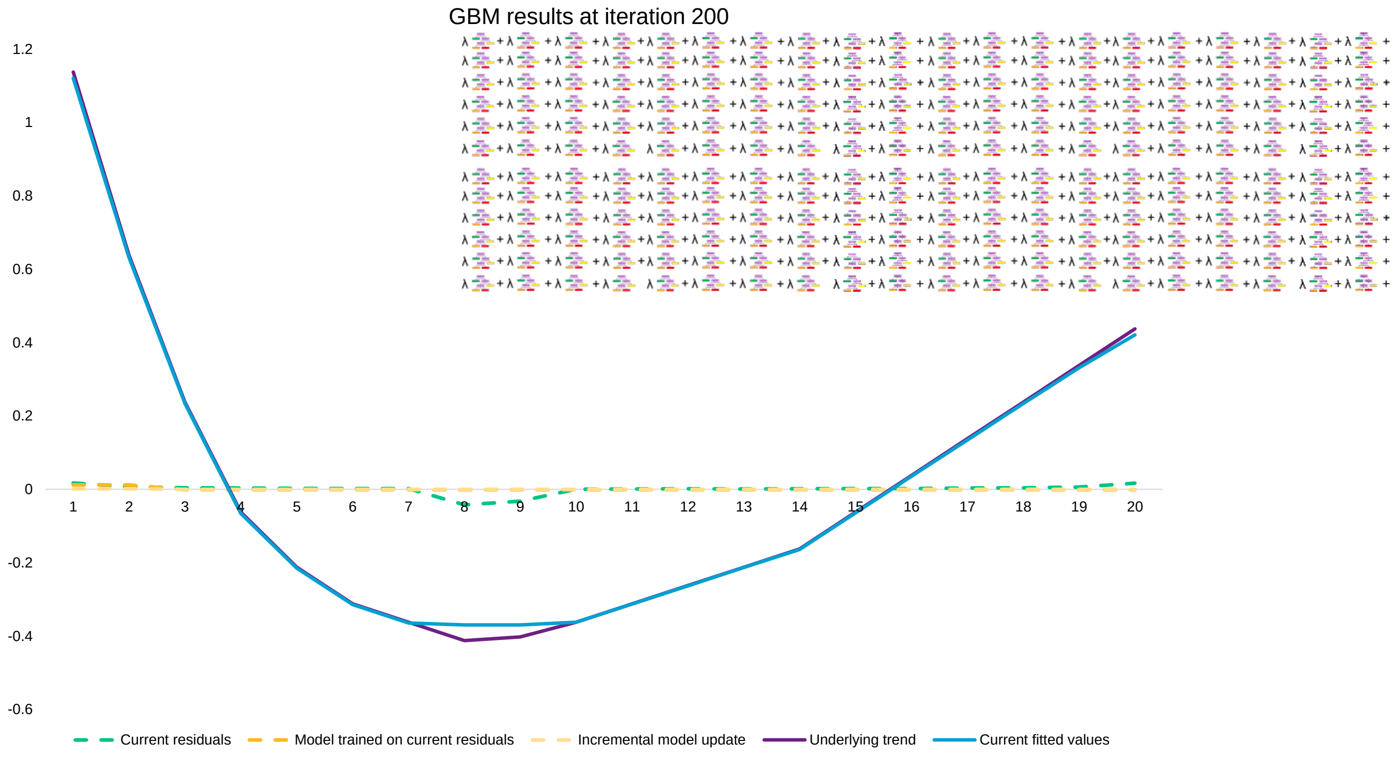



Institute and Faculty of Actuaries

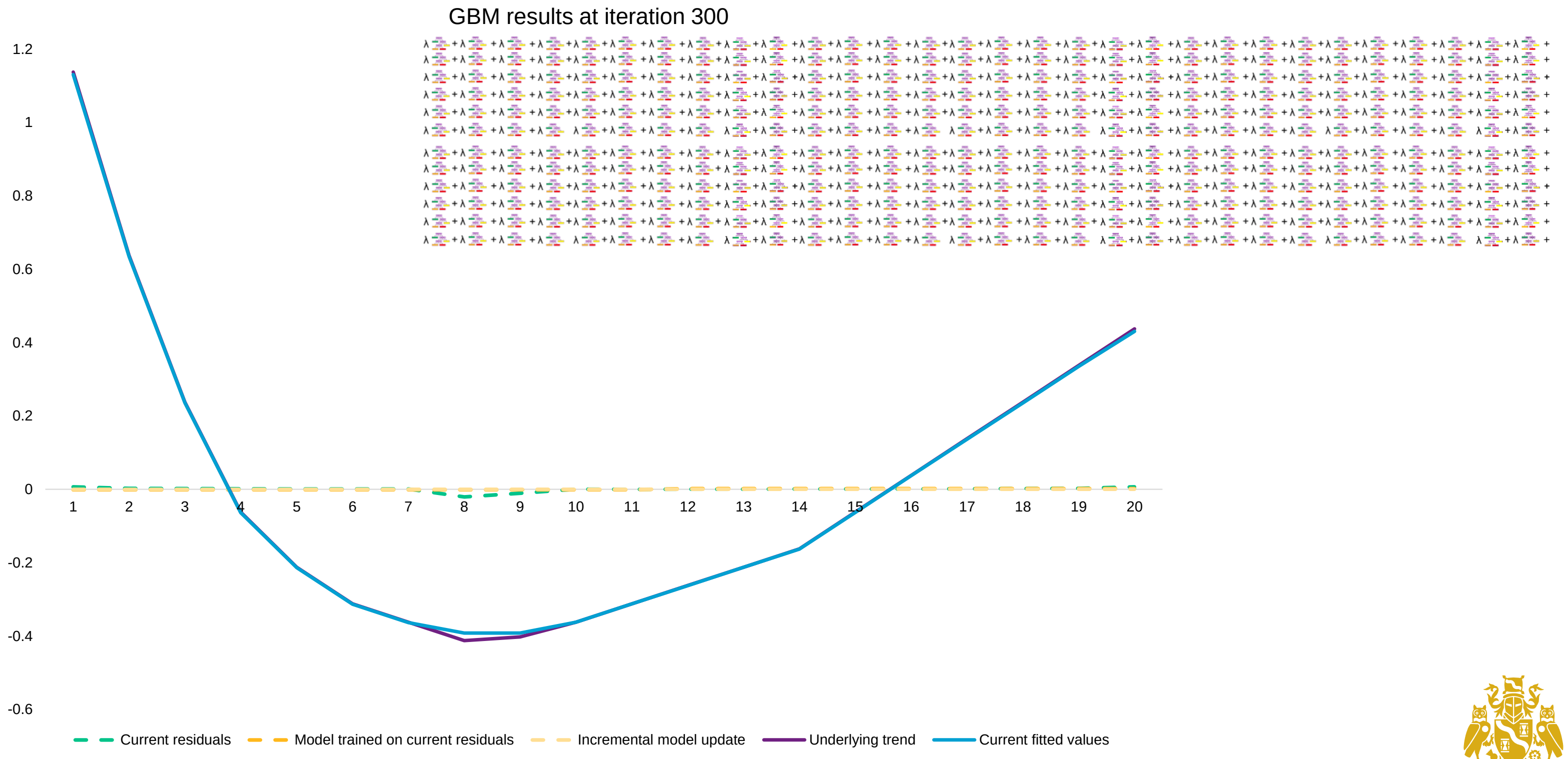
# A simple GBM example



# A simple GBM example

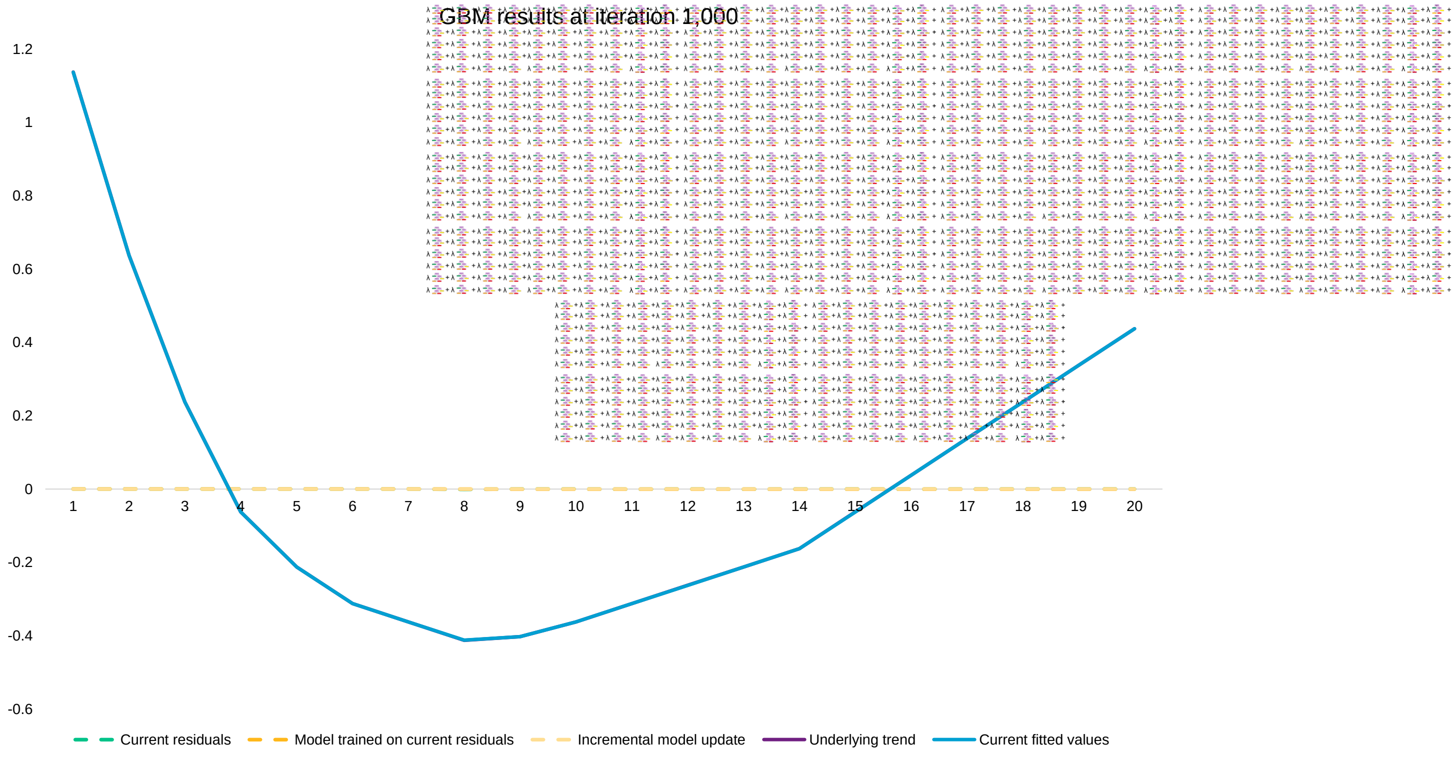


# A simple GBM example



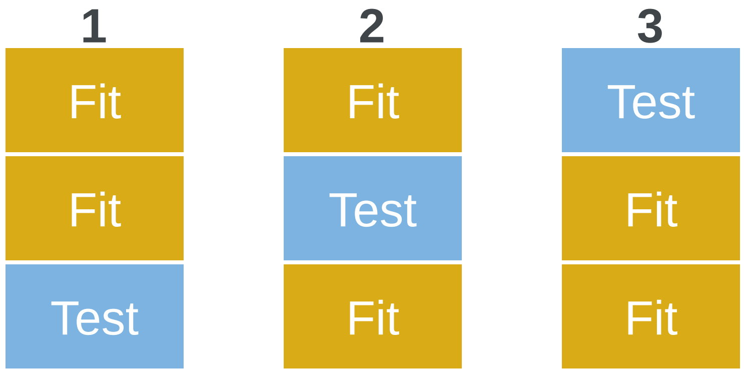

Institute and Faculty of Actuaries

# A simple GBM example



# Calibrating the assumptions

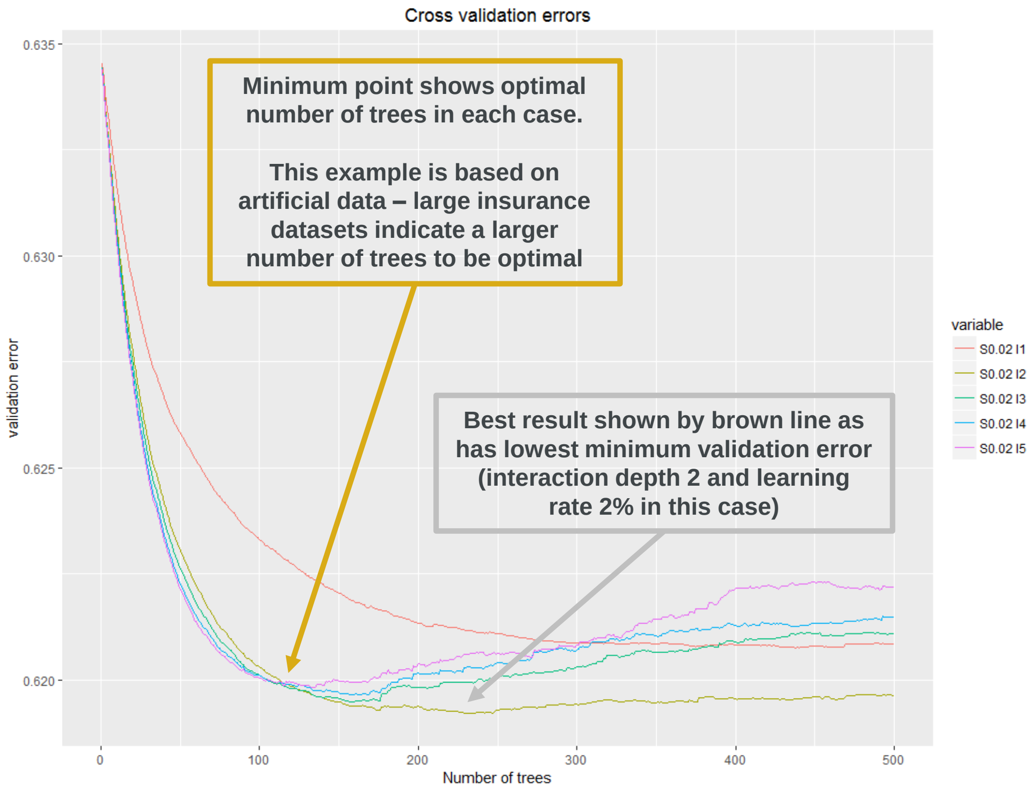
- *n*-fold cross validation used to develop the interaction depth and learning rate assumptions
  - Eg for 3-fold validation, split into 3, fit on gold, test on blue parts, take average




- Resulting plots can be used to determine the optimal assumption choice
  - Including how many trees to run

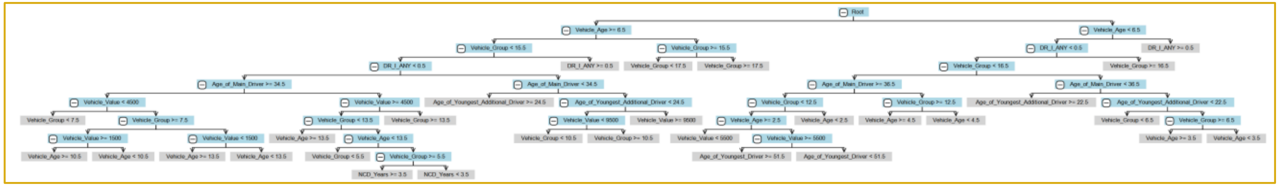
Institute and Faculty of Actuaries

# Example 5-fold cross validation




Institute and Faculty of Actuaries

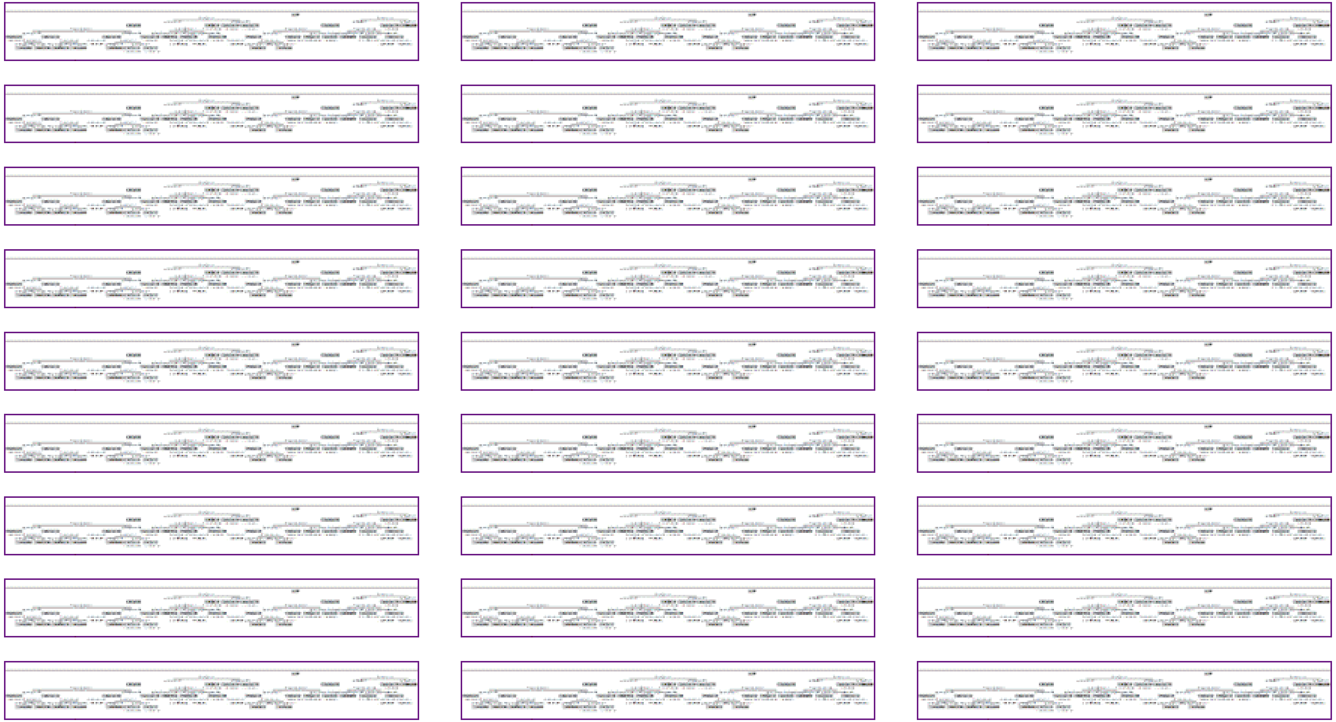
# What does a GBM look like?



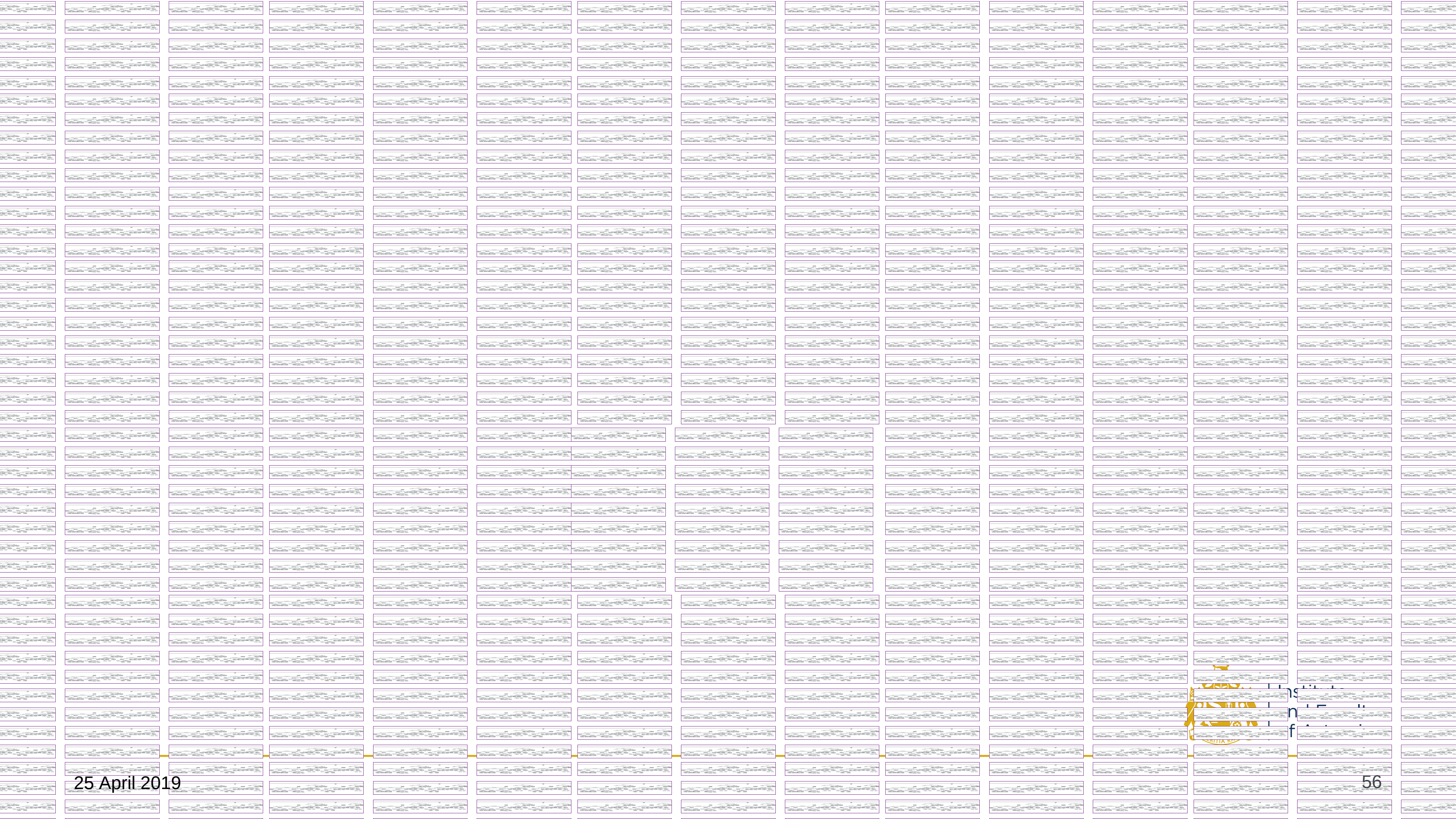

Institute and Faculty of Actuaries

# What does a GBM look like?

# What does a GBM look like?



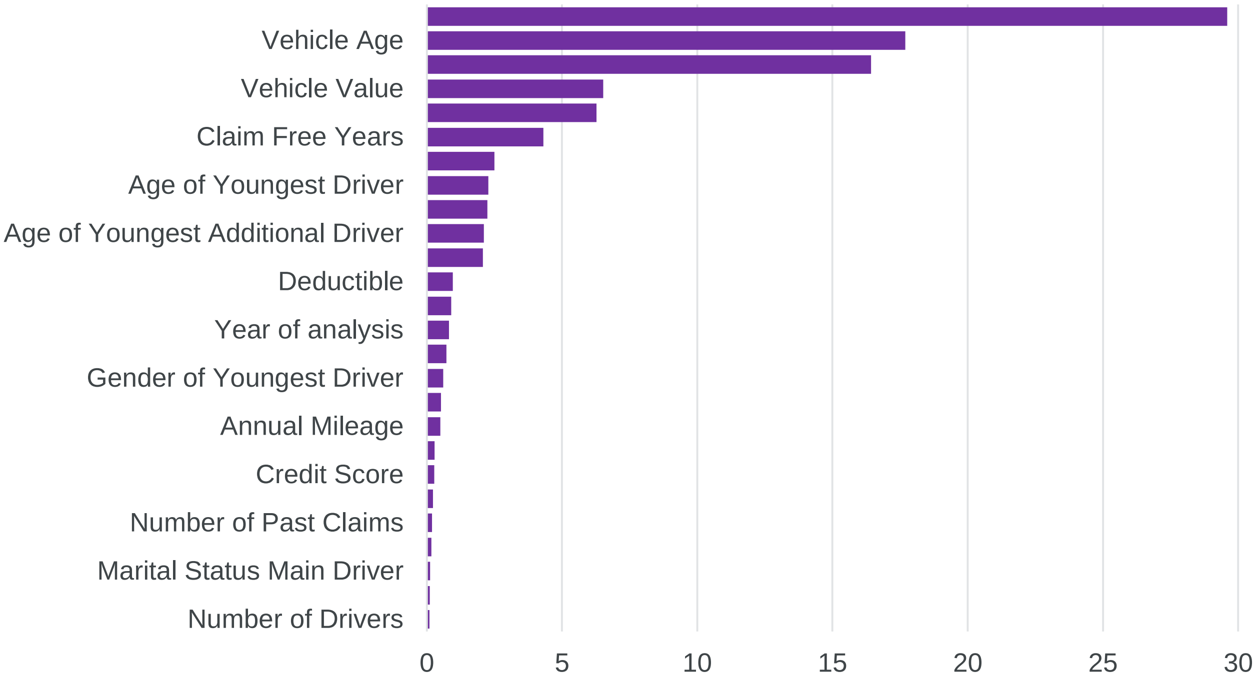
Institute and Faculty of Actuaries



- Does it work?
- How does it work?

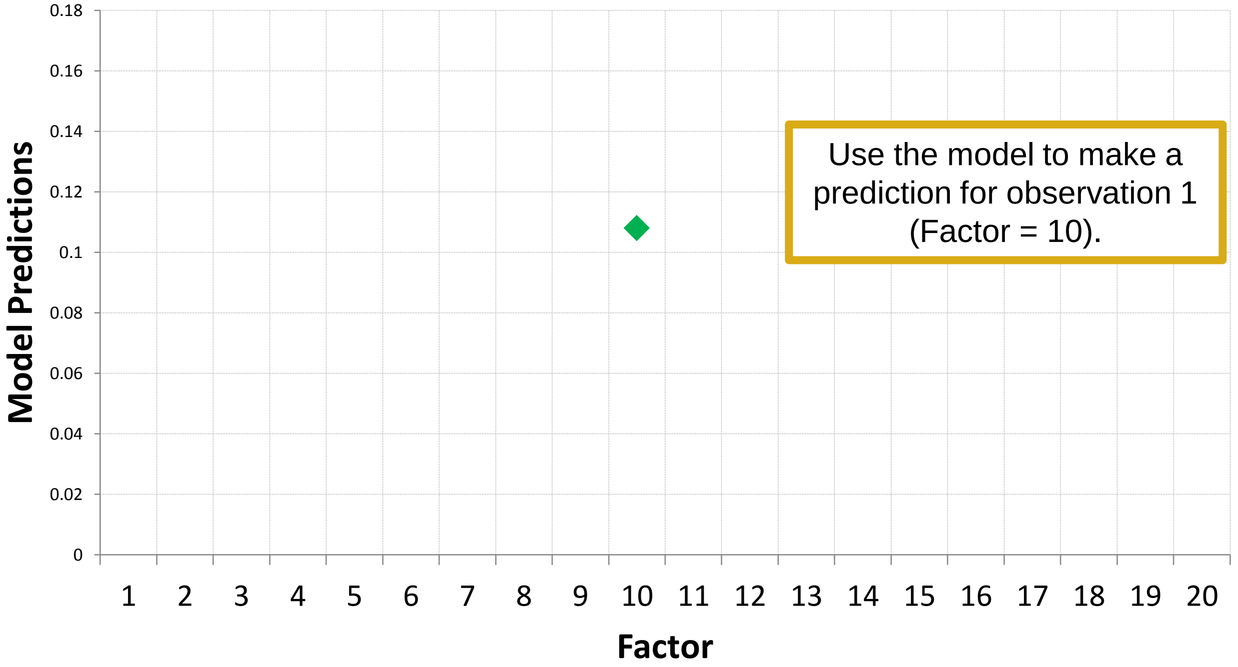
# Factor importance – relative influence

The relative influence of a factor can be measured as the total reduction in error attributable to splits by that factor, across all trees in the GBM



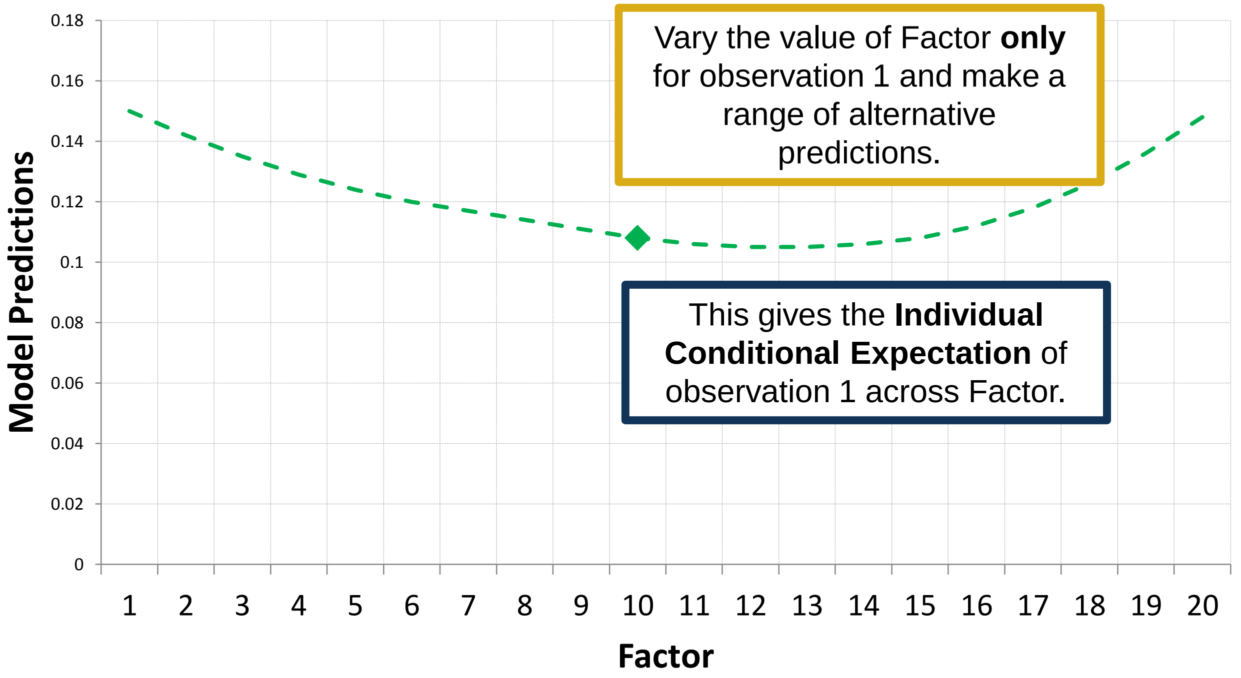
# Partial dependency plots

## Example



# Partial dependency plots

## Example

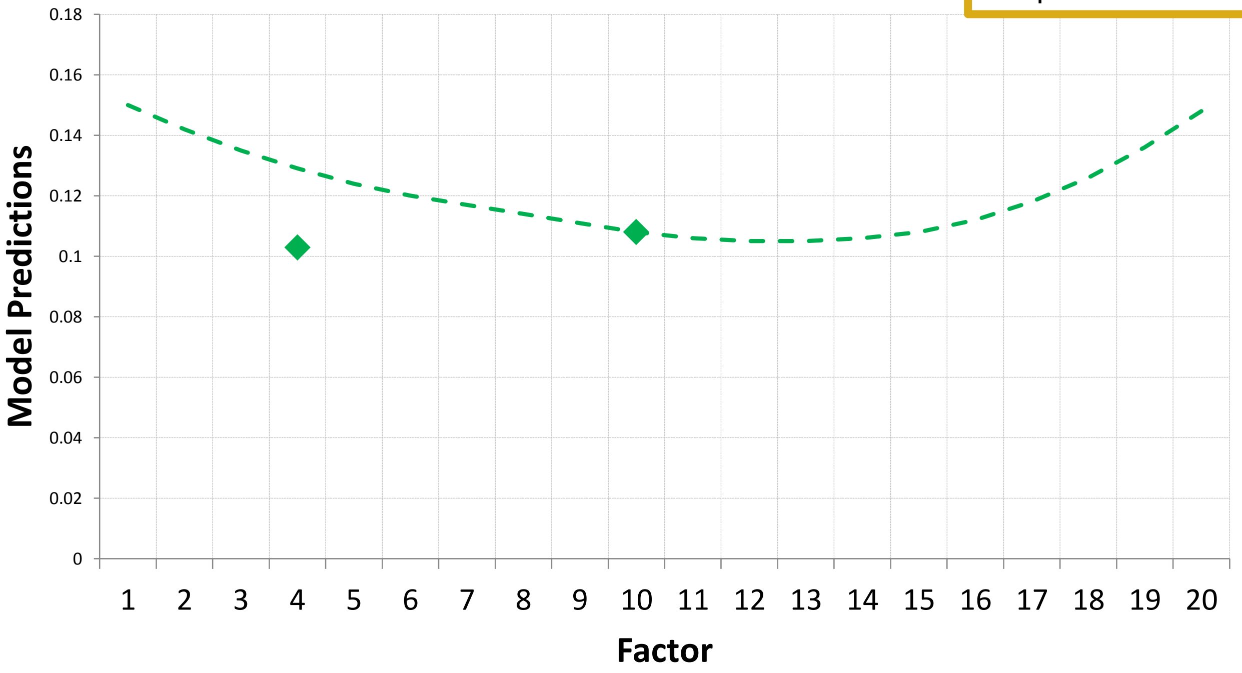



Institute and Faculty of Actuaries

# Partial dependency plots

## Example

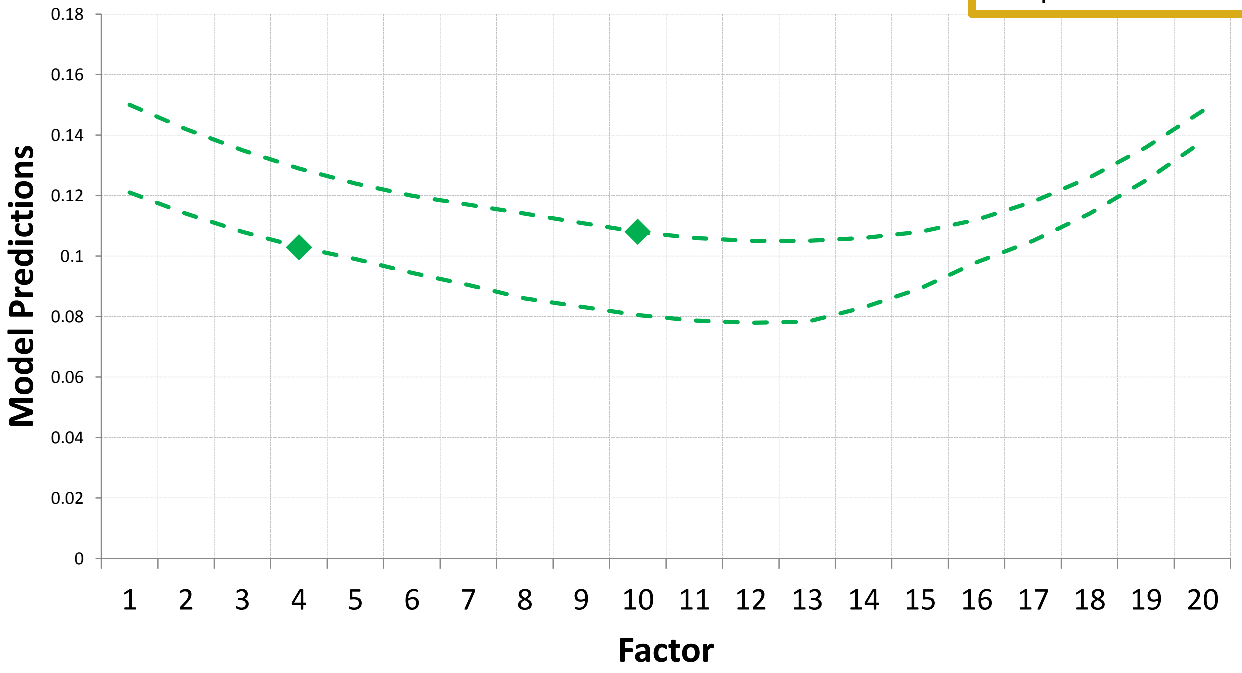
Repeat for all observations.



# Partial dependency plots

## Example

Repeat for all observations.

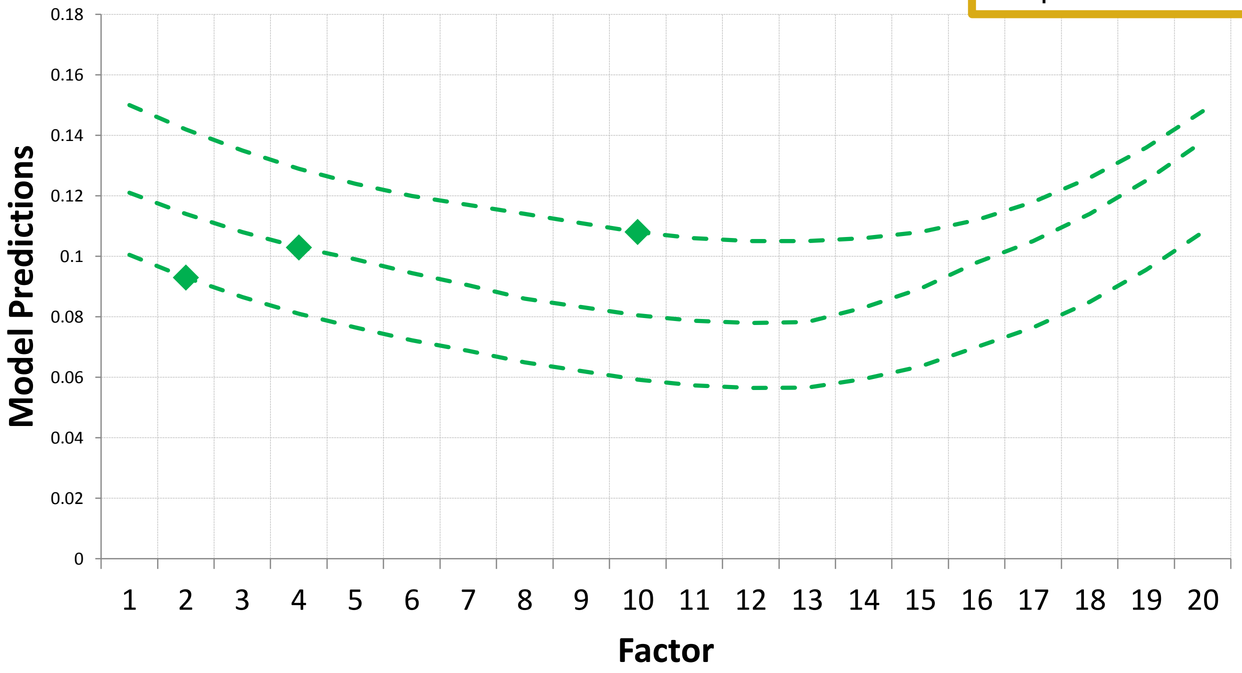



Institute and Faculty of Actuaries

# Partial dependency plots

## Example

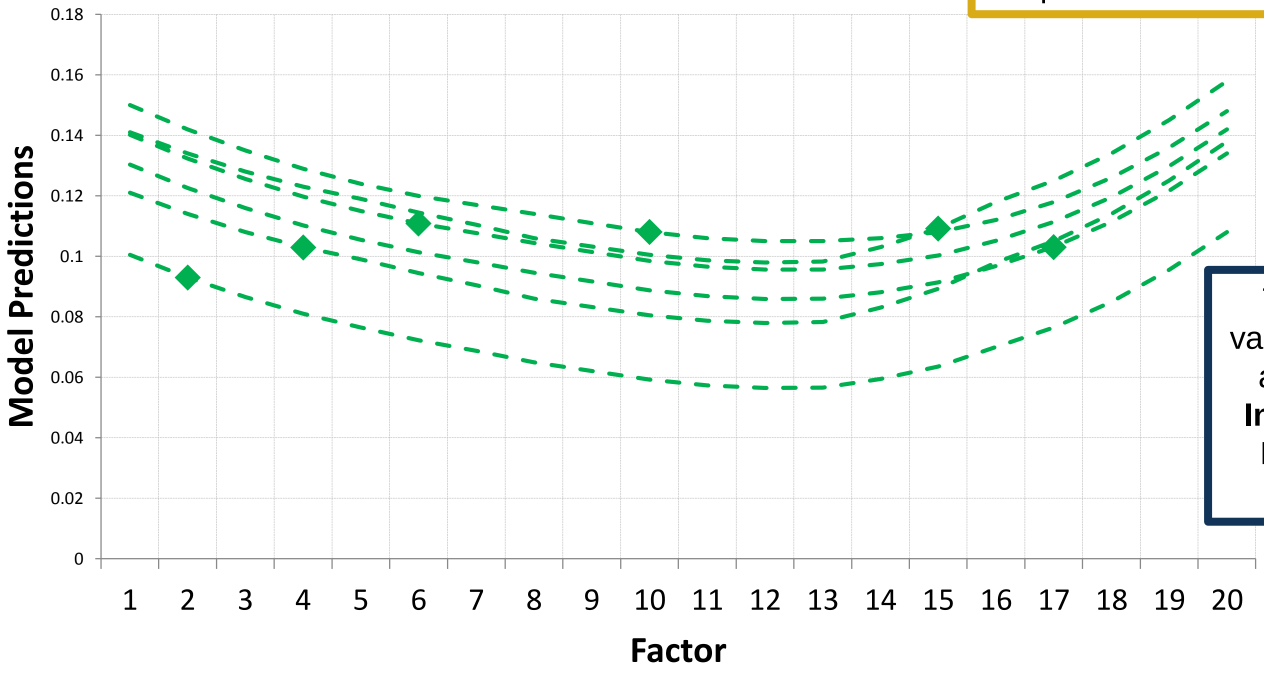
Repeat for all observations.



# Partial dependency plots

## Example

Repeat for all observations.

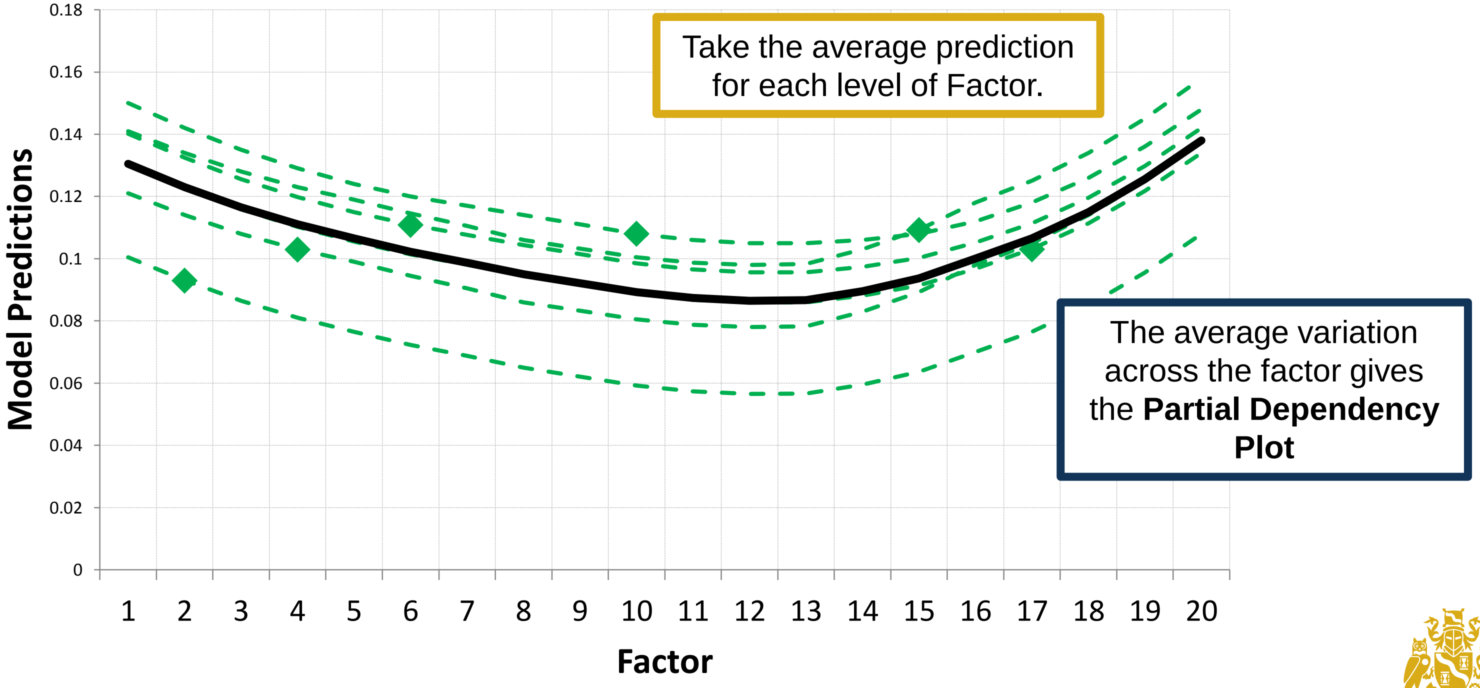



The full picture of the variation in predictions for all observations is the **Individual Conditional Expectation** (or **ICE**) plot.

Institute and Faculty of Actuaries

# Partial dependency plots

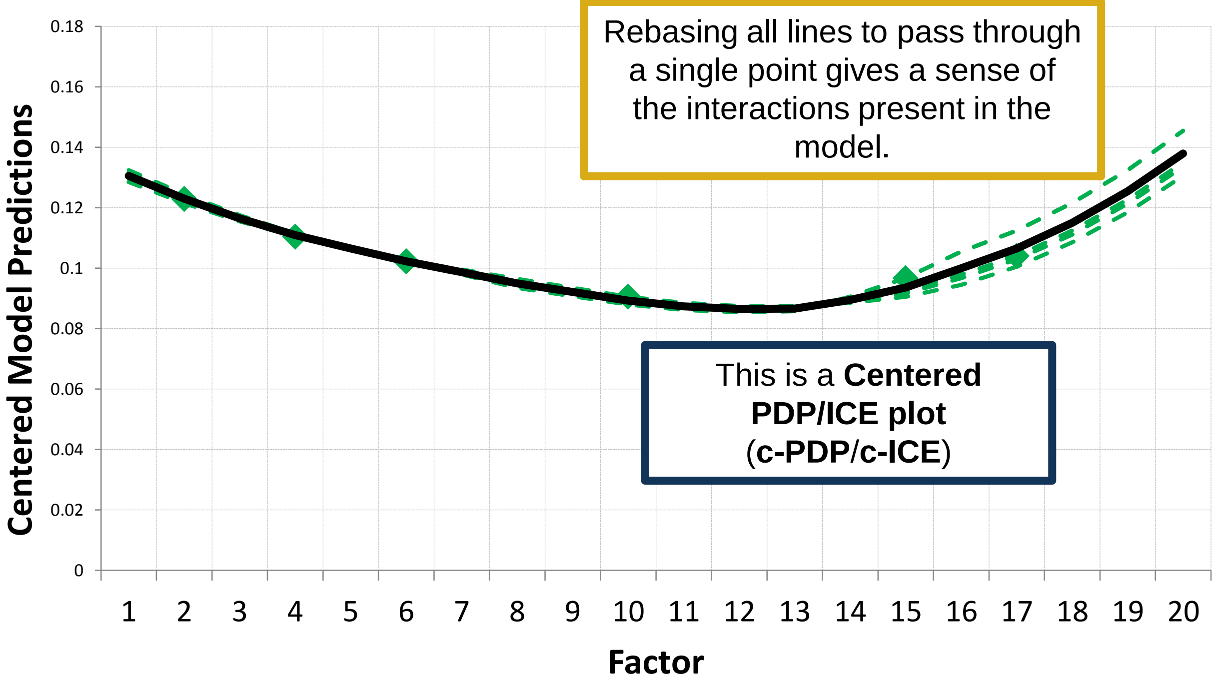
## Example




Institute and Faculty of Actuaries

# Partial dependency plots

## Example

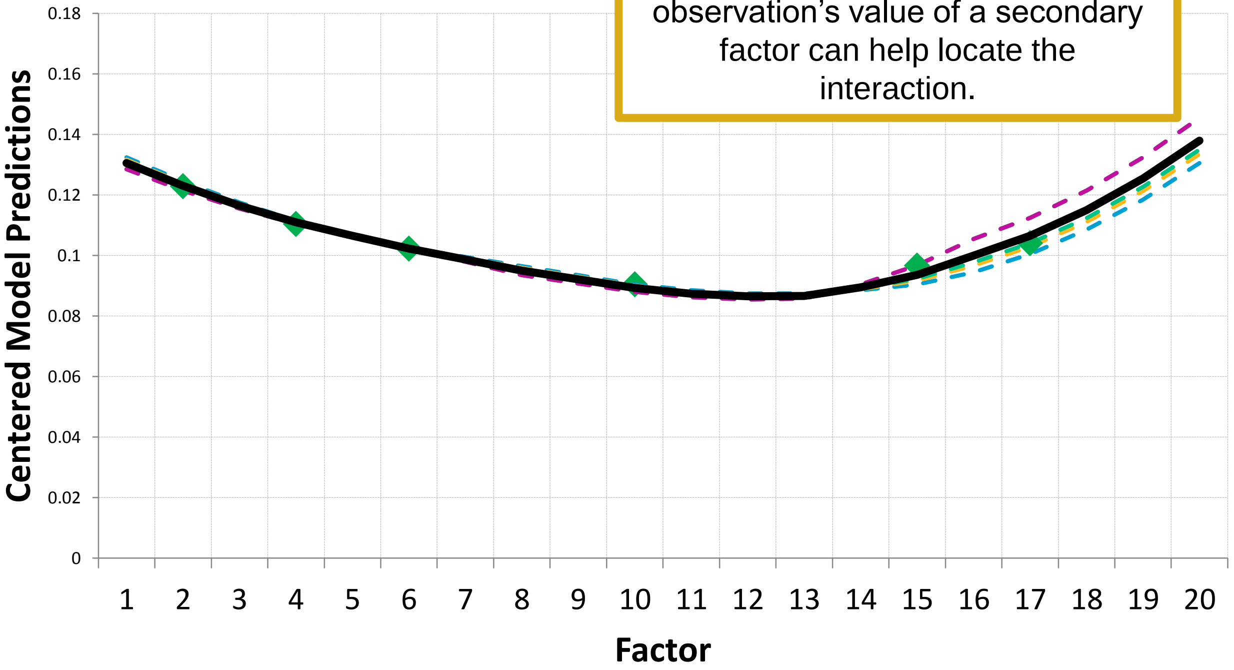



Institute and Faculty of Actuaries

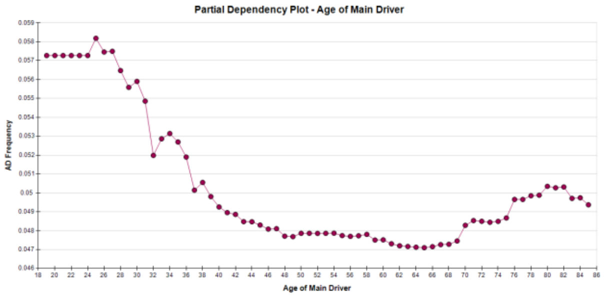
# Partial dependency plots

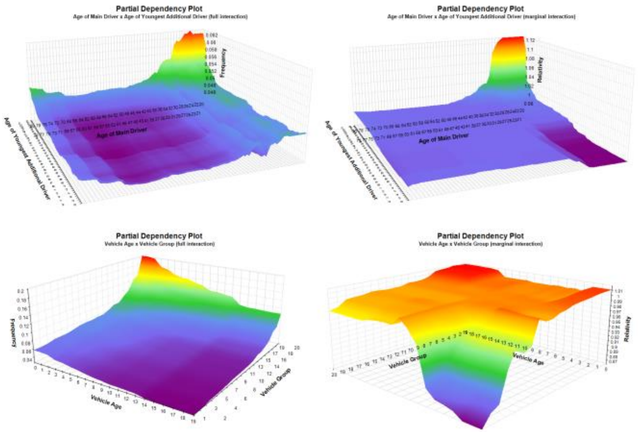
## Example

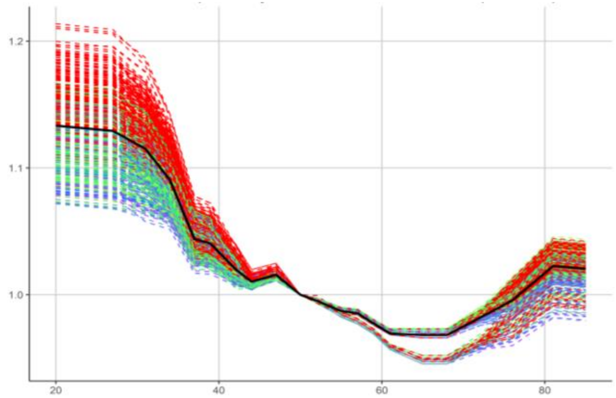
Colouring the c-ICE plots by each observation's value of a secondary factor can help locate the interaction.

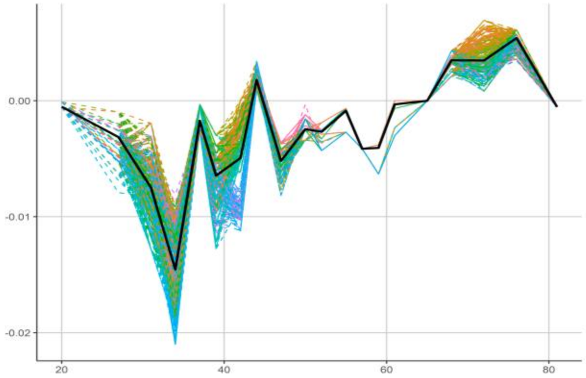


# Partial dependency plots etc



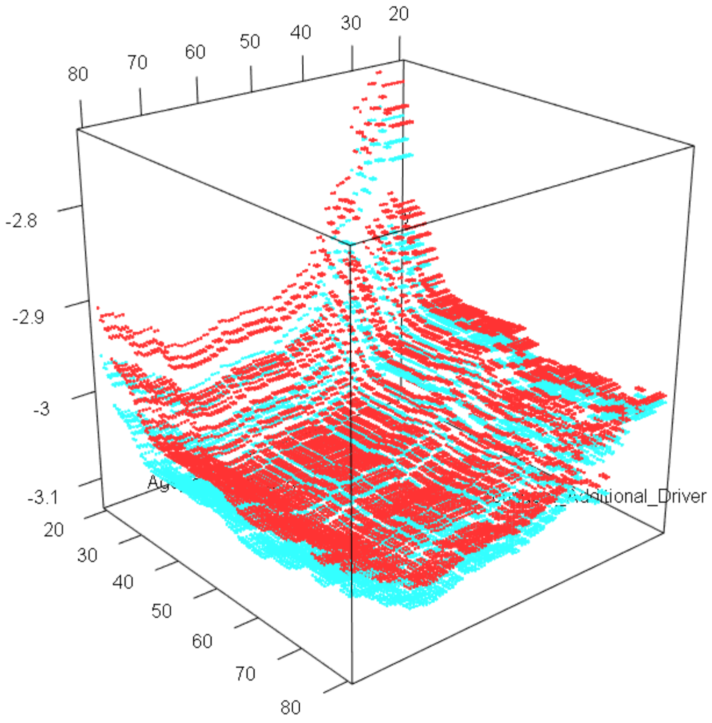




Institute and Faculty of Actuaries

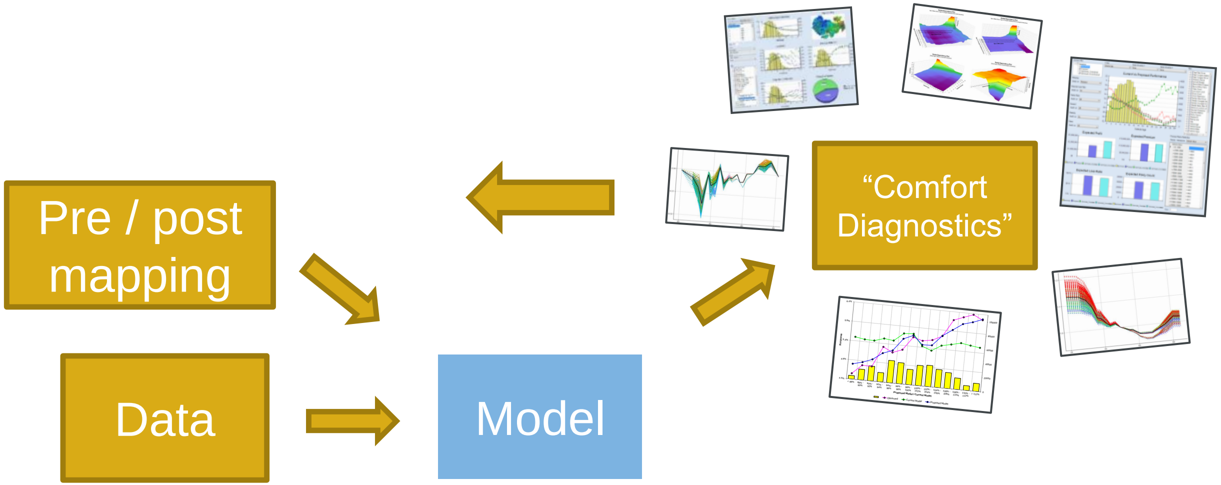
# Partial dependency plots




- Advantages
  - Qualitative description of properties of relationships
  - Most revealing of additive and multiplicative relationships
- Disadvantages
  - “GLM view of a non-GLM thing”
  - Interaction effects outside of the chosen subset may be obfuscated
  - eg if X1X2 is important and X2 is averaged out in the partial dependence plot, X1 may show as being heterogeneous, thus obfuscating the complexity of the modelled relationships

Institute and Faculty of Actuaries

# Model build process




Deploy directly

Institute and Faculty of Actuaries

# Deploying GBMs

Model down into multiplicative tables via GLMs

| | Age | Exposure | Burning Cost |
|---|---|---|---|
| 1 | <=20 | 1,720 | 179 |
| 2 | 21-30 | 34,893 | 122 |
| 3 | 31-50 | 118,182 | 102 |
| 4 | 51+ | 127,054 | 70 |
| 5 | Age Total | 281,849 | 91 |

| | Vehicle Group | Exposure | Burning Cost |
|---|---|---|---|
| 1 | 1-10 | 164,107 | 77 |
| 2 | 11-14 | 84,859 | 101 |
| 3 | 15-18 | 28,952 | 116 |
| 4 | 19-20 | 3,931 | 272 |
| 5 | VG Total | 281,849 | 91 |

| | Gender | Exposure | Burning Cost |
|---|---|---|---|
| 1 | Male | 197,339 | 92 |
| 2 | Female | 84,510 | 87 |
| 3 | Gender Total | 281,849 | 91 |

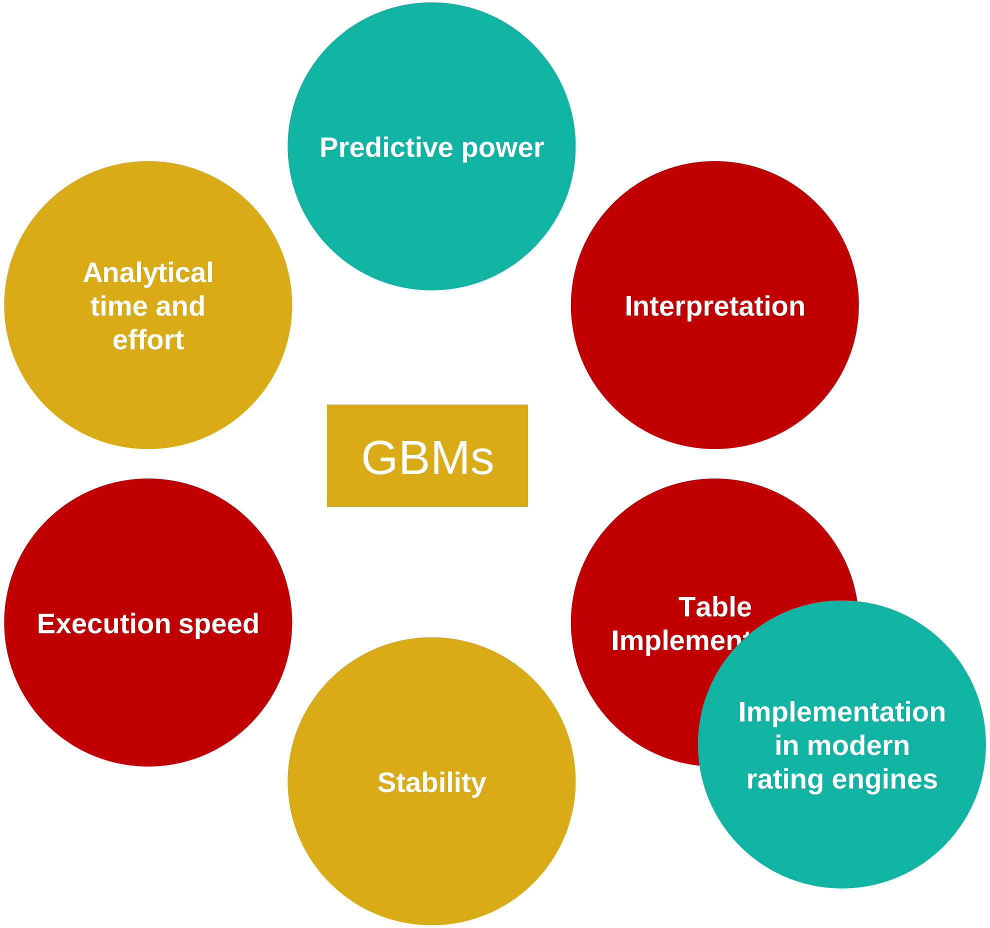
**Analytical environ-ment**

**Next gen rating engine**

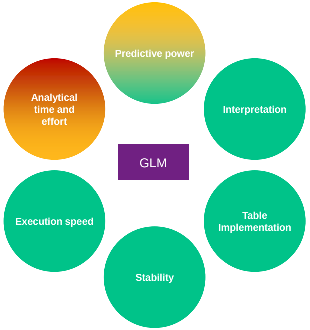
**Main Policy Admin System**

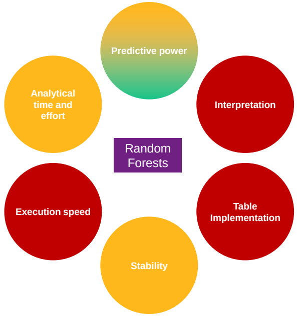
Deploy directly

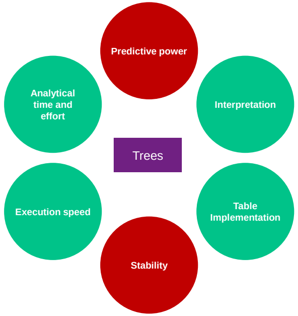
Institute and Faculty of Actuaries

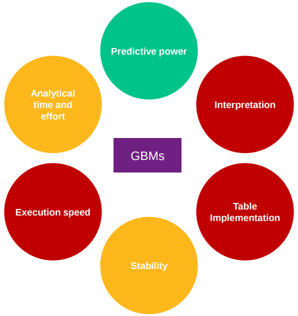

Predictive power
Analytical time and effort
Interpretation
GBMs
Execution speed
Table
Implement
Implementation in modern rating engines
Stability

# A summary…




Institute and Faculty of Actuaries

# Machine Learning in Pricing

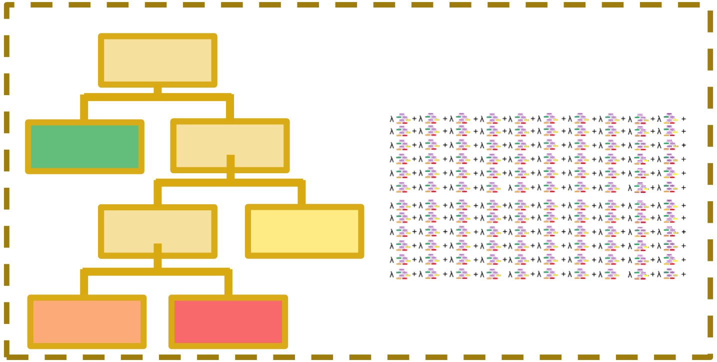
## Conclusions

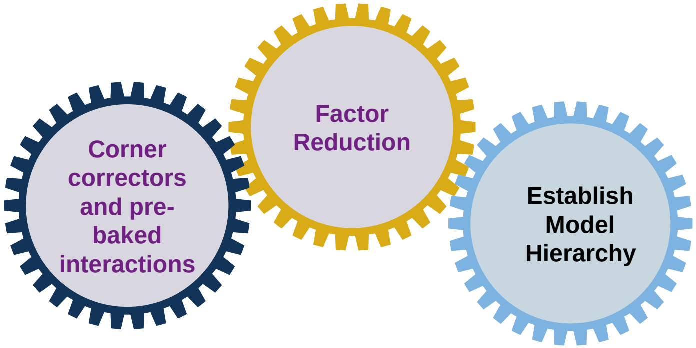
- There are many forms of ML models
- New data and feature/response engineering generally add more value than new methods BUT we need to continuously explore which methods work on which problems
- Traditional measures of prediction value may not reflect applications in insurance
- And it's not all about predictive power anyway – other criteria are important

- GBMs can provide predictive lift benefits by capturing higher order effects … BUT
  - Can you cope with not seeing the model and instead use broad diagnostics
  - Effort is required to expose/understand higher order effects in an expeditious manner
  - How will business leaders and regulators respond to this method?
  - Do you have the software and hardware to fit to large dataset
  - Do you have a rating engine that can implement a GBM

Institute and Faculty of Actuaries

# Practical applications of tree based methods in pricing






Institute and Faculty of Actuaries

# Questions

?

Institute and Faculty of Actuaries

Institute
and Faculty
of Actuaries

# Thank you