Opening the Black Box with R

JB Crozet & James Simmons

8th June 2017
We have a problem.

“Computers Are Useless. They Can Only Give You Answers”

Pablo Picasso
We created a man-made monster.
How we did it (1).
How we did it (2).

Contracts

- 2k contracts
- 1m location/contract
- 100k scenarios/locations
- 100k stochastic simulation

Technical Price

Catastrophe Modelling
How we did it (3).

Contracts

Marginal Pricing

- tail metrics
- Δ portfolio
- percentile layer capital allocation
- non-linear outwards reinsurance

Technical Price
How we did it (4).

- Contracts
- Localised movements
- Contract and portfolio effects
- Δ moving portfolio

Technical Price
Where we got to.

“Trying to make sense of crazy will drive you insane.”
R to the Rescue.

Reverse Engineer  Simplify
1- European Wind exposures drive increases.

2- North American Wind exposures drive decreases, with small pocket driving increases.

3- Higher NAWS layers are experiencing higher decreases.

Does that make intuitive sense?
Telling Stories with Pruned Trees (2).
Making Predictions (1).
Making Predictions (2).

Stepwise Regression

- ELoL
- Peak EMF
- AEP1
- AEP59

Predicted: -11.8%
(Actual: -13.5%)
MAE: 17%
Do we have a solution?

"Life is painting a picture, not doing a sum."

~ Oliver Wendell Holmes
Appendix
# Fit regression tree
library(rpart)
f <- rpart(TPMov ~ ELoL + PeakEMF + AEP4 + AEP5 + AEP6 + AEP7 + AEP8 + AEP10 + AEP12 + AEP13 + AEP14 + AEP15 + AEP18 + AEP25 + AEP26 + AEP28 + AEP30 + AEP56, data=Data, method="anova")
library(rattle)
library(rpart.plot)
library(RColorBrewer)
fancyRpartPlot(f)

# Optimising pruning by cross-validation
plotcp(f)
printcp(f)
ptree <- prune(f, cp=fit$cptable[which.min(fit$cptable[, 'xerror']), 'CP'])
fancyRpartPlot(ptree, uniform=TRUE, main="Pruned Classification Tree")

library(MASS)
RegressionData <- Data

# Remove outliers
RegressionData <- RegressionData[-c(1199, 1403, 1404)]

# Perform multiple regression
regressionFit <- lm(TPMov ~ ELoL + AEP4 + AEP5 + AEP6 + AEP7 + AEP8 + AEP10 + AEP12 + AEP13 + AEP14 + AEP15 + AEP18 + AEP25 + AEP26 + AEP28 + AEP30 + AEP56, data = RegressionData)
print(regressionFit)
summary(regressionFit)
plot(regressionFit)

# Stepwise regression
step <- stepAIC(regressionFit, direction="both")
step$anova # display results
summary(step)
plot(step)

MAE <- function(actual, predicted) { mean(abs(actual - predicted)) }

# Regression tree
p.part <- predict(f, Data)
MAE(p.part, Data$TPMov)

# Stepwise regression
p.step <- predict(step, RegressionData)
MAE(predict(step, RegressionData), RegressionData$TPMov)

# Uniform allocation
MAE(0, Data$TPMov)
MAE(mean(Data$TPMov), Data$TPMov)

# Uniform allocation
MAE(0, Data$TPMov)
MAE(mean(Data$TPMov), Data$TPMov)