Bayesian Finance

Christa Cuchiero, Irene Klein, Josef Teichmann

Obergurgl 2017

1. Calibrating a Bayesian model: a first trial

Calibration problem:

- Bayesian models tend to be highly parametrized.
- Ad-hoc choice of support of the prior Θ.
- Optimizing over different supports would possibly lead to an unfeasible optimization problem by classical methods.

Calibration by Machine learning following Andres Hernandez

We shall provide a brief overview of a procedure introduced by Andres Hernandez (2016) as seen from the point of view of Team 3's team challenge project 2017 at UCT:

Algorithm suggested by A. Hernandez

- Getting the historical price data.
- Calibrating the model, *a single factor Hull-White extended Vasiček model* to obtain a time series of (typical) model parameters, here the yield curve, the rate of mean reversion α, and the short rate's volatility σ.
- Pre-process data and generate new combinations of parameters.
- With a new large training data set of (prices,parameters) a neural network is trained.
- The neural network is tested on out-of-sample data.

The data set

- The collected historical data are ATM volatility quotes for GBP from January 2nd, 2013 to June 1st, 2016. The option maturities are 1 to 10 years, 15 years and 20 years. The swap terms from 1 to 10 years, plus 15, 20 and 25 years.
- The yield curve is given 44 points, i.e. it is discretely sampled on 0, 1, 2, 7, 14 days; 1 to 24 months; 3-10 years; plus 12, 15, 20, 25, 30, 40 and 50 years. Interpolation is done by Cubic splines if necessary.

Classical calibration a la QL

Historical parameters

- a Levenberg-Marquardt local optimizer is first applied to minimize the equally-weighted average of squared yield or IV differences.
- calibration is done twice, with different starting points:
  - at first, $\alpha = 0.1$ and $\sigma = 0.01$ are the default choice
  - second the calibrated parameters from the previous day (using the default starting point) are used for the second stage of classical calibration.

Calibration results along time series

The *re-calibration problem* gets visible ... and it is indeed a feasible procedure.

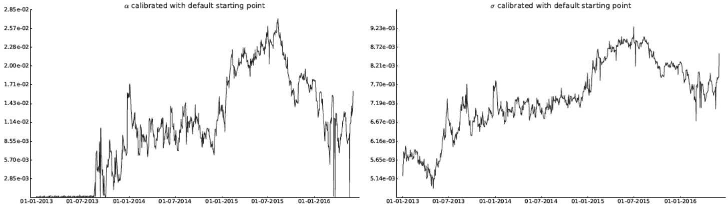


Figure: Calibration using default starting point

How do neural networks enter calibration?

Universal approximation of calibration functionals

- Neural networks are often used to approximate functions due to the universal approximation property.
- We approximate the calibration functional (yields,prices) $\mapsto$ (parameters) which maps (yields, prices) to optimal model parameters by a neural network.

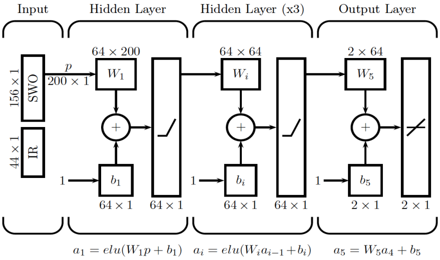
Neural Networks : Training Set Generation

With the calibration history A. Hernandez proceeds by generating the training set

- obtain errors for each calibration instrument for each day,
- take logarithms of of positive parameters, and rescale parameters, yield curves, and errors to have zero mean and variance 1,
- apply a principal component analysis and an appropriate amount of the first modes,
- generate random normally distributed vectors consistent with given covariance,
- apply inverse transformations, i.e. rescale to original mean, variance and exponentiate,
- apply random errors to results.

Neural Networks: Training the network

- With a sample set of 150 thousand training data points, A. Hernandez suggests to train a feed-forward neural network.
  - The architecture is chosen feed-forward with 4 hidden layers, each layer with 64 neurons using an ELU (Exponential Linear Unit)



Neural Networks: testing the trained network

- two neural networks were trained using a sample set produced where the covariance matrix was estimated based on 40% of historical data.
- the second sample set used 73% of historical data.
- for training, the sample set was split into 80% training set and 20% cross-validation.
- the testing was done with the historical data itself (i.e. a backtesting procedure was used to check the accuracy of the data).

Results of A. Hernandez

The following graphs illustrate the results. Average volatility error here just means

$$\frac{\sum_{n=1}^{156} \left| impvol^{mkt} - impvol^{model} \right|}{156} \tag{1}$$

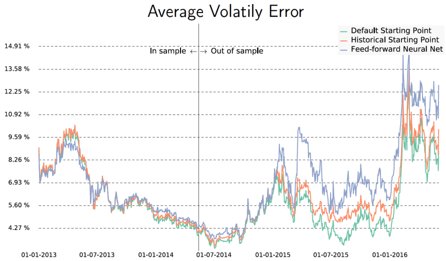


Figure: Correlation up to June 2014

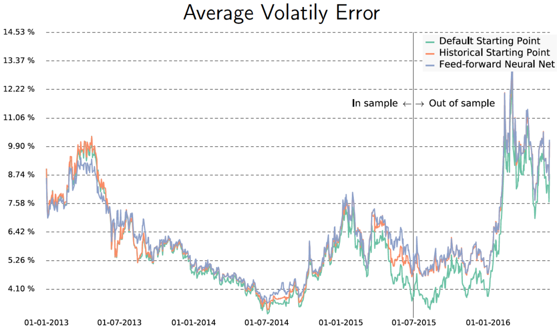


Figure: Correlation up to June 2015

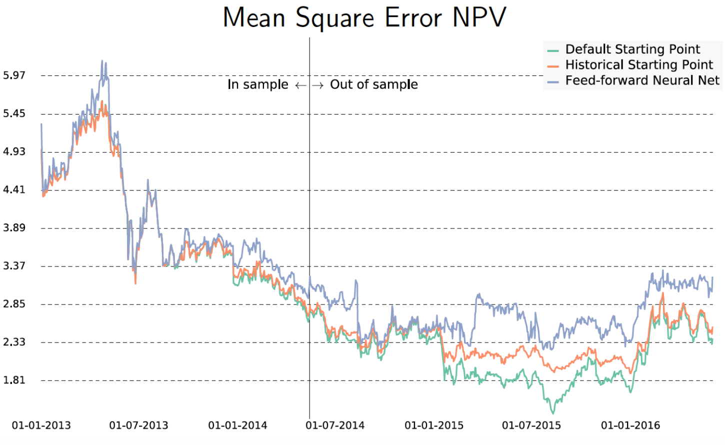


Figure: Correlation up to June 2014

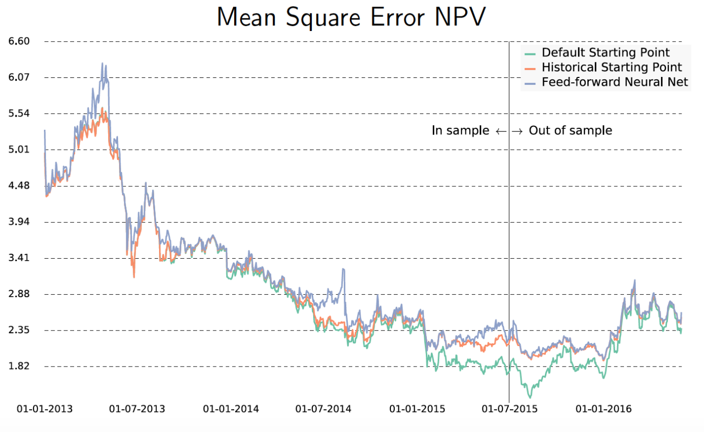


Figure: Correlation up to June 2015

Towards a Bayesian model

Consider the Hull-White extended Vasiček models (on a space $(\Omega, \mathcal{F}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$):

$$dr_t^{(1)} = (\beta_1(t) - \alpha_1 r_t^{(1)})\, dt + \sigma_1\, dW_t\,,$$

$$dr_t^{(2)} = (\beta_2(t) - \alpha_2 r_t^{(2)})\, dt + \sigma_2\, dW_t\,.$$

We assume that r is is a mixture of these two models with constant probability $\pi \in [0, 1]$, i.e.

$$\mathbb{P}(r_t \leq x) = \pi \mathbb{P}\left(r_t^{(1)} \leq x\right) + (1 - \pi)\mathbb{P}\left(r_t^{(2)} \leq x\right).$$

Of course the observation filtration generated by daily ATM swaption prices and a daily yield curve is smaller than the filtration $\mathbb{G}$, hence the theory of the first part applies.

Bayesian model: setup

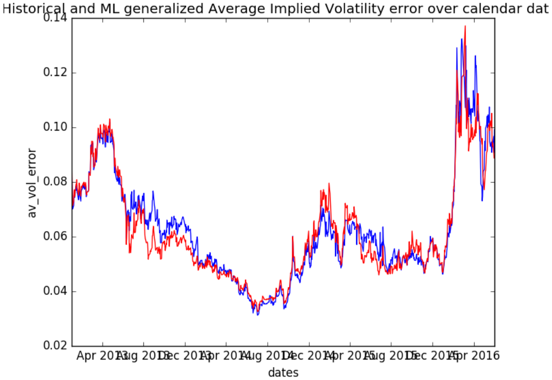
We still have the same set-up (in terms of data):

- $N = 156 + 44 = 200$ input prices (swaptions + yield curve)
- $n = 44 + 4 + 1 = 49$ parameters to estimate. These are $\alpha_1, \alpha_2, \sigma_1, \sigma_2, \pi$ and $\beta_1(t)$ (or, equivalently, $\beta_2(t)$) at 44 maturities.
- Hence, the calibration function is now

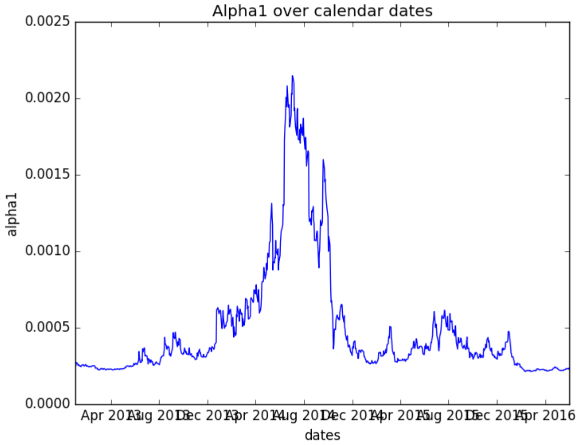
$$\Theta : \mathbb{R}^{200} \longrightarrow \mathbb{R}^{49}, \quad \begin{pmatrix} SWO_1 \\ SWO_2 \\ \dots \\ yield(0) \\ yield(1) \\ \dots \end{pmatrix} \mapsto \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \sigma_1 \\ \sigma_2 \\ \pi \\ \beta_1(0) \\ \beta_1(1) \\ \dots \end{pmatrix}$$

Bayesian model: training

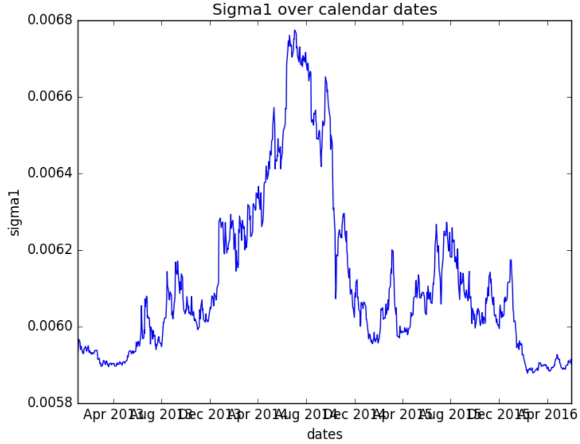
We generated a new training set and trained, tested another neural network with a similar architecture: the quality of the new calibration is the same as the QuantLib calibration and better than previous ML results, in particular out of sample.



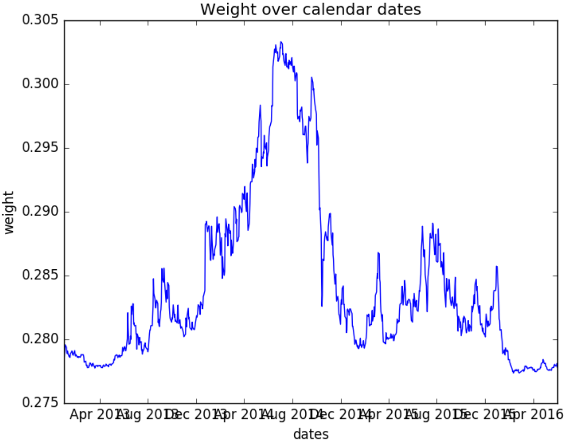
Mixture Model: α_1



Mixture Model: σ_1



Mixture Model: π



Conclusion

- Machine Learning for calibration of Bayesian models works, even where classical calibration would have difficulties.
- Improvements in parameter stability through a Bayesian model.
- Proof of concept that a combined Bayesian-updating, ML calibration approach is feasible and might lead to very stable modelling approaches.

References

- C. Cuchiero, I. Klein, and J. Teichmann:
  A fundamental theorem of asset pricing for continuous time large financial markets in a two filtration setting, Arxiv, 2017.
- A. Hernandez:
  Model Calibration with Neural Networks, SSRN, 2016.
- C. Cuchiero, A. Marr, M. Mavuso, N. Mitoulis, A. Singh, and J. Teichmann:
  Calibration with neural networks, report on the FMTC 2017, working paper, 2017.