Pairs Trading: Estimation Using the Kalman Filter

Heather E. Dempsey

www.hedempsey.com

FN912 Empirical Methods in Finance
Prof. W. Keener Hughen, Ph.D.

Outline

- Introduction
- Research Questions
- Literature Review
- Background
- Hypothesis
- Methodology
  - Data
  - Signals
  - Theoretical model
  - Kalman Filter
- Results
- Conclusion

Introduction

The purpose of this study is to determine whether profitable pairs trading strategies can be used against securities that don’t have the capitalization and average trading volume that are required by large hedge funds and banks. There is a niche in the market for the individual investor to profit.

Research Questions

- Best way to model highly-dynamic, noisy systems with inaccurate data to estimate an *unobservable* variable?
- *Kalman Filter* (1960), designed to track a moving target—airplanes, satellites, and park cars?
- Is *pairs trading* still profitable for the individual investor?

Literature Review

Methods commonly used to model cointegration between equities:

- Engle – Granger/Augmented Dickey Fuller (simple OLS regression - pair)
- Johansen – eigenvalues – (baskets of stocks)

State-Space Models –invented for the Apollo space missions (1961)

- Hidden Markov Model (Baum et al., 1960s)
    - Related to stochastic Markov Process (Markov, A. , 1906)
- Extended Kalman Filter (EKF) non-linear zero-mean, multivariate Gaussian noises–'Jacobean' partial derivatives matrix
- Unscented Kalman Filter
- Gordon (1993) Particle Filtering (Sequential Monte Carlo) – algorithm used in non-linear systems with non-Gaussian noise

Background – Pairs Trading

Hypothesis

$H_0: \neq H_1$

H_1: Pairs of cointegrated stocks which *depart from equilibrium* can be *profitable* for the *individual* investor when an appropriate position is taken before the pair mean-reverts.

Methodology – Data 1

Daily data

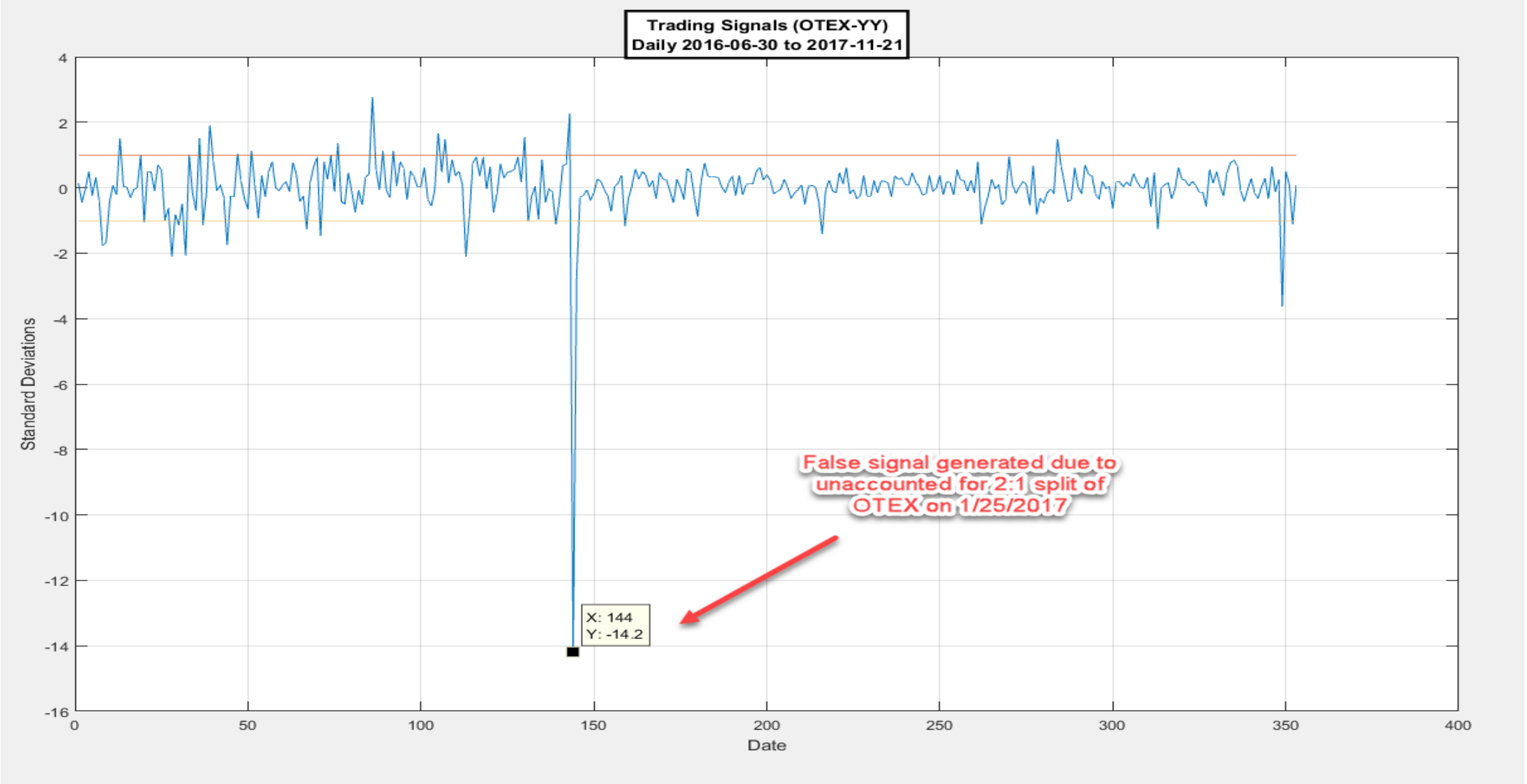
- 2016-06-30 to 2017-12-15

ETF/ETNs

- Average Daily Volume > 200K shares
- Minimum Price > $2
- *Adjusted* for splits and dividends
- After minimum liquidity measure met a total of 66,795 possible pairs were identified from the total of 366 screened ETF/ETNs

Data source: Commodity Systems Inc. (CSI) LLC

Methodology – Pairs Trading Signals 2



Methodology – Theoretical Model 3

Pairs Trading (statistical arbitrage)

$$Y_t = \beta_t * X_t + e_t$$

Y_t current price of first stock

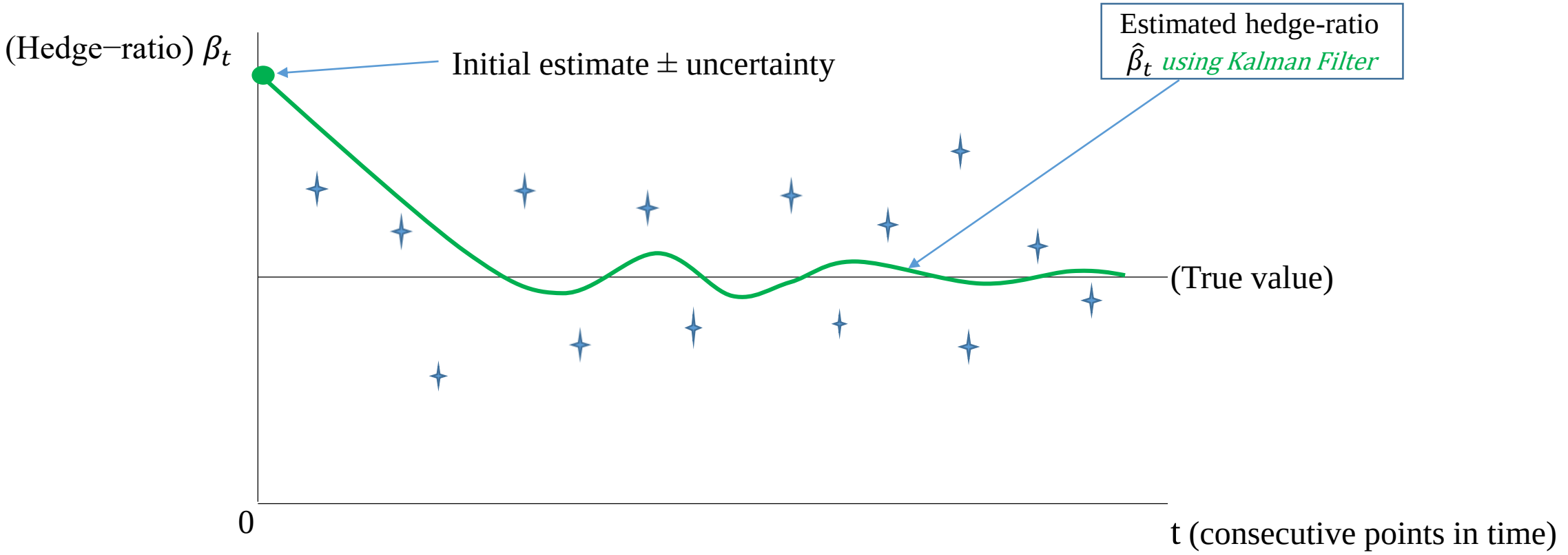
X_t current price of the second stock

β_t current hedge ratio

e_t current spread price

Methodology – Kalman Filter 4

- Needs only a few data points to “filter” out noise
- Estimates in real time as each successive datum arrives

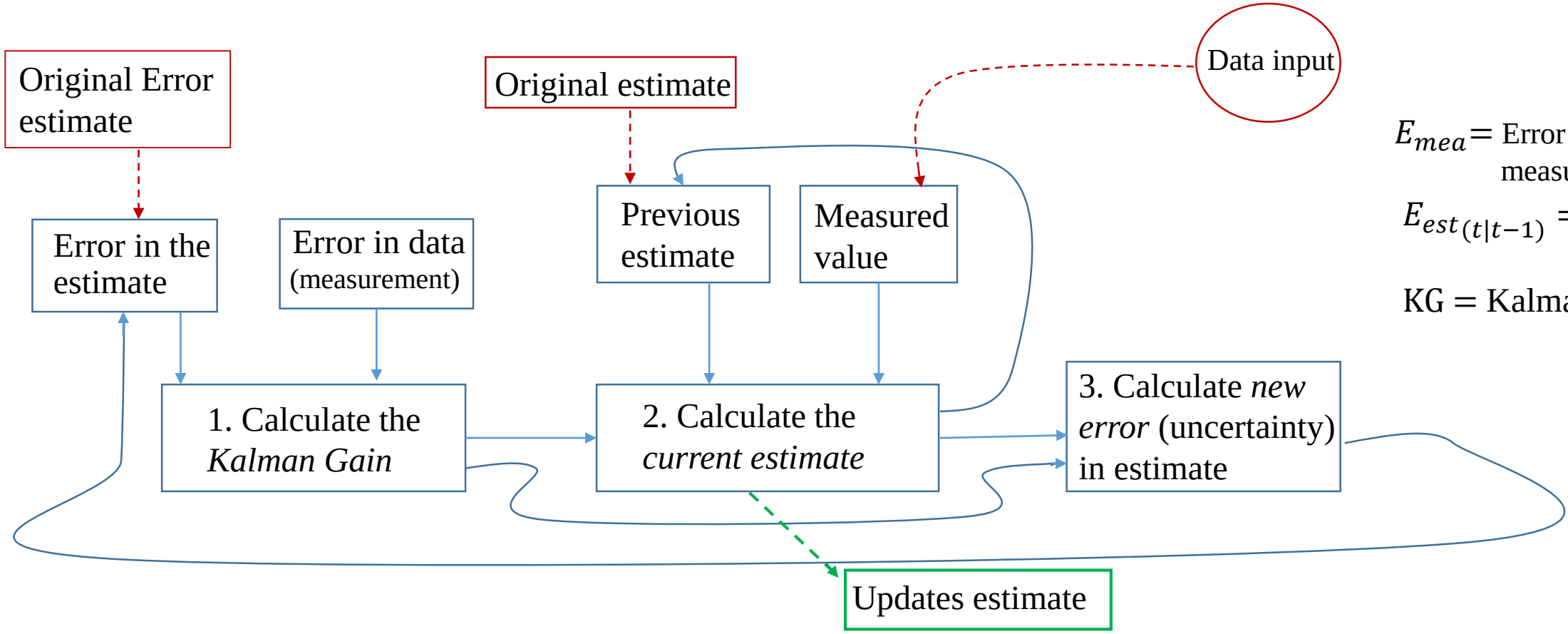


Van Biezen, M. (2015, October 27). Special Topics - The Kalman Filter (9 of 55) The Multi-Dimension Model 3: The State Matrix.

Methodology – Kalman Filter

- Uses two equations to estimate “hidden state”
- Observation equation $\mathbf{y}_t = \mathbf{x}_t \boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t$
- State-transition equation $\boldsymbol{\beta}_t = \mathbf{I}\boldsymbol{\beta}_{t-1} + \boldsymbol{\omega}_t$

Methodology – How Kalman Works



E_{mea} = Error in measurement

$E_{est_{(t|t-1)}}$ = Error in the estimate

KG = Kalman gain

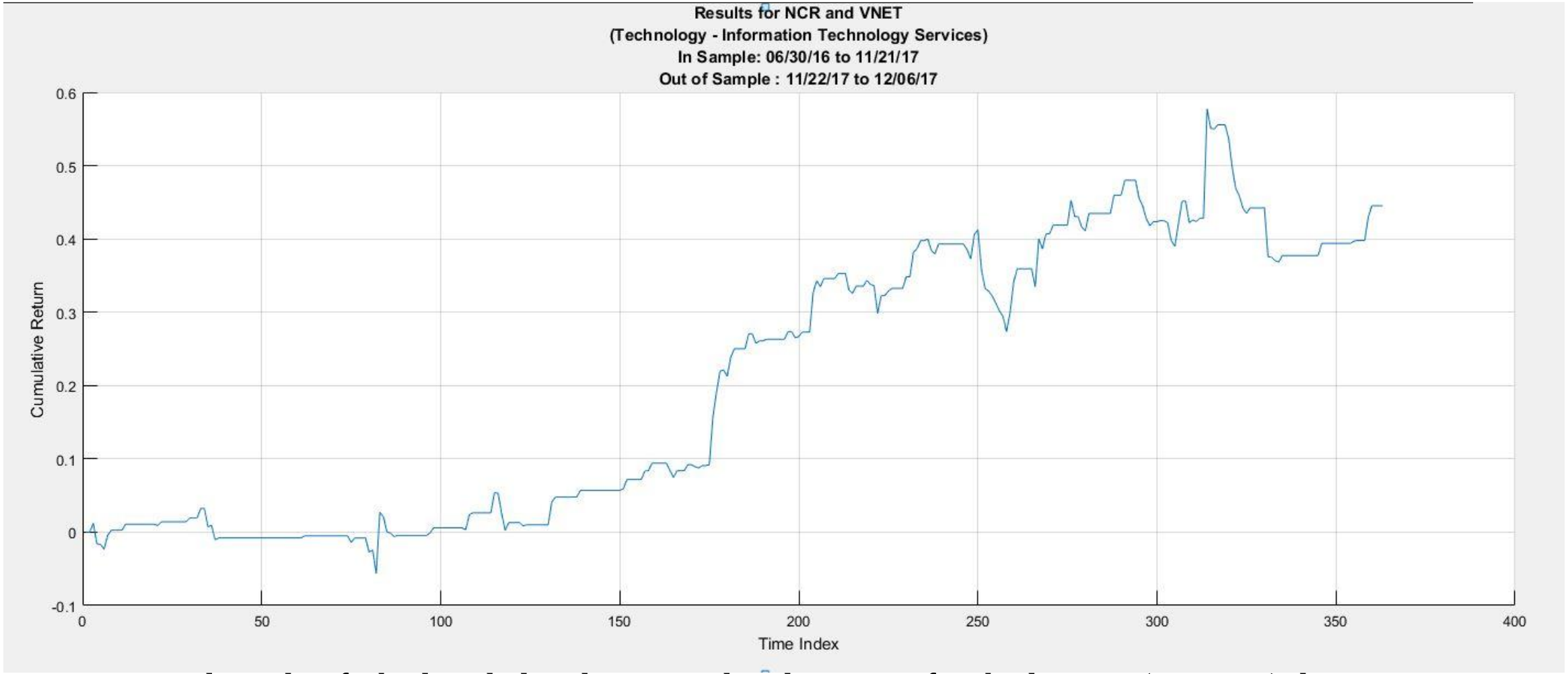
1. $\mathrm{KG} = \frac{E_{est}}{E_{est} + E_{mea}}$

2. $Est_t = Est_{t-1} + \mathrm{KG}[\mathrm{mea} - Est_{t-1}]$

3. $E_{est_t} = [1 - \mathrm{KG}]\, E_{est_{t-1}}$

Source: Van Biezen, M. (2015)

Results



- OOS actual results if I had traded and not simulated strategy for the last ten (353-363) days
- Total return In Sample 46% | APR 31% | Sharpe ratio 1.5 | MaxDD -9.8% | MaxDDD 80
- Transaction costs = 10 bps per trade including commissions, fees and slippage

Conclusion for this study

- Future recommendations-use of particle filters (sequential monte carlo) variable order Markov trees
- Dynamic asset allocation or pairs trading with use of improved Genetic algorithms and Neural Networks

References

De Moura, C. E., Pizzinga, A., & Zubelli, J. (2016). A Pairs Trading Strategy Based on Linear State Space Models and the Kalman Filter. *Quantitave Finance, 16*(10), 1559-1573. doi:10.1080/14697688.2016.1164886

Doucet, A., & Johansen, A. M. (2008). A Tutorial on Particle Filtering and Smoothing: Fifteen years later. Tokyo, Japan: Institute of Statistical Mathematics.Elliot, R. J., Van Der Hoek, J., & Malcolm, W. P. (2005). Pairs Trading. *Quantitative Finance*, 271-276. doi:10.1080/146976680500149370

Engle, R. F., & Granger, C. W. (1987, March). Cointegration and Error Correction: Representation, Estimation, and Testing. - *Econometrica, 55*(2), 251-276. Retrieved from http://www.jstor.org/stable/1913236

Johansen, S., & Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand For Money. *Oxford Bulletin of Economics and Statistics, 52*(2), 169-210.

PennState Elberly College of Science. (2017, December 8). State Space Models. Retrieved from Applied Time Series Analysis | STAT 510: https://onlinecourses.science.psu.edu/stat510/node/27

Van Biezen, M. (2015, October 27). Special Topics - The Kalman Filter (9 of 55) The Multi-Dimension Model 3: The State Matrix.

Vidyamurthy, G. (2004). Pairs Trading: Quantitative Methods and Analysis. Hoboken: John Wiley & Sons, Inc.