Multi-step-ahead prediction of volatility proxies

Jacopo De Stefani, Ir. - jdestefa@ulb.ac.be
Prof. Gianluca Bontempi - gbonte@ulb.ac.be
Olivier Caelen, PhD - olivier.caelen@worldline.com
Dalila Hattab, PhD - dalila.hattab@equensworldline.com

Benelearn 2017
Eindhoven University of Technology, Eindhoven, Nethelands

Friday 9th June, 2017

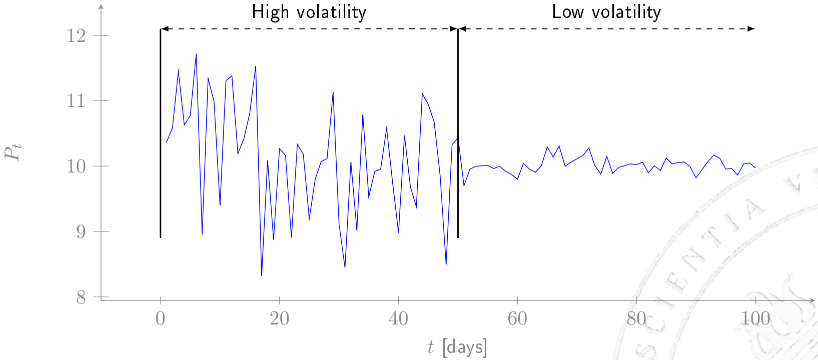
UNIVERSITÉ LIBRE DE BRUXELLES

Problem overview

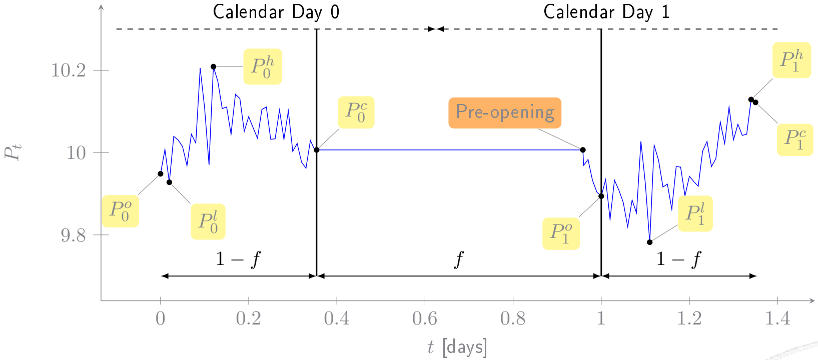
What is volatility?

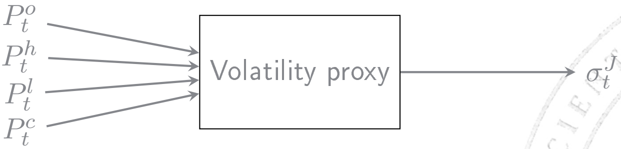
Definition

Volatility is a statistical measure of the dispersion of returns for a given security or market index.



A closer look on data





Time series forecasting - Taieb [2014]

Definition

Given a univariate time series $\{y_1, \cdots, y_T\}$ comprising T observations, forecast the next H observations $\{y_{T+1}, \cdots, y_{T+H}\}$ where H is the forecast horizon.

Hypotheses:

- Autoregressive model $y_t = f(y_{t-1}, \cdots, y_{t-d}) + \varepsilon_t$ with lag order d
- ε is a stochastic iid model with $\mu_\varepsilon = 0$ and $\sigma_\varepsilon^2 = \sigma^2$

Multistep ahead forecasting for volatility

State-of-the-art

$$
\begin{tikzcd}
{[\sigma^J_{t-d} \ \cdots \ \sigma^J_{t-1}]} \arrow[d] \\
\boxed{m(\sigma)} \arrow[d] \\
{[\hat{\sigma}^J_{t} \ \cdots \ \hat{\sigma}^J_{t+H}]}
\end{tikzcd}
$$

1 Input
1 Output

Multistep ahead forecasting for volatility

State-of-the-art

$$[\sigma^J_{t-d} \quad \cdots \quad \sigma^J_{t-1}]$$

$$\downarrow$$

$$\boxed{m(\sigma)}$$

$$\downarrow$$

$$[\hat{\sigma}^J_{t} \quad \cdots \quad \hat{\sigma}^J_{t+H}]$$

1 Input
1 Output

Proposed method

$$\begin{matrix} [\sigma^J_{t-d} & \cdots & \sigma^J_{t-1}] \\ [\cdots & \cdots & \cdots] \\ [\sigma^{X_M}_{t-d} & \cdots & \sigma^{X_M}_{t-1}] \end{matrix}$$

$$\downarrow$$

$$\boxed{m(\sigma)}$$

$$\downarrow$$

$$[\hat{\sigma}^J_{t} \quad \cdots \quad \hat{\sigma}^J_{t+H}]$$

$M+1$ inputs
1 output

Multistep ahead forecasting for volatility

State-of-the-art

$$[\sigma^J_{t-d} \quad \cdots \quad \sigma^J_{t-1}]$$

$$\downarrow$$

$$\boxed{m(\sigma)}$$

$$\downarrow$$

$$[\hat{\sigma}^J_{t} \quad \cdots \quad \hat{\sigma}^J_{t+H}]$$

1 Input
1 Output

Proposed method

$$\begin{matrix} [\sigma^J_{t-d} & \cdots & \sigma^J_{t-1}] \\ [\cdots & \cdots & \cdots] \\ [\sigma^{X_M}_{t-d} & \cdots & \sigma^{X_M}_{t-1}] \end{matrix}$$

$$\downarrow$$

$$\boxed{m(\sigma)}$$

$$\downarrow$$

$$[\hat{\sigma}^J_{t} \quad \cdots \quad \hat{\sigma}^J_{t+H}]$$

$M+1$ inputs
1 output

Future work

$$\begin{matrix} [\sigma^J_{t-d} & \cdots & \sigma^J_{t-1}] \\ [\cdots & \cdots & \cdots] \\ [\sigma^{X_M}_{t-d} & \cdots & \sigma^{X_M}_{t-1}] \end{matrix}$$

$$\downarrow$$

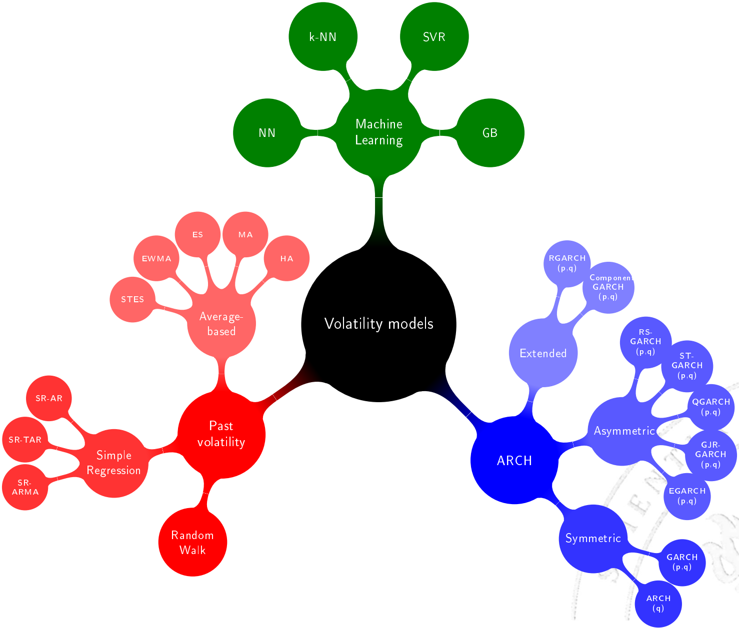
$$\boxed{m(\sigma)}$$

$$\downarrow$$

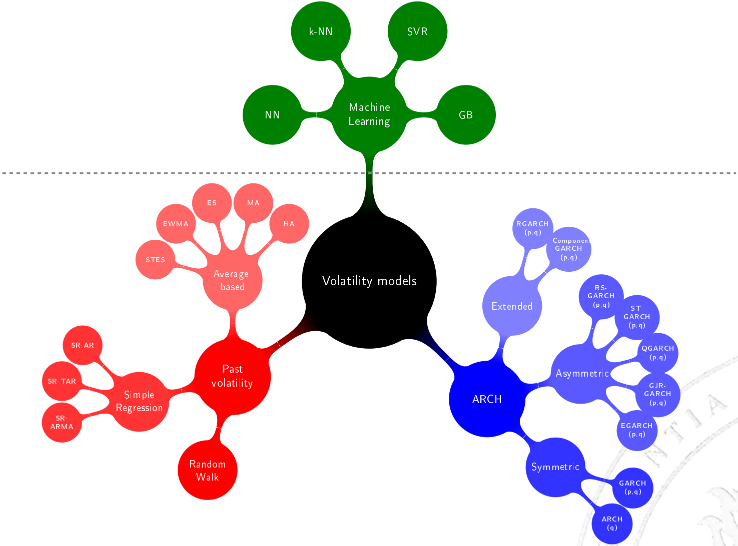
$$\begin{matrix} [\hat{\sigma}^J_{t} & \cdots & \hat{\sigma}^J_{t+H}] \\ [\cdots & \cdots & \cdots] \\ [\hat{\sigma}^{X_M}_{t} & \cdots & \hat{\sigma}^{X_M}_{t+H}] \end{matrix}$$

$M+1$ inputs
$M+1$ outputs

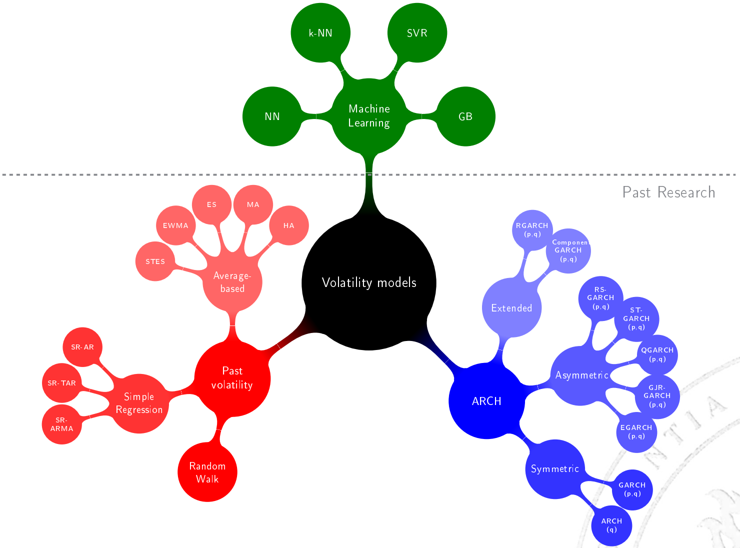
Models for volatility



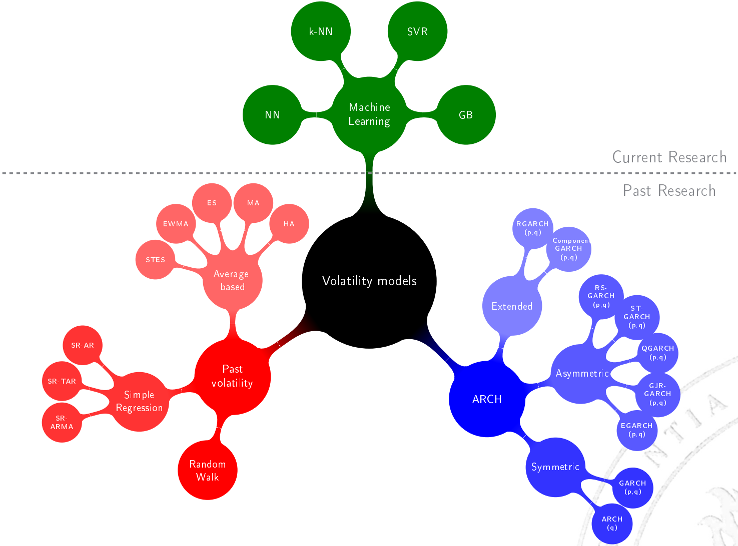
Models for volatility



Models for volatility



Models for volatility



Proposed model

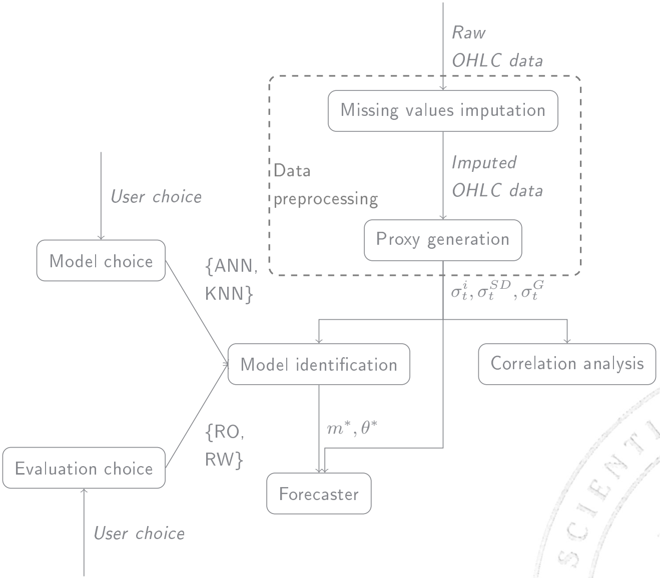
$$
\begin{array}{c}
[\sigma^J_{t-d} \quad \cdots \quad \sigma^J_{t-1}] \\
[\sigma^X_{t-d} \quad \cdots \quad \sigma^X_{t-1}] \\
\downarrow \\
\boxed{\quad m(\sigma) \quad} \\
\downarrow \\
[\hat{\sigma}^J_{t} \quad \cdots \quad \hat{\sigma}^J_{t+H}]
\end{array}
$$

2 TS Input
1 TS Output

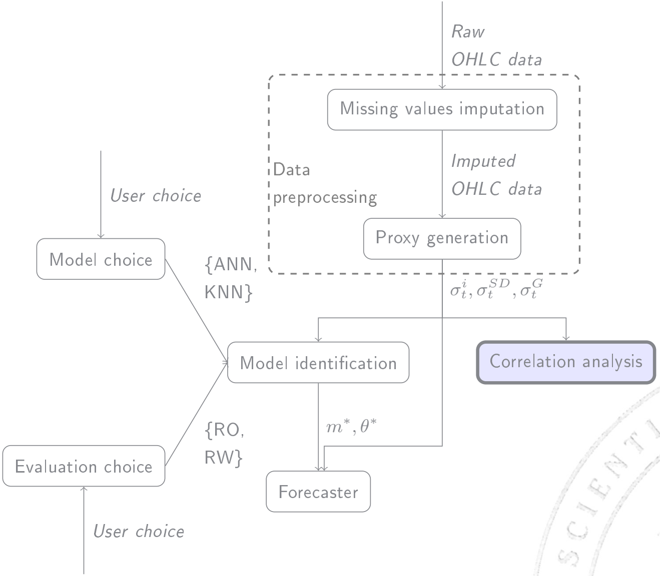
Volatility proxies σ^X, σ^J:

- $\sigma_{\mathbf{i}}$ **family** - Garman and Klass [1980]
- **GARCH (1,1) model** - Hansen and Lunde [2005]
- **Sample standard deviation**

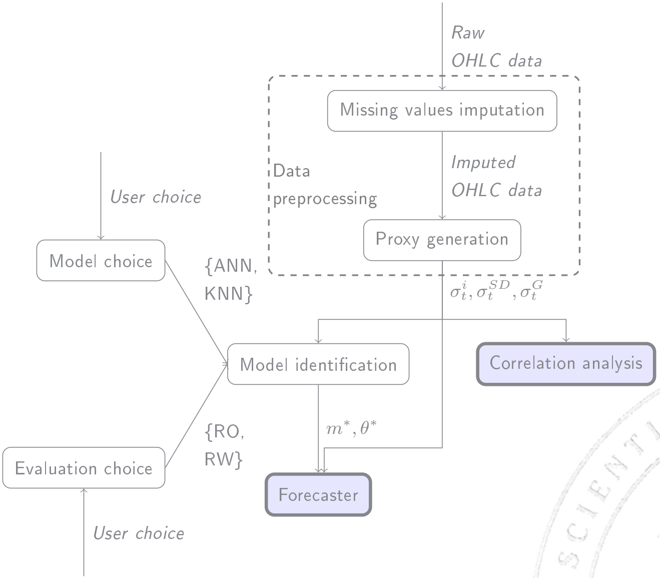
System overview



System overview

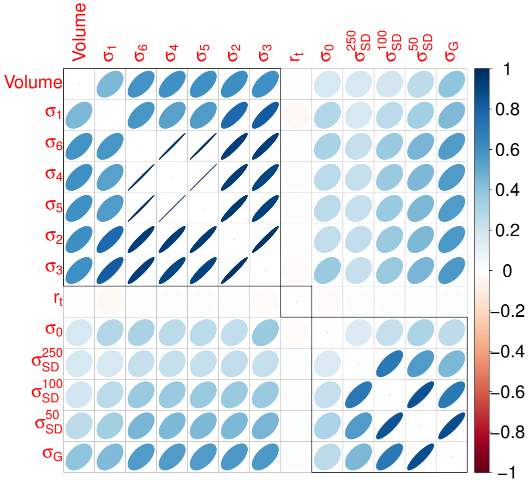


System overview



Correlation analysis - CAC40 Time series

Meta-analysis (cf. Field [2001]) across 40 time series (CAC40)



- Hierarchical clustering using Ward Jr [1963]
- Time range: 05-01-2009 to 22-10-2014
- 1489 OHLC samples per TS
- All the correlations are statistically significant

NARX forecaster - Results

Naive normalized MASE

σ^X	ANN	kNN	ANN_X	kNN_X	GARCH(1,1)
σ^6	0.07	0.08	0.06	0.11	1.34
$Volume$	0.07	0.08	0.07	0.14	1.34
$\sigma^{SD,5}$	0.07	0.08	0.07	0.09	1.34
$\sigma^{SD,15}$	0.07	0.08	0.06	0.10	1.34
$\sigma^{SD,21}$	0.07	0.08	0.06	0.10	1.34

Single CAC40 stock

- $\sigma_t^J = \sigma_t^G$
- 10-step ahead
- 10-fold CV
- 05-01-2009 $\Rightarrow$ 22-10-2014

Naive normalized MASE

σ^X	ANN	kNN	ANN_X	kNN_X	GARCH(1,1)
σ^6	0.58	0.49	0.53	0.56	1.15
$Volume$	0.58	0.49	0.57	0.66	1.15
$\sigma^{SD,5}$	0.58	0.49	0.58	0.58	1.15
$\sigma^{SD,15}$	0.58	0.49	0.65	0.65	1.15
$\sigma^{SD,21}$	0.58	0.49	0.56	0.65	1.15

S&P500 Index

- $\sigma_t^J = \sigma_t^G$
- 10-step ahead
- 10-fold CV
- 01-04-2012 to 30-07-2013 as in Dash and Dash [2016]

Conclusions

- ⚠ Preliminary results
- Correlation
  - Correlation clustering among proxies belonging to the same family, i.e. σ_t^i and $\sigma_t^{SD,n}$.
- Forecasting
  - Both machine learning methods outperform the benchmark methods (naive and GARCH).
  - ANN can take advantage of the additional information provided by the exogenous proxy better than k-NN
- Combination of proxies coming from different families could improve forecast accuracy
- We are currently assessing the performances of the models for different forecasting horizons h and model orders d.
- Inclusion of a greater number of input TS as a future research direction.

Thank you for your attention! Any questions/comments?

Find the paper at:

Bibliography I

References

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

Rajashree Dash and PK Dash. An evolutionary hybrid fuzzy computationally efficient egarch model for volatility prediction. *Applied Soft Computing*, 45:40–60, 2016.

Andy P Field. Meta-analysis of correlation coefficients: a monte carlo comparison of fixed-and random-effects methods. *Psychological methods*, 6(2):161, 2001.

Bibliography II

Mark B Garman and Michael J Klass. On the estimation of security price volatilities from historical data. *Journal of business*, pages 67–78, 1980.

Peter R Hansen and Asger Lunde. A forecast comparison of volatility models: does anything beat a garch (1, 1)? *Journal of applied econometrics*, 20(7):873–889, 2005.

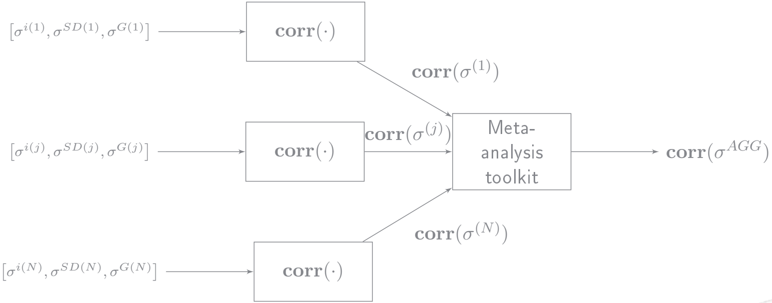
Rob J Hyndman and Anne B Koehler. Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4): 679–688, 2006.

Souhaib Ben Taieb. *Machine learning strategies for multi-step-ahead time series forecasting*. PhD thesis, Ph. D. Thesis, 2014.

Joe H Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58 (301):236–244, 1963.

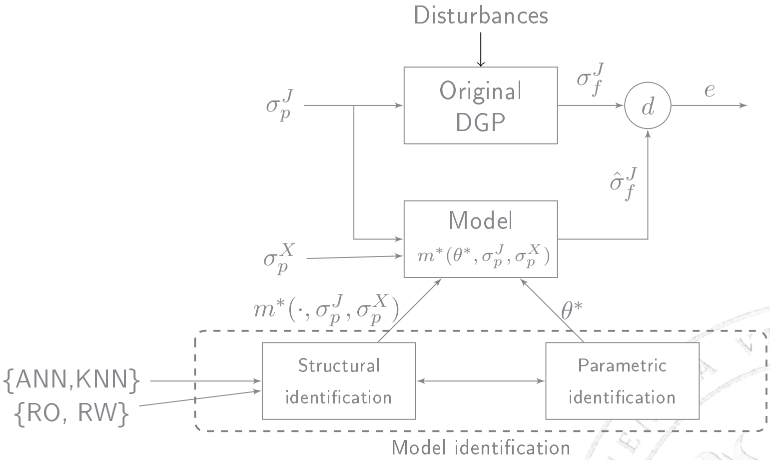
Appendix

Correlation analysis - Methodology



- 40 Time series (CAC40)
- Time range: 05-01-2009 to 22-10-2014 $\Rightarrow$ 1489 OHLC samples per TS

NARX forecaster - Methodology



Volatility proxies (1) - Garman and Klass [1980]

- Closing prices

$$\hat{\sigma}_0(t) = \left[\ln\left(\frac{P_{t+1}^{(c)}}{P_t^{(c)}}\right)\right]^2 = r_t^2 \tag{1}$$

- Opening/Closing prices

$$\hat{\sigma}_1(t) = \underbrace{\frac{1}{2f} \cdot \left[\ln\left(\frac{P_{t+1}^{(o)}}{P_t^{(c)}}\right)\right]^2}_{\text{Nightly volatility}} + \underbrace{\frac{1}{2(1-f)} \cdot \left[\ln\left(\frac{P_t^{(c)}}{P_t^{(o)}}\right)\right]^2}_{\text{Intraday volatility}} \tag{2}$$

- OHLC prices

$$\hat{\sigma}_2(t) = \frac{1}{2\ln 4} \cdot \left[\ln\left(\frac{P_t^{(h)}}{P_t^{(l)}}\right)\right]^2 \tag{3}$$

$$\hat{\sigma}_3(t) = \underbrace{\frac{a}{f} \cdot \left[\ln\left(\frac{P_{t+1}^{(o)}}{P_t^{(c)}}\right)\right]^2}_{\text{Nightly volatility}} + \underbrace{\frac{1-a}{1-f} \cdot \hat{\sigma}_2(t)}_{\text{Intraday volatility}} \tag{4}$$

Volatility proxies (2) - Garman and Klass [1980]

- OHLC prices

$$u = \ln\left(\frac{P_t^{(h)}}{P_t^{(o)}}\right) \qquad d = \ln\left(\frac{P_t^{(l)}}{P_t^{(o)}}\right) \qquad c = \ln\left(\frac{P_t^{(c)}}{P_t^{(o)}}\right) \tag{5}$$

$$\hat{\sigma}_4(t) = 0.511(u-d)^2 - 0.019[c(u+d) - 2ud] - 0.383c^2 \tag{6}$$

$$\hat{\sigma}_5(t) = 0.511(u-d)^2 - (2\ln 2 - 1)c^2 \tag{7}$$

$$\hat{\sigma}_6(t) = \underbrace{\frac{a}{f} \cdot \log\left(\frac{P_{t+1}^{(o)}}{P_t^{(c)}}\right)^2}_{\text{Nightly volatility}} + \underbrace{\frac{1-a}{1-f} \cdot \hat{\sigma}_4(t)}_{\text{Intraday volatility}} \tag{8}$$

Volatility proxies (3)

- **GARCH (1,1) model** - Hansen and Lunde [2005]

$$\sigma_t^G = \sqrt{\omega + \sum_{j=1}^{p} \beta_j (\sigma_{t-j}^G)^2 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2}$$

where $\varepsilon_{t-i} \sim \mathcal{N}(0,1)$, with the coefficients $\omega, \alpha_i, \beta_j$ fitted according to Bollerslev [1986].

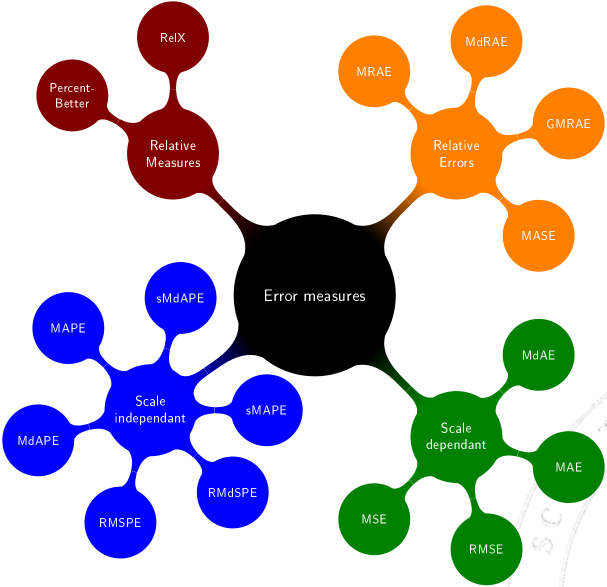
- **Sample standard deviation**

$$\sigma_t^{SD,n} = \sqrt{\frac{1}{n-1} \sum_{i=0}^{n-1} (r_{t-i} - \bar{r})^2}$$

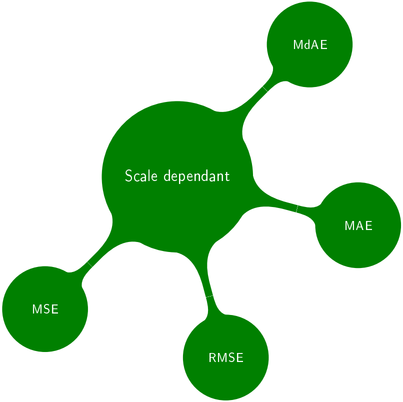
where

$$r_t = \ln\left(\frac{P_t^{(c)}}{P_{t-1}^{(c)}}\right) \qquad \bar{r_n} = \frac{1}{n} \sum_{j=t-n}^{t} r_j$$

Hyndman and Koehler [2006] - Error measures



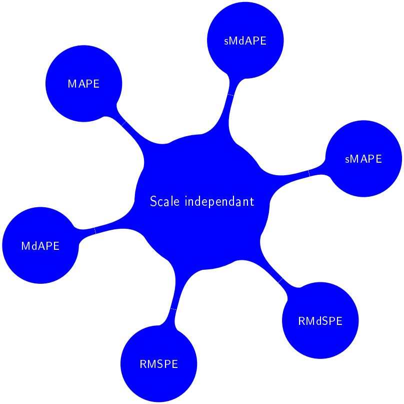
Hyndman and Koehler [2006] - Scale dependant



$$e_t = y_t - \hat{y}_t$$

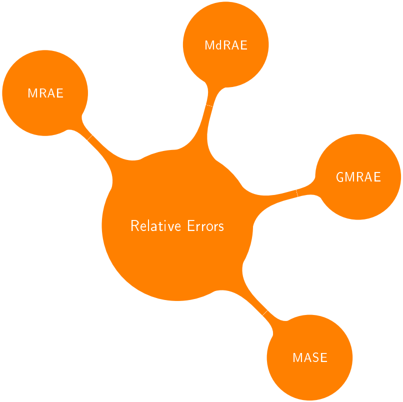
- **MSE** : $\frac{1}{n}\sum_{t=0}^{n}(y_t - \hat{y}_t)^2$
- **RMSE** : $\sqrt{\frac{1}{n}\sum_{t=0}^{n}(y_t - \hat{y}_t)^2}$
- **MAE** : $\frac{1}{n}\sum_{t=0}^{n}|y_t - \hat{y}_t|$
- **MdAE** : $Md_{t\in\{1\cdots n\}}(|y_t - \hat{y}_t|)$

Hyndman and Koehler [2006] - Scale independant



- **MAPE** : $\frac{1}{n}\sum_{t=0}^{n} \mid 100 \cdot \frac{y_t - \hat{y}_t}{y_t} \mid$
- **MdAPE** : $Md_{t\in\{1\cdots n\}}(\mid 100 \cdot \frac{y_t - \hat{y}_t}{y_t} \mid)$
- **RMSPE** : $\sqrt{\frac{1}{n}\sum_{t=0}^{n}(100 \cdot \frac{y_t - \hat{y}_t}{y_t})^2}$
- **RMdSPE** : $\sqrt{Md_{t\in\{1\cdots n\}}((100 \cdot \frac{y_t - \hat{y}_t}{y_t})^2)}$
- **sMAPE** : $\frac{1}{n}\sum_{t=0}^{n} 200 \cdot \frac{|y_t - \hat{y}_t|}{y_t + \hat{y}_t}$
- **sMdAPE** : $Md_{t\in\{1\cdots n\}}(200 \cdot \frac{|y_t - \hat{y}_t|}{y_t + \hat{y}_t})$

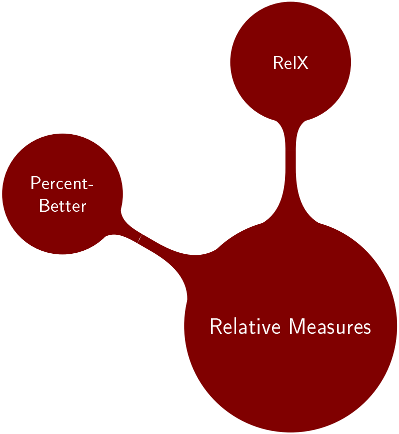
Hyndman and Koehler [2006] - Relative errors



$$r_t = \frac{e_t}{e_t^*}$$

- **MRAE** : $\frac{1}{n}\sum_{t=0}^{n} \mid r_t \mid$
- **MdRAE** : $Md_{t \in \{1 \cdots n\}}(\mid r_t \mid)$
- **GMRAE** : $\sqrt[n]{\frac{1}{n}\prod t = 0^n \mid r_t \mid}$
- **MASE** : $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{|e_t|}{\frac{1}{T-1}\sum_{i=2}^{T}|Y_i - Y_{i-1}|}\right)$

Hyndman and Koehler [2006] - Relative measures



- **RelX** : $\frac{X}{X_{\text{bench}}}$
- **Percent Better** : $PB(X) = 100 \cdot \frac{1}{n} \sum_{\text{forecasts}} I(X < X_b)$

where

- X: Error measure of the analyzed method
- X_b: Error measure of the benchmark