Eliza BUSZKOWSKA

Faculty of Mathematics and Computer Science
Adam Mickiewicz University, Poznań

COMPARING VOLATILITY FORECASTING MODELS AND THEIR COMBINATIONS FOR THE WIG20 INDEX USING THE MODEL CONFIDENCE SET METHOD

Introduction

According to Stock and Watson [2004] the combination of the models generates better forecast then the single model. A combination of forecasts is a good choice when it is not possible to distinguish one dominant model [Timmermann 2006]. Another argument for a combination is the lack of stationarity in data generating process, what can lead to the lack of stability in combination weights.

Searching for the best forecasting models for Polish WIG20 index among GARCH under the assumption of similar dynamic of the market comprises the subject of the article by Buszkowska [2008]. The point of the following paper is the comparison of the best volatility forecasting models for WIG20 index received in the paper Buszkowska [2008] with the optimal volatility forecasting models for various ARMA specification. The final purpose is the comparison of the previous results with the forecasts of optimal linear and nonlinear combinations of the best models. The aim is to verify if there exist the better prognostic model, when comparing mean squared error for WIG20 then received in Buszkowska [2008]. Then we investigate if optimal linear combination of forecasts in a sense of mean squared error, for WIG20 outperform the nonlinear one. We assign the optimal coefficients for linear and nonlinear combinations of the twoo forecasts solving the nonlinear least squares problem and presuming the patterns introduced by Timmermann [2006] for the linear case. We compare the volatility forecasts with daily realized volatility, calculated as the sum of the squared intraday returns. We investigate the results obtained with the Model Confidence Set (MCS) method of

Hansen et al [2003] for different measures of the realized volatility. Agreebly with the the conclusion of the paper Buszkowska [2008] we assume one 5 minute frequency if intraday quotations . We compare MCS sets for MSE known as the robust function of error [Patton, Sheppard 2007).

1. The specyfication of the conditional volatility models

In the article we consider the various types of GARCH models. The choice results from the fact that GARCH are the most popular in aplications volatility models of financial instruments by reason of the simple construction, the easy estimation and the natural interpretations. We use the following GARCH specifications

- GARCH(p, q)

$$y_t = \sigma_t \varepsilon_t,$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

$$\varepsilon_t \sim idd(0,1), \quad \omega > 0, \quad \beta_j \geq 0, \quad \alpha_i \geq 0.$$

- RISCMETRICS

RISCMETRICS is a GARCH model where the ARCH and GARCH coefficients are fixed.

$$\sigma_t^2 = \omega + (1-\lambda)\, y_{t-1}^2 + \lambda \sigma_{t-1}^2.$$

In practice, one assumes that $\omega = 0$ and $\lambda = 0.94$ for daily data and $\lambda = 0.97$ for weekly data.

We establish different distributions of error: Gauss, Student-t, skewed-Student-t and GED.

- EGARCH

EGARCH is the first model of a type GARCH, which describes the effect of asymmetry, proposed by Nelson in 1991, but modified by Bollerslev and Mikkelsen in 1996 to the following form.

Let ε_t be an independently and identically distributed process with $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) = 1$. Define

$$g(\varepsilon_t) \equiv \gamma_1 \varepsilon_t + \gamma_2 \left[|\varepsilon_t| - E|\varepsilon_t| \right],$$

then

$$\log \sigma_t^2 = \omega + [1-\beta(L)]^{-1}[1-\alpha(L)]g(\varepsilon_{t-1}),$$

where

$$\alpha(L) = 1-\alpha_1 L- \ldots -\alpha_q L^q,$$

$$\beta(L) = \beta_1 L+\beta_2 L^2+ \ldots +\beta_p L^p.$$

For the Normal distribution:

$$E([\varepsilon_t]) = \sqrt{\frac{2}{\pi}}.$$

For the skewed Student-t distribution:

$$E(|\varepsilon_t|) = \frac{4\xi^2}{\xi+\frac{1}{\xi}} \cdot \frac{\Gamma\left(\frac{1+v}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\Gamma\left(\frac{v}{2}\right)},$$

where $\xi = 1$ for the symmetric Student distribution. The parameter v is called the number of degrees of freedom.

For the GED distribution

$$E(|\varepsilon_t|) = 2^{\left(\frac{1}{v}\right)}\lambda_v \frac{\Gamma\left(\frac{2}{v}\right)}{\Gamma\left(\frac{1}{v}\right)},$$

where $0 < v < \infty, \lambda_v \equiv \sqrt{\frac{\Gamma\left(\frac{1}{v}\right)2^{\left(-\frac{2}{v}\right)}}{\Gamma\left(\frac{3}{v}\right)}}.$

- GJR

GJR is the model by Glosten, Jagannathan and Runkle from 1993 defined by

$$\sigma_t^2 = \omega + \sum_{i=1}^{q}\left(\alpha_i y_{t-i}^2 + \gamma_i S_{t-1}^{-} y_{t-i}^2\right) + \sum_{j=1}^{p}\beta_j \sigma_{t-j}^2,$$

where S_t^-, is the auxiliary variable, which is one when the value of the y_t is negative and 0 when it is positive. It is assumed that the impact of ε_t^2 on the conditional variance σ_t^2 is different when ε_t is positive or negative.

2. Forecasts combinations

The combination is the good alternative when it is impossible to identify one predominant model [Timmermann 2006]. Combinations of forecasts are more stable then individual forecasts [Stock, Watson 2004].

The simplest combination is linear with the identical coefficients and the sum of the weights equels one.

$$g\left(\hat{y}_{t+h};\omega_{t+h,t}\right)=\frac{1}{N}\sum_{j=1}^{N}\hat{y}_{t+h,t,j},$$

where $\hat{y}_{t+h,t}$ is the forecast, and $\omega_{t+h,t}$ is the weight.

The forecast error is defined by

$$e_{t+h,t}^{c}=y_{t+h}-g\left(\hat{y}_{t+h,t};\omega_{t+h,t}\right),$$

y_{t+h} is a realization of some variable.

The parametrs of the optimal combinations of the forecasts in this case are the solution of the following problem

$$\omega^{*}=\arg\min_{\omega\in W_t}E\left[L\left(e^{c}\left(\omega\right)\right)\right]. \tag{1}$$

where L denotes mean squared error (MSE) loss.

Under MSE the combination weights only depend on the first two moments of the joint distribution of y_{t+h} and $\hat{y}_{t+h,t}$

$$\begin{bmatrix} y_{t+h} \\ \hat{y}_{t+h} \end{bmatrix}\sim\begin{bmatrix} \mu_{y_{t+h,t}} \\ \mu_{\hat{y}_{t+h,t}} \end{bmatrix}\begin{bmatrix} \sigma^{2}_{y_{t+h,t}} & \sigma'_{y\hat{y}_{t+h,t}} \\ \sigma_{y\hat{y}_{t+h,t}} & \Sigma_{\hat{y}\hat{y}_{t+h,t}} \end{bmatrix}.$$

For MSE Timmermann [2006] obtained the following optimal weights:

$$\omega_0^{*}=\mu_{y_{t+h,t}}-\omega^{*\prime}\mu_{\hat{y}_{t+h,t}},\qquad \omega^{*}=\Sigma^{-1}_{\hat{y}\hat{y}_{t+h}}\sigma_{y\hat{y}_{t+h,t}}. \tag{2}$$

The combination which doesn't contain the correlation between forecasts outperform more sophisticated schemes. The assumption that the individual forecasts are unbiased implies unbiased forecast of the combination under condition that the weight's sum is one and the constant is correct [Timmermann 2006].

Consider the combination of two forecasts $\hat{y}_1$, $\hat{y}_2$. Let e_1 i e_2 denote the forecast errors. Assume $e_1 \sim \left(0, \sigma_1^2\right)$, $e_2 \sim \left(0, \sigma_2^2\right)$, where σ_1^2 = Var (e_1), σ_2^2 = Var (e_2) and $\sigma_{12} = \rho_{12}\sigma_1\sigma_2$ is the covariance between e_1 and e_2 and ρ_{12} is their correlation.

The optimal weights for this combination by Timmermann (2005) have the form

$$\omega^* = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}, \qquad 1 - \omega^* = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}. \tag{3}$$

The identical weights are optimal if the forecast variances are the same independently of the correlation between forecasts on condition that the forecasts are unbiased [Timmermann 2006]. The natural example is the following scheme of two forecasts:

$$\left(\frac{1}{2}\right) \cdot \left(\hat{y}_1 + \hat{y}_2\right) \tag{4}$$

When the forecast are unbiased Timmermann [2006] propose the combination that gives the inverse weights to the forecasts with the assumption that the correlation is zero:

$$\omega_{inv} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}, \qquad 1 - \omega_{inv} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}. \tag{5}$$

For N forecasts one can assume $0 \leq \omega_{ni} \leq 1$, $i = 1, \ldots, N$ to make the values of the combination forecasts be in the interval of values of the individual forecasts.

Let

$$\hat{y}_c = \omega y_1 + (1-\omega)\hat{y}_2, \quad y - \hat{y}_1 = e_1 \sim N(0, \sigma^2), \quad y - \hat{y}_2 = e_2 \sim \left(\mu_2, \sigma^2\right),$$

so $\hat{y}_2$ is the biased forecast and assume cov $(e_1, e_2) = \sigma_{12} = \rho_{12}\sigma^2$.

Timmermann obtained

$$MSE(\hat{y}_c) - MSE(\hat{y}_1) = (1-\omega)\sigma^2\left((1-\omega)\left(\frac{\mu_2}{\sigma}\right)^2 - 2\omega(1-\rho_{12})\right).$$

So if

$$\left(\frac{\mu_2}{\sigma}\right)^2 > \frac{2\omega(1-\rho_{12})}{1-\omega^2}, \quad \text{then} \quad MSE(\hat{y}_c) > MSE(\hat{y}_1).$$

The condition allways holds for $\rho_{12} = 1$. In this case the forecast of the combination of models doesn't outperform the unbiased forcest of the simple model. What is more the bigger is the bias of the forecast the smaller is the advantage of the combination. If the forecasts are biased then identical weights are optimal when the forecast errors have the same variance and identical correlation between forecasts [Timmermann 2006].

The optimal weights problem may be formulated as the optimalization task of minimalization of expected forecast error variance $\Sigma_e = E[ee']$ where $e = t\hat{y} - y$ with the condition that the sum of weights is one and the individual forecasts are unbiased:

$$\min \omega' \Sigma_e \omega,$$
$$\omega' t = 1,$$

where t is the vector of ones.

For the invertible covariance matrix Σ_e Timmerman [2005] obtains the following optimal weights:

$$\omega^* = \left(l' \Sigma_e^{-1} l\right)^{-1} \Sigma_e^{-1} l. \quad (6)$$

The problem of the optimal combination can be solved as the following test

$$\mathrm{H}_0 : E\left[L\left(\hat{\sigma}_t^2, h_t^A\right)\right] = E\left[L\left(\hat{\sigma}_t^2, f\left(h_t^A, h_t^B, \theta\right)\right)\right],$$
$$\mathrm{H}^A : E\left[L\left(\hat{\sigma}_t^2, h_t^A\right)\right] > E\left[L\left(\hat{\sigma}_t^2, f\left(h_t^A, h_t^B, \theta\right)\right)\right].$$

The test statistic of Diebold-Mariano and West (*DMW*) can be used in the test. Let define the difference

$$d_t = L\left(\hat{\sigma}_t^2, h_t^A\right) - L\left(\hat{\sigma}_t^2, f\left(h_t^A, h_t^B, \theta\right)\right).$$

Then the *DMW* test statistic is the following:

$$DMW_T = \frac{\sqrt{T}\,\bar{d}_T}{\sqrt{a\hat{\mathrm{V}}\mathrm{ar}\left(\sqrt{T}\,\bar{d}_T\right)}},$$

where

$$d_T \equiv \frac{1}{T}\sum_{t=1}^{T} d_t.$$

Under the null hypothesis the test statistic has normal distribution.

If

$$\sigma(y-\hat{y}_1) > \sigma(y-\hat{y}_2),$$

$$\operatorname{cov}(y-\hat{y}_1, y-\hat{y}_2) \neq \sigma(y-\hat{y}_2)\sigma(y-\hat{y}_1),$$

the optima model is the combination of forecasts, Timmermann [2006].

Another scheme can be created on the base of the ranking of models by Aiolfi and Timmermann [2006]. Let R_i be the position of the i-model in ranking. The weighs od the combination are the following:

$$\hat{\omega} = \frac{R_i^{-1}}{\left(\sum_{i=1}^{N} R_i^{-1}\right)}. \tag{7}$$

The combination doesn't allow for correlations between forecasts. It is insusceptible on exstremal values.

There may occur the two types of nonlinear combinations. The nonlinear function of the forecasts and the nonlinear function of weights.

The nonlinear function of error:

$$\hat{y}^c = \omega_0 + \omega' C(\hat{y})$$

and the nonlinear combinations

$$\hat{y}^c = C(\hat{y}, \omega).$$

Kamsatra [1996] proposes the nonlinear combination with nonlinear weights with the model of logistic function, which distinguish extermal values of the forecasts. The changes of the value of the logistic function are minimal when the values of variables are smaller then the fixed value and the function increases to one when the variables surpass the value.

$$\hat{y}^c = \beta_0 + \sum_{j=1}^{N} \beta_j y_j + \sum_{i=1}^{p} \delta_i g(z_i, \gamma_i), \tag{8}$$

$$g(z_i, \gamma_i) = \left(1 + \exp\left(-1 \cdot \left(\gamma_{0i} + \sum_{j=1}^{N} \gamma_{1,j} z_j\right)\right)\right)^{-1},$$

$$z_j = \frac{\hat{y}_j - \bar{y}}{\hat{\sigma}_y}, \quad p \in \{0, 1, 2, 3\},$$

where

$\bar{y}$ - the estymator of the mean value of the forecast.

$\hat{\sigma}$ - the estimator of the standard deviation of the forecast.

Donaldson and Kamsatra proved that this model applied to twoo forecasting models: moving average variance model and GARCH(1, 1) surpass other traditional combination schemes. For $\hat{\varepsilon}_t = R_t - \hat{\rho}_0 - \hat{\rho}_1 R_{t-1}$ the moving average variance model is

$$MAV_t = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \hat{\varepsilon}_{t-i}^2 .$$

The optimal coefficients of the nonlinear combination can not be assign analitically from the equation (1). The nonlinear least square estimation is needed. The same solution is in the case of the geometric mean of the forecasts, which allow to avoid a problem of the negative volatilities, preposed by Patton and Sheppard [2007]. The formula is the following:

$$y^c = \exp(\omega_1 \ln(\hat{y}_1) + \omega_2 \ln(\hat{y}_2)).$$

To assign the weights the Gauss-Newton algorithm may be used which gives the solution to the problem (1).

3. The realized volatility

The realized volatility can be calculated by summing the squares of intraday returns. With the use of the equation which allow for the night return it is defined as follow.

$$\sigma_{2,t}^2 = \sum_{i=0}^{N} r_{t,i}^2 , \tag{9}$$

where the intraday return in the day n and in the moment d is :

$$r_{n,d} = 100(\ln P_{n,d} - \ln P_{n,d-1}), \quad r_{n,0} = 100(\ln P_{n,1} - \ln P_{n-1,N}),$$

N is the numer of periods in a day.

The alternative approach was proposed by Andersen and Bollerslev in 1997. They suggested reprezenting the daily volatility as the sum of intraday returns

$$\sigma_{1,t}^2 = \sum_{i=1}^{N} r_{t,i}^2 . \tag{10}$$

They suggest multiplying $\sigma_{1,t}^2$ by $(1 + c)$, where c is the positive constant [Martens 2002]. They choose $\left(\sigma_{co}^2 + \sigma_{oc}^2\right)/\sigma_{oc}^2$ as the constant c, where $\sigma_{co}^2 = \mathrm{Var}(r_{t,0})$ and $\sigma_{oc}^2 = \mathrm{Var}\left(\sum_{t=1}^{N} r_{t,n}\right)$ [Koopman et al 2005]. Then the realized volatility can be expressed:

$$\sigma_{3,t}^2 = \frac{\sigma_{oc}^2 + \sigma_{co}^2}{\sigma_{oc}^2} \sum_{t=1}^{N} r_{t,i}^2 . \tag{11}$$

In the article MSE means the mean squared error, where N is the number of forecasts.

$$\mathrm{MSE} = N^{-1} \sum_{t=1}^{N} \left(\sigma_{l,t}^2 - \hat{\sigma}_{k,t}^2\right)^2,$$

where $l \in \{1, 2, 3\}$, $k \in \{1, \ldots, m\}$ is the numer of models from the considered set. In the following formula $\hat{\sigma}_{k,t}^2$ is the forecast of volatility from the model k on the moment t, $\sigma_{l,t}^2$ is the value of the realized volatility of the type l in the moment t.

4. The Model Confidence Set (MCS)

The MCS procedure consists of the test and the elimination rule. One check if the null hypothesis is true (if the loss functions for the models are the same to the expected value).

Let M_0 be a set of forecasting models, $\{i = 1, \ldots, m_0\}$. The objects from the set are evaluated out of sample in terms of the loss function $L_{i,t}$. $L_{i,t}$ – denotes the loss that is associated with the object i in the period t. One defines the relative performance of the models by

$$d_{ij,t} \equiv L_{i,t} - L_{j,t} \quad \text{for} \quad j, m \in M.$$

The null hypothesis is:

$$\mathrm{H}_{0,M} : E\,(d_{ij,t}) = 0, \qquad t = 1, \ldots, n.$$

The procedure is being repeated until the null hypothesis is accepted. The MCS is the set of models after the elimination. We denote it $\hat{M}^{*}_{1-\alpha}$, where α is the assumed level of confidence.

The set of superior objects is defined as:

$$M^{*} = \left\{ i \in M_0 : E\left(d_{ij,t}\right) \leq 0 \text{ , for all } j \in M_0 \right\}.$$

The set of inferior objects for $j \in M_0$ is defined as:

$$M^{+} \equiv \left\{ i \in M_0 : E\left(d_{ij,t}\right) > 0 \text{ , for some } j \in M_0 \right\}.$$

Let M be a set of models being reduced in the process of elimination. We use the following algorthm:

Step 1: $M = M_0$.

Step 2 : If $H_{0,M}$ is accepted define $\hat{M}^{*}_{1-\alpha} = M$.

Otherwise define the loss of the model i to the average loss from models from M

$$\bar{d}_{i.} = \frac{1}{m} \sum_{j \in M} \bar{d}_{ij} ,$$

and the worst model

$$i^{+} = \arg\max_{i \in M} \frac{\bar{d}_{i\cdot}}{\sqrt{\hat{\mathrm{V}}\mathrm{ar}\left(\bar{d}_{i\cdot}\right)}} .$$

Then eliminate i^{+} from M and repeat the procedure beginning with Step 2.

5. Data

We consider the models indicated as the best under the assumption of the similar dynamic of the market in the article by Buszkowska [2008]. There are

(1) GARCH(1, 1) with Gaussian distribution of error, $\omega = 0.015(0.0066)$, $\alpha_1 = 0.403(0.0067)$, $\beta_1 = 0.953(0.0078)$.

(2) RiskMetrics with Student distribution of error, Student DF = 9.4884(2.0357) and $\lambda = 0.94$.

(3) RiskMetrics with GED for $\lambda = 0.94$ and GED = 1.4367(0.0728).

(4) RiskMetrics with skewed–Student-t distribution of error for $\lambda = 0.94$, Student DF = 9.4524 (2.0108), Asymmetry = 0.0432(0.0357), Tail = 9.4527(2.0108).

(5) RiskMetrics with Gaussian distribution of error, for $\lambda = 0.94$
and 446 GARCH-type models estimated with different types of ARMA:
ARMA(0, 0)-GARCH(1, 1), ARMA(1, 0)-GARCH(1, 1),
ARMA(0, 1)-GARCH(1, 1), ARMA(1, 1)-GARCH, ARMA(1, 2)-GARCH(1, 1),
ARMA(2, 1)-GARCH(1, 1), ARMA(2, 2)-GARCH(1, 1)
with Gauss, Student-*t*, Skewed-Stident-*t* and GED since the type of ARMA model affects the volatility forecasts from GARCH. We rejected the models with not significant parameters. In the empirical investigation we use 1739 daily observations of the WIG20 index, from October 12, 2000 till September 14, 2007 for model estimation.

Table 1. Descriptive statistics for the return series (October 12, 2000 till September 14, 2007)

Max	Min	Mean	St. Deviation	Skewness	Kurtosis
5.4829	-6.4418	0.04833	1.4864	0.03617	4.0927

The next 265 data from August 28, 2006 till September 14, 2007 were exploited for calculating volatility forecasts. To evaluate the quality of our forecasts we compared them with the daily realized volatility calculated for 5 -minute intraday returns. The realized volatility was calculated using formulas (9), (10) and (11).

For GARCH-type models forecasts as examples from the paper by Buszkowska [2008], we consider:
(1) For 5 GARCH-type models

$$y_1^c = \left(\frac{1}{5}\right)\hat{y}_1 + \left(\frac{1}{5}\right)\hat{y}_2 + \left(\frac{1}{5}\right)\hat{y}_3 + \left(\frac{1}{5}\right)\hat{y}_4 + \left(\frac{1}{5}\right)\hat{y}_5.$$

The identical weights may be optimal as the investigated forecast erros have with the accuracy 0.01 the same variance and the same correlation between forecasts.
(2) The kombination on the base of the MCS according to the formula (7), that doesn't describe the correlation between forecasts. Ranking by using MCS.

$$\text{for } \sigma_{1,t}^2,\ (10),\ y_2^c = 0.4379\,\hat{y}_1 + 0.219\,\hat{y}_3 + 0.1460\,\hat{y}_2 + 0.1095\,\hat{y}_5 + 0.0876\,\hat{y}_4,$$

$$\text{for } \sigma_{2,t}^2,\ (9),\ y_3^c = 0.4379\,\hat{y}_1 + 0.219\,\hat{y}_4 + 0.1460\,\hat{y}_5 + 0.1095\,\hat{y}_2 + 0.0876\,\hat{y}_3,$$

$$\text{for } \sigma_{3,t}^2,\ (11),\ y_4^c = 0.4379\,\hat{y}_1 + 0.219\,\hat{y}_3 + 0.1460\,\hat{y}_2 + 0.1095\,\hat{y}_5 + 0.0876\,\hat{y}_4,$$

$$\omega_1 = \frac{1}{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}} = 0.4379,$$

$$\omega_2 = \frac{1}{2\cdot\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right)} = 0.219,$$

$$\omega_3 = \frac{1}{3\cdot\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right)} = 0.1460,$$

$$\omega_4 = \frac{1}{4\cdot\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right)} = 0.1095,$$

$$\omega_5 = \frac{1}{5\cdot\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right)} = 0.0876.$$

The sum of the coefficients is 1.

(3) The combination of the two best forecasting models according to (2), since the forecasts are biased.

for $\sigma^2_{2,t}$, (9), $y^c_5 = -0.6953 + 0.5158\,\hat{y}_1 + 0.7287\,\hat{y}_4$,

for $\sigma^2_{1,t}$, (10), $y^c_7 = -0.5841 + 0.4173\,\hat{y}_1 + 0.5231\,\hat{y}_3$,

for $\sigma^2_{3,t}$, (11), $y^c_8 = -0.5921 + 0.4204\,\hat{y}_1 + 0.5266\,\hat{y}_3$.

(4) The nonlinear combinations for the optimal proportions , according to Patton and Sheppard, [2006], assigned by using Gauss-Newton method.

for $\sigma^2_{1,t}$, (10), $y^c_9 = \exp\left(-1.073\cdot\ln\left(\hat{y}_1\right) + 1.538\cdot\ln\left(\hat{y}_3\right)\right)$,

$y^c_{10} = \exp\left(-1.069\cdot\ln\left(\hat{y}_1\right) + 1.538\cdot\ln\left(\hat{y}_2\right)\right)$,

for $\sigma^2_{2,t}$, (9), $y^c_{11} = \exp\left(0.256\cdot\ln\left(\hat{y}_1\right) + 0.66\cdot\ln\left(\hat{y}_4\right)\right)$,

for $\sigma^2_{3,t}$, (11), $y^c_{12} = \exp\left(-1.07\cdot\ln\left(\hat{y}_1\right) + 1.54\cdot\ln\left(\hat{y}_3\right)\right)$,

$y^c_{13} = \exp\left(-1.069\cdot\ln\left(\hat{y}_1\right) + 1.538\cdot\ln\left(\hat{y}_2\right)\right)$.

6. Empirical results

First, in the investigation we analized the selected 5 GARCH type models received as the best forecasting models, under the assumption of the similar dynamic of the market, in the paper by Buszkowska (2008). We consider also the combinations from the section 3,

$$y_1^C, y_2^C, y_3^C, y_4^C, y_5^C, y_6^C, y_7^C, y_8^C, y_9^C, y_{10}^C, y_{11}^C, y_{12}^C.$$

For 5 min frequency of quotations and MSE we received the following results:

Table 2. The estimates of the *p*-values MCS for the realized volatility $\sigma_{1,t}^2$, 5 min frequency and MSE

The MSE	*p*-value
$\exp(1.538 \cdot \ln(\hat{y}_3) - 1.073 \cdot \ln(\hat{y}_1))$	0.7624
$\exp(-1.069 \cdot \ln(\hat{y}_1) + 1.538 \cdot \ln(\hat{y}_2))$	1

Table 3. The estimates of the *p*-values MCS for the realized volatility $\sigma_{2,t}^2$, 5 min frequency and MSE

The MSE	*p*-value
$\frac{1}{3}\hat{y}_1 + \frac{2}{3}\hat{y}_4$	0.7790
$\exp(0.5\ln(\hat{y}_1) + 0.5\ln(\hat{y}_4))$	0.7790
$\exp(0.256 \cdot \ln(\hat{y}_1) + 0.66 \cdot \ln(\hat{y}_4))$	1

Table 4. The estimates of the *p*-values MCS for the realized volatility $\sigma_{3,t}^2$, 5 min frequency and MSE

The MSE	*p*-value
$\exp(-1.069 \cdot \ln(\hat{y}_1) + 1.538 \cdot \ln(\hat{y}_2))$	1

For σ_1^2 i σ_3^2 the sets are almost the same. For σ_2^2 the MCS is different then for σ_1^2 and σ_3^2. We notice that for σ_2^2 the linear combinations with the identical weights, according to formula (3) produce worse forecasts then the combination $(1/3)\hat{y}_1 + (2/3)\hat{y}_4$, although the condition on equal variances and corelations indicates equal weights of Timmermann is fulfilled. So the linear combination without the constans may outperform the optimal combination with the constans.

We compared the 446 series of forecasts from models of the types
ARMA(0, 0)-GARCH(1 ,1), ARMA(1, 0)-GARCH(1, 1),
ARMA(0, 1)-GARCH(1, 1), ARMA(1, 1)-GARCH, ARMA(1, 2)-GARCH(1, 1),
ARMA(2, 1)-GARCH(1, 1), ARMA(2, 2)-GARCH(1, 1)
with Gauss, *t*-Student, skewed-Student-*t* and GED with the linear and nonlinear combinations of their twoo serieses of foracasts of the previous types. We rejected

the models with no convergence of model as FIEGARCH with produce very good forecasts and the FIAPARCH-CHUNG with also gives very good forecasts but it's specification is not introduced in the literature. We also rejected the models with not significant parameters.

We acheived the following model confidence sets.

Table 8. The estimates of the *p*-values of MCS for the realized volatility $\sigma^2_{1,t}$, 5 min frequency and MSE

The MCS	*p*-value
GARCH with Gaussian distribution of error	1
ARMA(2, 2)-GARCH with Gaussian	0.7755

For ARMA(2, 2)-GARCH with Gaussian distribution of error we obtained the following parameter estimates

$$\text{Cst}(M) = 0.744(0.035), \quad \text{AR}(1) = -1.0577(0.04652),$$

$$\text{AR}(2) = -0.8295(0.04911), \quad \text{MA}(1) = 1.0546(0.0512),$$

$$\text{MA}(2) = 0.8164(0.0534), \quad \text{Cst(V)} = 0.0136(0.0062),$$
$$\alpha_1 = 0.0376(0.0064), \quad \beta_1 = 0.9566(0.0076)$$

Table 9. The estimates of the *p*-values of MCS for the realized volatility $\sigma^2_{2,t}$, 5 min frequency and MSE

The MCS	*p*-value
ARMA(2, 2)-RiskMetrics with GED	0.9903
GARCH with Gaussian	0.9903
EGARCH with Gaussian	1

For ARMA(2, 2)-RiskMetrics with GED distribution of error we obtained the following parameter estimates

$$\text{Cst}(M) = 0.0783(0.0295), \quad \text{AR}(1) = -1.1328(0.0372),$$

$$\text{AR}(2) = -\ 0.0685(0.0373), \quad \text{MA}(1) = 1.1378(0.04)$$

$$\text{MA}(2) = 0.8607(0.0403), \quad \text{G.E.D. DF} = 1.4258(0.0722), \quad \alpha_1 = 0.06, \quad \beta_1 = 0.94.$$

For EGARCH with Gaussian distribution of error we obtained the following parameter estimates

$$\text{Cst}(M) = 0.0874(0.0353), \quad \text{Cst}(V) = 1.075(0.22056), \quad \alpha_1 = 0.0602(0.3428),$$

$$\beta_1 = 0.989(0.0044), \gamma_1 = 0.00239(0.0094), \gamma_2 = 0.1012(0.0296).$$

Table 10. The estimates of the *p*-values of MCS for the realized volatility $\sigma^2_{3,t}$, 5 min frequency and MSE

The MCS	*p*-value
GARCH with Gaussian distribution of error	1
ARMA(2, 2)-GARCH with Gaussian	0.7755

We took into account the following linear combination from the formula (2).

for $\sigma^2_{1,t}$, $-0.8316 + 0.5201\hat{y}_{11} + 0.5299\hat{y}_{12}$,

for $\sigma^2_{3,t}$, $-0.84024 + 0.5237\hat{y}_{11} + 0.5332\hat{y}_{12}$,

where

$\hat{y}_{11}$ – GARCH(1, 1) with Gaussian distribution of error,

$\hat{y}_{12}$ – ARMA(2, 2)-GARCH(1, 1) with Gaussian distribution of error,

for $\sigma^2_{2,t}$, $-0.81566 + 0.621432\hat{y}_{13} + 0.7568353\hat{y}_{14}$

where

$\hat{y}_{13}$ – ARMA(2, 2)-RiskMetrics wirh GED distribution of error,

$\hat{y}_{14}$ – EGARCH(1, 1) with Gaussian distribution of error.

We received the following results:

Tabela 11. The estimates of the *p*-values of MCS for the realized volatility $\sigma^2_{1,t}$, 5 min, 10 min and 30 min frequencies and MSE

The MCS	*p*-value
$\exp\left(-1.073 \cdot \ln(\hat{y}_1) + 1.538 \cdot \ln(\hat{y}_3)\right)$	0.6064
$\exp\left(-1.078 \cdot \ln(\hat{y}_1) + 1.542 \cdot \ln(\hat{y}_2)\right)$	0.6064
$\exp\left(-1.069 \cdot \ln(\hat{y}_1) + 1.538 \cdot \ln(\hat{y}_4)\right)$	1
$\exp\left(9.6387 \cdot \ln(\hat{y}_{11}) - 9.1916 \cdot \ln(\hat{y}_{12})\right)$	0.6064

Tabela 12. The estimates of the *p*-values of MCS for the realized volatility $\sigma^2_{2,t}$, 5 min, 10min and 30 min frequencies and MSE

The MCS	*p*-value
$\exp\left(\frac{1}{2}\cdot\ln(\hat{y}_1)+\frac{1}{2}\cdot\ln(\hat{y}_4)\right)$	0.7141
$\exp(0.256\cdot\ln(\hat{y}_1)+0.66\cdot\ln(\hat{y}_4))$	0.7141
$\exp(0.49097\cdot\ln(\hat{y}_{13})+0.50022\cdot\ln(\hat{y}_{14}))$	1

Tabele 13. The estimates of the *p*-values of MCS for the realized volatility $\sigma^2_{3,t}$, 5 min, 10 min and 30 min frequencies and MSE

The MCS	*p*-value
$\exp(1.086\cdot\ln(\hat{y}_1)-1.55\cdot\ln(\hat{y}_4))$	1
$\exp(-1.069\cdot\ln(\hat{y}_1)+1.538\cdot\ln(\hat{y}_2))$	0.0758
$\exp(1.54\cdot\ln(\hat{y}_3)-1.07\cdot\ln(\hat{y}_1))$	0.0758
$\exp(9.68230\cdot\ln(\hat{y}_{11})-9.2306\cdot\ln(\hat{y}_{12}))$	0.0758

One can notice from the introduced tables that the sets for $\sigma^2_{1,t}$ and $\sigma^2_{3,t}$ are similar but very differ for $\sigma^2_{2,t}$. The results for the 5 min, 10 min and 30 min frequencies of quotations are almost the same. Another observation is that the nonlinear combinations of forecasts outperform the forecast from the single model and from the optimal linear combinations for the all measures of realized volatility, $\sigma^2_{1,t}$, $\sigma^2_{2,t}$ and $\sigma^2_{3,t}$. The nonlinear function of forecasts with the optimal coefficients of the form $\exp(\beta_1\cdot\ln(\hat{y}_A)+\beta_2\cdot\ln(\hat{y}_B))$ can be successfully used to generate better forecasts.

Conclusions

In the article, we compared the volatility forecasts from a set of ARMA-GARCH models by using the MCS method. Firstly we created a set of the best GARCH models for the WIG20 index and compared their forecasts with the optimal linear and nonlinear combinations of the two forecasts.The analizis was performed for the described types of the combinations of two forecasts. We received the MCS models with nonlinear combinations. We concluded that the nonlinear combinations of the two forecasts outperformed the optimal linear forecasts combinations of two forecasts and we introduced the optimal weights of the nonlinear combinations received by Gauss-Newton method. Next we created the set of ARMA-GARCH models for different ARMA specifications and selected the best forecasting models by the method MCS. For the best models from the set we created linear and nonliner optimal combinations of two forecasts. In the end we compared the forecasts of ARMA-GARCH models with the optimal combinations of two forecasts of the models. In the presented tables the sets for $\sigma^2_{1,t}$ and $\sigma^2_{3,t}$ are similar but very differ for $\sigma^2_{2,t}$. What is more, the nonlinear combinations of the two forecasts outperform the forecast from the single model and from the optimal linear combinations, for all the measures of realized volatility, $\sigma^2_{1,t}$, $\sigma^2_{2,t}$ and $\sigma^2_{3,t}$. For $\sigma^2_{2,t}$, the linear combination without the constans outperforms the optimal combination with the constans. The results for the three different frequencies 5 min, 10 min and 30 min, doesn't differ. The nonlinear function of forecasts with the optimal coefficients of the form $\exp\left(\beta_1 \cdot \ln\left(\hat{y}_A\right) + \beta_2 \cdot \ln\left(\hat{y}_B\right)\right)$ can be successfully used to generate better forecasts (with smaller mean squared error).

References

Aiolfi M., Timmermann A. [2006], *Persistence in Forecasting Performance and Conditional Combination Strategies.* Journal of Econometrics, 135, 31-53

Buszkowska E. [2008], *Porównanie zdolności prognostycznej modeli zmienności indeksu* WIG20 *za pomocą testu SPA*, Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Szczecińskiego 10, 383-393

Donaldson R.G., Kamsatra M. [1996], *Forecast Combining with Neural Networks.* Journal of Forecasting, 15, 49-61

Hansen P.R., Lunde A., Nason J.M. [2003], *Choosing the Best Volatility Models: The Model Confidence Set Approach.* Oxford Bulletin of Economics and Statistics, 65, 839-861

Koopman S.J., Jungbacker B., Hol E. [2005], *Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realized and Implied Volatility Measurements.* Journal of Empirical Finance, 12, 445-475

Martens M. [2002], *Measuring and Forecasting S&P 500 Index Futures Volatility Using High-Frequency Data.* Journal of Futures Markets, 22

Patton J., Sheppard K. [2007], *Evaluating Volatility and Correlation Forecasts*, Oxford-Man Institute of Quantitative Finance, Working Paper

Stock J.H., Watson M. [2004], *Combination Forecasts Output Growth in Seven-Country Data Set.* Journal of Forecasting 23, 405-430

Timmermann A. [2006], *Forecast Combinations*, in: G. Elliot, C.W.J. Granger, A. Timmermann, eds., *Handbook of Economic Forecasting*, North-Holland, Amsterdam

PORÓWNANIE PROGNOSTYCZNYCH MODELI ZMIENNOŚCI I ICH KOMBINACJI DLA INDEKSU WIG20 ZA POMOCĄ METODY ZBIORU UFNOŚCI MODELI

Streszczenie

Jak wiadomo, kombinacje prognoz mogą być lepszymi prognozami niż prognozy uzyskane za pomocą poszczególnych modeli. Celem pracy jest zastosowanie metodologii zbioru ufności modeli (MCS) do porównania zdolności prognostycznej różnych modeli zmienności dopasowanych do dziennych zwrotów notowań indeksu WIG20. Porównano wcześniejsze wyniki z prognozami otrzymanymi za pomocą optymalnych kombinacji liniowych i nieliniowych najlepszych modeli.