TU Dortmund University
Department of Statistics
Summer 2023

M.Sc. Karsten Reichold
M.Sc. Sven Pappert

Case Studies – Project 2
Application of the copula-GARCH model to commodity price time series

Deadline: June 27, 2023, at 2:00 pm
Submission: Submit your report (together with corresponding R code in a separate file) electronically via Moodle **and** as a printed version (without R code) in the course.

This report is concerned with modeling and forecasting first differences of commoditiy price time series using the copula-GARCH model. The multivariate time series consists of first differences of daily short-term future prices of natural gas, oil and coal starting from 2010-03-16 and ending in 2020-10-27. The natural gas price is given in EUR. The other two commodity prices are given in USD. You can download the data in the moodle room (`commodities.csv`). In the first part you will model and forecast each time series individually with an ARMA-GARCH model. In the second part, you will model and forecast the time series jointly by employing the copula-GARCH model. For all tasks you are asked to describe and interpret the results. The ARMA(p,q)-GARCH(r,s) model for a stationary time series $\{X_t\}_{t=1}^{N}$ is given as follows:

$$X_t = \mu + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sigma_t \varepsilon_t,\ \ \varepsilon_t \sim \mathrm{wn}(0,1), \tag{1}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_i \sigma_{t-j}^2, \tag{2}$$

For this project you can assume ε_t to be distributed with respect to the standard normal distribution.[1]

Tasks (First Part):

(a) Conduct a descriptive analysis of the time series at hand. Plot the time series and investigate the (partial-)autocorrelations of the first and second moments.

(b) Fit an ARMA-GARCH model to the first 2500 observations of each time series by maximum likelihood estimation (Reference: Francq and Zakoian, 2019, Chapter 7). Report the estimated parameters and interpret the results. The model orders of the ARMA-GARCH models should be determined by the AIC.

(c) Extract the estimated conditional variances from each time series and plot them. Interpret the plots. What are possible explanations for the behavior of the conditional variances?

(d) Create one-day-ahead probabilistic forecasts for each time series.[2] Start the forecasting study at the 2500th value (i.e. the first forecast is conducted for the 2501st observation).[3] Use the model

[1] Of course you also can work with more sophisticated distributions. But this is not mandatory.
[2] This means that you forecast the distribution of the next observation in the time series.
[3] Although it is better to re-fit the model for each forecast as done in the last report, in this report it is not necessary to re-fit the model for each forecast.

orders determined in part (b). Visualize and compare the probabilistic forecast from a volatile and a non-volatile phase.

(e) Use the continuous ranked probability score (CRPS) (Gneiting and Raftery, 2007) to evaluate the probabilistic forecasts. Compare the probabilistic forecasting performance of the ARMA-GARCH model with the performance of an AR(1) model with normally distributed innovations. Interpret the result.

(f) Extract the 0.05−quantile of each forecast and plot it together with the actual time series. What is the interpretation of the 0.05−quantile in this context? How often are the actual observations smaller than the 0.05−quantile? Based on your results, which of the time series would represent the safest asset to invest in?

(g) Name and explain at least three possible opportunities to improve the univariate probabilistic forecasts.

Familiarize yourself with copulas using e.g. the introduction by Schmidt (2007) and the copula-GARCH model as presented in e.g. Lu *et al.* (2014)[4]. Let $X_{i,t}, i = 1, \ldots, K, t = 1, \ldots, N$ be stationary and continuous time series. The basic idea amounts to decomposing the joint distribution (conditional on the σ-algebra spanned by all past values, indicated by $\mathcal{F}_{t-1}$) at time t into a copula and marginal distributions:

$$F_{X_{1,t},\ldots,X_{K,t}|\mathcal{F}_{t-1}}(a_1, \ldots a_K) = C[F_{X_{1,t}|\mathcal{F}_{t-1}}(a_1), \ldots, F_{X_{K,t}|\mathcal{F}_{t-1}}(a_K)]. \tag{3}$$

In the second part of the project, you will model the conditional joint distribution of the commodity price time series with the copula-GARCH model. This means that the univariate conditional distributions, $F_{X_{i,t}|\mathcal{F}_{t-1}}$ will be modeled with the ARMA-GARCH models from part 1) and then be coupled with an appropriate copula.

Tasks (Second Part):

(h) Extract the standardized residuals of the first 2500 observations from each time series from the univariate ARMA-GARCH models. Use the probability integral transformation (PIT) to transform the standardized residuals to be uniformly distributed.[5] Plot the respective bivariate dependence structures of the PIT-transformed standardized residuals. Describe the dependence structures using appropriate dependence measures. Interpret the results and the plots.

(i) Fit appropriate Copula models to the PIT-transformed standardized residuals by maximum likelihood estimation (as described in Lu *et al.*, 2014, Section 3.3) and decide which model is most suitable.

(j) Create one day-ahead multivariate probabilistic forecasts for the multivariate time series with the following copula specifications:

- the independence copula,
- the gaussian copula,
- the t-copula.

[4] In the paper, the model is presented with time varying dependence parameters. This is not necessary for this project. You can consider the dependence parameters as constant.

[5] You can either use the estimated distribution function for the PIT or apply the non-parametric distribution function. This amounts to assigning the rank to each observation and rescaling.

Start the forecasting study again at the 2500th value (i.e., the first forecast is conducted for the 2501st observation). You can simulate the forecast for the 2501st observation by first simulating $N = 1000$ 3-dimensional realizations from the respective copula. Then quantile transform the realizations with the respective quantile function $F^{-1}_{X_{i,2501}|\mathcal{F}_{2500}}$. Repeat this procedure for every forecast.

(k) Evaluate the multivariate probabilistic forecasts with the energy score (Gneiting and Raftery, 2007). Interpret the results. What are possible opportunities to improve the multivariate forecast performance?

References

Francq, C., Zakoian, J. M. (2019). GARCH models: structure, statistical inference and financial applications. John Wiley & Sons.

Gneiting, T., Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. Journal of the American statistical Association, 102(477), 359-378.

Lu, X. F., Lai, K. K., Liang, L. (2014). Portfolio value-at-risk estimation in energy futures markets with time-varying copula-GARCH model. Annals of operations research, 219(1), 333-357.

Schmidt, T. (2007). Coping with copulas. Copulas – From theory to application in finance, 3, 1-34.