Wroclaw University of Science and Technology
Faculty of Pure and Applied Mathematics
Department of Applied Mathematics

Ph.D. Thesis

Application of stochastic processes for modelling market risk factors in a mining company

Łukasz Tomasz Bielak

Supervisor:
Agnieszka Wyłomańska, PhD, DSc, Professor WUST

Assistant Supervisor:
Joanna Janczura, PhD

Wroclaw, September 2022

Politechnika Wrocławska
Wydział Matematyki
Katedra Matematyki Stosowanej

Rozprawa doktorska

Zastosowanie procesów stochastycznych do modelowania czynników ryzyka rynkowego w przedsiębiorstwie górniczym

Łukasz Tomasz Bielak

Promotor:
dr hab. inż. Agnieszka Wyłomańska, prof. uczelni

Promotor pomocniczy:
dr Joanna Janczura

Wrocław, wrzesień 2022

Acknowledgements

First and foremost, I am deeply grateful for the continuous support, insight and extreme patience of my supervisors, Prof. Agnieszka Wyłomańska and Dr Joanna Janczura: without their constant motivation and an outstretched helping hand, this thesis would not have been completed. I would also like to thank Dr Paweł Miśta for his helpful supervision of the project.

I wish to acknowledge my colleagues and co-authors of all publications, who shared with me their knowledge and skills: Dr Aleksandra Grzesiek, Dawid Szarek, Dr Grzegorz Sikora, Anna Michalak, Julia Adamska, Andrzej Puć and Tomasz Serafin.

Last, but not least, I could not have undertaken this journey without the unconditional support and love of my family. I dedicate this thesis to my wife Marta and our children: Jakub, Filip, Aleksander and Joanna.

Contents

Chapter 1

Introduction

Market risk factors forecasting is one of the most difficult tasks which has to be conducted by a company in the face of a current volatile and globalised financial world. Depending on the profile of the company, it may deal with commodity risk, foreign exchange risk, interest rate risk and equity price risk. There is wide range of methods which can be used for forecasting, but they should be well fitted to the purpose of their applications.

The probability of forecasting market risk prices without any errors in the longer horizon is close to zero, therefore determining price assumptions rather as ranges or probability distributions than single forecasts seems to be an effective way to address the high volatility of the market risk as well as a concept which enables a company to be able to prepare better not only for base case scenario but also for the range of more pessimistic or optimistic scenarios.

While stochastic modelling of market risk prices is a very useful tool, which allows for the simulation of future price distributions, it requires a proper understanding of market risk factors dynamics, which in turn enables selection of an appropriate model and its correct calibration and estimation. Traditionally, advanced methods of forecasting and modelling prices have been used in financial sector. However, an increasingly dynamic financial environment has begun to spur industrial companies to start implementing it as well.

One of the business areas which is very unique in terms of characteristics, impact and scale of market risk is the mining business. Mining companies are exposed mainly to commodity prices (metals and energy), exchange rates and interest rates. What makes this risk exposure exceptional is the fact that the sensitivity of production and financial plans of the company to market risk is far larger than in other business areas. Moreover, the business and investment plans of mining companies are usually of a multi- year nature. Hence, correctly assessing market risk prices behaviour has an enormous and fundamental impact on a company's strategic, technical and financial plans.

The higher volatility of market risk factors, time-inhomogeneous and non-Gaussian characteristics, means that classical processes used for stochastic modelling may not ideally reflects the properties of market data. In order to accommodate the above caveats, we propose in this thesis novel models and methods which could improve the quality of modelling and consequently improve the entire market risk management process.

The dissertation is organised as follows. Chapter 2 presents the foundational motivation behind this doctoral thesis and describes problems, which could be addressed or solved if the results of the study are applied. Chapter 3 contains a summary of published research articles related to forecasting and modelling market risk factors, with special attention to metals prices. The final section describes author's contribution to the field and lists all publications which serve as reference material to the thesis of the work. Chapter 4 is focused on modelling metals prices, with primary emphasis on the non-Gaussian behaviour and time-inhomogeneous character of data, as well as changes of regimes. Chapter 5 describes the CKLS/SGT model applied to modelling currency exchange rates. This model has been also used for presenting and addressing the problem of the calibration data length used for the prediction. Chapter 6 should be treated as an introduction for multidimensional view on a given problem. In this part we have proposed a two-dimensional VAR model with the $\alpha-$stable distribution that reflects the changing dependence structure of the analysed assets. Chapter 7 presents a product of two components for a two-dimensional time series. The theoretical results are applied to the energy market case study. Chapter 8 summarises the thesis and shows potential applications of the methods proposed in previous chapters for market risk management.

The presented doctoral dissertation has been prepared as a part of the Implementation Doctorate Programme in cooperation with KGHM Polska Miedź S.A. Given the fact, that the market risk factors modelling can be used by any manufacturing company, consideration of the results presented in this thesis may have wider business application.

Chapter 2

Problem formulation

A production company, whose financial results to a large extent depend on market prices, is obliged to properly manage its exposure to the risk of fluctuations of these prices. Market risk management has become a fundamental element of the corporate strategy and a source of enhancing a company's value. Effective risk management primarily involves identifying, measuring risk and deciding on the scale of risk acceptance, as well as the use of risk mitigating instruments.

KGHM Polska Miedź S.A. (KGHM) is a company active in the mining business, where the aforementioned properties are in an enhanced way valid. Based on uncertain metals prices and currency rates assumptions, multiyear technical and financial plans are prepared, investment decisions are undertaken and liquidity risk is managed. Market risk factors which have the greatest impact on the KGHM's financial results are: the copper price, the silver price, energy prices and exchange rates (especially USDPLN).

One of the key steps in the market risk management process in KGHM is risk measurement, which in some way quantifies the size of the company's current exposure to risk and helps to answer the question of whether this exposure is optimal for the company in the context of the current, external market conditions and the internal situation or the long-term plans of the company. Correctly calculated risk measures are also used to estimate the impact of the company's potential actions aimed at shaping the desired risk profile. Such actions include using natural hedging or actively managing the derivatives portfolio and preparing simulations of future financial results as support for the market risk management process.

Calculating such risk measures as Earnings-at-Risk, Cash-Flow-at-Risk or Net Debt / EBITDA-at-Risk among others, requires generating numerous paths of market risk factors. Simulation utilises probing from a predetermined predicted probability distribution. For this reason it is extremely important to choose an appropriate model on the basis of which the distribution is created,

especially when the market environment is constantly changing. In the scenario generation stage, it is necessary to use stochastic models to simulate market variables of up to even a 5-year time horison. Modelling market risk factors for such a long period is complex, and the literature on this issue is limited.

For modelling purposes KGHM has been assuming that the prices of metals (copper and silver) follow the dynamics of the Schwartz model for commodities (geometric Ornstein-Uhlenbeck process). It is a model commonly used in commodity markets, and often the forward curves show a similar shape to the expected value of this model. It has desirable properties such as a return to the mean and a stabilising variance in the long term. In the case of the USDPLN exchange rate simulation, an assumption has been made about the dynamics of this exchange rate in accordance with the standard Black-Scholes model (geometric Brownian motion).

The objective of this dissertation is to focus on analysis and selection of stochastic processes used in the simulation process which could better address properties of market data.

Areas of potential improvements are grouped below:

1. Financial variables are characterized by **non-Gaussian distributions** and some authors stressing that assuming the Gaussian distribution of the prices in modelling is inappropriate. The reason for this is that very often extreme events in financial markets occur more frequently than expected and their deviation from mean or median is far bigger than is by definition assumed in Gaussian distribution. Thus, in the literature, many researchers have proposed to apply the modification of the classical models by using processes different from the Brownian motion as the noise.

2. Additionally, a continuously changing macroeconomic situation means that financial data are not stable over time, which in other words can be defined as their having an in-homogeneous character. The implication of this is that a model with constant coefficients could be inappropriate to model data with a time-dependent mean and time-dependent scale parameter (in particular variance). Therefore, one can find in the literature different modifications of the classical stochastic processes by taking, instead of the constant, the **time-dependent coefficients**. However in the literature there is small number of examples showing how the above-mentioned characteristics could be applied to market risk factors modelling, especially in the longer horison.

3. Another important phenomenon that should also be addressed, based on a non-homogeneous nature in financial data, is the **change of price regimes**. The regimes captured by regime-switching models very often correspond to fundamental changes in specific markets. In empirical analysis this means that volatilities, autocorrelations, and cross-covariances of asset returns very often differ across regimes, which allow regime-switching models to properly capture the dynamics of

financial series (especially for commodity prices). In modelling market risk factors based on long term data history, the ability to recognise and apply the change of regimes can be of benefit rather assuming that the whole data history represents one modelling structure.

4. One of the most important problem in stochastic modelling can be formulated as follows: **how to properly model the distribution or range of values for an asset which is a combination of two (or more) other assets**. In the previously mentioned context of market risk management in a mining company, the problem can be formulated as the modelling of the range of values for the metal price expressed in the currency of a given country. In the one-dimensional modelling approach, it can be simply analysed as a product of two independent processes (metal price and exchange rate), while in two-dimensional models (such as the vector autoregressive model) the relation between these two assets is also included.

The situation becomes more complex when the relation between the metal price and the exchange rate (two assets) is changing over time. Where that relation, over time, to turn out to be apparent or unstable, it could lead to misleading conclusions and, as a consequence, wrong decisions in market risk management or more generally business process. Moreover, although including the above-mentioned properties, as non-Gaussian and time-inhomogenous behaviour of assets, makes such multidimensional modelling even more complex, nevertheless this can improve the quality of the market risk factors modelling and more adequately measure the market risk level for a company.

5. Finally, when using stochastic processes, one additional question very often appears, namely **how to choose the appropriate length of historical data for stochastic modelling**. The most common approach to the process of forecasting is to choose all of the available historical data and then calibrate the forecasting model on the selected sample and predict future values. The length of the historical data sample is most often chosen in an 'ad-hoc' fashion. It has been however shown in the literature on short-term forecasting that by using various approaches, forecasting errors can be significantly lowered. The first popular method used here involves averaging predictions obtained from different models calibrated to the same historical data sample. Another technique that brings gains in terms of forecasting accuracy involves averaging forecasts obtained from the same model but calibrated to historical samples of different lengths. The above-mentioned approaches have been proved to be effective and extensively used in the context of short-term forecasting but the problem of the optimal calibration sample length selection remains practically overlooked when it comes to long-term modelling.

The motivation for this thesis is to propose potential improvements to stochastic processes used for modelling which could address the problems described above and, in consequence,

improve the quality of the modelling process.These problems are not entirely new, but considering them in the context of a company active in the mining business adds an additional dimension to the discussion.

Chapter 3

State of the art

This chapter summarises the important theory and research direction in the considered field of modelling market risk factors.

3.1 State of the art

Advanced mathematical modelling is required for forecasting market risk factors distributions and simulating prices scenarios, which is an integral part of the market risk measurement process. A wide variety of forecasting methods are available and it is worth stressing that there is no single model which is universally applicable. As some authors suggest, it is even desirable to use different forecasts under diverse sets of assumptions so that the results can be compared under alternative scenarios [1].

3.1.1 Market risk forecasting methods

Although methods based on the continuous- and discrete-time processes seem to be appropriate for market risk price modelling, one can find many different approaches considered for this problem. As commodities prices are one of most important from KGHM's point of view, our review is based on this class of market risk factors; however it can be generalised to apply to other market risk factors. The methods proposed in the literature to commodities price forecasts, in general, divided into a few groups: qualitative, trend-based, econometric, stochastic processes- and time series-based. One can also find a modern approach where machine learning techniques are applied [2–5]. Various combinations of these methods are also considered in [6–11]. For more details, it is worth mentioning, for instance, to the review the article [12].

In the econometric approach, most of the research is based on the relationship between market prices and economic factors, and functions under the assumption that supply and demand are fundamental drivers of economic growth and price balance. The dependence between price and economic factors is the basis for the economic models. This approach is, for instance, demonstrated in [13–15], where the similarity between the price movements for different commodities was discussed. The econometric-based methods were also proposed in [16, 17].

In the stochastic-based approach for market price modelling, Gaussian-based models have mainly been discussed. The authors argue that predicted market prices based on the stochastic approach are a result of a widely held conviction that market fluctuations have random sources [12, 14, 18]. The commonly-used stochastic models for the commodity price modelling are ordinary Brownian motion, geometric Brownian motion, and mean reversion models, like the Vasicek model or its extension for the non-constant coefficients. In the following papers various Gaussian stochastic processes are applied for the commodity price modelling [19–25]. In the next section, stochastic models discussed in the literature in the context of modelling market risk prices are described in a more comprehensive way.

The time series-based approach was presented, for instance, in [10, 26, 27] where the authors proposed the autoregressive moving average (ARMA) models for a real commodity price. Moreover, in [10, 26, 27] more complicated time series models were applied, such as autoregressive conditional heteroscedastic (ARCH), the generalized ARCH (GARCH), or the autoregressive integrated moving average models (ARIMA). The error correction model was proposed in [28], while the unequal-interval contour lines and contour time sequences filtration was discussed in [29] to predict the metal price. Finally, the improved wavelet–ARIMA time series is presented in [30] for commodity price modelling. Similarly to the stochastic processes-based approach, as well as in time series modelling, the researchers depart from the assumption of Gaussianity and propose a heavy-tailed class of distributions [31].

On the border of time series-based and econometric approaches there are multidimensional time series models. On the one hand, multidimensional systems can describe the dependence between varied factors, while on the other hand they take into account the possible dependence in time within one single process. The most classical multidimensional time series is the class of vector autoregressive (VAR) models. In [16, 17, 32] the multidimensional modelling of commodity prices and the financial markets in two countries was considered. This method is discussed in more detail in the following section.

3.1.2 Stochastic processes used for modelling market risk prices

The analysis of real data by using continuous-time models is based on the discrete-time approximation of the theoretical stochastic processes. This approach seems to be more effective, especially in modelling long-term market risk prices. One of the classical continuous-time stochastic processes used in the financial data description is the Ornstein-Uhlenbeck model introduced by Uhlenbeck and Ornstein [33] as the appropriate system for the velocity in the classical Brownian diffusion. The Ornstein-Uhlenbeck process, known also as the Vasicek model [34], was one of the earliest stochastic systems of the term structure. It exhibits the mean reversion property which indicates that, over time, the process tends to its long-term mean. This pattern is definitely visible in the commodity price data [23, 35], where the supply-demand relation should stay in balance in the long term. Some authors also use such an assumption for currency modelling [36, 37].

Non-Gaussian distributions

However, some researchers voice the opinion that classical stochastic models are inappropriate for commodity and, also more generally, for financial price description as they do not take into account possible large observations apparent in market data. In [12, 38] the authors stressed that assuming the Gaussian distribution for modelling market prices is inappropriate as this is not characteristics of financial data. Thus, in the literature, many researchers proposed a modification of the classical Orstein-Uhlenbeck process by using process different from the Brownian motion one as the noise [25, 39–43]. This approach can be also continued by applying in place of the ordinary Brownian motion the process of stationary independent increments having skewed generalized Student's t (SGT) distribution, [44]. The class of SGT distributions contains many known distributions such as Student's t, Gaussian and Cauchy. Thus, it can be considered as the general class useful for modeling light- and heavy-tailed data. Models based on the SGT distribution were used in modelling of various phenomena, [45–48]. An interesting summary of non-Gaussian processes used in the energy and commodities area can be found in [49], while the use of non-Gaussian models in the exchange rates area is described in [50].

Time-dependent coefficients

Moreover, stochastic models, where it is assumed that the coefficients are time-dependent, are becoming meaningful. These models are especially useful for data where a changing (in time) mean and scale parameter (in particular variance) are apparent. The best-known examples of such models that are especially useful in financial data description are Ho-Lee [51], Hull-White [52],

Black-Derman-Toy [53] and Black-Krasinski [54]. One can find in the literature the techniques used for the estimation of the time-dependent parameters of the stochastic models [55–59]. A discussion about the use of time-dependent coefficients in exchange rates modelling can be found in [60] whereas in the energy market in [61].

Changing of price regimes

Another important modelling feature is the fundamental changes of price behaviour characteristics, captured by regime-switching models. In empirical analysis means, volatilities, autocorrelations, and cross-covariances of asset returns very often differ across regimes, which allow regime-switching models to capture properly the dynamics of financial series (especially commodity prices) including fat tails, heteroskedasticity, skewness, and time-varying correlations. Especially in modelling financial data, based on long-term data history, the ability to recognize and apply the change of regimes is extremely important to better forecast potential price distributions. In the literature, stochastic models are proposed that take into consideration the fact that the real data behave differently for different periods. This is related to the so-called regimes' switching behaviour [62, 63]. Interesting examples of such models, for financial data description, are presented for instance in [64–66]. A regime switching approach can be also found in research which only concentrates on a chosen asset class: energy [67, 68], metals [69], exchange rates [70, 71].

Multidimensional modelling

We have indicated in the previous section the VAR model as one of the methods used for metals' forecasting. This classical multidimensional model is used in various applications, especially in economy and finance [72–75]. The popularity of such models is related to the fact that on one hand they can describe the dependence between various factors and, on the other hand, they take into account the possible dependence in time within the one single process. There are many interesting research studies devoted to the theoretical analysis of VAR time series, see e.g. [76, 77]. Generally, the VAR model has been used in many different applications across asset classes. In the research paper [78] VAR models have been used for forecasting WTI, Brent, natural gas and heating oil prices. VAR models used for exchange rates modelling are proposed in [79–81].

Another approach to multidimensional modelling is analysing of time series (or general stochastic process) being a product of other time series. This topic has been rarely discussed in the literature, but it seems to be interesting, especially from the practical point of view. In [82], the basic statistics were discussed for time series being a product of two stationary models. In the mentioned bibliography position, the author highlighted that the product time series is crucial in

nonlinear time series analysis or in the theory of time series with random parameters. In [83] the effects of algebraic operations were discussed, for example the product, on trends in time series. An analysis related to the product of stochastic (time dependent) components is presented by [84], where the authors discussed the product of stochastic iterated integrals associated with general Lévy processes. It is also worth considering [85, 86] for other studies on the products of stochastic integrals.

Particular emphasis in studying the interdependence of data is placed on the product of two random variables that comes from the same class of distributions, see eg. [87–90]. In this case one can find interesting approaches used in the analysis of the distribution and probabilistic properties of the product random variable. Special attention is paid to a case where the two considered random variables are Gaussian or Student's t distributed, see e.g. [91–95] as well as [96–99].

One can also find other interesting cases in [100], where exponentially distributed random variables are considered or in [101], where Dirichlet distributed random variables are examined. For other references, see also [102–105]. Different classes of distributions are also considered in the literature and the product of such random variables is analysed, see eg. [106–111] and references therein. The theoretical results related to the product of random variables were used in various applications, including finance, risk management, the economy, but also the physical sciences, reliability theory, hydrology, and many others, see e.g. [112–120].

Optimal length of calibration data

The most common approach to the process of forecasting is to choose a sufficient portion of historical data and then calibrate the forecasting model on the selected sample and predict future values. The length of the historical data sample is most often chosen in an 'ad-hoc' fashion; usually the authors decided to use as many historical values as possible. It has been however shown in the literature on the short-term forecasting that by using various approaches, forecasting errors can be significantly lowered. The first popular method involves averaging predictions obtained from various models calibrated to the same historical data sample [121, 122]. Although it turned out that the averaged forecast usually outperforms each individual prediction, this approach has not resolved the problem of the optimal calibration sample length. Another technique that brings gains in terms of forecasting accuracy involves averaging forecasts obtained from the same model but calibrated to historical samples of different lengths [123–126]. The rationale behind this approach is that, when calibrating the model to a longer sample, we capture the long-term trends whereas in using a short calibration sample we take into consideration only recent, short-term price behaviour.

3.2 Author's contribution to the field

As was stated earlier, although there are variety of stochastic processes proposed by the literature, there is a limited availability of research describing their use in the market risk factors modelling process for business purposes.

In this thesis, we have been been focused on stochastic processes which can be used for modelling prices in a horison adequate to the requirements of a manufacturing company or, more specifically, a mining company. The proposed models also take into account the specific characteristics of the financial market.

We have introduced new stochastic models with time-dependent parameters, regime-switching and non-Gaussian, namely SGT, behaviour. Moreover, we have presented some theoretical properties of the analysed processes and described the step-by-step procedure of estimation of the introduced new model's parameters.

We have also considered the CKLS model based on the SGT distribution which exhibits non-Gaussian behaviour. We have described the estimation technique for the model parameters and checked its efficiency using Monte Carlo simulations. We also defined a new validation factor useful in the selection of the appropriate model to real data.

Further, we have described the problem of the selection of the calibration window length of the long–term prediction for the currency exchange rates data. We have proposed here a new averaging approach. As a result we have concluded that the long term averaging of different models can give better and more stable results in modelling exchange rates.

Moreover, we have examined the stability of the relation between two assets, namely metal price and exchange rate, using different correlation metrics. Basing on the results, we have proposed a two-dimensional VAR model with an α-stable distribution that reflects the changing dependence structure of the analysed assets in two identified regimes.

Finally, we have discussed the properties of the product of two time series. We have derived general formulas for the autocovariance function and study its properties for various cases of cross-dependence between the VAR(1) model components. A simulation study has been conducted for two types of bivariate distributions of the residual series, namely the Gaussian and Student's t. The results obtained were applied to the electricity market case study.

The results obtained in this thesis have been published in the following articles:

1. G. Sikora, A. Michalak, Ł. Bielak, P. Miśta, and A. Wyłomańska, "Stochastic modeling of currency exchange rates with novel validation techniques," *Physica A: Statistical Mechanics and its Applications*, vol. 523, pp. 1202 – 1215, 2019

2. Ł. Bielak, P. Miśta, A. Michalak, and A. Wyłomańska, "The application of correlation models for the analysis of market risk factors in KGHM Capital Group," *Proceedings in Earth and Geosciences vol. 3, Mining Goes Digital, Mueller et al*, vol. 523, pp. 38–46, 2019

3. T. Serafin, A. Michalak, Ł. Bielak, and A. Wyłomańska, "Averaged-calibration-length prediction for currency exchange rates by a time-dependent Vasicek model," *Theoretical Economics Letters*, vol. 10, pp. 579–599, 2020

4. D. Szarek, Ł. Bielak, and A. Wyłomańska, "Long-term prediction of the metals' prices using non-Gaussian time-inhomogeneous stochastic process," *Physica A: Statistical Mechanics and its Applications*, vol. 555, p. 124659, 2020

5. D. Szarek, Ł. Bielak, and A. Wyłomańska, "Non-Gaussian regime-switching model in application to the commodity price description," *Chaari F.et al. (eds) Nonstationary Systems: Theory and Applications. WNSTA 2021. Applied Condition Monitoring, Springer*, vol. 18, pp. 108–126, 2021

6. Ł. Bielak, A. Grzesiek, J. Janczura, and A. Wyłomańska, "Market risk factors analysis for an international mining company. Multi-dimensional heavy-tailed-based modelling," *Resources Policy*, vol. 74, p. 102308, 2021

7. J. Adamska, Ł. Bielak, J. Janczura, and A. Wyłomańska, "From multi- to univariate: A product random variable with an application to electricity market transactions: Pareto and student's t-distribution case.," *Mathematics*, vol. 10, no. 18, p. 3371, 2022

8. J. Janczura, A. Puć, Ł. Bielak, and A. Wyłomańska, "Dependence structure for the product of bi-dimensional finite-variance VAR(1) model components. an application to the cost of electricity load prediction errors.," submitted 2021

Chapter 4

Metals prices modelling

Modelling metals prices is a very difficult task, especially if it is focused on a longer, even several years horison. Metals prices are driven not only by fundamental factors like supply-demand balance but also by the macroeconomic environment, investor sentiment and other factors, which in practice very often implies high volatility and limited predictability of behaviour. While using stochastic modelling helps to better measure and understand the scale and probability of potential price movements, nevertheless the selection of the right model is extremely important to achieve the desired effect. Commodities prices, including metals, usually exhibit the mean reversion property, which indicates that over time, the process tends to its long-term mean. There is a consensus on this matter on the market. However, when it comes to further selection of a model, there are opinions that traditional stochastic models, with Gaussian distribution and time-homogeneous parameters, are not perfect in the description of commodities prices.

In this chapter, we introduce two novel stochastic models which could be used for modeling metals prices. Both models are based on the SGT distribution and stochastic differential equation, thus before we define the analysed stochastic process, first will first introduce the SGT distribution and general model framework.

Skewed Generalized t-Student distribution (described in paper [121]) is a family of a wide variety of well-known distributions such as Gaussian, uniform, Laplace or Student's t-distribution. Setting SGT distribution parameters:

$$\{(\mu, \sigma, \lambda, p, q) : \mu \in \mathbb{R}, \sigma > 0, -1 < \lambda < 1, p > 0, q > 0\};$$

accordingly to what is presented in Fig. 4-1, specific distribution can be obtained.

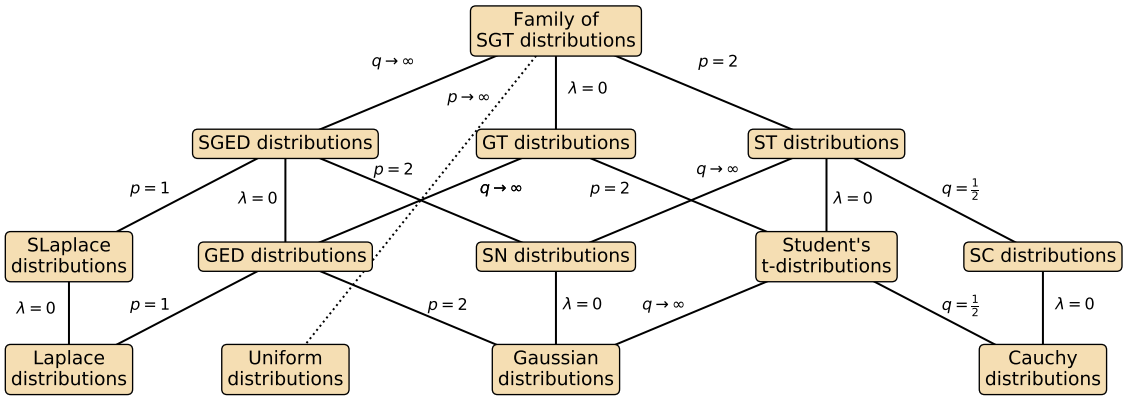


Figure 4-1: The SGT family of distributions, source: [129]

SGT distribution is defined by probability density function (PDF) [134] in the following way:

$$f_{SGT}(x;\mu,\sigma,\lambda,p,q) = \frac{p}{2\upsilon\sigma q^{\frac{1}{p}} B\left(\frac{1}{p},q\right)\left(\frac{|x-\mu+m|^p}{q(\upsilon\sigma)^p(\lambda \mathrm{sgn}(x-\mu+m)+1)^p}+1\right)^{\frac{1}{p}+q}}$$

where:

$$m = \frac{2\upsilon\sigma\lambda q^{\frac{1}{p}} B(\frac{2}{p},q-\frac{1}{p})}{B(\frac{1}{p},q)}; \tag{4.1}$$

$$\upsilon = q^{-\frac{1}{p}}\left[(3\lambda^2+1)\left(\frac{B\left(\frac{3}{p},q-\frac{2}{p}\right)}{B\left(\frac{1}{p},q\right)}\right) - 4\lambda^2\left(\frac{B\left(\frac{2}{p},q-\frac{1}{p}\right)}{B\left(\frac{1}{p},q\right)}\right)^2\right]^{-\frac{1}{2}}.$$

Moreover, we assume that $pq > 2$.

Let $S_1 \sim SGT(\mu = 0,\ \sigma = \tilde{\sigma},\ \lambda,\ p,\ q)$ and $S_2 \sim SGT(\mu = \mu,\ \sigma = \tilde{\sigma}\sigma,\ \lambda,\ p,\ q)$. Then:

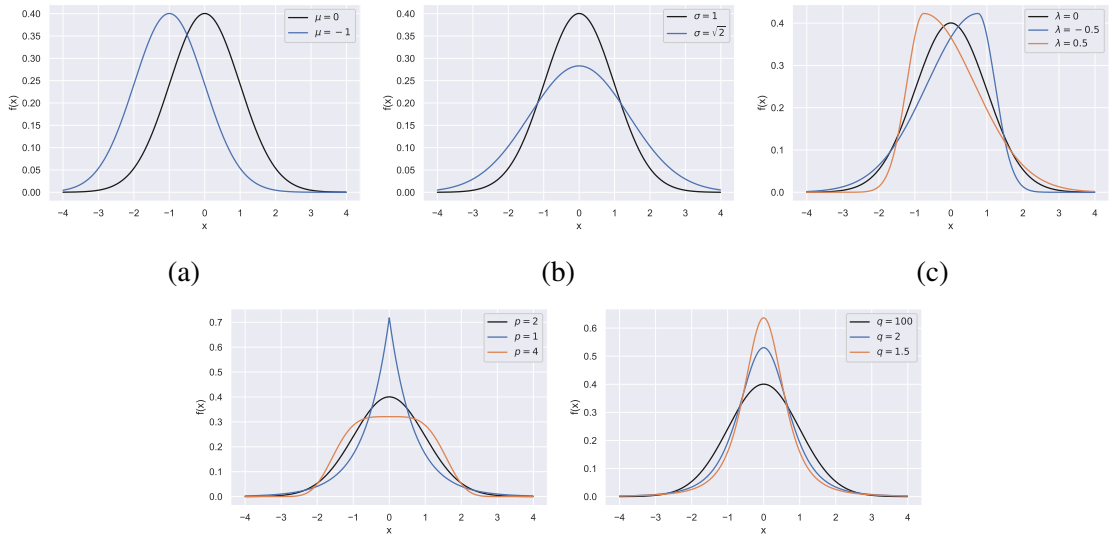
$$\forall_{\sigma>0,\ \mu\in\mathbb{R}}\ S_1\sigma + \mu \overset{d}{=} S_2. \tag{4.2}$$

Indeed, let us define

$$S_1 \sim SGT(\mu = 0,\ \sigma = \tilde{\sigma},\ \lambda,\ p,\ q),\ S_2 \sim SGT(\mu = \mu,\ \sigma = \tilde{\sigma}\sigma,\ \lambda,\ p,\ q)$$

with $\sigma > 0,\ \mu \in \mathbb{R}$. We will show that:

$$P(S_1\sigma + \mu < x) = P(S_2 < x). \tag{4.3}$$



(a) (b) (c)

(d) (e)

Figure 4-2: The PDF of the SGT distribution for different values of the parameters in comparison with the PDF of standard Gaussian distribution, source: [129]

We use the formula for the PDF of SGT distribution given in formula (4.1):

$$
\begin{aligned}
P(S_1\sigma+\mu<x) &= P\left(S_1<\frac{x-\mu}{\sigma}\right)=\int_{-\infty}^{\frac{x-\mu}{\sigma}} f_{SGT}(u;0,\tilde{\sigma},\lambda,p,q)du \\
&= \int_{-\infty}^{x}\frac{1}{\sigma}f_{SGT}\left(\frac{u-\mu}{\sigma};0,\tilde{\sigma},\lambda,p,q\right)du \\
&= \int_{-\infty}^{x}\frac{p}{2\upsilon(\tilde{\sigma}\sigma)q^{\frac{1}{p}}B\left(\frac{1}{p},q\right)\left(\frac{|\frac{u-\mu}{\sigma}+m|^p}{q(\upsilon\tilde{\sigma})^p(\lambda\mathrm{sgn}(\frac{u-\mu}{\sigma}+m)+1)^p}+1\right)^{\frac{1}{p}+q}}du \\
&= \int_{-\infty}^{x}\frac{p}{2\upsilon(\tilde{\sigma}\sigma)q^{\frac{1}{p}}B\left(\frac{1}{p},q\right)\left(\frac{|u-\mu+\sigma m|^p}{q(\upsilon(\tilde{\sigma}\sigma))^p(\lambda\mathrm{sgn}(u-\mu+\sigma m)+1)^p}+1\right)^{\frac{1}{p}+q}}du \\
&= \int_{-\infty}^{x} f_{SGT}\left(u;\mu,\tilde{\sigma}\sigma,\lambda,p,q\right)du = P(S_2<x).
\end{aligned}
$$

The above equality is true because of the fact following from the Eq. (4.1):

$$
\sigma m = \sigma\frac{2\upsilon\tilde{\sigma}\lambda q^{\frac{1}{p}}B(\frac{2}{p},q-\frac{1}{p})}{B(\frac{1}{p},q)} = \frac{2\upsilon(\tilde{\sigma}\sigma)\lambda q^{\frac{1}{p}}B(\frac{2}{p},q-\frac{1}{p})}{B(\frac{1}{p},q)}.
$$

In Fig. 4-2 we present the PDF of the SGT distribution for different vales of the parameters. The PDFs are compared with the PDF of standard Gaussian distribution.

The general model considered in this chapter is given by the following stochastic differential equation [135]:

$$dX_t = \alpha(X_t, t)dt + \beta(X_t, t)dB_t, \tag{4.4}$$

where $\{B_t\}$ is the Brownian motion.

Most known examples that fit the model (4.4) are presented in the Tab. 4.1.

Table 4.1: Classical models given by the Eq. (4.4).

Model	$\alpha(X_t, t)$	$\beta(X_t, t)$
Merton	α_1	β_1
Vasicek	$\alpha_1 + \alpha_2 X_t$	β_1
Dothan	$\alpha_1 X_t$	$\beta_1 X_t$
Brennan-Schwartz	$\alpha_1 + \alpha_2 X_t$	$\beta_1 X_t$
Cox-Ingersoll-Ross (CIR)	$\alpha_1 + \alpha_2 X_t$	$\beta_1\sqrt{X_t}$
CIR-VR	$\alpha_1 + \alpha_2 X_t$	$\beta_1 X_t\sqrt{X_t}$
Ho-Lee	$\alpha_1(t)$	$\beta_1(t)$
Hull-White	$\alpha_1(t) + \alpha_2(t)X_t$	$\beta_1(t)$
Black-Derman-Toy	$\alpha_1(t) + \alpha_2(t)\ln X_t$	$\beta_1(t)$
Black-Krasiński	$\alpha_1(t) + \alpha_2(t)X_t \ln X_t$	$\beta_1(t)$

However, in many business applications (like metals' prices modelling) the Gaussian distribution in the model (4.4) seems to be insufficient. Thus, we propose to modify the model and assume a process which satisfies the following stochastic differential equation:

$$dX_t = \alpha(X_t, t)dt + \beta(X_t, t)dS_t, \tag{4.5}$$

where similarly as previously, in the general case $\alpha(\cdot)$ and $\beta(\cdot)$ are appropriate functions. Here we assume $\{S_t\}$ is a process of stationary independent incremental having SGT distribution. In this case the increment process $\{dS_t\} = \{S_{t+dt} - S_t\}$ constitutes a sequence of independent identically distributed (iid) random variables of Skewed Generalized t-Distribution [121] with the assumption $\mathbb{E}(dS_t) = 0$ and $\mathbb{E}(dS_t^2) = dt$.

4.1 Non-Gaussian time-inhomogeneous model

The most important metals' risk factors for KGHM are copper, silver and gold prices. Modelling the behaviour of metal prices in the medium- and long-term horison is necessary for the market risk management process. As mentioned before, stochastic models traditionally used to describe metals' prices have proved not to be suitable to represent the dynamic behaviour and time-related nature of metal markets. Thus, the model (4.5) with the constant coefficients is inappropriate to model data with a time-dependent mean and time-dependent scale parameter (in particular variance). In this section we implement and describe a modification of the classical Ornstein-Uhlenbeck process (4.6) by taking, instead of the constant, the time-dependent coefficients α and β.

In the further calculations we assume that the functions $\alpha(\cdot)$ and $\beta(\cdot)$ have specific form, and finally the analysed process is given by the stochastic differential equation:

$$dX_t = (\alpha_1(t) + \alpha_2(t)X_t)dt + (\beta_1(t) + \beta_2(t)X_t)dS_t \tag{4.6}$$

for the general functions $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and $\beta_1(\cdot)$, $\beta_2(\cdot) : [0, T] \to \mathbb{R}$. Moreover, we consider only the case when $pq > 2$. The additional limitations of the functions are given in the next section.

The results presented in this section are published in [129].

4.1.1 Estimation

In this part, we present a procedure of estimating the parameters of the stochastic process given by Eq. (4.6).

Let us assume we have a vector of realization of the stochastic process (given by Eq. (4.6)) $x_0, x_1, \ldots, x_n$ with corresponding time points $t_0, t_1, \ldots, t_n$ such that $\forall_{i\in\{1,2,\ldots,n\}} t_i - t_{i-1} = \Delta$. For the sake of simplicity we assume $\Delta = 1$. Then the increments of the considered observations we denote as $y_0, y_1, \ldots, y_{n-1}$ where $y_i = x_{i+1} - x_i$ for $i = 0, 1, ..., n-1$.

To accomplish this, we first rewrite Eq. (4.6) to its discrete version:

$$y_i = x_{i+1} - x_i = \alpha_1(t_i) + \alpha_2(t_i)x_i + (\beta_1(t_i) + \beta_2(t_i)x_i)s_i,\ i = 0, 1, 2, \ldots, n-1; \tag{4.7}$$

where $\{s_i\}$ is a time series of iid random variables with the SGT distribution with parameters $\mu = 0,\ \sigma = 1$, unrestricted λ, and $p,\ q$ such that $pq > 2$.

In addition we apply the local regression approach [136], similarly as in [55], in order to obtain the estimates of $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions in the model (4.6). We assume here that $\alpha_1(\cdot) \in$

$\mathcal{C}^{d_1^\alpha}, \alpha_2(\cdot) \in \mathcal{C}^{d_2^\alpha}$ so that they can be expanded to Taylor's polynomials [137] in every time point $t^* \in \{t_0, \ldots, t_{n-1}\}$ of d_1^α and d_2^α degrees:

$$\alpha_w(t_i) = \sum_{k=0}^{d_w^\alpha} \frac{\alpha_w^{(k)}(t^*)}{k!}(t_i - t^*)^k + R_{d_w^\alpha}(t_i), \;\; w = 1, 2; \tag{4.8}$$

where $R_{d_w^\alpha}(\cdot)$ is Peano's remainder which we, in further considerations, neglect. After expanding (4.8) and gathering constants for common t_i^k, we arrive at the following approximation:

$$\alpha_w(t_i) \approx \sum_{k=0}^{d_w^\alpha} {}_k\alpha_w t_i^k, \;\; w = 1, 2. \tag{4.9}$$

In order to find ${}_k\alpha_w$ estimates from Eq. (4.34) for all t_i in b surrounding, we define the loss function as the weighted sum of squared errors. It should be noted that $\{{}_k\alpha_w\}$ are estimated separately for every timepoint $t^* \in \{t_0, \ldots, t_{n-1}\}$. From Eq. (4.7) we obtain the following:

$$s_i = \frac{y_i - (\alpha_1(t_i) + \alpha_2(t_i)x_i)}{\beta_1(t_i) + \beta_2(t_i)x_i} \approx \frac{y_i - \left(\sum_{k=0}^{d_1^\alpha} {}_k\alpha_1 t_i^k + \sum_{k=0}^{d_2^\alpha} {}_k\alpha_2 t_i^k x_i\right)}{\beta_1(t_i) + \beta_2(t_i)x_i} =: \tilde{s}_i, \; i = 0, 1, 2, \ldots, n-1. \tag{4.10}$$

We assume here that the loss function, applied in the estimation algorithm for each $t^* \in \{t_0, \ldots, t_{n-1}\}$, takes the following form:

$$L^*_{t^*,\beta,\tau,d_1^\alpha,d_2^\alpha,b^\alpha,b_r^\alpha}(\{x_i\}, \{t_i\}; \{{}_k\alpha_w\}) = \sum_{i=0}^{n-1} \tilde{s}_i^2 K_{b^\alpha,b_r^\alpha}(t_i - t^*) + \tau \left(\sum_{k=0}^{d_1^\alpha} {}_k\alpha_1^2 + \sum_{k=0}^{d_2^\alpha} {}_k\alpha_2^2\right). \tag{4.11}$$

The first component of the loss function, namely $\tilde{s}_i^2 K_{b^\alpha,b_r^\alpha}(t_i - t^*)$ is related to the fact that estimators are fitted locally (and not globally). Furthermore, similarly to Ridge regression [138], we have added to the loss function a second component - Tikhonov regularization [139] (with parameter τ). This regularization compensates for (possibly) not a unique solution and the high variance of the estimators. We used here a single-valued parameter τ however this can be swapped for a vector of $\{\tau_i\}$. This change yields better estimates but it needs the whole vector of $\{\tau_i\}$ that has to be found. We propose to use in Eq. (4.11) the asymmetric kernel function $K_{b,b_r}(\cdot)$ that we define as follows:

$$K_{b,b_r}(t) = 2\frac{K\left(\frac{t}{b-b_r}\right)1_{t\leq 0} + K\left(\frac{t}{b_r}\right)1_{t>0}}{b}, \tag{4.12}$$

where b is the width of the kernel function $K_{b,b_r}(\cdot)$ (distance from left root to right) and b_r is the distance to the right root from 0. This form of kernel enables to find a balance between the classical symmetric and causal kernel function such that estimators have smaller variance. In the estimation procedure the parameters $b^\alpha, b_r^\alpha, d_1^\alpha, d_2^\alpha$ and τ are called hyperparameters. In the applications, the most commonly used are three kernel functions $K(\cdot)$ in (4.12), [55, 136, 140, 141]:

- Epanechnikov kernel - $K(t) = \frac{3}{4}(1-t^2)1_{t\in(-1,1)}$;
- Tricube kernel - $K(t) = \frac{70}{81}(1-|t|^3)^3 1_{t\in(-1,1)}$;
- Gaussian kernel - $K(t) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)$.

Due to compact support, in problems related to financial data modelling, the Epanechnikov and tricube kernels are used [55, 136, 141]. In our applications we used the tricube kernel.
To simplify the calculations, in the first step of the estimation procedure we consider $\beta_1(\cdot)$ and $\beta_2(\cdot)$ functions in model (4.6) as they were known - we are using the iterative method of finding estimates with starting condition:

$$\xi_{i,\beta}^{(0)} := \hat{\beta}_1(t_i) + \hat{\beta}_2(t_i)x_i \equiv 1. \tag{4.13}$$

However, neither the optimal d_1^α, d_2^α values (see Eq. (4.34)) nor the kernel's widths $b^\alpha,\ b_r^\alpha$ are known. We find optimal values of hyperparameters d_1^α, d_2^α, b^α, b_r^α and τ (see Eq. (4.11)) to be the ones that entail the lowest values of the mean squared error (MSE) statistics:

$$\text{MSE}_y = \sum_{i=0}^{n-1}\left(y_i - \left(\sum_{k=0}^{d_1^\alpha} {}_k\hat{\alpha}_1 t_i^k + \sum_{k=0}^{d_2^\alpha} {}_k\hat{\alpha}_2 t_i^k x_i\right)\right)^2 \omega_i; \tag{4.14}$$

$$\text{MSE}_x = \sum_{i=1}^{n}\left(x_i - x_0 - \sum_{j=1}^{i}\left(\sum_{k=0}^{d_1^\alpha} {}_k\hat{\alpha}_1 t_j^k + \sum_{k=0}^{d_2^\alpha} {}_k\hat{\alpha}_2 t_j^k x_j\right)\right)^2 \omega_i \tag{4.15}$$

and Augmented Dickey–Fuller test [142] statistic (with null hypothesis that unit root is present in a time series data) for vector $\left\{y_i - \left(\sum_{k=0}^{d_1^\alpha} {}_k\hat{\alpha}_1 t_i^k + \sum_{k=0}^{d_2^\alpha} {}_k\hat{\alpha}_2 t_i^k x_i\right)\right\}$. The weights $\{\omega_i\}$ (in Eq. (4.14) and (4.15)) are generated using the method of exponential smoothing [143].

After obtaining the optimal hyperparameters' values, let us rewrite the loss function $L^*(\cdot)$ defined in (4.11) using the following matrices:

$$\boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} \quad \boldsymbol{\alpha} = \begin{pmatrix} {}_0\alpha_1 \\ {}_1\alpha_1 \\ {}_2\alpha_1 \\ \vdots \\ {}_{d_1^\alpha}\alpha_1 \\ {}_0\alpha_2 \\ {}_1\alpha_2 \\ \vdots \\ {}_{d_2^\alpha}\alpha_2 \end{pmatrix} \quad \boldsymbol{T} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ t_0 & t_1 & \dots & t_{n-1} \\ t_0^2 & t_1^2 & \dots & t_{n-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ t_0^{d_1^\alpha} & t_1^{d_1^\alpha} & \dots & t_{n-1}^{d_1^\alpha} \\ x_0 & x_1 & \dots & x_{n-1} \\ t_0 x_0 & t_1 x_1 & \dots & t_{n-1} x_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ t_0^{d_2^\alpha} x_0 & t_1^{d_2^\alpha} x_1 & \dots & t_{n-1}^{d_2^\alpha} x_{n-1} \end{pmatrix}$$

$$\boldsymbol{K}_{t^*} = \begin{pmatrix} \frac{K_{b^\alpha, b_r^\alpha}(t_0 - t^*)}{\left(\xi_{0,\beta}^{(0)}\right)^2} & 0 & \dots & 0 \\ 0 & \frac{K_{b^\alpha, b_r^\alpha}(t_1 - t^*)}{\left(\xi_{1,\beta}^{(0)}\right)^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{K_{b^\alpha, b_r^\alpha}(t_{n-1} - t^*)}{\left(\xi_{n-1,\beta}^{(0)}\right)^2}) \end{pmatrix}$$

Then the loss function (4.11) takes the form:

$$L^* = (\boldsymbol{y} - \boldsymbol{T}^\intercal \boldsymbol{\alpha})^\intercal \boldsymbol{K}_{t^*} (\boldsymbol{y} - \boldsymbol{T}^\intercal \boldsymbol{\alpha}) + \tau \boldsymbol{\alpha}^\tau \boldsymbol{\alpha}$$

which we minimize with respect to the vector $\boldsymbol{\alpha}$:

$$\frac{\partial L^*}{\partial \boldsymbol{\alpha}} = -2\boldsymbol{T}\boldsymbol{K}_{t^*} (\boldsymbol{y} - \boldsymbol{T}^\intercal \boldsymbol{\alpha}) + 2\tau \boldsymbol{I}\boldsymbol{\alpha} = 0.$$

We get:

$$(\boldsymbol{T}\boldsymbol{K}_{t^*}\boldsymbol{T}^\intercal + \tau \boldsymbol{I})\, \boldsymbol{\alpha} = \boldsymbol{T}\boldsymbol{K}_{t^*}\boldsymbol{y}.$$

Then, the $\boldsymbol{\alpha}$ estimates for t^* can be calculated from the following:

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{T}\boldsymbol{K}_{t^*}\boldsymbol{T}^\intercal + \tau \boldsymbol{I})^{-1} \boldsymbol{T}\boldsymbol{K}_{t^*}\boldsymbol{y}. \tag{4.16}$$

Having $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions from the model (4.6) estimated, we can estimate $\beta_1(\cdot)$ and $\beta_2(\cdot)$ functions. Similarly to $\alpha_1(\cdot), \alpha_2(\cdot)$ functions estimation, we will approximate $\beta_1(\cdot)$ and $\beta_2(\cdot)$

functions from the model (4.6) by Taylor's polynomials [137]:

$$\beta_w(t_i) \approx \sum_{k=0}^{d_w^\beta} {}_k\beta_w t_i^k, \;\; w = 1, 2. \tag{4.17}$$

Then, parameters $\{{}_k\beta_w\}$ can be found using the maximum likelihood method [136]. Exploiting properties proved in Eq. 4.3 and the fact that $\{s_i\}$ are iid, we see that the loglikelihood function can be written as follows:

$$l^*_{t^*,d_1^\beta,d_2^\beta,b^\beta,b_r^\beta}(\{\hat{e_i}\},\{t_i\};\{{}_k\beta_w\},\lambda,p,q) = \tag{4.18}$$
$$\sum_{i=0}^{n-1} \ln\left(f_{SGT}\left(\hat{e_i}; 0, \left(\sum_{k=0}^{d_1^\beta} {}_k\beta_1 t_i^k + \sum_{k=0}^{d_2^\beta} {}_k\beta_2 t_i^k x_i\right), \lambda, p, q\right)\right) K_{b^\beta,b_r^\beta}(t_i - t^*),$$

where $\{\hat{e_i}\}$ is obtained by transformation of the Eq. (4.7) in the following way:

$$\hat{e_i} := y_i - (\hat{\alpha_1}(t_i) + \hat{\alpha_2}(t_i)x_i) \approx (\beta_1(t_i) + \beta_2(t_i)x_i)s_i, \; i = 0, 1, 2, \ldots, n-1. \tag{4.19}$$

In this case, the additional optimal hyperparameters', namely d_w^β $(w = 1, 2)$, b^β and b_r^β, need to be found. We propose the use of the Breusch–Pagan test [144] statistic with null hypothesis that variance is independent of descriptive (independent) variables (a time series is homoscedastic). We prefer the set of hyperparameters d_w^β $(w = 1, 2)$, b^β and b_r^β that entail lowering the test statistic calculated for the time series:

$$\{\hat{s}_i\} = \left\{ \frac{y_i - \left(\sum_{k=0}^{d_1^\alpha} {}_k\hat{\alpha}_1 t_i^k - \sum_{k=0}^{d_2^\alpha} {}_k\hat{\alpha}_2 t_i^k x_i\right)}{\sum_{k=0}^{d_1^\beta} {}_k\hat{\beta}_1 t_i^k + \sum_{k=0}^{d_2^\beta} {}_k\hat{\beta}_2 t_i^k x_i} \right\}. \tag{4.20}$$

Having hyperparameters picked, we maximize the loglikelihood function (4.18) with respect to $\{{}_k\beta_w\}$ (defined in (4.17)) and unknown residuals' parameters $\lambda,\ p,\ q$. Since the analytical solution to the maximization of the function (4.18) does not exist, it is needed to find a maximum of the function using numerical algorithms. To simplify calculations we will benefit from the invariance property of maximum likelihood estimators [145]. Let us note that $\lambda \in (-1, 1)$, $p > 0$ and $q > 0$. We will maximize function $l^*(\cdot)$ (4.18) with respect to parameters: ${}_k\beta_w \in \mathbb{R}$ $(w = 1, 2;\ k = 0, \ldots, d_w^\beta)$, $\lambda^* \in \mathbb{R},\ p^* \in \mathbb{R},\ q^* \in \mathbb{R}$, using the substitution:

$$\lambda = \frac{2}{\pi}\text{arctan}(\lambda^*),\ p = \log(|p^*| + 1),\ q = \log(|q^*| + 1). \tag{4.21}$$

Then optimization can be done with a wider and simpler class of algorithms like Broyden–Fletcher–Goldfarb–Shanno algorithm [146].

The earlier proposition of $\xi_{i,\beta}^{(0)} \equiv 1$ (see equation (4.13)) can be highly doubtful, especially when the time series is obviously heteroskedastic. To compensate for this, we use the iterative method of $\alpha_1(\cdot)$, $\alpha_2(\cdot)$, $\beta_1(\cdot)$, $\beta_2(\cdot)$ estimation. In the next step of the estimation we set:

$$\left\{\xi_{i,\alpha}^{(1)}\right\} = \{\hat{\alpha}_1(t_i) + \hat{\alpha}_2(t_i)x_i\}$$

and

$$\left\{\xi_{i,\beta}^{(1)}\right\} = \left\{\hat{\beta}_1(t_i) + \hat{\beta}_2(t_i)x_i\right\}$$

and repeat the whole estimation procedure until the changes in the estimated functions are insignificant, i.e. while

$$\exists_i \left|\xi_{i,\alpha}^{(\gamma)} - \xi_{i,\alpha}^{(\gamma-1)}\right| > \epsilon_\alpha \text{ or } \exists_i \left|\xi_{i,\beta}^{(\gamma)} - \xi_{i,\beta}^{(\gamma-1)}\right| > \epsilon_\beta,$$

where $\epsilon_\alpha, \epsilon_\beta$ are some defined thresholds and γ is the number of the current iteration (or after a given number of iterations).

After estimating the $\alpha_1(\cdot)$, $\alpha_2(\cdot)$, $\beta_1(\cdot)$, and $\beta_2(\cdot)$ functions, we can finally estimate the global parameters of residuals $\{\hat{s}_i\}$ (defined in (4.20)) that are modeled by SGT distribution (in the previous steps of the estimation procedure only the local estimates of the parameters are found). We find $\hat{\lambda}$, $\hat{p}$, $\hat{q}$ by numerically maximizing the likelihood function with respect to λ^*, p^* and q^* (see the substitution proposed in (4.21)):

$$L(\{\hat{s}_i\}; \lambda^*, p^*, q^*) = \prod_{i=0}^{n-1} f_{SGT}(\hat{s}_i; 0, 1, \lambda(\lambda^*), p(p^*), q(q^*)). \tag{4.22}$$

A schematic algorithm of the parameters' estimation procedure is presented in Fig. 4-3.

4.1.2 Simulation study

Using the methodology presented in Section 4.1.1 we have checked the performance of the estimation procedure using simulated data. Utilizing Euler's method [147] we have generated the trajectory of the process given by the following stochastic differential equation:

$$dX_t = [0.1 + 0.025t - 0.015X_t]dt + [0.03 + 0.001t]dS_t, \tag{4.23}$$

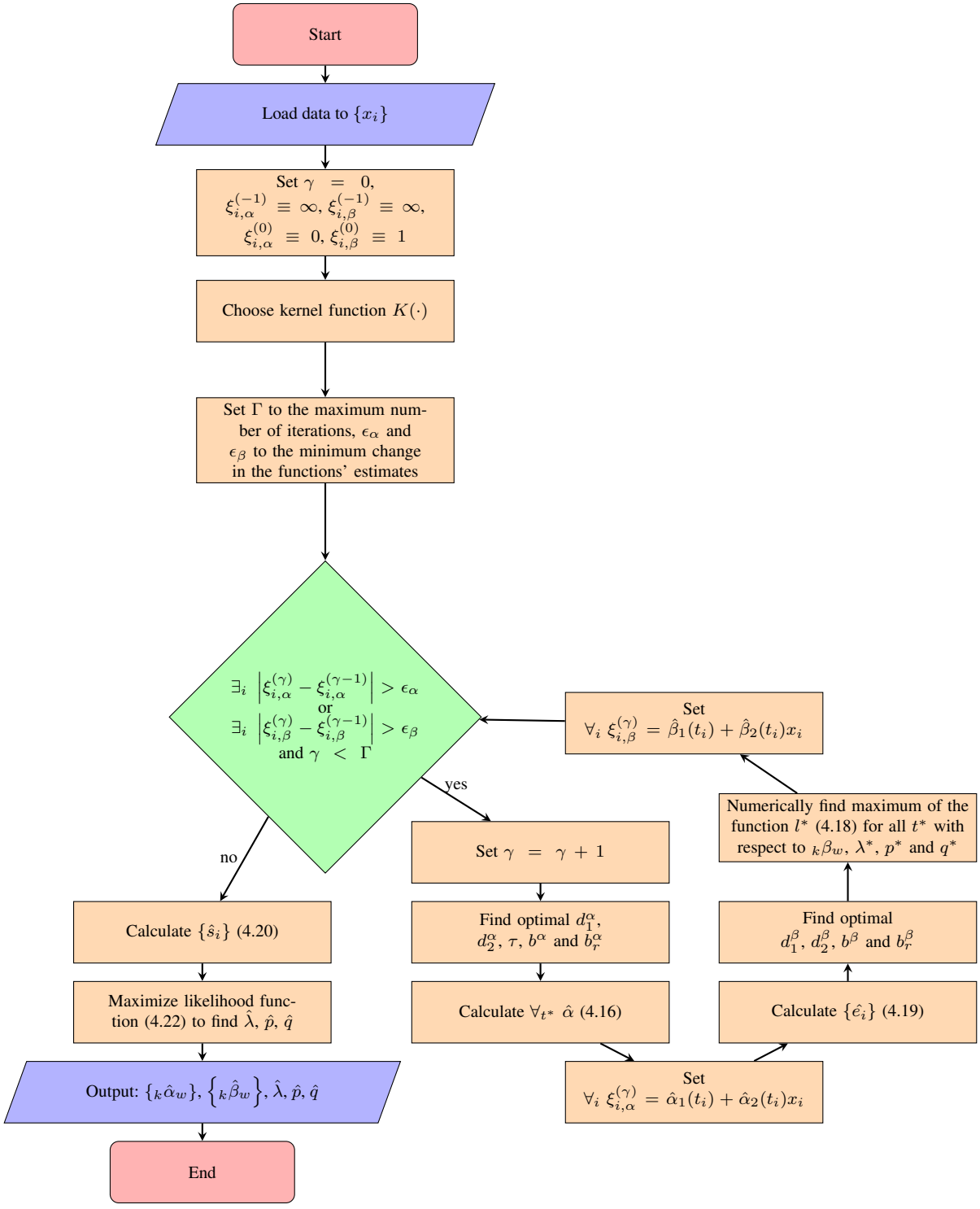
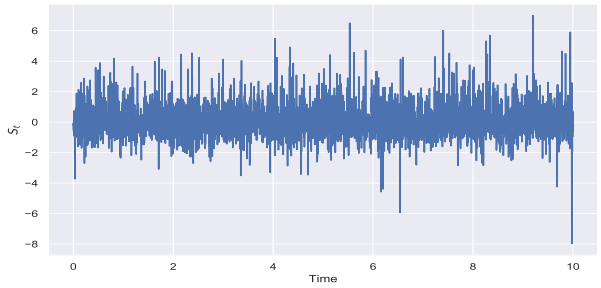
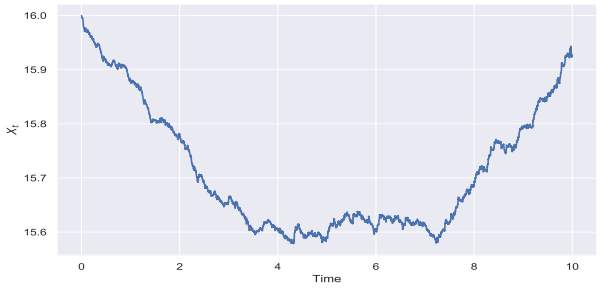


Figure 4-3: The schematic algorithm of the estimation procedure, source: [129]

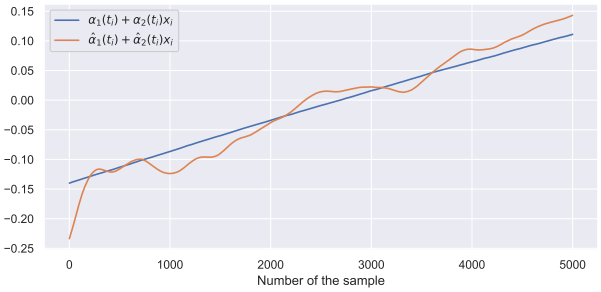


(a) Exemplary realization of the process $\{dS_t\}$.

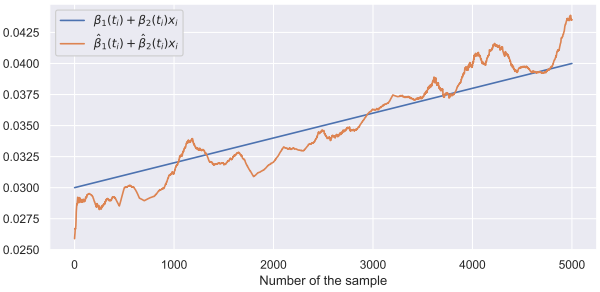


(b) Exemplary realization of the process $\{X_t\}$.

Figure 4-4: The exemplary realizations of the stochastic process defined by the equation (4.44) with residuals having SGT distribution (4.24), source: [129]



(a) Comparison of the theoretical function $\alpha_1(t) + \alpha_2(t)x_t = 0.1 + 0.025t - 0.015x_t$ with its estimate.



(b) Comparison of the theoretical function $\beta_1(t) + \beta_2(t)x_t = 0.03 + 0.001t$ with its estimate.

Figure 4-5: Estimation of the model's parameters for simulated time series, source: [129]

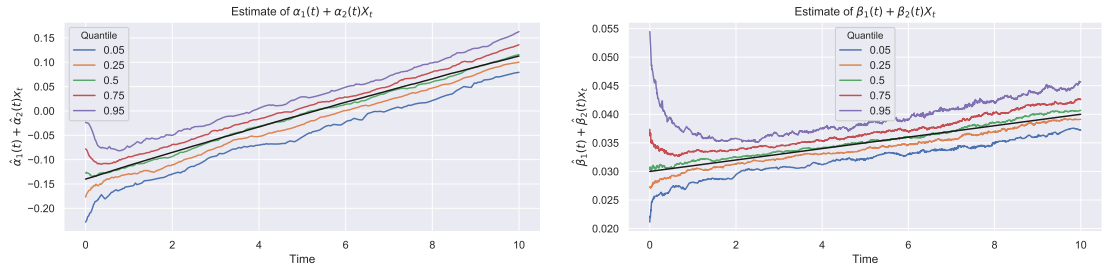
under the assumption:

$$dS_t \sim SGT(\mu = 0, \sigma = \sqrt{dt}, \lambda = 0.25, p = 1.5, q = 3), \tag{4.24}$$

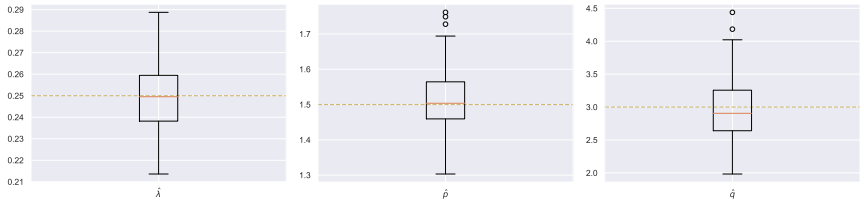
and $t \in [0, 250 \cdot 10]$ (10 years with 250 working days). The exemplary realization of the processes $\{dS_t\}$ and $\{X_t\}$ are presented in Fig. 4-4. The parameters of the model are chosen arbitrarily in order for them to be in some sense comparable to the parameters obtained for the real data.

Using the presented methodology, we have estimated $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and $\beta_1(\cdot)$, $\beta_2(\cdot)$ functions from the model (4.6). In Fig. 4-5 we present the estimated functions and the theoretical ones from the model (4.44). We observe that the estimates correspond to the theoretical functions.

The estimated parameters of the SGT distribution are $\hat{\lambda} = 0.246$, $\hat{p} = 1.475$ and $\hat{q} = 3.213$ - they are close to the theoretical parameters $\lambda = 0.25$, $p = 1.5$ and $q = 3$. Furthermore, based on the Kolmogorov–Smirnov test [148] statistic $K = 0.00715$ and p-value $= 0.960$, we conclude that the model's parameters were correctly estimated.



(a) Estimation of the $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions. (b) Estimation of the $\beta_1(\cdot)$ and $\beta_2(\cdot)$ functions.



(c) Box-plots of SGT distribution's parameters' estimates.

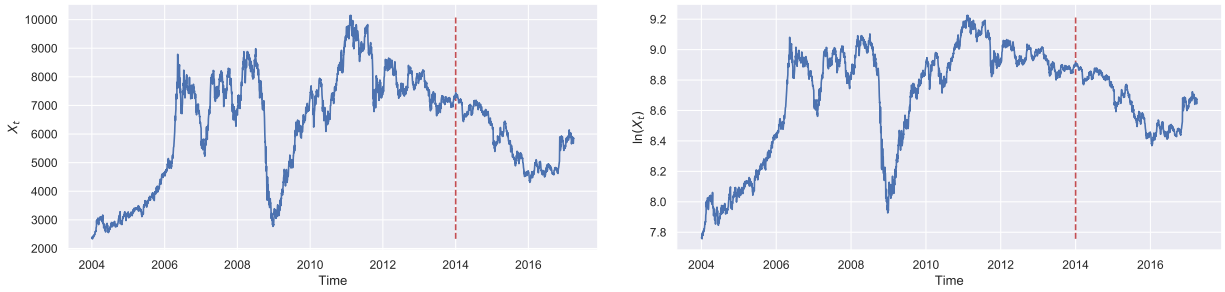
Figure 4-6: Estimation of the parameters for the model (4.44) for 100 Monte Carlo simulations, source: [129]

Furthermore, we have performed the Monte-Carlo [149] simulations for the process given by Eq. (4.44). More precisely, we have simulated 100 realizations of the process and performed the estimation methodology presented in the previous section. Finally, we found estimates of $[0.05, 0.25, 0.5, 0.75, 0.95]$-quantiles for $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and $\beta_1(\cdot)$, $\beta_2(\cdot)$ functions' estimators. In Figs. 4-6a and 4-6b we present the result. In both Figs. 4-6a and 4-6b we see that the first few points of the functions' estimators demonstrate significant variance. This is due to the low number of samples taken into the functions' estimation. We have also plotted the box-plots of the estimated parameters of the SGT distribution, see Fig. 4-6c. For every estimated parameter, we observe that medians are close to theoretical values and the variance of the estimated parameters is low.

4.1.3 Real data application

In this part, we analyse the real data series describing the price of copper, while the results for other metals (silver and gold) are presented in [129].

For the considered time series we will demonstrate that the proposed model (4.6) based on the SGT distribution is acceptable. Finally, we will show the results of the long-term prediction based on the fitted model. Moreover, in the estimation procedure, we take $\Gamma = 2$ (two iterations of estimation) and applied the tricube to be the kernel function $K(\cdot)$ in (4.12). It should be noted that the real data related to the copper prices are only used here for the illustration of the introduced methodology. At the end of this part we also include conclusions from a similar application of



(a) Copper prices data.

(b) Copper data after Box-Cox transformation.

Figure 4-7: Copper price data, source: [129]

the model to two other metals (silver and gold), which should prove that the proposed model is universal and can also be applied to real data from different areas.

In the first real example, we analyse the time series corresponding to the copper price. The data set consists of 3348 observations from the period 02.01.2004 to 31.03.2017. In Fig. 4-7 we demonstrate the considered data. Due to unstable variance, visible in the vector of observations, we transformed the data taking the Box-Cox transformation [150]:

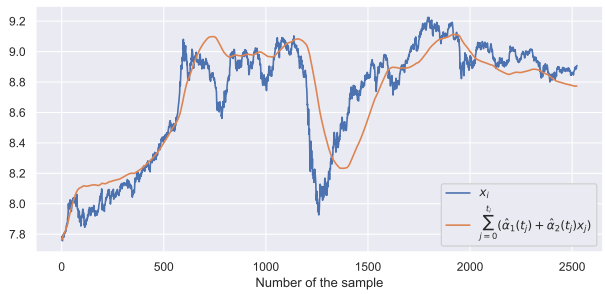
$$\forall_t \; x_t = \ln(x_t^*),$$

where $\{x_t^*\}$ is the vector of the copper price data. We present the transformed vector in Fig. 4-7b. We split the data into training time series (used for the estimation of the model's parameters) - the period from the beginning of 2004 to the end of 2013 and a testing time series (used for the validation of the model) - the period from 2014 up to 31.03.2017. The vector of training time series consists of 2525 observations while the time series used for validation has 823 observations.

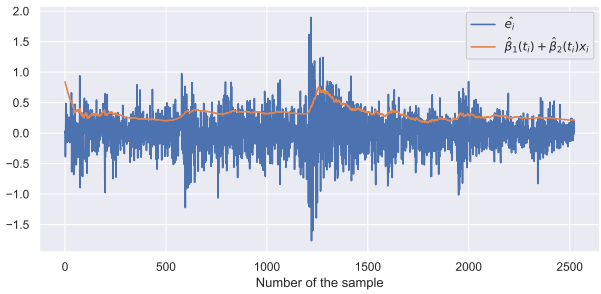
In the first step, we find the optimal hyperparameters $d_1^\alpha,\ d_2^\alpha,\ b^\alpha,\ b_r^\alpha,\ \tau$ needed for minimization of the loss function (4.11). As proposed in Section 4.1.1, we use MSE_x (4.15), MSE_y (4.14) and Augmented Dickey–Fuller test statistic. We get weights $\{\omega_i\}$ (for statistics MSE_x and MSE_y - see Eq. (4.15) and (4.14)) using the exponential smoothing method [143] with smoothing parameter $\phi = 8 \times 10^{-4}$. Weights $\{\omega_i\}$ are calculated using the following formula:

$$\omega_i = \frac{1 - \exp(-\phi)}{1 - \exp(-n\phi)} \exp(\phi(i - n)),\ i = 1, \ldots, n. \tag{4.25}$$

With selected hyperparameters $d_1^\alpha = 0$, $d_2^\alpha = 1$, $b^\alpha = 750$, $b_r^\alpha = 237.5$, $\tau = 0.7$ we estimate $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions from the model (4.6) using Eq. (4.16). Acquired estimates are presented in Fig. 4-8a. We see that the estimates are well fitted to the data. However, this might



(a) Comparison of estimated trend line to the examined time series.



(b) Comparison of composition of $\{\hat{\beta}_1(t_i)+\hat{\beta}_2(t_i)x_i\}$ with $\{\hat{e}_i\}$ series.

Figure 4-8: Copper price data: resulting estimates from the first iteration of estimation of $\alpha_1(\cdot),\ \alpha_2(\cdot)$ and $\beta_1(\cdot),\ \beta_2(\cdot)$ functions, source: [129]

also indicate that they are overfitted. Due to the heteroskedasticity property of the time series, in the first iteration, the change in the process's variance could be wrongly assigned to the $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions. To compensate for this effect we take advantage of the iterative estimation method (discussed in Section 4.1.1).

We calculated the vector $\{\hat{e}_i\}$ using formula (4.19) and we found optimal hyperparameters for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ functions' estimation $d_1^\beta = 0$, $d_2^\beta = 0$, $b^\beta = 1000$, $b_r^\beta = 10$. The appropriate estimates were found by the maximizing function (4.18). Then based on acquired estimates, their composition was calculated and shown in Fig. 4-8b with a series of $\{\hat{e}_i\}$ for comparison. We see that the composition of $\beta_1(\cdot)$ and $\beta_2(\cdot)$ functions estimates is a good approximation of the standard deviation of the examined time series.

Next, we repeat the previously presented steps taking $\xi_{i,\beta}^{(1)} = \hat{\beta}_1(t_i) + \hat{\beta}_2(t_i)x_i$. We get the following results:

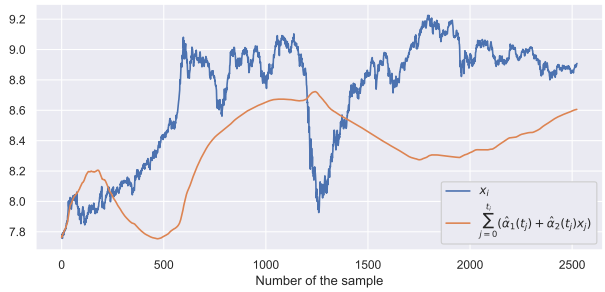
1. Selected hyperparameters for $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions estimation - $d_1^\alpha = 0$, $d_2^\alpha = 0$, $b^\alpha = 937.5$, $b_r^\alpha = 1.25$, $\tau = 8 \times 10^{-3}$;

2. The estimated $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions - see Fig. 4-9a;

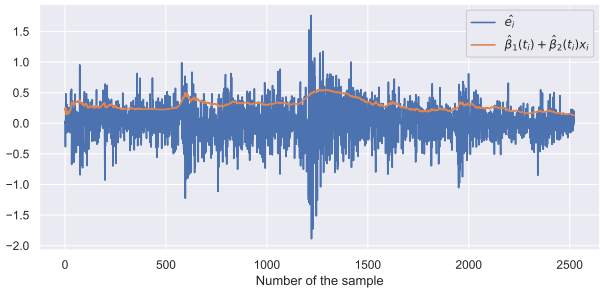
3. Picked hyperparameters for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ functions estimation - $d_1^\beta = 1$, $d_2^\beta = 0$, $b^\beta = 1437.5$, $b_r^\beta = 60$;

4. The estimated $\beta_1(\cdot)$ and $\beta_2(\cdot)$ functions - see Fig. 4-9b.

Having the $\alpha_1(\cdot),\ \alpha_2(\cdot)$ and $\beta_1(\cdot),\ \beta_2(\cdot)$ functions estimated, the estimation of residuals' distribution's parameters is performed maximizing the likelihood function (4.22). The obtained estimates of those parameters are $\hat{\lambda} = -0.08$, $\hat{p} = 2.32$ and $\hat{q} = 3.11$. From the analysis of the



(a) Comparison of estimated trend line to the examined time series.



(b) Comparison of composition of $\{\hat{\beta}_1(t_i)+\hat{\beta}_2(t_i)x_i\}$ to $\{\hat{e}_i\}$ series.

Figure 4-9: Copper price data: resulting estimates from the second iteration of estimation of $\alpha_1(\cdot),\ \alpha_2(\cdot)$ and $\beta_1(\cdot),\ \beta_2(\cdot)$ functions, source: [129]

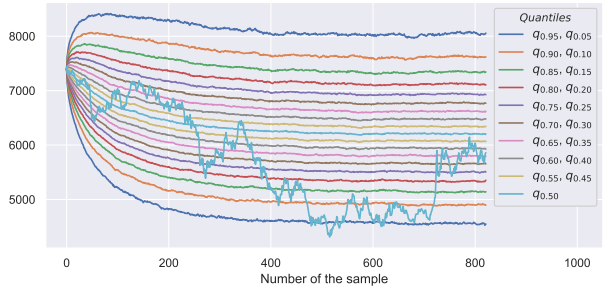
parameters' estimates, using the SGT distributions family tree (see Fig. 4-1), we conclude that the time series $\{s_i\}$ can be approximated by the Generalized Student's t-distribution. Furthermore, we observe that it can be roughly approximated by standard Student's t-distribution, considering the following equality:

$$SGT(\mu = 0, \sigma = 1, \lambda = 0, p = 2, q = \nu/2) \stackrel{d}{=} \mathcal{T}(\nu),$$

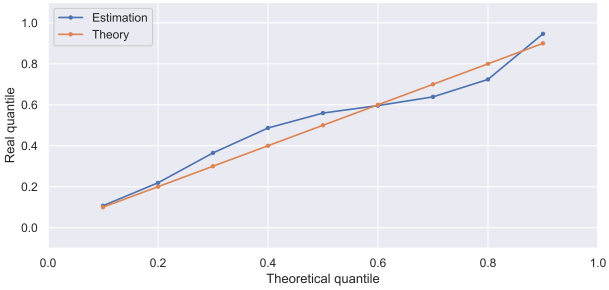
where $\mathcal{T}(\nu)$ denotes the Student's t-distribution with ν degrees of freedom. Then distribution with estimated parameters $SGT(0, 1, \hat{\lambda}, \hat{p}, \hat{q})$ is Student's t-distribution with 6 degrees of freedom.

We have checked if the fitted model properly describes the data by constructing the quantile lines using the Monte-Carlo method performing 10 000 simulations of the process given by (4.6) with the estimated parameters. We have plotted the vector of the real data (from the testing set) with marked quantile lines (see Fig. 4-10a). As can be seen, the time series used for validation perfectly falls into the constructed quantile lines. To confirm that the realizations of the process (from the validation period) fall into the constructed confidence intervals (from constructed quantile lines) with the right probabilities, we use here the method described in [127], where the simple visual validation factor for proper model recognition was proposed. For perfect models, $1-\alpha$ confidence intervals should include data with probability $1-\alpha$. Thus, we present the percent of the data falling inside of the estimated confidence intervals. For our model (see Fig. 4-10b) we see that the real and theoretical probabilities are similar. Therefore, we conclude that the model correctly predicted fluctuations in the process. We also have shown the whole vector of the data with quantile lines for the tested period, see Fig. 4-10c.

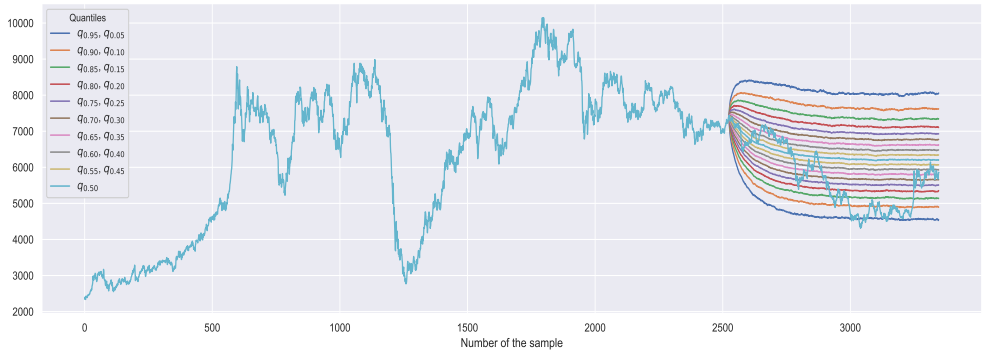
Finally, we have tested the residuals $\{s_i\}$ using the Kolmogorov–Smirnov (KS) test [148] - we obtain the value of the statistic $K = 0.00888$ with p-value = 0.989 thus the null hypothesis,



(a) Copper prices data with the constructed quantile lines.



(b) Visual validation factor that compares the estimated and theoretical probabilities of the data falling inside confidence intervals.



(c) Copper prices data with constructed quantile lines by using Monte Carlo simulations.

Figure 4-10: Validation of the fitted model for copper price data, source: [129]

that $\{s_i\}$ has the SGT distribution with estimated parameters, can not be rejected. We compare the efficiency of the model with SGT distribution with the corresponding model with Gaussian distribution. In the Gaussian case the KS test yields the test statistic $K = 0.02622$ with p-value = 0.062.

Thus, we can conclude that the considered model based on the SGT distribution (4.6) is fitted correctly and accurately describes the copper price time series characteristics. Furthermore, it outperforms the model with residuals modeled with Gaussian distribution.

Similar analysis for the silver and gold prices data have been presented in [129].

4.2 Non-Gaussian regime-switching model

In this section we describe the second considered stochastic model. Similarly, as the model discussed in Section 4.1, it is based on the SGT distribution introduced in (4.1) and stochastic differential equation (4.5). However this particular model additionally takes into account one very important characteristics of commodities markets - the change of regimes. However, considering the fact that it is difficult to perfectly indicate the moment of regimes change, the model could be more useful for post-factum analysis.

We start by introducing the stochastic process that is often used for a description of financial data [34, 51, 54, 151]:

$$dX_t = (\alpha_1(t) + \alpha_2(t)X_t)\,dt + (1 - H_t)\sigma_1 dS_t^1 + H_t\sigma_2 dS_t^2, \tag{4.26}$$

where $\{S_t^1\}$ and $\{S_t^2\}$ are independent processes of stationary independent increments having SGT distribution.

Factors $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ appearing in Eq. (4.26) are some functions dependent on time - $\alpha_1(\cdot),\ \alpha_2(\cdot) : [0, T] \to \mathbb{R}$. The process $\{H_t\}$ is defined as follows:

$$H_t = \begin{cases} 0, & \text{if } X_t \in C_1; \\ 1, & \text{if } X_t \in C_2. \end{cases} \tag{4.27}$$

We use the notation $X_t \in C_1$ to indicate that the process $\{X_t\}$ at time t is in a state/case C_1. We assume that we model data that we suspect of having two states of volatility. For example, in metals prices we might discern periods when a metal keeps its value with small deviations or periods with sudden rises and drops of value. The scale of these deviations is controlled by parameters $\sigma_1, \sigma_2 \in \mathbb{R}_+$ for the states C_1 and C_2, respectively.

The results presented in this section have been published in [130].

4.2.1 Estimation

For the estimation procedure, we will first use the discretized version of Eq. (4.26) with $dt \to 1$ ($\forall_i\ t_{i+1} - t_i = 1$):

$$Y_{t_i} := X_{t_{i+1}} - X_{t_i} = \alpha_1(t_i) + \alpha_2(t_i)X_{t_i} + (1 - H_{t_i})S^1_{t_i} + H_{t_i}S^2_{t_i}, \tag{4.28}$$

where $\{S^k_{t_i}\}$ ($k = 1, 2$; $i = 1, 2, \ldots, n$) are independent random variables and have the following distributions:

$$\begin{aligned} S^1_{t_i} &\sim SGT(\mu_1, \sigma_1, \lambda_1, p_1, q_1), \\ S^2_{t_i} &\sim SGT(\mu_2, \sigma_2, \lambda_2, p_2, q_2). \end{aligned} \tag{4.29}$$

In the business applications we actually consider the realization of the process $\{X_t\}$ given in (4.28). In the further analysis we denote the vector of realisations of the process $\{X_t\}$ as $\boldsymbol{x} = \{x_i\}$ while its increments as $\boldsymbol{y} = \{y_i\} = \{x_{i+1} - x_i\}$, $i = 1, 2, \ldots, n$.

For the estimation of $\alpha_1(t_i) + \alpha_2(t_i)x_{t_i}$ we will use a modified method of least squares. Let us define the loss function of the classic least squares method [152] for our model (4.28):

$$L(\boldsymbol{x}, \boldsymbol{y}; \{{}_k\alpha_w\}) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - (\alpha_1(t_i) + \alpha_2(t_i)x_{t_i}))^2 .$$

However, this choice of loss function is not well suited for noise with heavy tails. Thus we will use a loss function which makes estimates less sensitive to outliers. We could utilize the weighted least squares method [153], but there is a problem in choosing the proper weights, especially when we assume that residuals do not have to have a finite first moment. Another loss function that tries to handle outliers is the Huber loss function [154]:

$$L(\boldsymbol{y}, \hat{\boldsymbol{y}}; \delta) = \begin{cases} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 , & \text{for } |y_i - \hat{y}_i| \le \delta; \\ \sum_{i=1}^{n} 2\delta|y_i - \hat{y}_i| - \delta^2, & \text{for } |y_i - \hat{y}_i| > \delta. \end{cases} \tag{4.30}$$

Merging together the mean absolute and mean squared loss functions, results in better performance by Huber loss function when dealing with heavy tailed data. The parameter δ in Eq. (4.30) defines the distance from $\hat{y}$ that is needed for y_i to be treated as an outlier.

In our research, we will use the Charbonnier/pseudo-Huber loss function [155]. This function is a modification of the Huber loss function that avoids assembling $|x|$ and x^2 explicitly - for small values of x the Charbonnier loss function behaves like x^2 and for large values like $|x|$. It is given by the following formula:

$$L(\boldsymbol{y}, \hat{\boldsymbol{y}}; \delta) = \sum_{i=1}^{n} \delta^2 \left(\sqrt{1 + \left(\frac{y_i - \hat{y}_i}{\delta} \right)^2} - 1 \right). \tag{4.31}$$

Here, the δ parameter has a similar meaning to the one in (the previously presented) Huber loss function standard (4.30).

To capture the local changes in trend in the data we will use the local regression method [136]. First let us define the pseudo-Huber function as follows:

$$f_H(y - \hat{y}; \delta) = \delta^2 \left(\sqrt{1 + \left(\frac{y - \hat{y}}{\delta} \right)^2} - 1 \right). \tag{4.32}$$

Then, to find local estimates of the parameters, we will minimize the loss function of the following form:

$$L_{t^*}(\boldsymbol{x}, \boldsymbol{y}; \{{}_k\alpha_w\}) = \sum_{i=1}^{n} f_H(y_i - \alpha_1(t_i) + \alpha_2(t_i)x_{t_i}) K_{b,b_r}(t_i - t^*) + \tau \sum_{w=1}^{2} \sum_{k=0}^{d_w} {}_k\alpha_w^2. \tag{4.33}$$

The functions $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ (from Eq. (4.28)) are locally approximated by Taylor's polynomials [137] of d_1 and d_2 degree (we assume that $\alpha_1 \in \mathcal{C}^{d_1}$ and $\alpha_2 \in \mathcal{C}^{d_2}$):

$$\alpha_w(t_i) \approx \sum_{k=0}^{d_w} \frac{\alpha_w^{(k)}(t^*)}{k!} (t_i - t^*)^k = \sum_{k=0}^{d_w} {}_k\alpha_w t_i^k, \;\; w = 1, 2. \tag{4.34}$$

The function $K_{b,b_r}(\cdot)$ is an asymmetric kernel function (proposed in [129]). We will use here the asymmetric tricube kernel function [141] given by the following Eq.:

$$K_{b,b_r}(t) = 2 \frac{K\left(\frac{t}{b - b_r}\right) 1_{t \leq 0} + K\left(\frac{t}{b_r}\right) 1_{t > 0}}{b}, \tag{4.35}$$

where $K(\cdot)$ is tricube kernel function [141]:

$$K(t) = \frac{70}{81} (1 - |t|^3)^3 1_{t \in (-1,1)}.$$

Hyperparameters b and b_r are (in order) the width of the kernel and the distance from 0 to the right root of the kernel.

The last term in the loss function (4.33) is Tikhonov regularization [139]. We will exploit this regularization (similarly to ridge regression [138]) with one hyperparameter $\tau \in \mathbb{R}_+$.

We minimize the loss function (4.33) using numeric minimization methods such as the Broyden-Fletcher-Goldfarb-Shanno algorithm [146]. To ease computations, we can pass Jacobian of the loss function (4.33). Knowing the derivative of the pseudo-Huber function (4.32):

$$f'_H(y-\hat{y};\delta) = \frac{y-\hat{y}}{\sqrt{1+\left(\frac{y-\hat{y}}{\delta}\right)^2}},$$

we can easily compute the Jacobian:

$$\begin{aligned}
\frac{\partial L_{t^*}}{\partial_k \alpha_1} &= \sum_{i=1}^{n} -f'_H(y_i-\hat{y};\delta) K_{b,b_r}(t_i-t^*) t_i^k + 2\tau_k \alpha_1; \\
\frac{\partial L_{t^*}}{\partial_k \alpha_2} &= \sum_{i=1}^{n} -f'_H(y_i-\hat{y};\delta) K_{b,b_r}(t_i-t^*) x_i t_i^k + 2\tau_k \alpha_2.
\end{aligned} \tag{4.36}$$

Let us define:

$$W_{t_i} := H_{t_i} S^1_{t_i} + (1-H_{t_i}) S^2_{t_i}$$

as the detrended time series (4.28). Then, using estimates of $\{{}_k\alpha_w\}$, we remove a drift from the data:

$$\hat{w}_i = y_i - \sum_{k=0}^{d_1} {}_k\hat{\alpha}_1 t_i^k - \sum_{k=0}^{d_2} {}_k\hat{\alpha}_2 x_i t_i^k \approx h_i s_i^1 + (1-h_i) s_i^2.$$

Assuming that $\{H_{t_i}\}$ is a Markov chain [156] and for any measurable set $\mathbb{A}$ the following Eq. holds [157]:

$$P(W_{t_i} \in \mathbb{A} | H_{t_1} = h_1, \ldots, H_{t_i} = h_i) = P(W_{t_i} \in \mathbb{A} | H_{t_i} = h_i),$$

then we can use estimation methodology for Hidden Markov Models for continuous distributions with two hidden states.

Let us define:

$$\zeta_{t_i|t_j}(k) = P(H_{t_i} = k | W^{(t_j)}; \mathbb{M}). \tag{4.37}$$

Namely, it is a probability of $\{W_{t_i}\}$ being in a state k at a time t_i under condition of data up to the time t_j ($W^{(t_j)} = \{W_t\}_{t=t_1}^{t_j}$) and the set of the model's parameters $\mathbb{M} = \{\mu_1, \sigma_1, \lambda_1, p_1, q_1, \mu_2, \sigma_2,$

$\lambda_2, p_2, q_2, \boldsymbol{P}, \boldsymbol{\eta}\}$. In the set $\mathbb{M}$, $\boldsymbol{P}$ is a matrix of transition probabilities $\boldsymbol{P} = \{p_{ij}\}$:

$$\forall_k \, p_{ij} = P(H_{t_{k+1}} = i | H_{k_i} = j), \; i, j = 0, 1$$

and $\boldsymbol{\eta}$ is a vector of initial distribution, namely $\eta_i = P(H_{t_1} = i)$, $i = 0, 1$. The other parameters that need to be estimated correspond to the SGT distribution for the C_1 and C_2 states.

To estimate $\boldsymbol{\zeta}_{t_i|t_i}$ $(t_i = 1, 2, \ldots, T)$ we use the following Eq. [158]:

$$\hat{\boldsymbol{\zeta}}_{t_i|t_i} = \frac{\hat{\boldsymbol{\zeta}}_{t_i|t_{i-1}} \odot \boldsymbol{\xi}_{t_i}}{1'(\hat{\boldsymbol{\zeta}}_{t_i|t_{i-1}} \odot \boldsymbol{\xi}_{t_i})}, \tag{4.38}$$

$$\hat{\boldsymbol{\zeta}}_{t_{i+1}|t_i} = \boldsymbol{P}\hat{\boldsymbol{\zeta}}_{t_i|t_i}; \tag{4.39}$$

with starting condition $\hat{\boldsymbol{\zeta}}_{t_1|t_0} = \boldsymbol{\eta}$. The $1'$ indicates the transposition of the matrix 1. In the Eq. (4.38), the $\odot$ symbol stands for the Hadamard product (element-wise multiplication), 1 is a vector of ones, namely: $1 = \begin{bmatrix} 1 & 1 \end{bmatrix}'$ and $\boldsymbol{\xi}_{t_i}$ is given by:

$$\boldsymbol{\xi}_{t_i} = \{f_{SGT}(w_i | h_i = j - 1; \mu_j, \sigma_j, \lambda_j, p_j, q_j)\}_{j=1,2}.$$

Using those matrices we can compute the log-likelihood function [159]:

$$l(\mathbb{M}) = \sum_{t=t_1}^{t_n} \ln \left(1'(\hat{\boldsymbol{\zeta}}_{t_i|t_{i-1}} \odot \boldsymbol{\xi}_{t_i}) \right). \tag{4.40}$$

Using Kim's algorithm [160] we can compute probabilities conditioned on a whole data set (calculate starting from t_{n-1} down to t_1):

$$\hat{\boldsymbol{\zeta}}_{t_i|t_n} = \hat{\boldsymbol{\zeta}}_{t_i|t_i} \odot \left[\boldsymbol{P} \left(\hat{\boldsymbol{\zeta}}_{t_i|t_n} \oslash \hat{\boldsymbol{\zeta}}_{t_{i+1}|t_i} \right) \right], \tag{4.41}$$

where $\oslash$ is the Hadamard division operator.

Then using Hamilton's estimator [161] for transition probabilities we have:

$$\hat{p}_{ij} = p_{ij} \frac{\sum_{t=t_2}^{t_n} P(h_t = j | w^{(t_n)}; \hat{\mathbb{M}}) \frac{P(h_{t-1} = i | w^{(t-1)}; \hat{\mathbb{M}})}{P(h_t = j | w^{(t-1)}; \hat{\mathbb{M}})}}{\sum_{t=t_2}^{t_n} P(h_{t-1} = i | w^{(t_n)}; \hat{\mathbb{M}})}. \tag{4.42}$$

Afterwards, we estimate parameters of SGT distributions for both states using the maximum likelihood estimation method. For each subset:

$$\mathbb{W}_0 = \{\hat{w}_i : \boldsymbol{\zeta}^{(0)}_{t_i|t_n} > \boldsymbol{\zeta}^{(1)}_{t_i|t_n}\} \text{ and } \mathbb{W}_1 = \{\hat{w}_i : \boldsymbol{\zeta}^{(0)}_{t_i|t_n} \leq \boldsymbol{\zeta}^{(1)}_{t_i|t_n}\}$$

($\boldsymbol{\zeta}^{(k)}_{t_i|t_n}$ indicates k$^{\text{th}}$ element of vector $\boldsymbol{\zeta}_{t_i|t_n}$) we maximize log-likelihood function:

$$l(\mathbb{W}_{k-1}; \mu_k, \sigma_k, \lambda_k, p_k, q_k) = \sum_{w \in \mathbb{W}_{k-1}} \ln\left(f_{SGT}(w; \mu_k, \sigma_k, \lambda_k, p_k, q_k)\right), \; k = 1, 2, \tag{4.43}$$

using e.g. Nelder-Mead algorithm [162].

The whole procedure is repeated until convergence i.e. until a change in the log-likelihood function (4.40) is less than some small selected constant ε. All of the estimation steps are gathered in pseudo-code (Algorithm 1).

Algorithm 1: Estimation algorithm

[1]

$\delta \leftarrow$ selected hyperparameter defined alongside the Huber loss function (4.30)
$d_1, d_2 \leftarrow$ selected hyperparameters for $\alpha_1(\cdot), \alpha_2(\cdot)$ functions (4.34)
$b, b_r \leftarrow$ selected hyperparameters defined for the asymmetric kernel function (4.35)
$\boldsymbol{\eta} \leftarrow$ selected initial distribution ie. $\begin{bmatrix}1 & 0\end{bmatrix}'$
$\boldsymbol{P} \leftarrow$ selected initial transition matrix ie. $\begin{bmatrix}12 & 12\\ 12 & 12\end{bmatrix}$
$\varepsilon \leftarrow$ Some small constant greater than zero $n \leftarrow length(\boldsymbol{y})$ $\hat{\boldsymbol{\alpha}} \leftarrow [\,]$ **for**
$t^* \leftarrow [t_1, t_2, \ldots, t_n]$ **do**

$\{{}_k\alpha_w\} \leftarrow \underset{\{{}_k\alpha_w\}}{\arg\min}(L_{t^*}(\boldsymbol{x}, \boldsymbol{y}; \{{}_k\alpha_w\}))$ Minimize L (4.33) using Jacobian (4.36)
$\boldsymbol{\alpha}[t^*] \leftarrow \sum_{k=0}^{d_1} {}_k\hat{\alpha}_1 t^{*k} + \sum_{k=0}^{d_2} {}_k\hat{\alpha}_2 x_i t^{*k}$ $\hat{\boldsymbol{w}} \leftarrow \boldsymbol{y} - \hat{\boldsymbol{\alpha}}$
$\hat{\mu}_1, \hat{\mu}_2 \leftarrow$ random data point from $\boldsymbol{y}$ $\hat{\sigma}_1 \leftarrow std(\boldsymbol{y})$; $\hat{\sigma}_2 \leftarrow 2 \cdot std(\boldsymbol{y})$ $\hat{\lambda}_1, \hat{\lambda}_2 \leftarrow 0$;
$\hat{p}_1, \hat{p}_2 \leftarrow 2$; $\hat{q}_1, \hat{q}_2 \leftarrow \infty$ $L(\mathbb{M}) \leftarrow 0$; $L(\hat{\mathbb{M}}) \leftarrow -\infty$ **while** $|L(\mathbb{M}) - L(\hat{\mathbb{M}})| > \varepsilon$ **do**

$L(\mathbb{M}) \leftarrow L(\hat{\mathbb{M}})$; $L(\hat{\mathbb{M}}) \leftarrow 0$ $\hat{\zeta}_{t+1|t} \leftarrow [\eta]$; $\hat{\zeta}_{t|t} \leftarrow []$ **for** $i \leftarrow 1 : n$ **do**

$\hat{\zeta}_{t|t}[i] \leftarrow \hat{\boldsymbol{\zeta}}_{t_i|t_i}$ Calculate using Eq. (4.38) $\hat{\zeta}_{t+1|t}[i] \leftarrow \hat{\boldsymbol{\zeta}}_{t_{i+1}|t_i}$ Calculate using Eq. (4.39)
$L(\hat{\mathbb{M}}) \leftarrow L(\hat{\mathbb{M}}) + \ln\left(1'(\hat{\boldsymbol{\zeta}}_{t_i|t_{i-1}} \odot \boldsymbol{\xi}_{t_i})\right)$ $\hat{\zeta}_{t|T} \leftarrow [\,]$ $\hat{\zeta}_{t|T}[n] \leftarrow \hat{\zeta}_{t|t}[-1]$ Get the last element - $\hat{\zeta}_{t_n|t_n}$ **for** $i \leftarrow n-1 : 0$ **do**

$\hat{\zeta}_{t|T}[i] \leftarrow \hat{\zeta}_{t_i|t_n}$ Calculate using Eq. (4.41).

Utilize previously calculated elements of $\zeta_{t|T}$ $\boldsymbol{P} \leftarrow \hat{\boldsymbol{P}}$ Update transition matrix $\boldsymbol{P}$ using Eq. (4.42) **for** $k \leftarrow 1 : 2$ **do**

$\hat{\mu}_k, \hat{\sigma}_k, \hat{\lambda}_k, \hat{p}_k, \hat{q}_k \leftarrow \underset{\mu_k, \sigma_k, \lambda_k, p_k, q_k}{\arg\max}\; l(\mathbb{W}_{k-1}; \mu_k, \sigma_k, \lambda_k, p_k, q_k)$ See Eq.(4.43)

(a) Exemplary realization of the process $\{X_t\}$ given in (4.44).

(b) Exemplary realization of the process $\{H_t dS_t^1 + (1 - H_t)dS_t^2\}$.

Figure 4-11: The exemplary realization of the stochastic process defined by Eq. (4.44), source: [130].

4.2.2 Simulation study

To check the effectiveness of the proposed estimation procedure, we have analysed the simulated data that resemble the real data analysed in the further parts of this thesis.

The trajectory of the model was generated using the Euler method [147] for the following stochastic differential equation:

$$\begin{aligned} Y_t &= \alpha_1(\gamma t) + \alpha_2(\gamma t)X_t + (1 - H_t)S_t^1 + H_t S_t^2,\ X_0 = 0,\ \gamma = 10^{-3}; \\ \alpha_1(t) &= 1.87t - 28.9t^2 + 104.8t^3 - 133.3t^4 + 55.5t^5,\ t \in [0, 1]; \\ \alpha_2(t) &= 0.1 - 0.2t + 0.15t^2,\ t \in [0, 1]; \end{aligned} \tag{4.44}$$

with initial distribution $P(H_1 = 0) = 1 - P(H_1 = 1) = 1$ and

$$\begin{aligned} S_t^1 &\sim SGT(0, 1, 0.1, 1, 4), \\ S_t^2 &\sim SGT(0, 3, -0.2, 2.5, 2), \end{aligned} \qquad P = \begin{bmatrix} 0.986 & 0.014 \\ 0.0225 & 0.9775 \end{bmatrix}.$$

For a such set of parameters the SGT distribution is skewed and has the property of heavy tails.

The vector of 1001 samples was generated for $t \in \{0, 1, \ldots, 1000\}$. The exemplary trajectory is shown in Fig. 4-11. We then used the estimation methodology proposed in section 4.2.1.

Using the grid search method, we picked the following hyperparameters:

- $\delta = 2$ (parameter of pseudo-Huber function (4.32));
- $d_1 = 2$ (degree of local approximation of $\alpha_1(\cdot)$ polynomial (4.34));
- $d_2 = 1$ (degree of local approximation of $\alpha_2(\cdot)$ polynomial (4.34));

- $b = 200$ (bandwidth of the asymmetric kernel function (4.35));
- $b_r = 75$ (distance from 0 to right boundary of the asymmetric kernel function (4.35));
- $\tau = 0.1$ (Tikhonov regularization).

For those hyperparameters, for every time point we calculated the local estimates of $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions. We show the results in Fig. 4-12a. The robustness of the estimation method to outliers (see the behavior near $t = 500$ where abnormal values can be found) can be seen, rising from the usage of tho pseudo-Huber loss function. Then we performed the estimation procedure for the Hidden Markov Chain part with the following results:

$$\hat{\mu}_1 = 0.0582,\ \hat{\sigma}_1 = 1.1268,\ \hat{\lambda}_1 = 0.1454,\ \hat{p}_1 = 1.3525,\ \hat{q}_1 = 2.0168,$$
$$\hat{\mu}_2 = -0.1841,\ \hat{\sigma}_2 = 2.9016,\ \hat{\lambda}_2 = -0.1892,\ \hat{p}_2 = 3.4400,\ \hat{q}_2 = 1.0538,$$
$$\hat{P} = \begin{bmatrix} 0.992646 & 0.007354 \\ 0.012827 & 0.987173 \end{bmatrix}.$$

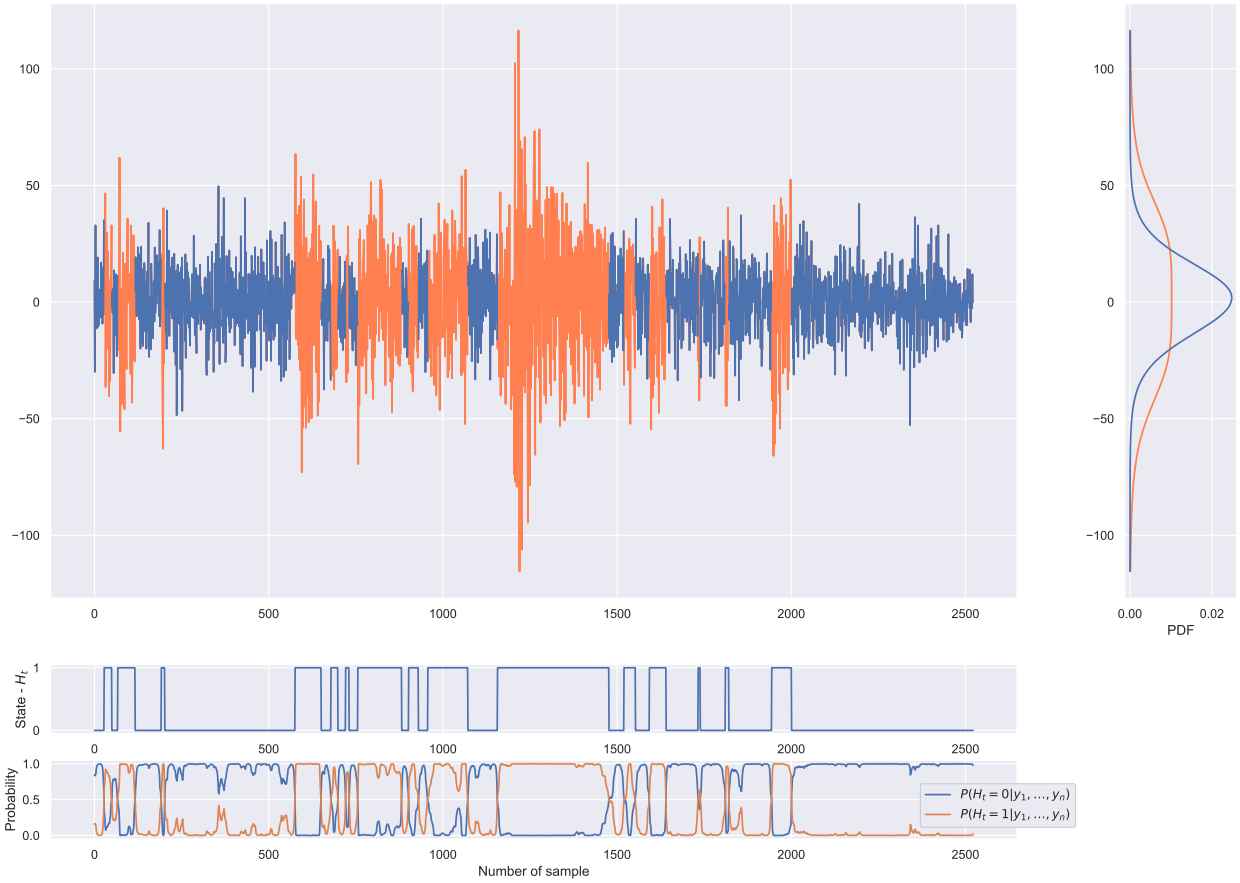
In Fig. 4-12b we also present the probabilities of the vector of the data being in a more violate and stable state with the accompanying vector of differences of the vector $\{x_{t_i}\}$. Let us note that we determine the process' state by selecting one that entails the highest conditional probability $\zeta_{t_i|t_n}$ (4.37).

On this single example, we see that the estimates are very close to the theoretical ones. To further validate the method we performed 250 Monte-Carlo simulations and visualized the distribution of the SGT distributions' parameters estimators and the distribution of the hidden state accuracy (ratio of when the estimated and theoretical states match) using box plots (presented in Fig. 4-13). We can see that the method estimates regimes with high accuracy. However, from the Fig. 4-13c, we find that sometimes the estimated parameter σ (responsible for the variance) is overestimated. This is caused by the fact that when many unlikely values ("heavy tail" property) occur, the likelihood of it being caused by large variance is larger than by the "heavy tail" property while the sample size is relatively small (then also parameters p and q are wrongly estimated as they are responsible for modeling the "heavy tail" property). The solution to this problem can be solved by using the trajectories of a larger sample size. Thus one can conclude that the estimation methodology is correct and provides reasonable results.

4.2.3 Real data application

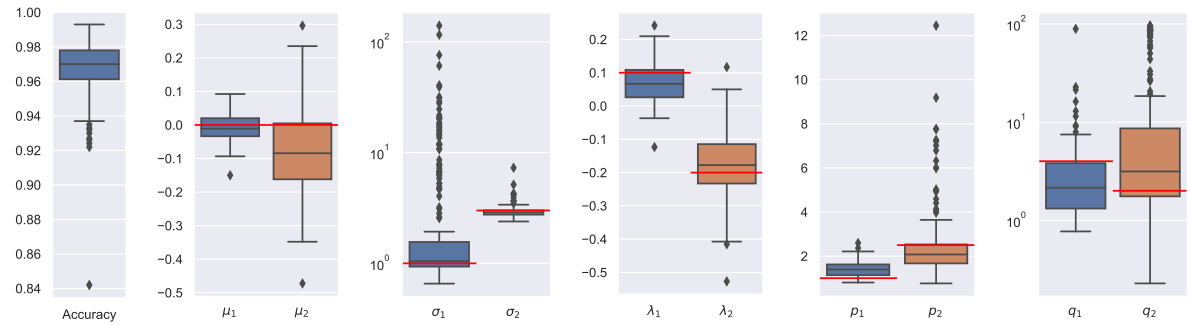
In this part, we check the performance of the proposed model by modelling daily copper price data. The data consists of 2525 data points from the beginning of 2004 until the end of 2013. In

(a) $\mathbb{E}[X_t|X_0]$ estimate for the process $\{X_t\}$ Eq. (4.44) with marked simulated states of the underlying Markov Process.



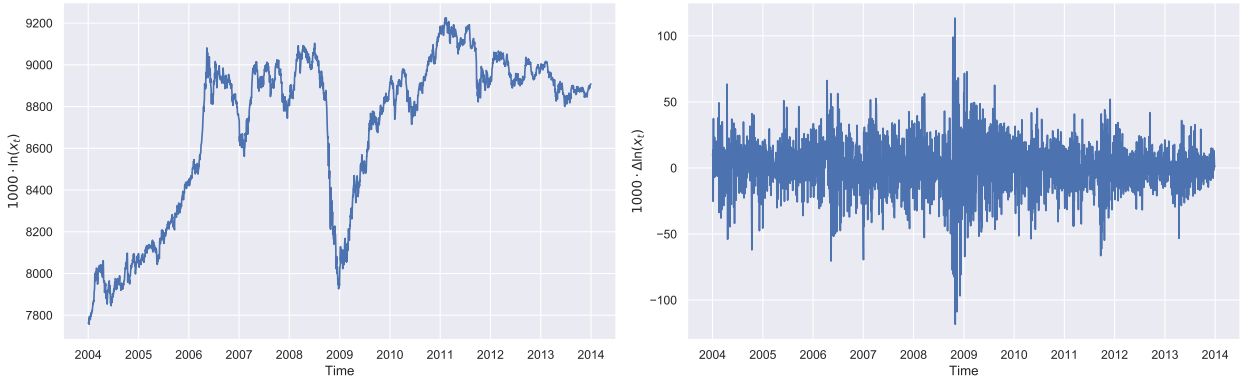
(b) Estimates of probabilities of the realization $\{x_{t_i}\}$ being in C_1 and C_2 state marked on $\{y_{t_i}\}$ ($y_{t_i} = x_{t_{i+1}} - x_{t_i}$ definition in Eq. (4.28)) with comparison of estimated PDFs.

Figure 4-12: Resulting estimates for simulated data for SDE given by Eq. (4.44, source: [130]).



(a) Accuracy (b) $\hat{\mu}$ (c) $\hat{\sigma}$ (d) $\hat{\lambda}$ (e) $\hat{p}$ (f) $\hat{q}$

Figure 4-13: Accuracy of regimes estimates with distributions of model's (4.28) parameters' estimators, source: [130]



(a) Copper price data.

(b) Differentiated copper price data.

Figure 4-14: Copper price data used for evaluation of the proposed methodology, source: [130]

Fig. 4-14a we see that variance changes over time. Due to this fact we transformed the vector taking the Box-Cox transformation [150] - i.e. the natural logarithm of the data. In the next step, we scaled the data by 1000 to reduce numerical errors. From Fig. 4-14b, where differentiated data is presented, we see periods when the variance is significantly higher than in other periods and it resembles the realization of a random variable with heavier tails than Gaussian distribution. Thus the proposition of using this model seems to be reasonable.

The estimation method requires the selection values of hyperparameters which we can find using grid search - we chose ones with the lowest loss. For the local estimation of $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ functions we picked the following values:

- $\delta = 75$ (parameter of pseudo-Huber function (4.32));
- $d_1 = 0$ (degree of local approximation of $\alpha_1(\cdot)$ polynomial (4.34))

(a) Trend estimates of copper price data.

(b) Local estimates $\{_0\hat{\alpha}_1 + (_0\hat{\alpha}_2 + {}_1\hat{\alpha}_2 t_i)x_i\}$ for the copper price data.

Figure 4-15: Resulting trend estimates for copper price data, source: [130]

- $d_2 = 1$ (degree of local approximation of $\alpha_2(\cdot)$ polynomial (4.34));
- $b = 750$ (bandwidth of the asymmetric kernel function (4.35));
- $b_r = 187.6$ (distance from 0 to the right boundary of the asymmetric kernel function (4.35));
- $\tau = 0.7$ (Tikhonov regularization).

Taking the above-mentioned parameters, we minimized the loss function (4.33). The results are presented in Fig. 4-15.

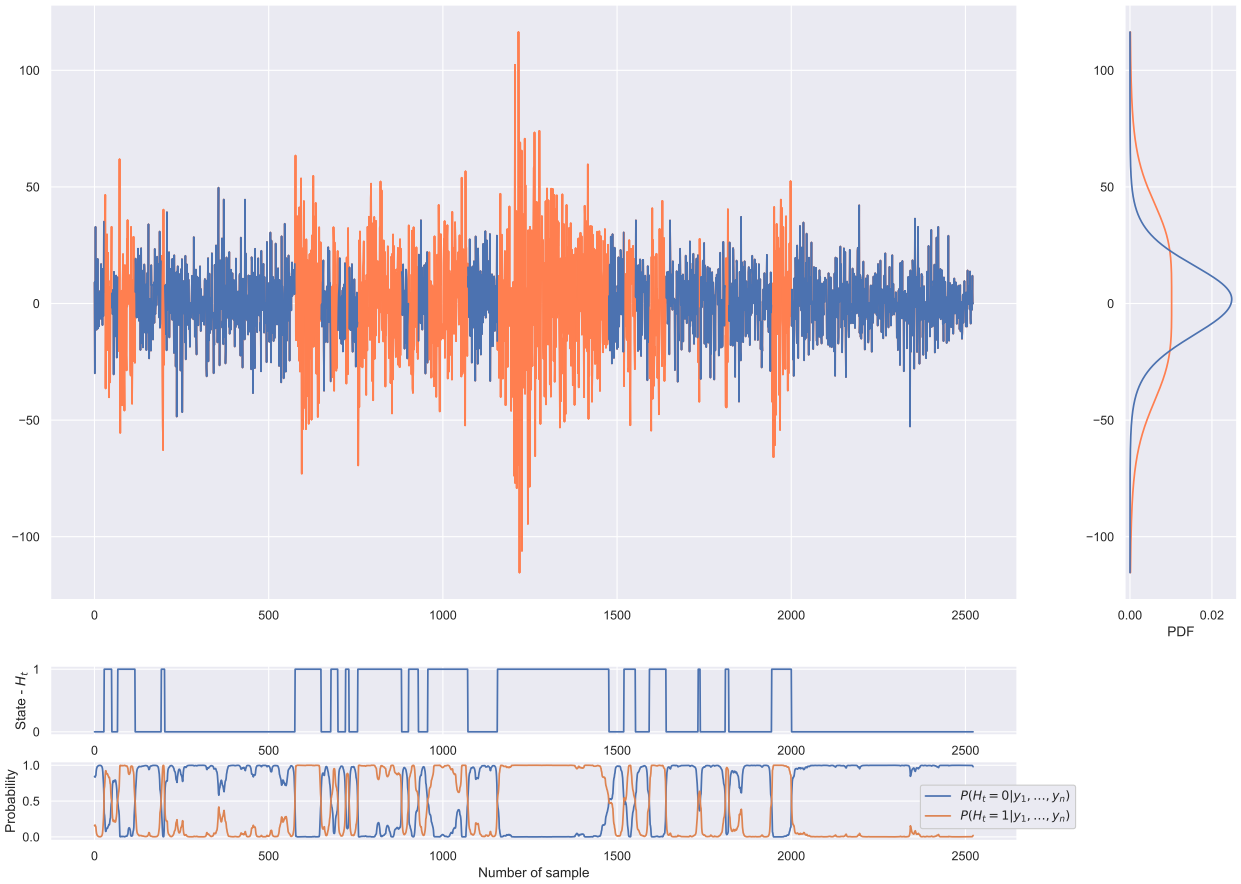
After that, we estimated parameters of the Hidden Markov Chain with SGT distribution as the noise. The estimated parameters are as follows:

$$\begin{gathered} \hat{\mu}_1 = 0.12,\ \hat{\sigma}_1 = 36.71, \\ \hat{\lambda}_1 = -0.0270,\ \hat{p}_1 = 3.8919,\ \hat{q}_1 = 1.1947, \\ \hat{\mu}_2 = 1.05,\ \hat{\sigma}_2 = 17.12, \\ \hat{\lambda}_2 = -0.0300,\ \hat{p}_2 = 2.0268,\ \hat{q}_2 = 5.8758, \end{gathered} \qquad \hat{P} = \begin{bmatrix} 0.97713 & 0.02287 \\ 0.01223 & 0.98777 \end{bmatrix}.$$

The final results are presented in Fig. 4-16.

We also tested the fit for residuals. Because we presumed that they are independent, we can test residuals separately for every case (orange and blue parts of the vector from plot 4-16a) of Hidden Markov Chain. We used KS to validate the null hypothesis that the vector of residuals constitutes a sample from SGT distribution with estimated parameters.

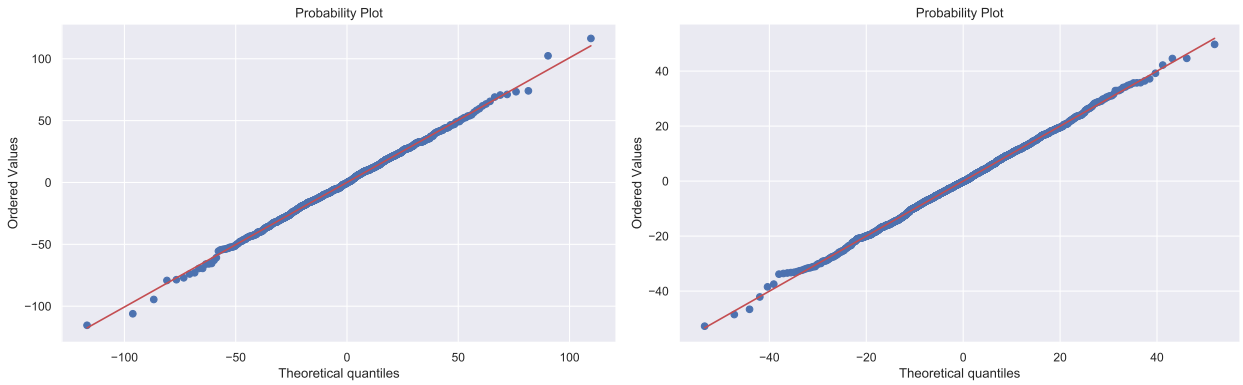
For both of the states, the test passes with a p-value greater than 0.95 which is significantly larger than the commonly-used significance level $\alpha = 0.05$. The KS test returned test statistics



(a) Estimates of probabilities of the copper price being in C_1 and C_2 state marked on $\{y_{t_i}\}$ with estimated PDFs.

(b) Marked estimated most probable states of the copper price data.

Figure 4-16: Resulting Hidden Markov Chain estimates for copper price data, source: [130]



(a) Q–Q plot of residuals for first state.

(b) Q–Q plot of residuals for second state.

Figure 4-17: Q–Q plots of residuals for each estimated hidden state, source: [130]

$K_1 = 0.01267$, $K_2 = 0.01267$ with p-values 0.9572 and 0.9601 for the state C_1 and C_2, respectively. Thus we can not reject the null hypothesis.

To further review the ability of the model to explain randomness in the data, especially the heavy tails, we performed a visual test of goodness of fit - Q–Q plot [163] presented in Fig. 4-17. The Q–Q plot is a visual test which compares the empirical quantiles of real data and the theoretical ones corresponding to the tested distribution.

Based on the results of both tests, we conclude that the model properly describes the data. From Fig. 4-16a we see that the estimation procedure found the presumed two states which can be easily labelled as calm and violate. This fact finally confirms our assumption of the existence of such hidden states and thus justifies the use of the model to the copper price data.

4.3 Discussion and summary

In this chapter we have proposed two stochastic models which address characteristics of financial data, that is the non-Gaussian behaviour and time-inhomogeneous character.

On top of this, these two models have a built-in inflation trend, unlike basic Orstein-Uhlenbeck which has constant parameters of the trend and return to the mean. In light of rising inflationary pressure around the world, the increasing cost of commodities extraction and the growing demand for metals from new applications, this may better reflect changing business dynamics.

The initially-proposed model improves the quality of modelling and outperforms models with Gaussian distributions.

The second model is more complex and difficult to apply as it assumes changes of regimes. The detection of switching points ex-ante, when regimes of metals prices are changing, is very

challenging. This model cannot be applied to all metals. According to the results in [130], model is a better fit to cyclical metals like copper (base metal), rather than gold or silver (precious metals), in case of which changes are of a disordered nature. The second model has more applicable when used for descriptive purposes rather than modelling future prices.

Chapter 5

Currency exchange rates modelling

For metal mining companies active in the international environment, not only metal prices have a substantial impact on their financial situation. Currency exchange rates are no less important, as they impact both cash flows as well as the fair value of balance sheet positions and, in fact, very often determine the business competitiveness of each mine as they are located in various jurisdictions around the world.

In modelling currency exchange rates, the random walk model has been very often assumed. However, the longer the forecast horizon, the more visible is the property of data showing that over time, the process based on data tends to drift towards its long-term mean. The methodology used in this chapter is very similar to the one used in the previous chapter, dedicated to metals prices modelling.

5.1 Non-Gaussian time-inhomogeneous model

In this section we present the time-inhomogeneous non-Gaussian model for currency exchange rates description. Similarly, as for the metal prices analysis, we utilize here the general class of distributions, namely SGT distribution (4.1) and the stochastic process described by the special case of the general process given by Eq. (4.4). The results presented in this section are published in [127].

An important prediction of the theoretical literature on target exchange rates is that we might expect mean reversion of the exchange rate when the central banks engage in intramarginal intervention and market participants expect the exchange rate band to be fully credible and engage in

stabilising speculation. This mean-reverting property is widely referred to in the literature [164–167]. The strength of the mean-reverting component of the process measures the robustness and the credibility of the exchange rate bands.

In our modelling study of the currency exchange rates data, we propose to apply the generalisation of the Chan-Karolyi-Longstaff-Sanders (CKLS) model. This model was introduced for describing the future evolution of the short term interest rate in [168]. Short term interest rates are a crucial feature of the monetary transmission mechanism, which enables a monetary authority's actions to influence short-term interest rates and also, indirectly, the exchange rate [169].

The general model under consideration is based on SGT distribution and it is given by the following form:

$$dX_t = (\alpha + \beta X_t)\, dt + \sigma X_t^d dS_t, \tag{5.1}$$

where $\{S_t\}$ is the process with stationary increments with SGT distribution. Such a model is widely applied in the validation of derivative instruments dependent on interest rates [170, 171]. It belongs to the class of one-factor stochastic processes, but one can consider multi-factor models [172, 173].

The formula (5.1) simplifies to the well-known, simpler models as presented in Table 4.1. The classical CKLS model is defined in a similar way as the formula (5.1), however $\{S_t\}$ is replaced by the ordinary Brownian motion. The most popular examples of the classical CKLS model are: the Merton [174], Vasicek [34], CIR-SR [175], Dothan [176], Brennan-Schwartz [177], CIR-VR [178] and CEV [179] models. As one can see, it is a natural extension of the classical Ornstein-Uhlenbeck process (also called the Vasicek model) widely discussed in the literature. The structure of the model (5.1) and its parameters allow modification of the rate of the return to the mean value [180]. The β parameter is the rate of the return and the ratio $-\alpha/\beta$ is the mean value to which the process returns. The volatility term not only has constant components $\sigma > 0$ and $d \in [0, 2)$, but also random sources X_t and dS_t.Therefore, the conditional mean and the conditional variance of the process $\{X_t\}$ depend on its values. This is the crucial assumption and an upgrade of previously studied models [181].

5.1.1 Model estimation

In this part we present the procedure of estimating the CKLS model parameters from empirical data. We may recall, that in the estimation algorithm as the output we obtain four estimators $(\hat{\alpha}, \hat{\beta}, \hat{\sigma}, \hat{d})$ of the CKLS model parameters. For that goal, we apply the generalized methods of moments (GMM), which is the procedure introduced by the Nobel Prize winner, economist Lars

Hansen in the paper [182]. The method is constantly being developed [183–185] and is widely applied in multi-fractals research [186], emission modelling [187], economic studies [188], time series modelling [189], queueing networks [190] and turbulence modelling [191], among others. As was mentioned, the classical CKLS model allows different distributions of the noise $\{S_t\}$. This does not affect the procedure of parameters estimation. Therefore we present the estimation procedure for the general model given in (5.1).

One can rewrite the discrete version of Eq. (5.1) for the process on the interval $[0, T]$ as a system of the following equation:

$$X_{n+1} - X_n = (\alpha + \beta X_n)\,\Delta t + \epsilon_{n+1}, \tag{5.2}$$

$$\epsilon_{n+1} = \sigma X_n^d \Delta S_n, \tag{5.3}$$

where Δt is a time step and $n \in \{0, \Delta t, \ldots, N\Delta t\}$ is an even division of the interval $[0, T]$ with $N = T/\Delta t$ subintervals of the same length. Moreover, innovations $\Delta S_n = S_{(n+1)\Delta t} - S_{n\Delta t}$ are independently identically distributed (i.i.d.) with the distribution of the process $\{S_t\}$ in (5.1).
The sequence $\{X_n\}$ defined by the above system of equations is a Markov chain [192] with a set of four moment functions:

$$f_n(\theta) = \begin{bmatrix} \epsilon_{n+1} \\ \epsilon_{n+1} X_n \\ \epsilon_{n+1}^2 - \sigma^2 X_n^{2d} \\ \left[\epsilon_{n+1}^2 - \sigma^2 X_n^{2d}\right] X_n \end{bmatrix},$$

where $\theta = (\alpha, \beta, \sigma, d)$. Under the hypothesis of the CKLS model form (5.2)-(5.3) and the orthogonality assumption for X_n and ϵ_{n+1}, the quantity $f_n(\theta)$ satisfies the following condition:

$$E\left[f_n(\theta)\right] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

which in the GMM theory is called moment conditions for moments $E\left[f_n(\theta)\right]$. The main idea behind the GMM methodology is to choose the CKLM model with parameters $\theta = (\alpha, \beta, \sigma, d)$ satisfying the above moment conditions. Since from the practical point of view the CKLM model is estimated from the real trajectory, instead of the function $E\left[f_n(\theta)\right]$ one considers its empirical

version $g_N(\theta)$:

$$g_N(\theta) = \frac{1}{N}\begin{bmatrix} \sum_{n=1}^{N}\left[x_n - \alpha\Delta t - (1+\beta\Delta t)x_{n-1}\right] \\ \sum_{n=1}^{N}\left[x_n - \alpha\Delta t - (1+\beta\Delta t)x_{n-1}\right]x_{n-1} \\ \sum_{n=1}^{N}\left[x_n - \alpha\Delta t - (1+\beta\Delta t)x_{n-1}\right]^2 - \sigma^2 x_{n-1}^{2d}\Delta t \\ \sum_{n=1}^{N}\left[\left(x_n - \alpha\Delta t - (1+\beta\Delta t)x_{n-1}\right)^2 - \sigma^2 x_{n-1}^{2d}\Delta t\right]x_{n-1} \end{bmatrix},$$

where $x_1, x_2, \ldots, x_N$ is the trajectory of the process defined in (5.1). The goal of the estimation procedure is the solution of the following equation:

$$g_N(\theta) = \vec{0}, \tag{5.4}$$

with respect to the parameters $\theta = (\alpha, \beta, \sigma, d)$. Such a problem could equally be considered as a minimization of the quadratic form $g_N(\theta)^T g_N(\theta)$ with respect to θ vector. Therefore, the estimator $\hat{\theta}$ of the true parameters $(\alpha, \beta, \sigma, d)$ is defined as follows:

$$\hat{\theta} = arg\min_{\theta}\left[g_N(\theta)^T g_N(\theta)\right].$$

Since the two top Eqs. in (7.2) are linear with respect to α and β, we can write their explicit solutions as follows:

$$\hat{\alpha} = \Delta t\frac{\sum_{n=1}^{N} x_n - (1+\beta\Delta t)\sum_{n=1}^{N} x_{n-1}}{N},$$

and

$$\hat{\beta} = \frac{1}{\Delta t}\left[\frac{\left(\frac{1}{N}-1\right)\sum_{n=1}^{N} x_n x_{n-1}}{\frac{\left(\sum_{n=1}^{N} x_{n-1}\right)^2}{N} - \sum_{n=1}^{N} x_{n-1}^2} - 1\right].$$

Moreover, in order to compute solutions of two other equations of the system (7.2) one can apply any numerical method to get $\hat{\sigma}$ and $\hat{d}$. In our empirical study, we used the `fminsearch` Matlab built-in function.

5.1.2 Simulation study

In Fig. 5-1 we present two example trajectories (see left and right panels) of the CKLS model with SGT distribution with: $\beta = -0.4$, $\sigma = 0.1$, $d = 0.5$ and $p = 1.5, q = 10, \lambda = -0.0082, \mu = 0, s = 0.0062$ and three different vales of α parameters. The parameters are selected to be close to the parameters estimated from the real data analysed in the next section. One can see the α parameter influences the response of the trajectories. The smaller the α, the higher the values of the process are observed. We expect the typical trajectories of the CKLS model with SGT

distribution exhibiting non-Gaussian behaviour and mean-reversion property. Such properties we can observe in Fig. 5-1.

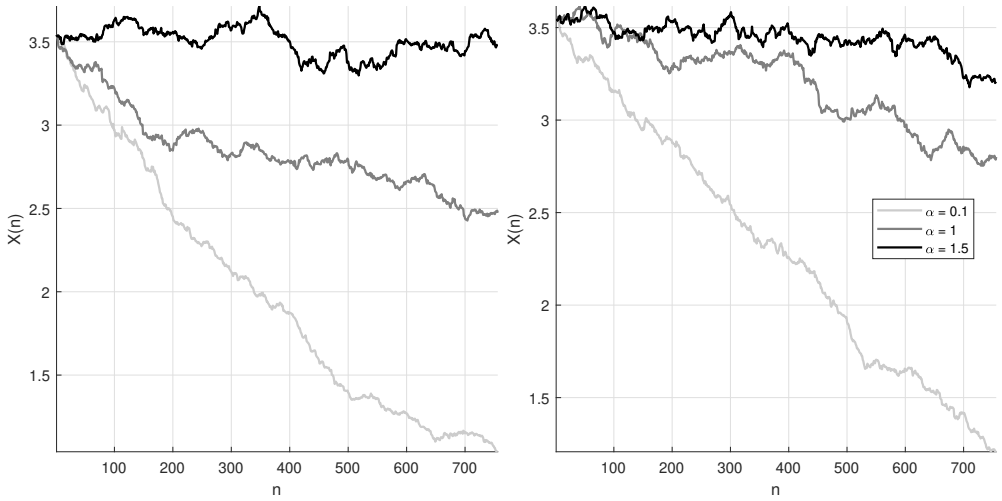


Figure 5-1: Example trajectories of the CKLS model with SGT distribution with: $\beta = -0.4$, $\sigma = 0.1$, $d = 0.5$ and $p = 1.5, q = 10, \lambda = -0.0082, \mu = 0, s = 0.0062$ and three different values of α parameter, source: [127].

In order to check the proposed estimation technique of the CKLS model with SGT distribution we simulate 3000 trajectories of this model, and using the procedures presented in the previous sections we estimate the parameters for each trajectory. Each simulated trajectory contains 756 $(2 * 252)$ observations. Finally, we create the boxplots of the obtained estimates and present them in Fig. 5-2. The parameters used for the simulations are selected to be close to the parameters obtained for the real data analysed in the next section. The mean absolute percentage errors (MAPE) of estimated parameters from the CKLS model are presented in the titles of each boxplot.

We have also studied the MAPE and its dependence on the changes of the α parameter. In our simulation study we have simulated 3000 trajectories of the CKLS model with $\mathcal{SGT}$ distribution with the following parameters: $\beta = -0.40, \sigma = 0.1, d = 0.5$. In Fig. 5-3 we present the MAPE with respect to the α value taken to the simulations. The range of the α parameter is selected to be close to the estimated α for the real data considered in this part. As one can see the MAPE is not very sensitive for the changes in α parameter which indicates the stability of the estimation method.

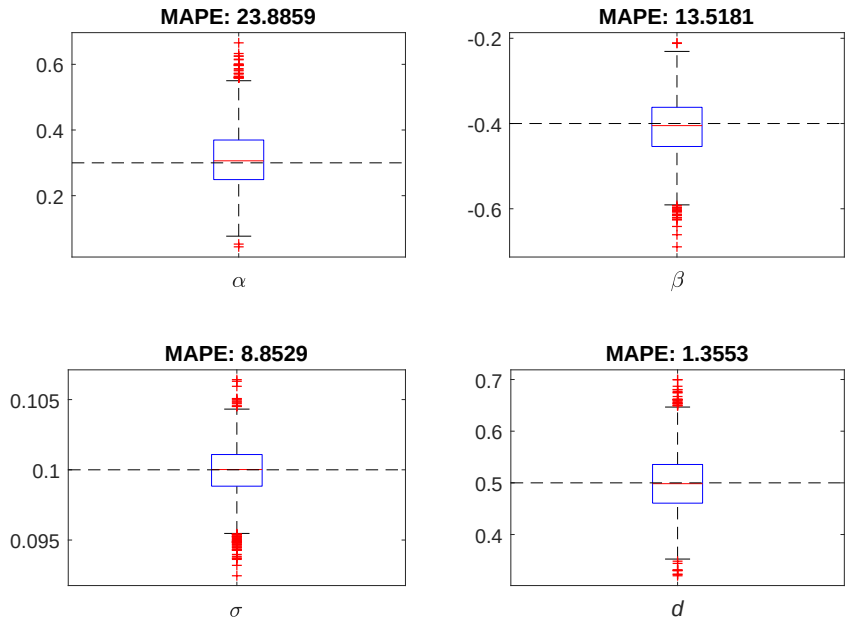


Figure 5-2: Estimators of the parameters of the CKLS model. The true parameters are: $\alpha = 0.3, \beta = -0.40, \sigma = 0.1, d = 0.5$ are marked by black dashed lines. To the simulations we take the following parameters of SGT distribution: $p = 1.5, q = 10, \lambda = -0.0082, \mu = 0, s = 0.0062.$, source: [127].

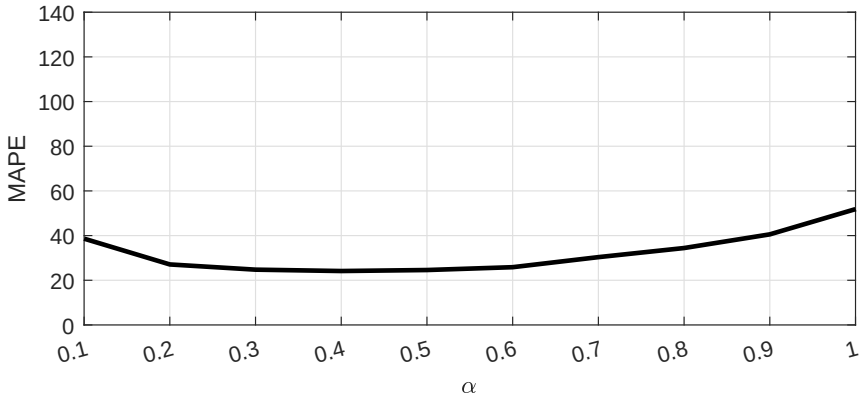


Figure 5-3: The MAPE for the estimator of the α parameter in the CKLS model with $\mathcal{SGT}$ distribution. The other parameters taken to the simulations are: $\beta = -0.40, \sigma = 0.1, d = 0.5$ and $p = 1.5, q = 10, \lambda = -0.0082, \mu = 0, s = 0.0062.$, source: [127].

5.1.3 Validation factor

In the classical validation procedure of fitting a given model to a real time series we compute the quantile lines of specific order based on the calibrated model and compare them with the real data.

The model is fitted to the historical data from the calibration period and then the validation is made for the data from the prediction period. In order to calculate the quantile lines we use the Monte-Carlo simulations. In this approach very often the quantile lines on the level 5% and 95% or 25% and 75% are calculated. If the model is appropriate to the analysed time series, we expect that the 90% or 50%, respectively, of the data from the prediction period falls into the constructed intervals. The open issue is always the choice of the proper order of such quantiles. Thus, we propose a novel validation technique which we call the space quantile-inclusion data plot. The idea of this validation factor is the following. The horizontal axis (x-axis) presents the $q\%$ difference between space quantile lines $(100-q)/2\%$ and $q+(100-q)/2\%$ for the whole spectrum of q. In our applications we take $q \in [10, 20, \ldots, 90]$. The vertical axis (y-axis) represents the fraction $\phi(q)$ of real data from the prediction period that lies between quantile lines $(100-q)/2\%$ and $q+(100-q)/2\%$ (equivalently, the fraction of real data included in $q\%$ quantile region). Hence, such a plot presents a line describing the goodness of fit and dependence between the historical data and simulated trajectories of the fitted model. For an ideally fitted model the plot should present the identity line. It is worth emphasising, such a factor bypasses the problem of choosing the right quantile lines. Moreover, by applying the whole spectrum of q, the space quantile-inclusion data plot can easily show the under-fitting or over-fitting of the calibrated model (too wide or too narrow quantiles lines from generated trajectories, respectively). It also serves as a great tool in comparison of different calibrated models from the same data set.

This graphical representation can be factorised as the average square distance between the space quantile-inclusion data plot and the identity line. Thus the validation factor for the given model M can be formulated as follows:

$$\Phi_M = \frac{1}{\#Q} \sum_{q \in Q} |\phi(q) - q|^2, \tag{5.5}$$

where Q is a set of possible values of q. In our case $Q = [10, 20, \ldots, 90]$, $\#Q$ is the number of elements of the set Q, $\phi(q)$ is the fraction of real data from the prediction period that lies between quantile lines $(100-q)/2\%$ and $q+(100-q)/2\%$ constructed using the model M fitted for the data from the calibration period. In an ideal situation this the factor Φ_M should tend to zero. In this case the model M properly describes the real data. This factor can be also used in order to compare two (or more) fitted models M_1 and M_2 to the same time series. We select the model with the smallest validation factor, i.e. if $\Phi_{M_1} < \Phi_{M_2}$, then model M_1 is selected as the more appropriate for the considered data. We used this approach in the next section when we compare

the CKLS model with SGT distribution and the classical GBM applied to the currency exchange rates.

5.1.4 Real data application

In this section we present the analysis of real data sets which contain the currency exchange rates. For the comparison, we examine two daily currencies pairs: EURUSD and USDPLN. These specific pairs have been chosen intentionally to verify the ability to applying the new proposed model to currency exchange rates with different characteristics: EURUSD – the major world's currency pair and USDPLN – an emerging market currency with lower liquidity. In order to prove that the CKLS model is universal for the currency exchange rates data, we compare the results of the CKLS model with the classical GBM [193]. Moreover, we propose to apply the CKLS with general SGT distribution, as the general class of distributions. The idea of estimation of the new model is presented in the previous sections. In order to confirm that the SGT distribution is more efficient in the analysed data modelling in Fig. 5-4 we present (in log-log scale) the right empirical tails of logarithmic returns of two considered currency exchange rates. We compare them with the empirical tails of the SGT distribution with the parameters estimated from the logarithmic returns of real data. In the same plot, we demonstrate the right tails of the Gaussian distribution with the parameters estimated as the empirical mean and empirical variance of the corresponding logarithmic returns. Fig. 5-4 shows that SGT distribution fits the data better than the Gaussian one. Thus it is reasonable to consider the CKLS model with innovations from SGT distribution. For

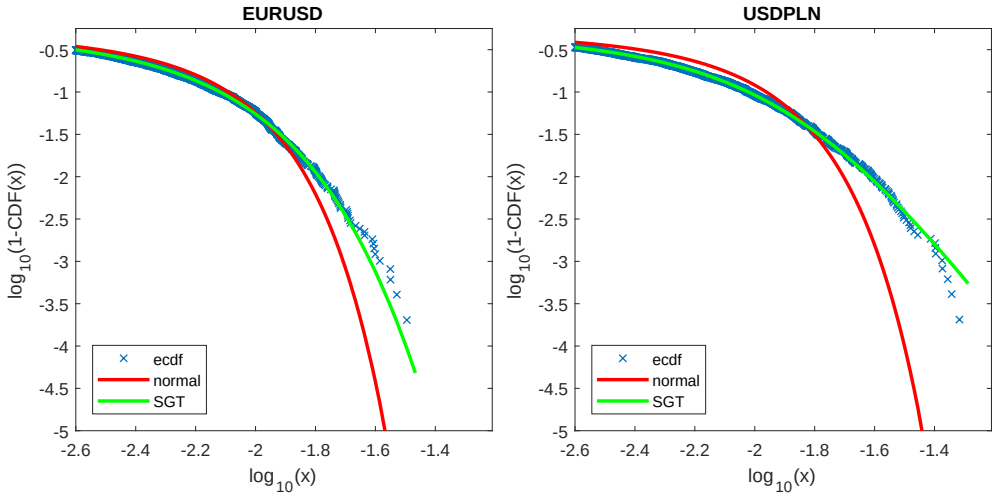


Figure 5-4: The right empirical tails of the logarithmic returns of EURUSD (left panel) and USDPLN (right panel) with the right tails of the SGT and Gaussian distributions with the parameters estimated based on the data (in double logarithmic scale), source: [127].

each examined data set we fit two models, namely CKLS with SGT distribution (CKLS/SGT) and GBM. Both models are fitted for the historical data. For each currency exchange rate, we consider 9 historical data sets. All of them contain data from the year 1997. The end date of historical data are the years 2007-2015.

	CKLS				SGT				
Year:	α	β	σ	d	μ	s	λ	p	q
2007	1.6397	-0.4406	0.0894	1.1247	-0.00009	0.0066	0.0223	1.6739	7.7393
2008	0.5195	-0.1534	0.0774	1.2295	-0.00007	0.0065	0.0471	1.6891	7.5221
2009	0.8957	-0.2521	0.2728	0.3173	0.00001	0.0075	0.0623	1.9143	2.0000
2010	1.0371	-0.2966	0.3740	0.1420	0.00001	0.0084	0.0503	1.6986	2.0022
2011	1.1230	-0.3225	0.4018	0.0961	-0.00015	0.0087	0.0283	1.6435	4.1954
2012	1.2620	-0.3570	0.4243	0.0657	0.00005	0.0089	0.0580	1.5991	4.5397
2013	1.2706	-0.3677	0.4204	0.0781	0.00003	0.0089	0.0498	1.5798	4.9101
2014	1.2605	-0.3679	0.3918	0.1244	0.00002	0.0088	0.0528	1.6074	4.8272
2015	1.3231	-0.3796	0.3547	0.1912	0.00006	0.0086	0.0520	1.6171	4.7127

Table 5.1: The parameters of the CKLS model with SGT distribution estimated for USDPLN data, source: [127].

	CKLS				SGT				
Year:	α	β	σ	d	μ	s	λ	p	q
2007	0.2662	-0.2370	0.1018	0.4349	0.00002	0.0062	-0.0083	1.4543	200.0036
2008	0.0741	-0.0498	0.1023	0.1824	0.00008	0.0060	-0.0061	1.4277	142.4036
2009	0.1809	-0.1471	0.1011	0.9015	0.00004	0.0063	0.0039	1.3238	99.5099
2010	0.1839	-0.1458	0.1022	0.9839	0.00005	0.0064	0.0049	1.3442	75.1792
2011	0.2518	-0.2083	0.1032	1.0073	0.00003	0.0065	-0.0007	1.3668	75.9255
2012	0.2769	-0.2295	0.1027	1.0957	0.00003	0.0066	-0.0112	1.3606	135.7517
2013	0.2824	-0.2319	0.1018	1.0722	0.00002	0.0065	-0.0072	1.4106	47.1199
2014	0.2731	-0.2203	0.1013	1.0100	0.00003	0.0064	-0.0126	1.4023	42.5706
2015	0.2938	-0.2439	0.1011	0.9229	-0.00003	0.0063	-0.0069	1.3204	131.7730

Table 5.2: The parameters of the CKLS model with SGT distribution estimated for EURUSD data, source: [127].

Finally, we make the prediction for the next three years. In Figs. 5-5-5-8 we demonstrate 9 panels corresponding to the 9 mentioned end years of the historical data. Moreover, we also present the empirical quantiles for the prediction period for the levels 5% and 95% (red lines) and 25% and 75% (green lines). In Figs. 5-5 and 5-6 we demonstrate the results for EURUSD while in Figs. 5-7 and 5-8 - for USDPLN.

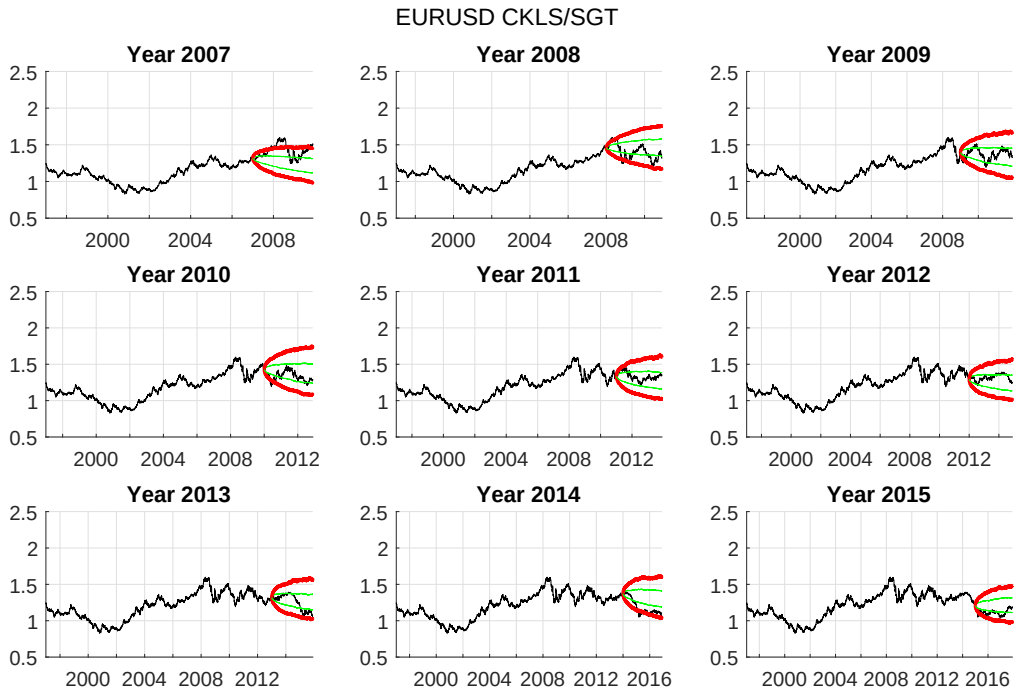


Figure 5-5: The CKLS/SGT modelling for EURUSD for 9 considered end years of historical data. The prediction is made for a three-year period. In each case, we present the empirical quantiles for the prediction period for the levels 5% and 95% (red lines) and 25% and 75% (green lines), source: [127].

In order to compare the obtained results, we analyse the space quantiles-inclusion data plot described in the previous sections for two considered exchange currency rates, 9 end years of historical data and two considered models. The appropriate plots are presented in Figs. 5-9 and 5-10. In Fig. 5-9 we present the results for EURUSD. We can see that in general the CKLS/SGT model seems to be superior with respect to the GBM. The space quantiles-inclusion data plots are closer to the identity line in case of CKLS/SGT for all end years of a three-year calibration period. However, for both models we can see that the constructed quantile lines are too narrow for the same end years of the calibration period. Namely, for the years 2007, 2008, 2014 and 2015 the space quantiles-inclusion data plots for both models are below the identity line, which suggests both models are overfitting. For the other end years of a three-years calibration period for both considered models the space quantiles-inclusion data plots are above the identity line, that suggests the constructed quantile lines for all considered q values are too wide and the models are underfitting. In Fig. 5-10 we present the results for USDPLN. Also in this case the CKLS/SGT

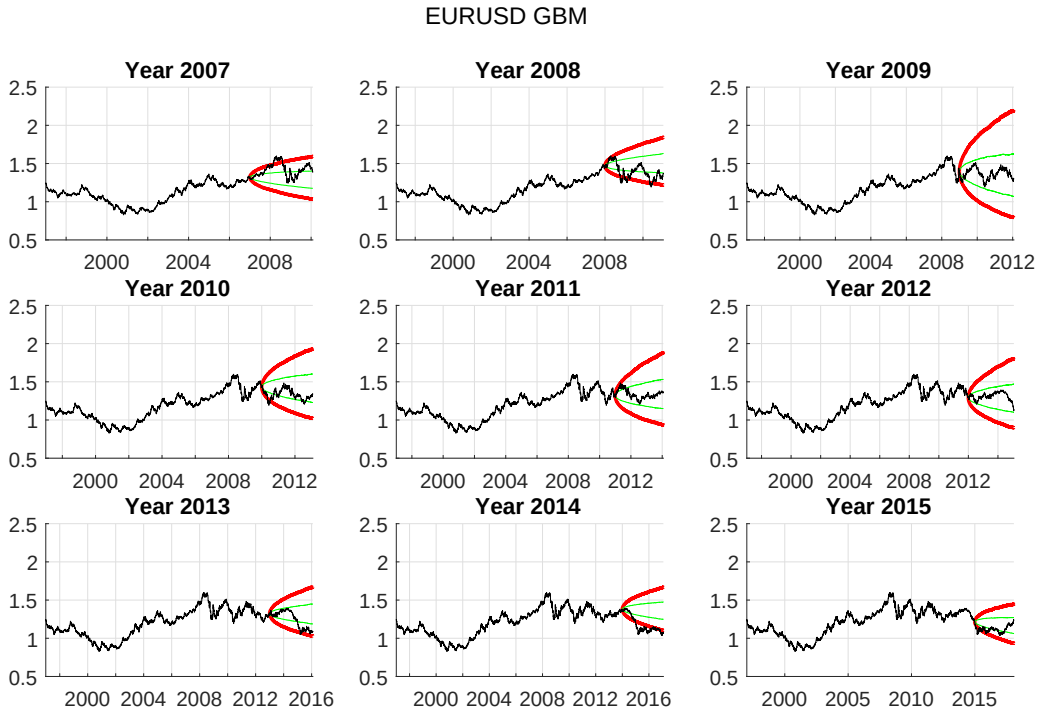


Figure 5-6: The GBM modelling for EURUSD for 9 considered end years of historical data. The prediction is made for a three-years period. In each case, we present the empirical quantiles for the prediction period for the levels 5% and 95% (red lines) and 25% and 75% (green lines), source: [127].

model, in general, seems to be the superior one with respect to the GBM. The space quantiles-inclusion data plots are closer to the identity line. For both considered currency exchange rates the CKLS/SGT model is more universal and can be useful in describing the currency exchange rates with different behaviour. In table 5.3 the validation factor calculated as the average square distance between the space quantile-inclusion data plot and the identity line for two considered models, two considered currency exchange rates and 9 end years of historical data. As we can see, this factor clearly indicates that the CKLS/SGT model is more appropriate for the considered cases than the classical GBM model.

5.2 Averaging calibration window for prediction

In this section we discuss the problem of the calibration length used for the prediction. The analysis will be performed based on the generalized (time-dependent) Vasicek model for the currency

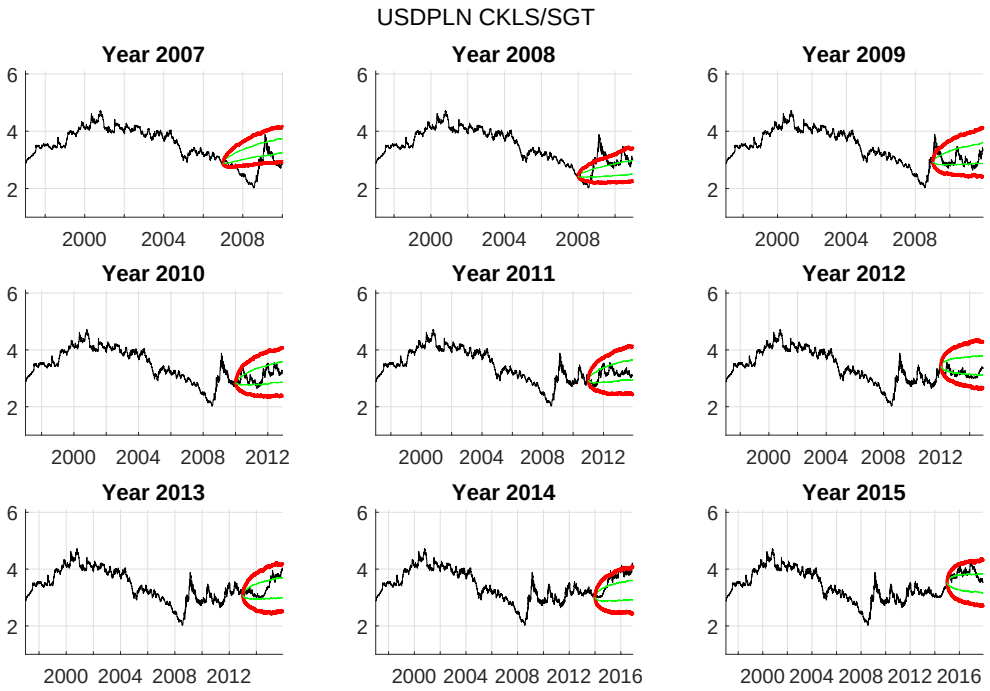


Figure 5-7: The CKLS/SGT modelling for USDPLN for 9 considered end years of historical data. The prediction is made for three years period. In each case, we present the empirical quantiles for the prediction period for the levels 5% and 95% (red lines) and 25% and 75% (green lines), source: [127].

exchange rate data description being a special case of (4.4). However, the proposed methodology is universal and can be applied for any model used for data analysis. The results presented in this section are published in [20].

We analyse in this section the extended Vasicek model by replacement of the constant coefficient in the classical process by coefficients dependent on time. The considered process seems to be perfect for modeling the analysed exchange rates data, under the condition that they are homogeneous. However, in the analysed data, we observe that some of their characteristics change over time, thus in the case of a long-term prediction problem this behaviour needs to be taken into consideration. In general, as was mentioned, one may expect that the longer the calibration length, the better the prediction performance. This statement is true, but only in the case of the homogeneous data. This problem is highlighted in this part of the thesis. We demonstrate that the analysed vectors of observations cannot be modeled by the same process, thus in the problem of a long-term prediction the non-homogeneous character of the time series should be taken into consideration.

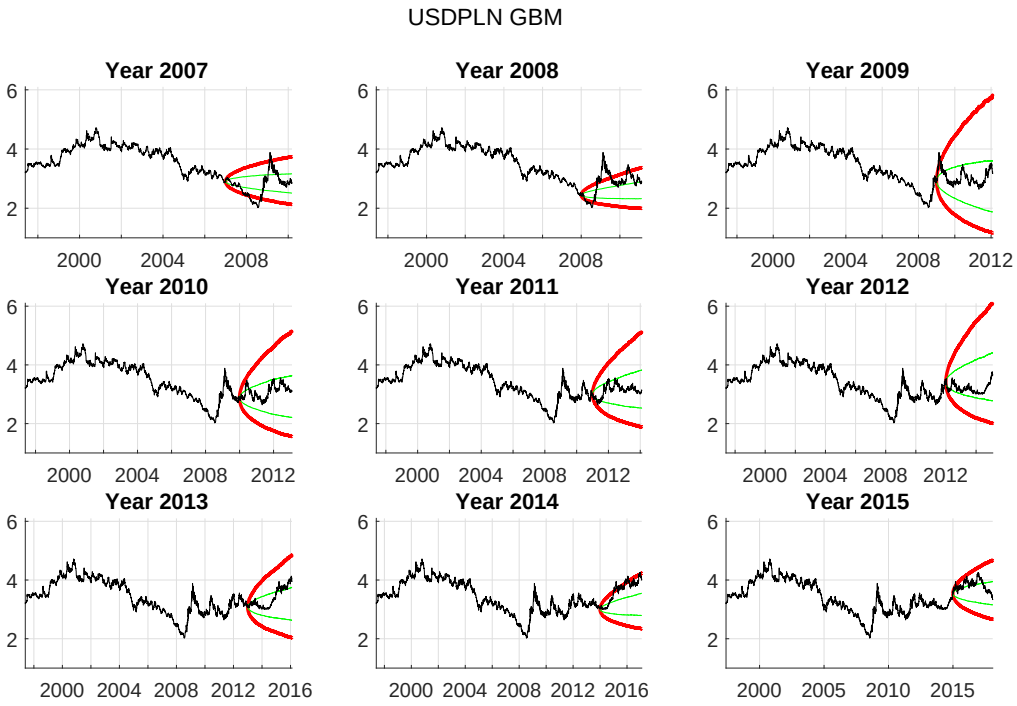


Figure 5-8: The GBM modelling for USDPLN for 9 considered end years of historical data. The prediction is made for three years period. In each case, we present the empirical quantiles for the prediction period for the levels 5% and 95% (red lines) and 25% and 75% (green lines), source: [127].

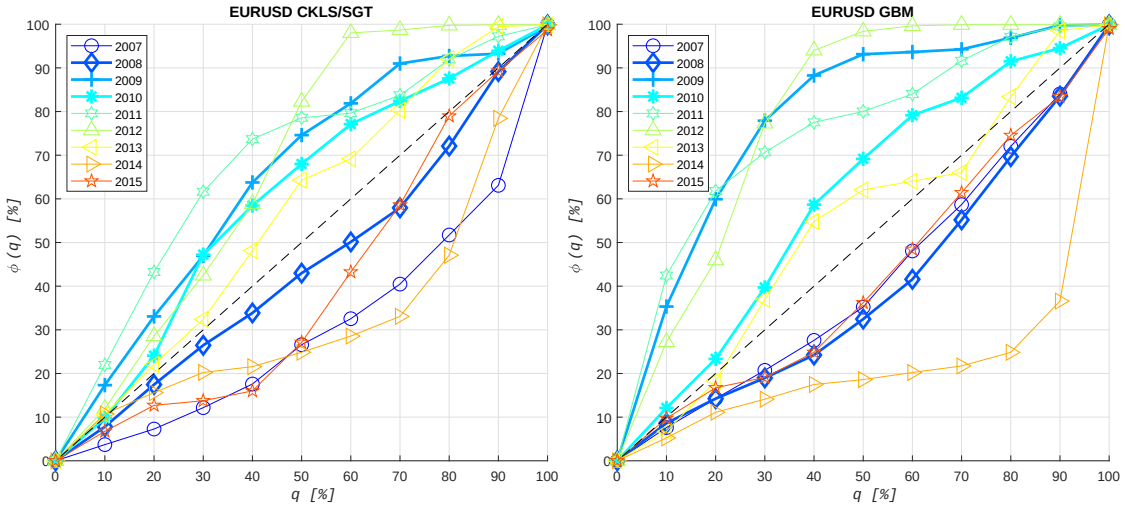


Figure 5-9: The space quantiles-inclusion data plots for EURUSD for 9 considered end years of historical data for CKLS/SGT model (left panel) and GBM (right panel), source: [127].

	EURUSD		USDPLN	
Year:	**CKLS/SGT**	**GBM**	**CKLS/SGT**	**GBM**
2007	428.74	74.76	1019.08	66.84
2008	48.59	126.64	426.31	670.8
2009	222.93	1022.34	59.43	676.57
2010	128.16	140.23	383.47	693.89
2011	403.49	761.75	413.94	690.29
2012	321.10	1046.75	94.69	600.87
2013	60.69	46.43	245.79	322.71
2014	394.03	1077.77	297.43	616
2015	196.22	58.58	108.73	53.75
Mean	**244.88**	**483.91**	**338.76**	**487.97**

Table 5.3: The validation factor calculated as the square distance between the space quantile-inclusion data plot and the identity line for CKLS/SGT and GBM two considered currency exchange rates and 9 end years of historical data, source: [127].

Financial data, including exchanges rates, are characterised by high volatility and uncertainty driven by interest rates changes, inflation rates fluctuation, a three-year monetary policy conducted by central banks and obviously industrial production and other financial events. In light of such a volatile environment, models which are used for modelling should take into account not only a three-year changing parameters of the model through the time but also time frames of data based on which three-year model is calibrated. Averaging of calibration windows for modelling purposes is a good solution here, and enables the limition of potential outstanding forecasting errors, which could be very costly in real business application.

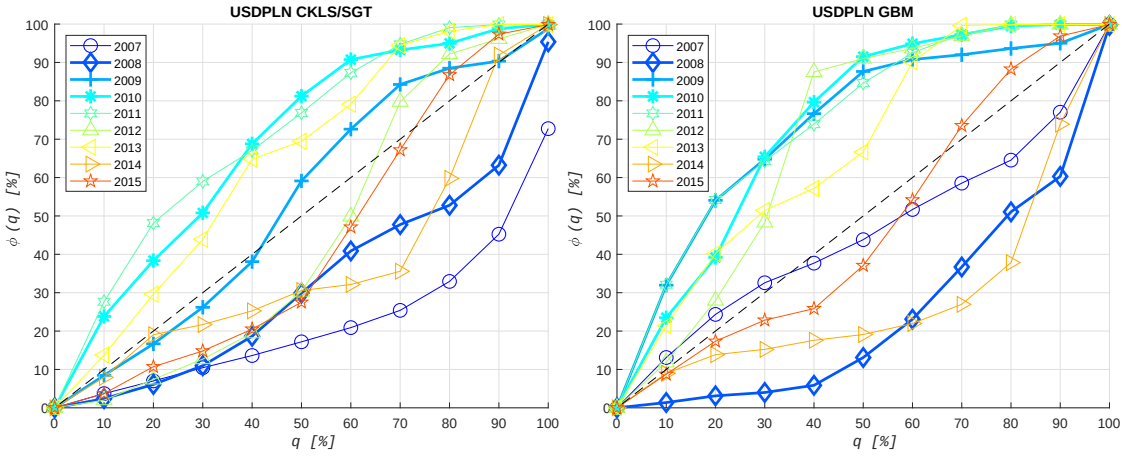


Figure 5-10: The space quantiles-inclusion data plots for USDPLN for 9 considered end years of historical data for CKLS/SGT model (left panel) and GBM (right panel), source: [127].

The extended Vasicek model (also called generalized Vasicek model) with time-dependent parameters belongs to the group of time-inhomogeneous models [52, 194]. This model has been extended by Hull and White by allowing both, the drift and variance coefficients, to be time-varying and its stochastic differential equation is given by the formula [55]:

$$dX_t = \{\alpha_0(t) + \alpha_1(t)X_t\}dt + \beta_0(t)dB_t, \tag{5.6}$$

where $\alpha_0(t)$, $\alpha_1(t)$, $\beta_0(t)$ are functions. The special case of the model (5.6) is the CKLS model with $d = 0$, (7.1). However, in practical applications, it is more reasonable to expect that at least some of the parameters of the extended model involve time-dependent coefficients. It is interesting to note that the model is also a special case of (4.6) with $\beta_2(t) = 0$ and Gaussian distribution.

The explicit solution of the formula (5.6) is given by [194]:

$$X_t = e^{-l(t)}\left(X_s + \int_s^t e^{l(u)}\alpha_0(u)du + \int_s^t e^{l(u)}\beta_0(u)dB_t\right), \tag{5.7}$$

where $l(t) = \int_s^t \alpha_1(u)du$. The conditional mean of X_t (with known X_s) is given by:
$e^{-l(t)}\left(X_s + \int_s^t e^{l(u)}\alpha_0(u)du\right)$, and the conditional variance is $e^{-2l(t)}\int_s^t e^{-2l(u)}\beta_0^2(u)du$.

5.2.1 Simulation study

We focus on the problem of the parameter estimation for the time-varying Vasicek model using different lengths of the historical data calibration sample. Then we forecast future values by simulating a number of trajectories from the extended Vasicek model with estimated parameters, and we investigate how the length of the calibration sample impacts the quality of our predictions. Finally, we try to find the 'optimal' calibration window length. Our general methodology consists of several steps which are presented by a flowchart in Fig. 5-11.

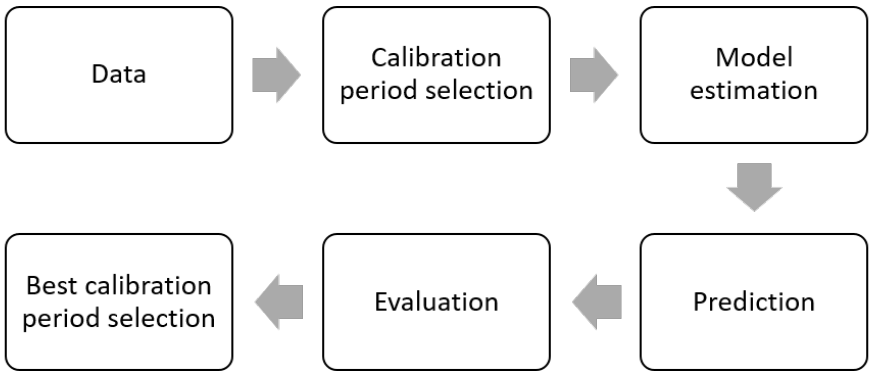


Figure 5-11: The schematic algorithm of the applied methodology, source: [20].

In order to illustrate the formulated problem, we first present our approach to the simulated data set. As our exemplary data, we consider a simulated trajectory (Fig. 5-12) of the model defined in (5.6) for constant coefficients. We use first $252 * 9 = 2268$ observations (on the left-hand side from the dashed vertical line) as a training period for estimation of the model parameters. Note that in the simulated trajectory, we can distinguish three different regimes, i.e., the trajectory consists of three realizations of the Vasicek process, simulated with three different sets of parameters (we denote these three parts X_1, X_2, X_3, respectively). The values of these parameters are shown in Tab. 5.4.

	X_1	X_2	X_3
α_0	0.004	0.007	0.005
α_1	-0.003	-0.004	-0.003
β_0	0.025	0.05	0.015

Table 5.4: Parameters of the Vasicek model for each part of the trajectory in Fig. 5-12, source: [127].

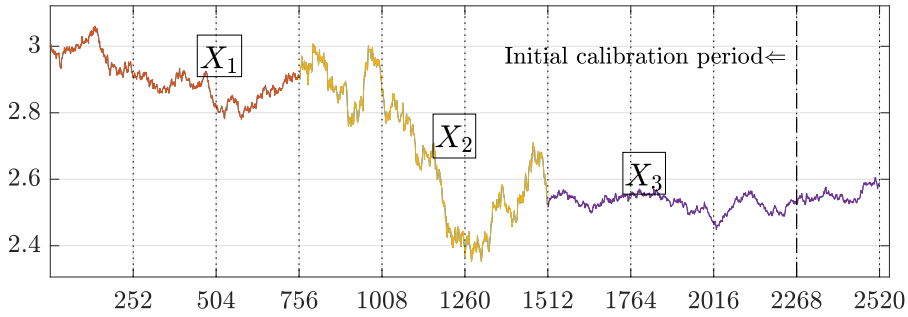


Figure 5-12: A sample simulated trajectory with three distinguishable regimes X_1, X_2 and X_3 (marked with different colors). The dashed, vertical line indicated the end of the training period, source: [20].

The next step requires the selection of a calibration period for the parameter estimation - a training sample consisting of historical data on which we estimate the parameters of the Vasicek model (to later use the estimated model to forecast the future values). A common belief is that the longer the calibration sample is, the better the results are. However, in our case, as can be seen from Fig. 5-12, the data that we want to forecast, i.e., observations from 2267 to 2520 (on the right-hand side from the vertical dashed line), was generated using the same set of parameters as for the last 3*252 observations in the calibration sample (observations from 1513 to 2268). Therefore, with this knowledge, intuitively, taking a long calibration period that takes into account the whole calibration sample should bring less satisfactory results than calibrating our model to

the last 3*252 observations. In our study we consider different lengths of the calibration period (calibration windows) - it ranges from 3*252 up to 9*252 observations for the simulated case. Note that all training samples are left-truncated so that they end on the same observation of the training period, i.e., when considering a 3*252 observation calibration window we take observations from 1513 to 2268, and for 7*252 calibration sample we take observations from 505 to 2268.

Once the parameters of the Vasicek model are estimated from historical data, we predict future values by simulating 10 000 trajectories (each containing 252 observations) using the calibrated model. Next, we evaluate the performance of our estimations using the following measures: mean absolute percentage error (MAPE) and a novel technique introduced in the previous sections, called a validation factor (VF).

The first measure, MAPE, is calculated in the following way:

$$\mathrm{MAPE(cal)} = \frac{1}{NT}\sum_{n=1}^{N}\sum_{t=1}^{T}\frac{|\hat{X}_{t,n}^{cal} - X_t|}{X_t}, \tag{5.8}$$

where $\hat{X}_{t,n}^{cal}$ is the predicted value at prediction time t from the n-th generated trajectory from the validated period, obtained from the Vasicek model calibrated to a training sample of length cal, X_t is the real value at time t, N is the number of generated trajectories and T is the length of each trajectory. We take $N = 10000$ and $T = 252$.

In the next step of assessing our performance, we follow [127] and evaluate the quality of our estimations with the novel method based on the quantile lines computed from the 10000 simulated trajectories. The quantile lines, $\mathcal{Q}(cal)$, are obtained by calculating the empirical quantiles of the simulated values at each time point:

$$\mathcal{Q}(cal) = \begin{bmatrix} Q_{0.05}^{cal}(t_1) & \dots & Q_{0.05}(t_{252}) \\ Q_{0.1}^{cal}(t_1) & \dots & Q_{0.1}(t_{252}) \\ \vdots & & \vdots \\ Q_{0.9}(t_1) & \dots & Q_{0.9}(t_{252}) \\ Q_{0.95}(t_1) & \dots & Q_{0.95}(t_{252}) \end{bmatrix}, \tag{5.9}$$

where $Q_q^{cal}(t)$ is a q-th quantile of the simulated prices from a certain calibration window cal, calculated for values of 10000 simulations at time t. The second error measure, VF, is calculated in the following way:

$$\mathrm{VF(cal)} = \frac{1}{\#\mathrm{P}}\sum_{p\in\mathrm{P}}|\phi(p) - p|^2, \tag{5.10}$$

where $\mathrm{P} = [0.1, 0.2, \ldots, 0.9]$ is a set of possible values of p, $\phi(p)$ is the fraction of real data from the validated period lying between calculated quantile lines $\frac{1-p}{2}$ and $p + \frac{1-p}{2}$ and $\#\mathrm{P}$ is the size of the set P. When evaluating our results with the use of this factor, we pay particular attention to the amount of real data placed between certain quantile lines - if the model is correctly estimated, 90% of data should lie between the 0.05 and 0.95 quantile lines, 50% between the 0.25 and 0.75 quantile lines etc. Similarly to MAPE, the lower the value of the VF is, the better the results of the model parameter estimation are.

The results of the evaluation for the simulated data case are presented in Tab. 5.5. It turns out that, indeed, taking the shortest, 3-year (3*252 observations) calibration sample produced the best results in terms of both MAPE and VF. This simple analysis is a motivation and a starting point to the real data consideration.

	Calibration length						
	3*252	4*252	5*252	6*252	7*252	8*252	9*252
MAPE	0.0161	0.0176	0.0306	0.0276	0.0167	0.0221	0.0266
VF	0.0068	0.0138	0.0376	0.0851	0.0184	0.0181	0.0581

Table 5.5: Results in terms of MAPE and VF for the simulated data case. Columns refer to lengths of the model calibration window. Best performing model is marked in blue, source: [20].

5.2.2 Real data application

In this section, we present the results of our investigation concerning the selection of the calibration window for the estimation of the generalized Vasicek model parameters for real data. We propose here a new approach in the considered issue. We consider daily-frequency of the EUR/USD and USD/PLN currency exchange rates, both spanning from 2nd January 1997 up to 2nd January 2015. We use the first 10 years of data, i.e., from 2 January 1997 to 29 December 2006 (dashed, vertical line in Fig. 5-13 and Fig. 5-14 marks the end of this period) as the initial period for the model calibration. For each currency pair, we examine the influence of the length of a training sample on the quality of our 1-year forecasts - the methodology is similar to the one used in case of simulated data. For each test-year (from 2007 to 2014) we first calibrate the generalized Vasicek model's parameters on the different lengths of historical data and then generate 10 000 exchange rate trajectories for the corresponding year. Then we use the real data to evaluate the predictions in terms of both MAPE and VF - results are shown in Tables 5.6 - 5.7. The columns are labeled from 2 to 10 and correspond to the length of the historical data sample used to estimate generalized Vasicek model's parameters. For each row (which corresponds to a certain test-year) we mark the

best performing calibration window, i.e., we mark the calibration window length for which the generated trajectories produced the best results in terms of a certain error measure (MAPE or VF). This allows us to clearly depict which calibration window length was the 'optimal' one in a certain year.

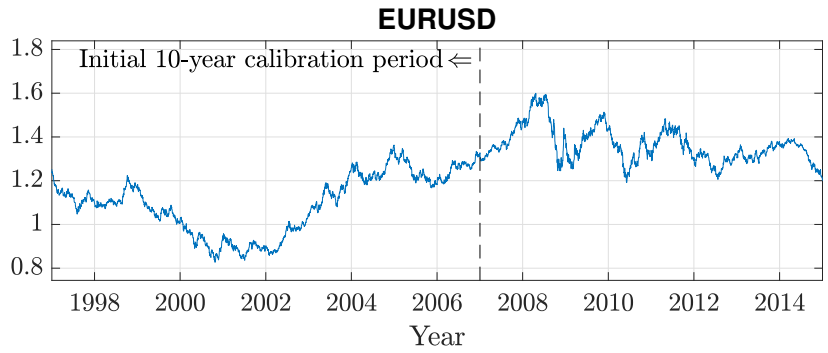


Figure 5-13: EUR/USD daily exchange rates from from 2 January 1997 to 2 January 2015. The dashed, vertical line indicates the end of the initial 10-year calibration period, source: [20].

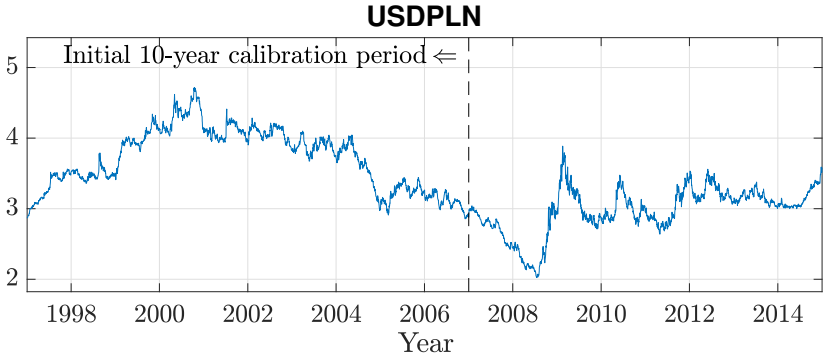


Figure 5-14: USD/PLN daily exchange rates from from 2 January 1997 to 2 January 2015. The dashed, vertical line indicates the end of the initial 10-year calibration period, source: [20].

As we have seen in the previous section, the choice of just one optimal calibration window length for estimating the extended Vasicek model parameters is, in our case, a cumbersome and hardly doable task. The performance of certain windows changes significantly over time, and an inappropriate choice can bring extremely disappointing results by significantly lowering the accuracy of our predictions. Also, it is worth mentioning that as selection of the optimal calibration window was made ex-post, we would not be able to indicate the best calibration sample length in advance. Therefore, in order to tackle this issue, analogously to [195] we examine several combinations of different calibration window lengths and average the predictions obtained from these windows. Then we evaluate the accuracy of the averaged forecast in terms of MAPE and VF.

We consider five different combinations of calibration windows; we denote them by **Avg(·)** and the argument describes the windows that we choose for averaging. We use MATLAB notation in order to refer to a certain combination, e.g. Avg(2:10) refers to averaging predictions from all window lengths from 2 to 10 years while Avg(2:4,8:10) to selecting predictions from the three shortest (2, 3, 4-years) and three longest (8, 9, 10-years) windows.

We use two different approaches to average the predictions - the choice depends on the selection of the evaluation measure.

		Calibration length								
		2	3	4	5	6	7	8	9	10
MAPE	2007	0.0500	0.0512	0.0538	0.0570	0.0534	0.1060	0.0526	0.0567	0.0574
VF		0.0240	0.0052	0.0017	0.0017	0.0149	0.1472	0.0064	0.0025	0.0063
MAPE	2008	0.0668	0.0666	0.0643	0.0632	0.0648	0.0641	0.0675	0.0666	0.0676
VF		0.0694	0.0623	0.0802	0.0870	0.0725	0.0641	0.0436	0.0522	0.0453
MAPE	2009	0.1231	0.1158	0.1066	0.0903	0.0713	0.0708	0.0702	0.0717	0.0688
VF		0.0443	0.0359	0.0106	0.0163	0.0017	0.0016	0.0293	0.0018	0.0037
MAPE	2010	0.0970	0.0946	0.0939	0.1821	0.1096	0.1171	0.1187	0.0890	0.0880
VF		0.0710	0.0699	0.0628	0.0901	0.0066	0.0073	0.0051	0.0341	0.0286
MAPE	2011	0.0689	0.0889	0.0918	0.0527	0.0528	0.0611	0.0651	0.0568	0.0695
VF		0.0039	0.0203	0.0198	0.0208	0.0130	0.0095	0.0040	0.0267	0.0973
MAPE	2012	9.2977	0.0664	0.0895	0.0855	0.0525	0.0479	0.0513	0.0533	0.0643
VF		0.1043	0.0769	0.1048	0.1063	0.0200	0.0313	0.0388	0.0457	0.0685
MAPE	2013	0.0461	0.0305	0.0359	0.0617	0.0619	0.0367	0.0523	0.0426	0.0427
VF		0.0583	0.0157	0.0329	0.1027	0.1085	0.0355	0.0017	0.0559	0.0835
MAPE	2014	0.0486	0.0592	0.0615	0.0552	0.0792	0.0522	0.0330	0.0505	0.0539
VF		0.0085	0.0157	0.0191	0.0098	0.0568	0.0055	0.0275	0.0117	0.0071

Table 5.6: Results in terms of MAPE and VF for the EUR/USD data. The columns refer to the lengths of the model calibration window (in years). The best performing model for each year is marked in blue, source: [20].

Averaging approach when evaluating MAPE

When evaluating the forecasts in terms of MAPE, we take every simulated trajectory from each calibration window length and average their values with corresponding trajectories from different windows in the chosen combination. As an example we take Avg(2:4) - we average predictions of currency pairs of exchange rates obtained from the extended Vasicek model calibrated to 2, 3, and 4-year historical data samples. For each calibration window lengths, we obtain 10000 simulated trajectories of the underlying' price. We take the first trajectory for each of the calibration window length and average their values - we obtain a new, averaged trajectory of the same length (1-year).

		Calibration length								
		2	3	4	5	6	7	8	9	10
MAPE	2007	0.0950	22.9650	0.1073	0.0942	0.1044	0.1042	0.1019	0.1029	0.0982
VF		0.0022	0.0191	0.0038	0.0087	0.0044	0.0032	0.0036	0.0027	0.0043
MAPE	2008	0.1169	0.1125	0.1087	0.1178	0.1118	0.1226	0.1377	0.1345	0.1360
VF		0.0631	0.0807	0.0934	0.1414	0.1366	0.1201	0.0269	0.0357	0.0363
MAPE	2009	0.2582	0.1774	0.2134	0.1184	0.1223	0.1127	0.1109	0.1155	0.1179
VF		0.0409	0.0132	0.0826	0.0130	0.0315	0.0156	0.0293	0.0101	0.0124
MAPE	2010	0.1423	0.0928	0.1269	0.1741	0.2043	0.1779	0.1016	0.1040	0.0984
VF		0.0673	0.0095	0.0513	0.0487	0.1004	0.1139	0.0466	0.0229	0.0404
MAPE	2011	0.1110	0.1507	0.1479	0.0733	0.0757	0.0977	0.1080	0.0828	0.0782
VF		0.0260	0.0438	0.0560	0.0488	0.0101	0.0070	0.0274	0.0110	0.0129
MAPE	2012	0.1370	0.1052	0.1277	0.1421	0.1239	0.0744	0.0797	0.0849	0.0752
VF		0.0303	0.0081	0.0198	0.0258	0.0155	0.0338	0.0254	0.0072	0.0176
MAPE	2013	0.0656	0.0811	0.0809	0.1251	0.0716	0.1171	0.1123	0.0640	0.0703
VF		0.0702	0.0740	0.0988	0.1316	0.0901	0.1515	0.1358	0.0538	0.0608
MAPE	2014	0.0728	0.0770	0.0835	0.0951	0.1120	0.0954	0.0727	0.0674	0.0654
VF		0.0414	0.0480	0.0390	0.0690	0.1193	0.0695	0.0241	0.0180	0.0245

Table 5.7: Results in terms of MAPE and VF for the USD/PLN data. The columns refer to lengths of the model calibration window (in years). The best performing model for each year is marked in blue, source: [20].

This procedure is repeated for all trajectories and for each test-year. Then, MAPE is calculated for new, 10 000 averaged trajectories in the same way as was shown in Section 5.2.1.

Averaging approach when evaluating VF

The second error measure, a validation factor, assesses the performance of predictions based on the quantile lines obtained from simulated trajectories. When averaging predictions from a certain combination of windows, for each calibration window length, we calculate the values of quantile lines from the generated samples. Then we average their values (across different window lengths) for each of the quantiles $0.05, \ldots, 0.95$ separately and obtain new, averaged quantile lines. Then we evaluate the performance of a certain combination by calculating the value of the VF for averaged quantile lines. An exemplary procedure of averaging quantile lines for two calibration windows is shown in Fig. 5-15.

5.3 Discussion and summary

In this chapter we have proposed a stochastic model which address the characteristics of financial data, that is the non-Gaussian behaviour and time-inhomogeneous character. The results presented

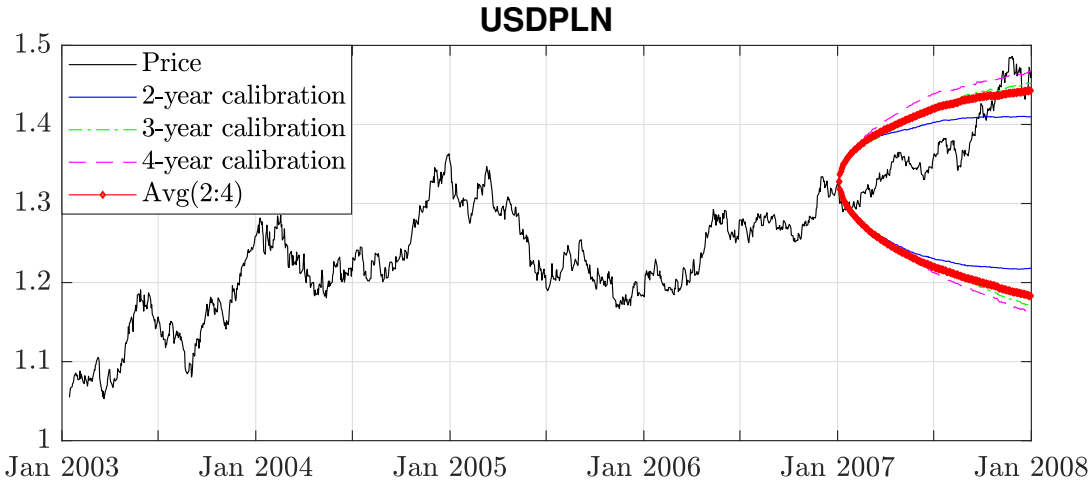


Figure 5-15: Method of obtaining averaged quantile lines, shown only for 0.05 (lower curves) and 0.95 quantiles (upper curves). The red, dotted lines, which represent averaged quantile lines, are obtained by averaging the values of three other quantile lines, source: [20].

show that the CKLS/SGT model is universal and can be useful in describing currency exchange rates with different behaviour and is more appropriate for the considered cases than the classical GBM model.

In addition, we discussed the problem of the calibration data length used for the prediction. We have demonstrated that, depending on the test-year data length and starting point, one might obtain different results and in the long term, the averaging of different models can give better and more stable results in modelling exchange rates.

Chapter 6

Multidimensional market risk prices modelling

In the previous chapters, we have concentrated on one-dimensional market risk factors modelling. However, in business reality, very often there is a necessity to monitor and model the behaviour of two or more risk factors at the same time, because they are interlinked and only a holistic approach gives a proper overview of the market dynamics and the useful information required for market risk management.

Adding a business context to multidimensional modelling, an example of the most important market risk factors for KGHM are the price of copper (Cu in US Dollars) and the USDPLN exchange rate. From the market risk measurement and modelling perspective, the copper price in Polish zloty (Cu in PLN) effectively drives revenues, cash flows, earnings and therefore, in fact, the financial situation of the company. Therefore, the main problem considered can be formulated as follows: how to properly model the range of Cu in PLN, taking into account the changing in time dependence between the USDPLN exchange rate and the copper price in USD, under the additional assumption that both assets have non-Gaussian impulsive behaviour. Although the formulated problem seems to be dedicated to a specific case, it can be considered as a general one, namely the modelling of the range of values for the metal price expressed in the currency of a given country, when the relationship between the metal price and the exchange rate (between the original currency and the national currency) is changing over time. In the considered case, it is important to capture the non-Gaussian characteristics, changing regimes, and non-constant relation between risk factors, as such instability can create additional risk for the company if it occurs.

Based on the relation and distribution analysis, we propose in this chapter a two-dimensional VAR model with the $\alpha-$stable distribution for which the non-homogeneity of the data is reflected

in two identified regimes. The proposed VAR(1) model is a generalization of the AR(1) model, which in turn is a generalisation of the Orstein-Uhlenbeck process presented in Chapter 4. The VAR(1) model has however a discrete character, whereas the models introduced in Chapter 4 are continuous. Additionally, the $\alpha-$stable distribution has a common part with the class of SGT distributions, meaning Cauchy distribution. The use of the $\alpha-$stable distribution is advantageous due to the easy definition of a stable multivariate distribution and the fact that it is a generalization of multivariate normal distribution. The results presented in this chapter are published in [131].

6.1 Relation between assets

Fig. 6-1 presents weekly copper prices in USD (top panel) and the USDPLN exchange rate (bottom panel) from the period January 7th, 2000 - October 2nd, 2020. The logarithmic returns of the analysed data are demonstrated in Fig. 6-2 in the top and middle panels, respectively. One can see the specific behaviour of the time series. On one hand, it is clearly seen that both data sets exhibit a non-Gaussian behaviour with visible large observations that may suggest a heavy-tailed distribution of the time series. This behaviour is especially apparent in the logarithmic returns of the examined assets. On the other hand, one can see that the data have a non-homogeneous structure - there are visible regimes of the time series that do not fit the overall pattern. Special attention should be paid to the period from the year 2008 to 2012, where the large observations are more frequent than in the other periods. This behaviour is especially visible for copper price logarithmic returns (see the top panel of Fig. 6-2). The non-homogeneity of the time series indicates that the parameters switch in time and one stationary model can not be used for the whole data description.

Despite the fact that there is no direct economic relationship between the copper price (in USD) and the USDPLN exchange rate, even a brief look at the charts shows that there is a negative relation between these risk factors, see Fig. 6-2, the bottom panel. From the market risk management point of view, such a statement is very important, especially if these assets are the major risk factors in the portfolio. However, if that relation were to turn to be apparent or unstable during the time, it could lead to misleading conclusions and as a consequence incorrect decisions.

The mentioned relation between the considered assets can be expressed by different measures of dependence. Here we examine three metrics, namely Pearson, Spearman rank, and Kendall rank correlation coefficients. The definitions and properties of these measures as well as their sample versions are presented in the next section. Here we only mention that the Pearson correlation is effective for the data with light-tailed distribution (like the Gaussian one), while the two other correlation measures can also be used for the heavy-tailed distributed time series. In order

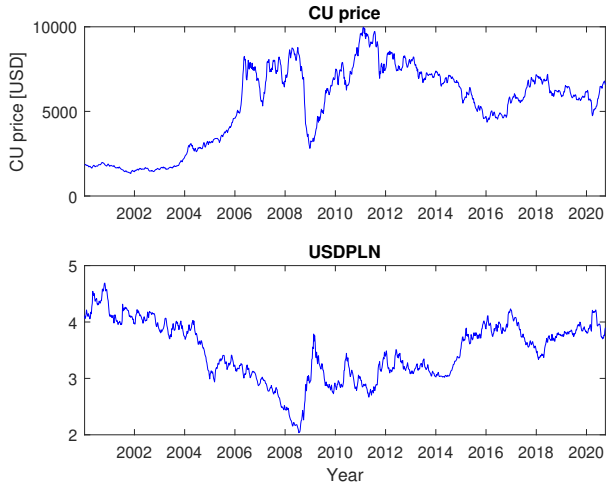


Figure 6-1: The weekly data corresponding to the copper price in USD (top panel) and USDPLN exchange rate (bottom panel) from the period Jan 7th, 2000 - October 2nd, 2020, source [131].

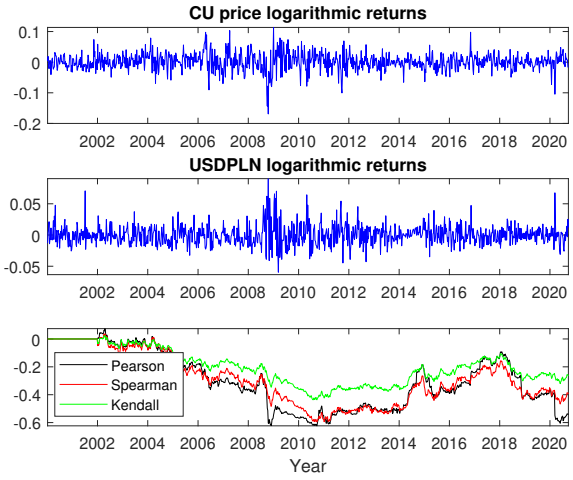


Figure 6-2: The logarithmic returns of weekly copper price in USD (top panel), USDPLN exchange rate (middle panel), and the corresponding sample Pearson, Spearman rank, and Kendall rank correlation coefficients calculated for two-yearly windows (104 observations), source [131].

to demonstrate the dynamics of the dependence structure, the sample correlation coefficients are calculated for the data from a moving window corresponding to a two-year period (104 observations). As one can see in Fig. 6-2, the sample correlation coefficients change over time, which indicates the dynamic structure of the relationship between the assets. Moreover, for the time period between the years 2006 and 2012, there is a clear difference between the Pearson correlation coefficient and the two other measures (especially the Kendall rank correlation coefficient). This

phenomenon may confirm the non-homogeneous structure of the data. Moreover, the heavy-tailed distribution may also influence the difference between the dependency measures.

The specific characteristics of the data described above are the motivation for using the non-Gaussian models for their description. Moreover, the visible relationship between the analysed assets implies that a multidimensional model needs to be applied. The dynamics of the correlation coefficients and the non-homogeneous behaviour of the data indicate that in the first step of the analysis the examined time series should be divided into regimes of the homogeneous structure.

6.2 Methodology

In this section, we present the general methodology used in the subsequent analysis. The demonstrated methods are known from the literature, thus we recall only the main concepts, definitions, and properties.

6.2.1 The dependence structure description for Gaussian and non-Gaussian time series

Different measures of dependence between variables may be used to describe the interplay of different elements in a complex system and the strength of their relationship. We focus on three of the most broadly used measures, namely, the Pearson correlation, the Spearman rank correlation, and the Kendall rank correlation [143].

The most common dependency measure between two random variables is the Pearson correlation coefficient ρ_P which for a random vector (X, Y) is defined as follows [196]:

$$\rho_P = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}, \tag{6.1}$$

where $cov(\cdot, \cdot)$ is the covariance function, σ_X is the standard deviation of X and σ_Y is the standard deviation of Y. The sample version of ρ_P for $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, denoted as r_P is defined as [196]:

$$r_P = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}, \tag{6.2}$$

where $\overline{x}$, $\overline{y}$ are sample means of the data vectors x and y, respectively.

The Pearson correlation is used to study the linear relationship between two variables. It takes values in the range $[-1, 1]$. Values close to 1 and -1 indicate a strong relationship (positive and

negative, respectively). A value of 0 means no linear dependence. The Pearson correlation coefficient is sensitive to outliers. Therefore, this measure is useful especially for Gaussian (or light-tailed) distributed variables.

The Spearman rank correlation coefficient for a random vector (X, Y) has the following form [197, 198]:

$$\rho_S = \frac{\text{cov}(Q, W)}{\sigma_Q \sigma_W}, \tag{6.3}$$

where $cov(\cdot, \cdot)$ is the covariance between X and Y, (Q, W) is a random vector of ranks corresponding to (X, Y), and σ_Q and σ_W are the standard deviations of variable Q and W, respectively. The sample version of the Spearman rank correlation coefficient for a random bi-dimensional sample $(x_1, y_1), \ldots, (x_n, y_n)$ being the realisation of the random vector (X, Y) is defined as follows:

$$r_S = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(q_i - \overline{q})(w_i - \overline{w})}{[\frac{1}{n-1}\sum_{i=1}^{n}(q_i - \overline{q})^2 \frac{1}{n-1}\sum_{i=1}^{n}(w_i - \overline{w})^2]^{1/2}}, \tag{6.4}$$

where q_i, w_i are the empirical counterpart of the random variables Q and W, whereas $\overline{q}$ and $\overline{w}$ are the sample means of the relevant rank samples. The Spearman rank correlation takes values between $[-1, 1]$ and verifies a monotonic relationship. This measure is especially useful in the case when the analysed variables are non-Gaussian. It is insensitive to large observations and thus in these cases indicates the relation more adequately than the Pearson correlation.

The last considered measure is the Kendall rank correlation. Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be a random sample corresponding to the random vectors (X, Y). The sample Kendall rank correlation coefficient is defined as follows [199]:

$$r_K = \frac{2}{n(n-1)} \sum_{1 \le i \le j \le n} J((x_i, y_i), (x_j, y_j)), \tag{6.5}$$

where $J((x_i, y_i), (x_j, y_j)) = \text{sgn}(x_i - y_i)\text{sgn}(x_j - y_j)$ and $J((x_i, y_i), (x_j, y_j)) = 1$, if a pair (x_i, y_i) is concordant with a pair (x_j, y_j), i.e. if $(x_i - x_j)(y_i - y_j) > 0$; $J((x_i, y_i), (x_j, y_j)) = -1$, if a pair (x_i, y_i) is discordant with a pair (x_j, y_j), i.e. if $(x_i - x_j)(y_i - y_j) < 0$. The Kendall rank correlation coefficient is based on the difference between the probability that two variables are in the same order (for the observed data vector) and the probability that their order is different. In the formula (6.5) it is required that the variable values can be ordered. This coefficient takes values between $[-1, 1]$. A value equal to 1 means full accordance, a value equal to 0 does not match orders, while a value equal to -1 means the complete opposite accordance. The Kendall rank correlation coefficient indicates not only the strength but also the direction of the dependence. Similarly to the Spearman rank correlation, it is resistant to outliers and is used especially for the non-Gaussian

distributed data [143]. Despite the mentioned correlation coefficients, other dependency measures adequate for heavy-tailed distributed data are also considered in the literature, see for instance [200–202].

6.2.2 The $\alpha-$stable distribution

The analysed data describing metals' price and currency exchange rates are non-Gaussian. To model such a specific behaviour, we propose to apply the $\alpha-$stable distribution. Below we recall the corresponding definition and the main properties.

For the $\alpha-$stable distribution the probability density function (PDF) and the cumulative distribution function (CDF) are given in close form only in a few special cases. Therefore, a common way to define the distribution of an $\alpha-$stable random variable Z is by determining its characteristic function [203]:

$$[\exp\{i\theta Z\}] = \exp\left\{-\sigma^{\alpha}|\theta|^{\alpha}\left(1 + i\beta w(t,\theta)\right) + i\mu\theta)\right\}, \tag{6.6}$$

where:

$$w(\theta,\alpha) = \begin{cases} -\operatorname{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right) & \text{if } \alpha \neq 1, \\ \frac{2}{\pi}\operatorname{sgn}(\theta)\ln|\theta| & \text{if } \alpha = 1, \end{cases} \tag{6.7}$$

and $\operatorname{sgn}(\cdot)$ denotes a sign function. The parameter $0 < \alpha \leq 2$ is called the stability index and regulates the rate at which the distribution tails converge. The other parameters are: the scale parameter $\sigma > 0$, the skewness parameter $-1 \leq \beta \leq 1$, and the shift parameter $\mu \in \mathbb{R}$. If $\beta = \mu = 0$, the distribution of Z is symmetric with respect to 0 and the characteristic function given in Eq. (6.6) simplifies to the following one:

$$[\exp\{i\theta Z\}] = \exp\left\{-\sigma^{\alpha}|\theta|^{\alpha}\right\}. \tag{6.8}$$

It is worth emphasizing that the $\alpha-$stable distribution with $0 < \alpha < 2$ constitutes a generalization of the Gaussian distribution corresponding to the case of $\alpha = 2$. For the non-Gaussian distribution, the properties differ significantly from the ones corresponding to $\alpha = 2$: the tails converge to zero according to a power function and the second moment is infinite. Additionally, for $0 < \alpha \leq 1$ the first moment is also infinite. The above-mentioned facts yield non-Gaussian $\alpha-$stable random variables that take extreme values more likely than is observed in the Gaussian case.

6.2.3 Vector autoregressive model with $\alpha-$stable distribution

The VAR model with the α-stable distribution is defined as an extension of the classical model where the innovations are assumed to be Gaussian distributed (or at least have finite second moments), see for example [204]. In the α-stable non-Gaussian case with infinite variance, the VAR system can be used to model the data exhibiting a higher likelihood of more extreme events.

A time series $\{\mathbf{X}(t)\} = \{(X_1(t), \ldots, X_m(t))^T\}$ is called a vector autoregressive model with α-stable distribution, if for each $t \in \mathbb{Z}$ it satisfies the following system of equations:

$$\mathbf{X}(t) - \Theta_1 \mathbf{X}(t-1) - \ldots - \Theta_p \mathbf{X}(t-p) = \mathbf{Z}(t), \tag{6.9}$$

where $\{\mathbf{Z}(t)\} = \{(Z_1(t), \ldots, Z_m(t))^T\}$ is an m-dimensional $\alpha-$stable random vector and $\Theta_1, \ldots, \Theta_p$ are $m \times m$ matrices with time-constant coefficients. For simplicity, we assume here that the noise vector $\mathbf{Z}(t)$ consists of independent α-stable distributed random variables, i.e. $Z_i(t)$ and $Z_j(t)$ are independent for any $t \in \mathbb{Z}$ when $i \neq j$, and the characteristic function of $Z_i(t)$ is given by Eq. (6.6) for all $t \in \mathbb{Z}$. Moreover, the vector $\mathbf{Z}(t)$ is assumed to be independent of the vector $\mathbf{Z}(s)$ for $t \neq s$, where $t, s \in \mathbb{Z}$.

The conditions for the existence and uniqueness of the bounded solution of a vector autoregressive time series with multidimensional $\alpha-$stable distribution are provided in [205]. Note that for $p = 1$, the bounded solution of the autoregressive system of order 1 takes the form:

$$\mathbf{X}(t) = \sum_{j=0}^{+\infty} \Theta^j \mathbf{Z}(t-j), \tag{6.10}$$

assuming that the elements of Θ^j are absolutely summable, i.e., if the eigenvalues of Θ are less than 1 in the absolute value, where $\Theta = \Theta_1$ in Eq. (6.9). It should be mentioned that in the case when the coefficients of the matrices in Eq. (6.9) responsible for the relationship between time series components are zero, then the VAR model reduces to m independent one-dimensional $\alpha-$stable autoregressive (AR) time series [31].

In the classical (Gaussian) version of the VAR system, the dependence structure of the process can be described using the covariance or the correlation. As a consequence, to estimate the parameters of the system, one often uses the multidimensional Yule-Walker method based on the auto-covariance function [206]. However, since for the VAR model with non-Gaussian $\alpha-$stable distribution the second moment is infinite, there is no theoretical justification for using the covariance-based method to estimate the unknown parameters. Therefore, in [207] the authors propose the modified Yule-Walker method, similarly to the one-dimensional case examined

in [208, 209], which is based on the covariation well defined for the $\alpha-$stable distribution with $\alpha > 1$. The covariation can be also used, instead of the auto-covariance, to quantify the interdependence within a time series. In this case, the measure is called auto-covariation and it is defined in the following way:

$$\mathrm{CV}(X_i(t), X_i(t-h)) = \frac{\mathbb{E}[X_i(t)X_i(t-h)^{\langle p-1\rangle}]}{\mathbb{E}[|X_i(t-h)|^p]}\sigma^{\alpha}_{X_i(t-h)}, \tag{6.11}$$

where:

$$x^{\langle a\rangle} = |x|^a \mathrm{sign}(x), \tag{6.12}$$

$1 < p < \alpha$ and $\sigma_{X_i(t-h)}$ is the scale parameter of the random variable $X_i(t-h)$. In practice, one often estimates the so-called normalized auto-covariation from the data, i.e., the auto-covariation divided by the parameter $\sigma_{X_i(t-h)}$. The appropriate estimators are presented in [208–210].

6.3 Two-dimensional analysis of the copper price (in USD) and USDPLN exchange rate

In this section, we present the analysis of two-dimensional modelling for data corresponding to the copper price (in USD) and the USDPLN exchange rate. The visual inspection of the logarithmic returns (see Fig. 6-2 top and middle panels) of the considered assets and the evident difference between the correlation coefficients (see Fig. 6-2 bottom panel) indicate that the data are related, however, they should be divided into parts of the homogeneous structure. To do this, we assume that the one-dimensional time series of logarithmic returns follows a symmetric $\alpha-$stable distribution with parameters σ and α switching between two values. These two parameter sets (σ_1, α_1) or (σ_2, α_2) correspond to an unobserved state process and, hence, reflect the changes in market conditions. To estimate the moments of switching, we apply a Hidden Markov Model approach [211] and assume that the state process is driven by a Markov chain with probabilities of changing the states given by a transition matrix. The HMM estimation procedure is based on the expectation-maximization algorithm, [212], designed to infer parameters in the models depending on latent variables (here the state process). As a by-product of the EM algorithm, we obtain the probabilities of the two states for each time point. These probabilities are then used for the identification of different regimes within the time series. Namely, for each value of the logarithmic returns, we assign the state that is more probable. The resulting regime classification for both variables is illustrated in Fig. 6-3. The obtained results are consistent with the market situation reflected in the

analysed data. For a few years before and after the peak of the great financial crisis (2008) volatility in the market among many different assets stayed at elevated levels, whereas in other periods the market moves were significantly weaker. One of the reasons for this issue could be that before and after the crisis the valuations of many assets achieved extreme levels, with very dynamic changes also reflected in currency markets. In the case of the copper price, the data showed that elevated volatility was observed even earlier as a result of substantial incremental Chinese demand growth dynamics. For further analysis, we decided to choose the overlapping regimes timing for both assets.

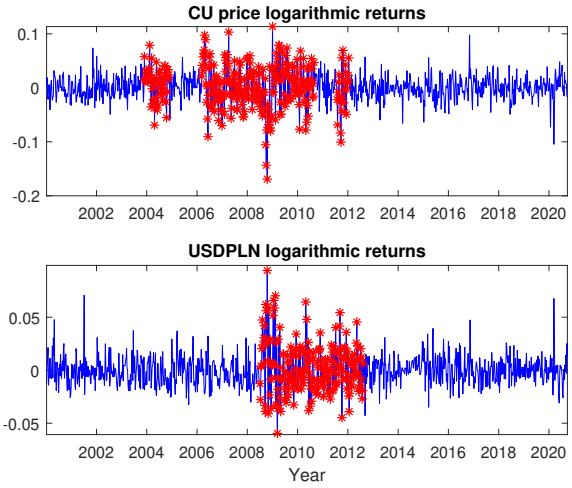


Figure 6-3: The logarithmic returns of the weekly copper price in USD (top panel) and USDPLN exchange rate (bottom panel) with regimes obtained via the HMM classification procedure, source: [131].

6.3.1 The $\alpha-$stable VAR modelling involving relationship between the considered assets

In this section, we model the logarithmic returns of the copper price (in USD) and the USDPLN exchange rate using the two-dimensional VAR time series with the α-stable distribution, described in Section 6.2. We assume the simplest version of the model, namely VAR(1). In this approach, we allow for a possible dependence between the considered assets, in contrast to the second approach presented in the next subsection. Taking into account the regime identification step, we assume that the parameters of the VAR(1) model change at a certain point in time. Therefore, we separately consider regime 1 and regime 2 marked in Fig. 6-4. In our analysis, we assume that regime 1 starts when first of the assets (copper price in USD or USDPLN exchange rate) falls into this regime

due to the HMM classification step (see in Fig. 6-3). Note that we omit the short period in 2004, where regime 1 was identified for the copper price, since there is no corresponding regime change in the USDPLN exchange rate. Thus, regime 1 starts in March of 2006. In the economical context of world exchanges, including commodities markets, there was a dynamic growth of assets value starting from 2006, which led eventually to a financial crisis outbreak two years later.

The end of regime 1 is specified as the second half of 2012 when the situation on the market started to stabilize and the classification results indicate the existence of the second regime for both assets.

The final regimes segmentation is plotted in Fig. 6-4. In practice, we separately fit the VAR(1) model for each regime.

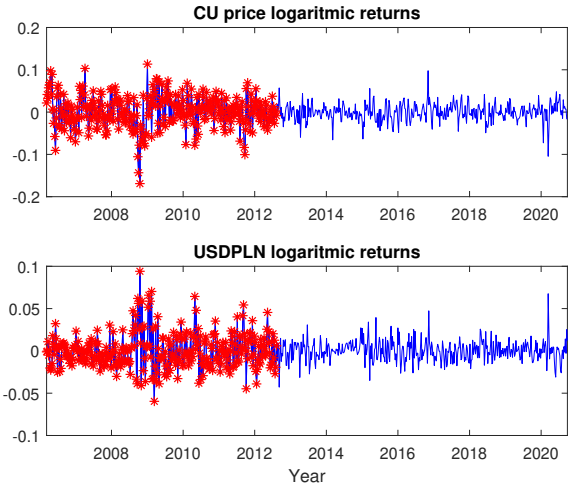


Figure 6-4: The logarithmic returns of the weekly copper price in USD (top panel) and USDPLN exchange rate (bottom panel) with marked regimes used for time series modelling. Regime 1 is marked with red stars, while regime 2 is marked in blue, source: [131].

The coefficients matrices of the two-dimensional VAR(1) models estimated based on the data corresponding to regime 1 and regime 2, respectively, are given by:

$$\widehat{\Theta}_{regime1} = \begin{bmatrix} 0.2706 & -0.0569 \\ 0.0063 & 0.2134 \end{bmatrix},$$
$$\widehat{\Theta}_{regime2} = \begin{bmatrix} 0.3119 & 0.1403 \\ 0.0010 & 0.1472 \end{bmatrix}. \tag{6.13}$$

The estimation results indicate that there exists a relation between the considered factors in both regimes. The parameters related to the dependence between the assets lie outside the main diagonal and one can see they take non-zero values. The corresponding residual time series are presented in Fig. 6-6. We recall that in the VAR(1) model the residual vectors are assumed to be independent and identically distributed. This also holds in a one-dimensional sense, i.e., for the components of the residual vectors treated separately. Therefore, in Fig. 6-5 we plot the corresponding auto-covariation functions which indicate a non-zero value only for $h = 0$. Recall that the auto-covariation function, similarly to the auto-covariance function for the Gaussian (or general light-tailed) case, indicates the interdependence of the examined time series.

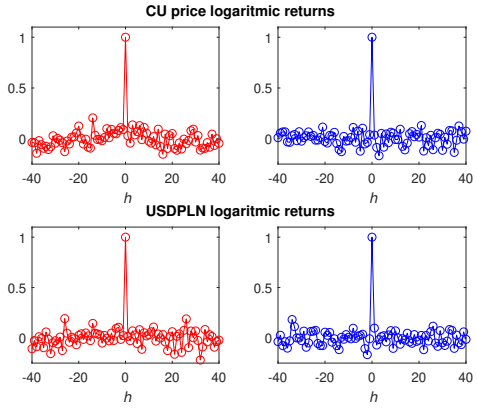


Figure 6-5: The normalized auto-covariation functions of the residuals corresponding to the VAR(1) models applied to the logarithmic returns of the weekly copper price in USD and USDPLN exchange rate for the data from regime 1 (left panels) and regime 2 (right panels), source: [131].

To demonstrate that the residuals of the model are not Gaussian distributed, we use five goodness-of-fit tests based on the distances between the empirical and theoretical cumulative distribution functions (CDF). The empirical CDF is calculated for the residual series, while the theoretical one is the CDF of the Gaussian distribution with parameters estimated from the residual series. Here we use the following statistical tests: Kolmogorov-Smirnov test (T1) [213], Kuiper test (T2) [214], Watson test (T3) [163], Cramer-von Mises test (T4) [215] and Anderson-Darling test (T5) [216]. In Table 6.1 we present the results of the tests: the statistics values and the p-values, respectively. As one can see, the p-values for the H_0 hypothesis of the Gaussian distribution are relatively small, which indicates that Gaussianity is rejected for most of the considered cases at the standard 0.05 significance level.

This result is a motivation for the $\alpha-$stable distribution testing. We use the same test procedures, namely T1-T5, however, the theoretical CDF is calculated for the $\alpha-$stable distribution with

the parameters fitted to the corresponding residual series, see Table 6.1 for the results of the tests. All considered tests indicate that there is no evidence to reject the null hypothesis that the residuals are α-stable distributed (the corresponding p-values are higher than the standard confidence level of 0.05). In Table 6.2 we present the results of fitting the one-dimensional α-stable distribution to each residual time series separately. For the estimation of the $\alpha-$stable distribution parameters, we applied the regression method [217]. Note that for both assets we obtained the higher σ values in the first regime, with the value being almost two times larger than in the second regime. It shows that there was a significant change in the scale of the market fluctuations between the regimes.

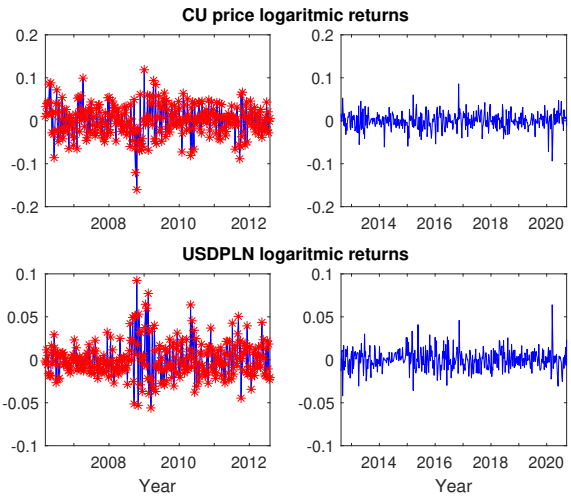


Figure 6-6: The residuals corresponding to the VAR(1) model applied to the logarithmic returns of weekly copper price in USD and USDPLN exchange rate for the data from regime 1 (left panels) and regime 2 (right panels), source: [131].

Using a two-dimensional $\alpha-$stable VAR(1) model is consistent with the economic reality, where the relation between apparently unrelated market factors, which are in our case the copper price (in USD) and the USDPLN exchange rate, is visible and even growing over time in recent times, due to the large amounts of money put by central banks into circulation.

6.3.2 The $\alpha-$stable VAR modelling involving no relationship between the considered assets

In this part, we present the results obtained under the assumption that the relation between the copper price (in USD) and USDPLN exchange rate is negligible and can be omitted. Thus, in this approach it is assumed that in the two-dimensional $\alpha-$stable VAR(1) model the coefficients outside the main diagonal are equal to zero, so the components are independent. Similarly to the

Table 6.1: The results of the statistical tests (statistics values and p-values given in brackets) verifying the null hypotheses that the residual time series corresponding to the two-dimensional VAR(1) models are Gaussian or α-stable distributed, source: [131].

T1	T2	T3	T4	T5
CU Price REGIME 1				
H_0: Residuals are Gaussian distributed				
0.7807	1.4246	0.1187	0.1151	0.6901
(0.1580)	(0.0730)	(0.0530)	(0.0460)	(0.0590)
H_0: Residuals are α-stable distributed				
0.5943	1.0813	0.0545	0.0553	0.3203
(0.4080)	(0.3230)	(0.2980)	(0.2260)	(0.3930)
USDPLN REGIME 1				
H_0: Residuals are Gaussian distributed				
1.3233	2.2381	0.5617	0.4688	3.4434
(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
H_0: Residuals are α-stable distributed				
0.6142	0.8775	0.0247	0.0179	0.4046
(0.5440)	(0.7760)	(0.8850)	(0.9520)	(0.4600)
CU Price REGIME 2				
H_0: Residuals are Gaussian distributed				
0.9345	1.6041	0.2428	0.2353	1.6185
(0.0380)	(0.0210)	(0.0030)	(0.0020)	(0.0010)
H_0: Residuals are α-stable distributed				
0.4141	0.7056	0.0146	0.0153	0.1326
(0.9090)	(0.9550)	(0.9910)	(0.9730)	(0.9680)
USDPLN REGIME 2				
H_0: Residuals are Gaussian distributed				
1.1238	1.9029	0.2293	0.2250	1.4983
(0.0010)	(0.0010)	(0.0020)	(0.0010)	(0.0010)
H_0: Residuals are α-stable distributed				
0.4645	0.8962	0.0389	0.0359	0.3061
(0.7880)	(0.6650)	(0.4680)	(0.4780)	(0.3560)

Table 6.2: The parameters of the α-stable distribution estimated for the residual time series corresponding to the two-dimensional VAR(1) models, source: [131].

α	σ	β	μ
CU Price REGIME 1			
1.9219	0.0236	-0.5714	0.0007
USDPLN REGIME 1			
1.7229	0.0114	1.0000	0.0016
CU Price REGIME 2			
1.8243	0.0119	-0.3416	-0.0004
USDPLN REGIME 2			
1.8424	0.0074	-0.0160	0.0000

previous case, we fit the two models to the data separately for regime 1 and regime 2, where the regimes are chosen in the same manner as previously. The coefficients of the models in the second considered approach are as follows:

$$\widehat{\Theta}_{regime1} = \begin{bmatrix} 0.2927 & 0 \\ 0 & 0.2100 \end{bmatrix}, \qquad \widehat{\Theta}_{regime2} = \begin{bmatrix} 0.2810 & 0 \\ 0 & 0.1386 \end{bmatrix} \tag{6.14}$$

and the residual time series are presented in Fig. 6-8.

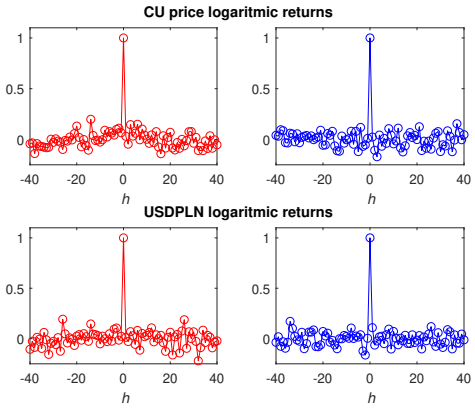


Figure 6-7: The normalized auto-covariation functions of the residuals corresponding to the VAR(1) model with independent components applied to the logarithmic returns of the weekly copper price in USD and USDPLN exchange rate for the data from regime 1 (left panel) and regime 2 (right panel), source: [131].

In Fig. 6-7 we additionally plot the corresponding auto-covariation functions of the residuals having non-zero values only for $h = 0$. Similarly to the first approach, in Table 6.3 we present the results for the Gaussian and $\alpha-$stable distribution testing for the residual series. All tests indicate no evidence in favor of rejecting the hypothesis about the α-stable distribution. On the other hand, the p-values for the Gaussian distribution testing are significantly smaller than for the $\alpha-$stable distribution. In Table 6.4 we present the estimated parameters of the α-stable distribution for the residual time series.

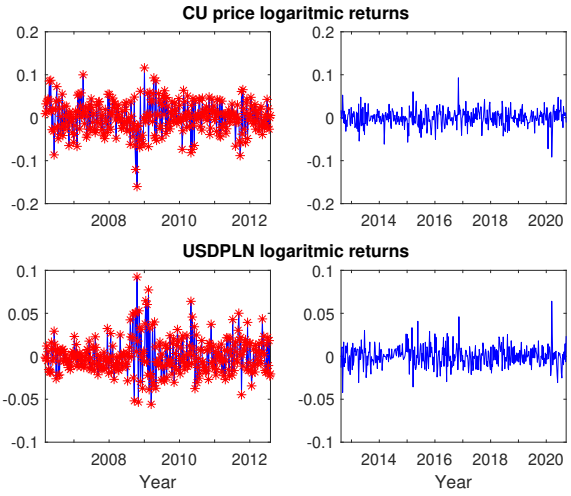


Figure 6-8: The residuals corresponding to the VAR(1) model with independent components applied to the logarithmic returns of the weekly copper price in USD and USDPLN exchange rate for the data from regime 1 (left panel) and regime 2 (right panel), source: [131].

6.4 Real data application. Modelling of the copper price in PLN. The comparative study

Based on the models fitted to the logarithmic returns of the market quotations of the copper price in USD and the USDPLN exchange rate we also infer the dynamics of the copper price in PLN, which is the main risk factor in the mining company KGHM. To this end, we simulate the trajectories of the copper price in USD and USDPLN exchange rate using the fitted two-dimensional $\alpha-$stable VAR(1) models with the parameters given in Eq. (6.13) and in Eq. (6.14), respectively, i.e., when the relationship between assets is taken under consideration or not. Then, the trajectories of the copper price in PLN are obtained as a product of the basic variables. The simulated trajectories are further used to derive the distribution of the copper prices in PLN. The obtained distributions are

Table 6.3: The results of the statistical tests (statistics values and p-values given in brackets) verifying the null hypotheses that the residual time series corresponding to the VAR(1) model with independent components are Gaussian or α-stable distributed, source: [131].

T1	T2	T3	T4	T5
CU Price REGIME 1				
H_0: Residuals are Gaussian distributed				
0.7223	1.3586	0.1189	0.1151	0.6882
(0.2510)	(0.1330)	(0.0730)	(0.0590)	(0.0860)
H_0: Residuals are α-stable distributed				
0.6459	1.1040	0.0574	0.0582	0.3299
(0.3120)	(0.2990)	(0.2730)	(0.2050)	(0.3800)
USDPLN REGIME 1				
H_0: Residuals are Gaussian distributed				
1.3065	2.2167	0.5633	0.4708	3.4536
(0.0010)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
H_0: Residuals are α-stable distributed				
0.5304	0.7368	0.0236	0.0171	0.4093
(0.7260)	(0.9530)	(0.8960)	(0.9510)	(0.4250)
CU Price REGIME 2				
H_0: Residuals are Gaussian distributed				
0.9037	1.6305	0.2274	0.2207	1.5643
(0.0490)	(0.0230)	(0.0030)	(0.0020)	(0.0010)
H_0: Residuals are α-stable distributed				
0.3283	0.6568	0.0143	0.0150	0.1273
(0.9970)	(0.9750)	(0.9820)	(0.9630)	(0.9720)
USDPLN REGIME 2				
H_0: Residuals are Gaussian distributed				
1.1281	1.9804	0.2319	0.2278	1.5169
(0.0000)	(0.0010)	(0.0020)	(0.0000)	(0.0000)
H_0: Residuals are α-stable distributed				
0.4640	0.8905	0.0398	0.0366	0.3121
(0.7630)	(0.6410)	(0.4050)	(0.4150)	(0.3470)

Table 6.4: The parameters of the α-stable distribution estimated for the residual time series corresponding to the VAR(1) model with independent components, source: [131].

α	σ	β	μ
CU Price REGIME 1			
1.9233	0.0236	-0.5953	0.0007
USDPLN REGIME 1			
1.7214	0.0114	1.0000	0.0016
CU Price REGIME 2			
1.8314	0.0120	-0.4291	-0.0005
USDPLN REGIME 2			
1.8411	0.0073	-0.0280	0.0000

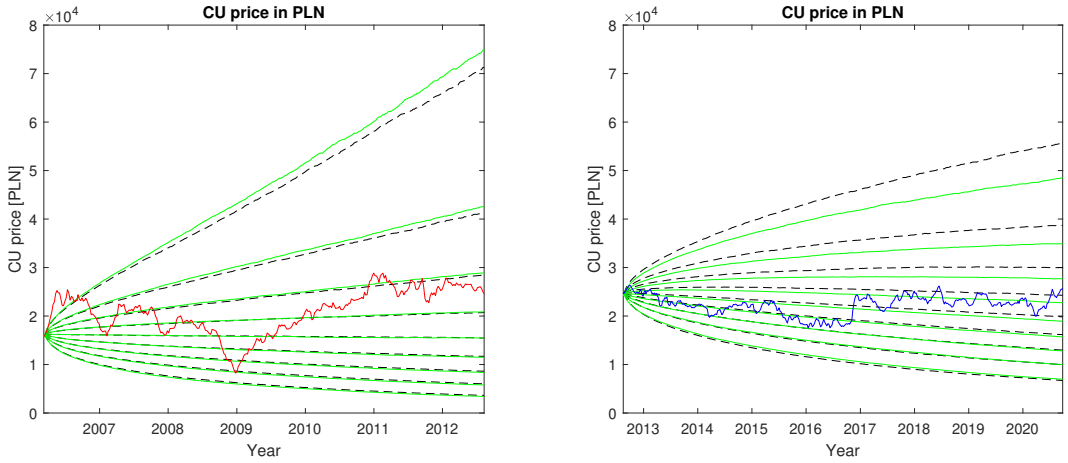


Figure 6-9: The copper (Cu) price in PLN in both analysed regimes (blue solid lines) together with the quantile lines for $10\%, 20\%, ..., 90\%$ confidence levels for the VAR(1) model (black dashed lines) and for the VAR(1) model with independent components (green solid lines), source: [131].

plotted in the form of quantile lines in Fig. 6-9 for both regimes and both models. The calculations were based on 100000 simulated trajectories. In the first regime, the copper price in PLN probability distribution is skewed to the upside reflecting the higher volatility of return rates both for the copper price and the USDPLN exchange rate and theoretically unlimited growth potential for the value of the assets in the extraordinary market environment. Using the $\alpha-$stable VAR(1) model taking into account the relationship between assets helps to narrow down the quantile lines, which should be supportive from the market risk management point of view, however, the improvement is not substantial. In regard to the second analysed regime, the probability distribution for the copper price in PLN is more symmetric and narrower than for the first period. The $\alpha-$stable VAR(1) model with dependent factors does not show narrower quantile lines than in the case of no relation

between both coordinates. Hence, for more stable market conditions the relationship between these two factors does not have to lead to narrowing price distributions. In a more stable situation on the market, often specific events, related exclusively to copper or the USDPLN exchange rate drive their prices, which can create some volatility of the assets with a harder to capture and modelling relation. In such circumstances, the two-dimensional model that assumes no dependence between components could give a similar outcome to the general VAR(1) model, when their relationship is taken into consideration.

6.5 Discussion and summary

The analysis presented in this section has been focused on the proper understanding of the dynamics of the analysed historical data corresponding to the two assets, in this case the copper price in USD and the USDPLN exchange rate, and its implications for modelling purposes.

We have proposed a two-dimensional VAR model with the $\alpha-$stable distribution that reflects the changing dependence structure of the analysed assets. This model takes into account the relationship between the analysed time series and allows the specific (heavy-tailed) behaviour of the data to be described. Moreover, the application of the regime-switching methodology enables the periods, when the data relatively have a homogeneous structure, to be considered separately. Indeed, for both assets, we obtained a significant difference in the scale of the market fluctuations among the regimes. For comparison, we have also applied an approach with no relation between the assets, i.e the VAR(1) model with independent components.

Using such a model improved the data description quality, especially in the volatile market environment where the negative correlation between the copper price in USD and USDPLN exchange rates has been higher and gave an advantage over using separate one-dimensional models for both assets.

From our analysis, we can conclude that in a regime with a higher volatility it is hard to control heavy tails even taking into account the higher negative correlation between the assets. This is extremely important information for the risk management process. Therefore, careful investigation of the real data and the proper selection of the used methods enables the building of more adequate forecasts, especially for stress test scenarios in multidimensional asset management.

Chapter 7

From multi- to univariate case. Risk modelling by product of components for bi-dimensional model in energy market.

In previous chapters we have described and analysed stochastic models for metals and exchange rates in one-dimensional and multi-dimensional cases. Due to the dynamically growing prices of energy factors, energy costs have become another market risk factor to which mining companies are exposed. In this light modelling energy risk, including electricity risk, is gaining more and more importance for mining companies.

Since electricity is a special commodity with very limited storage possibilities, electricity supply and demand must be constantly balanced. This makes operational planning a crucial issue for electricity trading companies. Production or cost management is usually based on electricity load or demand predictions, published on a day-ahead basis by the transmission system operators (TSO). However, these forecasts are burdened with prediction errors, which might cause large deviations of the actual energy cost from its projections. On the other hand, the financial cost of the load or demand prediction errors is equal to the product of the electricity price and the size of these errors.

From the mathematical point of view the problem analysed in this chapter can be considered as a generalisation of the known issue related to the analysis of the product of two (or more) random variables, which is an important branch of probability theory, statistics and applied mathematics. When one studies the product of random variables, the main attention is usually paid to its distribution and the analysis of how the probabilistic properties of the individual random variables influence the characteristics of their product. A simpler model, the product of two random

variables, can be found in [132]. In this chapter we present a product of two components for a two-dimensional time series, which is a more difficult one. The theoretical results are applied here to the electricity market case study. The results presented in this chapter are published in [133].

The multi-dimensional VAR time series is often used to model the impulse-response functions of macroeconomics variables. However, in some economical applications the variable of main interest is the product of time series describing market variables, like e.g. the cost, being the product of price and volume. In this chapter we analyse the product of the bi-dimensional VAR(1) model components. For the introduced time series we derive general formulas for the autocovariance function and study its properties for different cases of cross-dependence between the VAR(1) model components. The theoretical results are then illustrated in the simulation study for two types of bivariate distributions of the residual series, namely the Gaussian and Student's t. The obtained results are applied to the electricity market case study, in which we show that the financial cost of load prediction errors can be well described by time series being the product of the VAR(1) model components with the Student's t distribution.

In the following we recall the main properties of the VAR(1) model. Definition of the multidimensional VAR(p) model has already been presented in Eq. (6.9). Here we take $p = 1$ and $\Theta_1 = \Phi$. The bi-dimensional VAR(1) time series, $\{\mathbf{X}(t),\ t \in \mathbb{Z}\}$, satisfies the following equation [206]:

$$\mathbf{X}(t) - \Phi\mathbf{X}(t-1) = \mathbf{Z}(t), \tag{7.1}$$

where $\mathbf{X}(t) = (X_1(t), X_2(t))$, Φ is 2×2 matrix

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}, \tag{7.2}$$

and $\{\mathbf{Z}(t),\ t \in \mathbb{Z}\}$ is the zero-mean bi-dimensional residual series, i.e. for each $t \in \mathbb{Z}$, $\mathbf{Z}(t) = (Z_1(t), Z_2(t))$.

We assume here that $\{\mathbf{Z}(t)\}$ is a series of independent bi-dimensional random variables having the same distribution, i.e. for each $t \in \mathbb{Z}$, $(Z_1(t), Z_2(t)) \sim (Z_1, Z_2)$. Moreover, we consider only the finite-variance case, i.e. the covariance matrix of (Z_1, Z_2) (denoted further as Γ_Z) is properly defined. In the further analysis the variances of random variables Z_1 and Z_2 for any t are denoted as $\sigma^2_{Z,1}$ and $\sigma^2_{Z,2}$, respectively, while the correlation coefficient between them as ρ_Z.

Let us assume that the following condition for the model coefficients is satisfied:

$$\det(I - z\Phi) \neq 0 \quad \text{for all } z \in \mathbb{Z} \text{ such that } |z| \leq 1, \tag{7.3}$$

i.e. the eigenvalues of the matrix Φ (denoted further as ν_1 and ν_2) are less than 1 in the absolute value. Under this assumption, for each $t \in \mathbb{Z}$ one can express $\mathbf{X}(t)$ in the causal representation:

$$\mathbf{X}(t) = \sum_{j=0}^{+\infty} \Phi^j \mathbf{Z}\,(t-j)\,. \tag{7.4}$$

Let us note that, when the condition (7.3) is satisfied, then the coefficients Φ^j are absolutely summable. In this case, the time series given in Eq. (7.4) is the unique bounded stationary solution of Eq. (7.1) and it converges [206]. In this chapter, we consider only the case when the eigenvalues of the matrix Φ are the real numbers.

We take the following notation:

$$\Phi^j = \begin{bmatrix} \phi_{11}^{(j)} & \phi_{12}^{(j)} \\ \phi_{21}^{(j)} & \phi_{22}^{(j)} \end{bmatrix}, \;\; j = 0, 2 \ldots . \tag{7.5}$$

Obviously, for $j = 0$, $\phi_{11}^{(j)} = \phi_{22}^{(j)} = 1$ and $\phi_{12}^{(j)} = \phi_{21}^{(j)} = 0$. [218] shows that for a 2×2 matrix, the coefficients of Φ^j can be expressed in the following form depending on the eigenvalues of the matrix Φ:

- if ν_1, ν_2 are different eigenvalues of the matrix Φ, i.e. $(\phi_{11}-\phi_{22})^2 \neq -4\phi_{21}\phi_{12}$ (and $|\nu_1| < 1$, $|\nu_2| < 1$), then we have

$$\Phi^j = \begin{bmatrix} \frac{\nu_2\nu_1^j - \nu_1\nu_2^j}{\nu_2-\nu_1} + \frac{\nu_2^j-\nu_1^j}{\nu_2-\nu_1}\phi_{11} & \frac{\nu_2^j-\nu_1^j}{\nu_2-\nu_1}\phi_{12} \\ \frac{\nu_2^j-\nu_1^j}{\nu_2-\nu_1}\phi_{21} & \frac{\nu_2\nu_1^j - \nu_1\nu_2^j}{\nu_2-\nu_1} + \frac{\nu_2^j-\nu_1^j}{\nu_2-\nu_1}\phi_{22} \end{bmatrix}, \;\; j = 1, 2, \ldots, \tag{7.6}$$

- if the eigenvalues of the matrix Φ are equal $\nu_1 = \nu_2 = \nu$, i.e. $(\phi_{11} - \phi_{22})^2 = -4\phi_{21}\phi_{12}$ (and $|\nu| < 1$), then we have

$$\Phi^j = \begin{bmatrix} j\nu^{j-1}\phi_{11} - (j-1)\nu^j & j\nu^{j-1}\phi_{12} \\ j\nu^{j-1}\phi_{21} & j\nu^{j-1}\phi_{22} - (j-1)\nu^j \end{bmatrix}, \; j = 1, 2 \ldots . \tag{7.7}$$

Using Eq. (7.4) one can show that the components of the VAR(1) model can be expressed in the following form:

$$X_i(t) = \sum_{j=0}^{\infty} \sum_{k=1}^{2} \phi_{ik}^{(j)} Z_k(t-j), \;\; i = 1, 2 \tag{7.8}$$

and their distributions do not depend on t. Thus, from Eq. (7.8) one can obtain the formulas for variances $\sigma_{X,1}^2$, $\sigma_{X,2}^2$ of $X_1(t)$ and $X_2(t)$, respectively

$$\sigma_{X,i}^2 = \mathbb{V}\text{ar}(X_i(t)) = \sum_{j=0}^{\infty} \sum_{k,l=1}^{2} \phi_{ik}^{(j)} \phi_{il}^{(j)} \gamma_{Z,k,l}, \tag{7.9}$$

where $\gamma_{Z,k,l}$ is the (k,l) component of the covariance matrix Γ_Z. Recall that Γ_Z is is given by

$$\Gamma_Z = \Gamma_Z(t) = [\gamma_{Z,i,j}(t)]_{i,j=1}^2 = [\mathbb{E}[Z_i(t) Z_j(t)]]_{i,j=1}^2 \,, \tag{7.10}$$

where $\gamma_{Z,i,i} = \sigma_{Z,i}^2$ and $\gamma_{Z,1,2} = \gamma_{Z,2,1} = \rho_Z \sigma_{Z,1} \sigma_{Z,2}$.

Let us note that the covariance between $X_1(t)$ and $X_2(t)$ is also dependent on t and it is given by

$$\begin{aligned} \gamma_{X,1,2} &= \mathbb{E}[X_1(t) X_2(t)] = \mathbb{E}\left[\sum_{j=0}^{\infty} \sum_{k=1}^{2} \phi_{1k}^{(j)} Z_k(t-j) \sum_{i=0}^{\infty} \sum_{l=1}^{2} \phi_{2l}^{(i)} Z_l(t-i) \right] \\ &= \sum_{j=0}^{\infty} \sum_{k,l=1}^{2} \phi_{1k}^{(j)} \phi_{2l}^{(j)} \mathbb{E}[Z_k(t-j) Z_l(t-j)] = \sum_{j=0}^{\infty} \sum_{k,l=1}^{2} \phi_{1k}^{(j)} \phi_{2l}^{(j)} \gamma_{Z,k,l}. \end{aligned} \tag{7.11}$$

Thus, the correlation coefficient between $X_1(t)$ and $X_2(t)$ for each $t \in \mathbb{Z}$ is given by

$$\rho_X = \frac{\gamma_{X,1,2}}{\sigma_{X,1} \sigma_{X,2}} = \frac{\sum_{j=0}^{\infty} \sum_{k,l=1}^{2} \phi_{1k}^{(j)} \phi_{2l}^{(j)} \gamma_{Z,k,l}}{\sqrt{\sum_{j=0}^{\infty} \sum_{k,l=1}^{2} \phi_{1k}^{(j)} \phi_{1l}^{(j)} \gamma_{Z,k,l} \sum_{j=0}^{\infty} \sum_{k,l=1}^{2} \phi_{2k}^{(j)} \phi_{2l}^{(j)} \gamma_{Z,k,l}}}. \tag{7.12}$$

The autocovariance function of $\{X_i(t)\}$ for $i = 1, 2$ is dependent on t and takes the form:

$$ACVF_{X_i}(h) = \mathbb{E}[X_i(t) X_i(t+h)] = \sum_{j=0}^{\infty} \sum_{k,l=1}^{2} \phi_{ik}^{(j)} \phi_{il}^{(h+j)} \gamma_{Z,k,l}. \tag{7.13}$$

Using the same reasoning as in the above calculations, one can show that the cross-covariance between $\{X_1(t)\}$ and $\{X_2(t)\}$ is also dependent on t and it is given by:

$$CCVF_{X_1,X_2}(h) = \mathbb{E}[X_1(t)X_2(t+h)] = \sum_{j=0}^{\infty}\sum_{k,l=1}^{2}\phi_{1k}^{(j)}\phi_{2l}^{(h+j)}\gamma_{Z,k,l}. \tag{7.14}$$

7.1 Product of the components of bi-dimensional finite-variance VAR(1) model

Here, we introduce the time series $\{Y(t),\ t \in \mathbb{Z}\}$ that is a product of two components of the bi-dimensional VAR(1) model discussed in the previous section. Precisely, for each $t \in \mathbb{Z}$ we have:

$$Y(t) = X_1(t)X_2(t), \tag{7.15}$$

where the bi-dimensional time series $\{\mathbf{X}(t)\}$ satisfies Eq. (7.1). Assuming that condition (7.3) is fulfilled and applying Eq. (7.8) one can show that for each $t \in \mathbb{Z}$, $Y(t)$ can be represented as:

$$Y(t) = \sum_{j,i=0}^{\infty}\sum_{k,l=1}^{2}\phi_{1k}^{(j)}\phi_{2l}^{(i)}Z_k(t-j)Z_l(t-i). \tag{7.16}$$

Using the above representation, one can calculate the main characteristics of the $\{Y(t)\}$ time series. In the equations presented below, we assume that in general, $\mathbb{E}[Z_kZ_lZ_nZ_r] < \infty$ for $k,l,n,r = 1,2$. However, for some special cases (see Section 7.2), this assumption may be less restrictive depending on the relationship between components of the considered VAR(1) model.

Observe that if $\{Y(t)\}$ is the product time series defined in Eq. (7.15), where $\{X_1(t)\}$ and $\{X_2(t)\}$ are the two components of the finite-variance VAR(1) model given in Eq. (7.1) satisfying the condition (7.3), then the expected value and variance of $\{Y(t)\}$ exist and are given by:

$$\mathbb{E}(Y(t)) = \gamma_{X,1,2} = \rho_X\sigma_{X,1}\sigma_{X,2}, \tag{7.17}$$

$$\mathbb{V}\text{ar}(Y(t)) = \sigma_Y^2 = \mathbb{E}\left[\left(\sum_{j,i=0}^{\infty}\sum_{k,l=1}^{2}\phi_{1k}^{(j)}\phi_{2l}^{(i)}Z_k(t-j)Z_l(t-i)\right)^2\right] - \gamma_{X,1,2}^2, \tag{7.18}$$

where $\sigma_{X,i}$, $\gamma_{X,1,2}$, ρ_X are given in Eqs. (7.9), (7.11) and (7.12), respectively. These results can be directly derived from from Eqs. (7.15) and (7.16).

In the following we derive the formula for the ACVF of the time series $\{Y(t)\}$, $\text{ACVF}_Y(t, t+h) = \text{Cov}(Y(t), Y(t+h))$ for $t, h \in \mathbb{Z}$. Assume that $\{Y(t)\}$ is the product time series defined in Eq. (7.15), where $\{X_1(t)\}$ and $\{X_2(t)\}$ are the two components of the finite-variance VAR(1) model given in Eq. (7.1) satisfying the condition (7.3). The expectations $\mathbb{E}(Y(t))$ and $\mathbb{E}(Y(t+h))$ are given in Eq. (7.17). Thus, we need to calculate $\mathbb{E}(Y(t)Y(t+h))$ for any $h = 0, 1, \ldots$. Using Eq. (7.16) one obtains:

$$\begin{aligned}\mathbb{E}(Y(t)Y(t+h)) &= \sum_{j,i,m,p=0}^{\infty} \sum_{k,l,n,r=1}^{2} \phi_{1k}^{(j)} \phi_{2l}^{(i)} \phi_{1n}^{(m)} \phi_{2r}^{(p)} \\ &\quad \times \mathbb{E}\left[Z_k(t-j) Z_l(t-i) Z_n(t+h-m) Z_r(t+h-p)\right] \\ &= \sum_{j,i=0}^{\infty} \sum_{m,p=-h}^{\infty} \sum_{k,l,n,r=1}^{2} \phi_{1k}^{(j)} \phi_{2l}^{(i)} \phi_{1n}^{(m+h)} \phi_{2r}^{(p+h)} \\ &\quad \times \mathbb{E}\left[Z_k(t-j) Z_l(t-i) Z_n(t-m) Z_r(t-p)\right].\end{aligned}$$

Applying the formula for the ACVF of $\{Y(t)\}$ we get:

$$\text{ACVF}_Y(t, t+h) = \mathbb{E}(Y(t)Y(t+h)) - \mathbb{E}(Y(t))\mathbb{E}(Y(t+h))$$

we obtain that the autocovariance function of $\{Y(t)\}$ for $h = 0, 1, \ldots,$ has the following form:

$$\begin{aligned}\text{ACVF}_Y(t, t+h) &= \sum_{j,i=0}^{\infty} \sum_{m,p=-h}^{\infty} \sum_{k,l,n,r=1}^{2} \phi_{1k}^{(j)} \phi_{2l}^{(i)} \phi_{1n}^{(m+h)} \phi_{2r}^{(p+h)} \\ &\quad \times \mathbb{E}\left[Z_k(t-j) Z_l(t-i) Z_n(t-m) Z_r(t-p)\right] - \gamma_{X,1,2}^2, \qquad (7.19)\end{aligned}$$

where $\gamma_{X,1,2}$ is given in Eq. (7.11).

Let us emphasize that the ACVF given in Eq. (7.19) is dependent on t. Moreover, $\{Y(t)\}$ has a constant mean function. Thus, it is stationary in the weak sense. Therefore, in the further analysis it will be denoted as $\text{ACVF}_Y(h)$.

7.2 Special cases analysis

In this section, we consider the following special cases related to the dependence of the components of the finite-variance VAR(1) model given in Eq. 7.1. Let us emphasise that in the considered cases we do not consider any specific distribution of the residual series.

- Case 1: the time series $\{X_1(t)\}$ and $\{X_2(t)\}$ are independent. This is the case, when for each $t \in \mathbb{Z}$ the random variables Z_1 and Z_2 are independent and $\phi_{12} = \phi_{21} = 0$, where ϕ_{ij} $i, j = 1, 2$ are the coefficients of the matrix Φ given in Eq. (7.2). In this case, $\{X_1(t)\}$ and $\{X_2(t)\}$ are two independent autoregressive time series of order 1 (called AR(1)) satisfying the following equations:

$$X_1(t) - \phi_{11}X_1(t-1) = Z_1(t), \;\; X_2(t) - \phi_{22}X_2(t-1) = Z_2(t). \tag{7.20}$$

- Case 2: the time series $\{X_1(t)\}$ and $\{X_2(t)\}$ are dependent only through the residual components. In this case, we assume that the random variables Z_1 and Z_2 are dependent (and we assume their correlation coefficient $\rho_Z \neq 0$), however $\phi_{12} = \phi_{21} = 0$, where ϕ_{ij} $i, j = 1, 2$ are the coefficients of the matrix Φ given in Eq. (7.2). In this case, $\{X_1(t)\}$ and $\{X_2(t)\}$ also satisfy Eq. (7.20), but they are dependent.

- Case 3: the time series $\{X_1(t)\}$ and $\{X_2(t)\}$ are dependent only through the coefficients of the VAR(1) model. This is the case, when the random variables Z_1 and Z_2 are independent, however $\phi_{12} \neq 0$ or/and $\phi_{21} \neq 0$, where ϕ_{ij} $i, j = 1, 2$ are the coefficients of the matrix Φ given in Eq. (7.2). For simplicity we assume $\phi_{12} \neq 0$ and $\phi_{21} = 0$. In this case, $\{X_1(t)\}$ and $\{X_2(t)\}$ satisfy the following equations:

$$X_1(t) - \phi_{11}X_1(t-1) - \phi_{12}X_2(t-1) = Z_1(t), \;\; X_2(t) - \phi_{22}X_2(t-1) = Z_2(t), \tag{7.21}$$

  thus, the time series $\{X_2(t)\}$ is the AR(1) model while $\{X_1(t)\}$ does not satisfy the AR(1) equation.

Case 1

In this case we assume $\sigma^2_{Z,i} < \infty$ for $i = 1, 2$. The coefficients of the matrix Φ given in Eq. (7.2) that lie outside the main diagonal are equal to zero, i.e., $\phi_{12} = \phi_{21} = 0$. Thus, we have:

$$\phi^{(j)}_{11} = \phi^j_{11}, \; \phi^{(j)}_{22} = \phi^j_{22}, \; \phi^{(j)}_{12} = \phi^{(j)}_{21} = 0, \; j = 0, 1, \ldots. \tag{7.22}$$

Moreover, according to the condition given in Eq. (7.3), $|\phi_{11}| < 1$ and $|\phi_{22}| < 1$. Using Eq. (7.9) and (7.12) one can easily show that:

$$\sigma^2_{X,i} = \sum_{j=0}^{\infty} \phi^{2j}_{ii}\gamma_{Z,i,i} = \frac{\sigma^2_{Z,i}}{1 - \phi^2_{ii}}. \tag{7.23}$$

In the considered case $\rho_X = 0$.

Finally, using Eqs. (7.17) and (7.19) one can show that the following hold:

$$\mathbb{E}(Y(t)) = 0, \;\; \mathbb{V}\text{ar}(Y(t)) = \frac{\sigma^2_{Z,1}\sigma^2_{Z,2}}{(1-\phi^2_{11})(1-\phi^2_{22})}, \;\; \text{ACVF}_Y(h) = \frac{\sigma^2_{Z,1}\sigma^2_{Z,2}(\phi_{11}\phi_{22})^h}{(1-\phi^2_{11})(1-\phi^2_{22})}. \quad (7.24)$$

Case 2

In this case we assume $\mathbb{E}[Z_1^2 Z_2^2] < \infty$. Similarly as previously, condition (7.22) is satisfied. However, now, we assume that the components of the residual series are dependent and the correlation coefficient ρ_Z is non-zero. One can show that $\sigma^2_{X,i}$ has the same form as in Case 1 for $i = 1, 2$, i.e., it is given by Eq. (7.23). However, using Eq. (7.12) we obtain that the ρ_X coefficient is given by:

$$\rho_X = \frac{\rho_Z\sqrt{(1-\phi^2_{11})(1-\phi^2_{22})}}{1-\phi_{11}\phi_{22}}. \quad (7.25)$$

Using Eq. (7.17) one obtains:

$$\mathbb{E}(Y(t)) = \frac{\rho_Z\sigma_{Z,1}\sigma_{Z,2}}{(1-\phi_{11}\phi_{22})}. \quad (7.26)$$

On the other hand, using Eq. (7.19) we can calculate the ACVF for $\{Y(t)\}$ for $h = 0, 1, \ldots$. Indeed, we have:

$$\begin{aligned} \text{ACVF}_Y(h) &= \sum_{j,i=0}^{\infty}\sum_{m,p=-h}^{\infty} \phi_{11}^{j+m+h}\phi_{22}^{i+p+h}\mathbb{E}\left[Z_1(t-j)Z_2(t-i)Z_1(t-m)Z_2(t-p)\right] \\ &- \frac{\rho_Z^2\sigma^2_{Z,1}\sigma^2_{Z,2}}{(1-\phi_{11}\phi_{22})^2}. \end{aligned}$$

Now, we can calculate the value:

$$r_{1,2}(t,j,m,i,p) = \mathbb{E}\left[Z_1(t-j)Z_1(t-m)Z_2(t-i)Z_2(t-p)\right]$$

for all $t \in \mathbb{Z}$ and $i, j = 0, 1, \ldots, m, p = -h, -h+1, \ldots$. Using the fact that for each $t \in \mathbb{Z}$ the bi-dimensional residual series $\mathbf{Z}(t)$ is a zero-mean vector and for $t \neq s$, $\mathbf{Z}(t)$ is dependent on $\mathbf{Z}(s)$,

one obtains the following:

$$r_{1,2}(t,j,m,i,p) = \begin{cases} \mathbb{E}\left[Z_1^2(t-j)Z_2^2(t-j)\right], \text{ if } i=j=p=m; \\ \\ \mathbb{E}\left[Z_1^2(t-j)\right]\mathbb{E}\left[Z_2^2(t-i)\right], \text{ if } j=m, i=p, j\neq i; \\ \\ \mathbb{E}\left[Z_1(t-j)Z_2(t-j)\right]\mathbb{E}\left[Z_1(t-m)Z_2(t-m)\right], \text{ if } j=i, m=p, j\neq m; \\ \\ \mathbb{E}\left[Z_1(t-j)Z_2(t-j)\right]\mathbb{E}\left[Z_1(t-m)Z_2(t-m)\right], \text{ if } j=p, i=m, j\neq i. \end{cases} \tag{7.27}$$

Thus, we have:

$$r_{1,2}(t,j,m,i,p) = \begin{cases} m_Z, \text{ if } i=j=p=m;\ i,j,m,p=0,1,2\ldots; \\ \\ \sigma_{Z,1}^2\sigma_{Z,2}^2, \text{ if } j=m, i=p, j\neq i;\ \ i,j,m,p=0,1,2\ldots; \\ \\ \rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2, \text{ if } j=i, m=p, j\neq m;\ i,j=0,1,2\ldots,\ \ m,p=-h,-h+1\ldots; \\ \\ \rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2, \text{ if } j=p, i=m, j\neq i;\ i,j,m,p=0,1,2\ldots, \end{cases} \tag{7.28}$$

where the value $m_Z = \mathbb{E}\left[Z_1^2(t)Z_2^2(t)\right]$ is dependent on t. Therefore, we have:

$$\begin{aligned}
\text{ACVF}_Y(h) &= m_Z \sum_{j=0}^{\infty} \left(\phi_{11}\phi_{22}\right)^{2j+h} + \sigma_{Z,1}^2\sigma_{Z,2}^2 \sum_{j=0}^{\infty} \phi_{11}^{2j+h} \left[\sum_{i=0}^{\infty} \phi_{22}^{2i+h} - \phi_{22}^{2j+h}\right] \\
&\quad + \rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2 \sum_{j=0}^{\infty} \phi_{11}^{j+h} \left[\phi_{22}^{j+h} \sum_{m=0}^{\infty} \left(\phi_{11}\phi_{22}\right)^m - \phi_{22}^{j+h}\left(\phi_{11}\phi_{22}\right)^j\right] \\
&\quad + \rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2 \sum_{j=0}^{\infty} \phi_{11}^{j+h} \left[\phi_{22}^{j+h} \sum_{m=-h}^{\infty} \left(\phi_{11}\phi_{22}\right)^m - \phi_{22}^{j+h}\left(\phi_{11}\phi_{22}\right)^j\right] \\
&\quad - \frac{\rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2}{\left(1-\phi_{11}\phi_{22}\right)^2} \\
&= \frac{m_Z\left(\phi_{11}\phi_{22}\right)^h}{1-\left(\phi_{11}\phi_{22}\right)^2} + \sigma_{Z,1}^2\sigma_{Z,2}^2\left(\phi_{11}\phi_{22}\right)^h \left(\frac{1}{(1-\phi_{11}^2)(1-\phi_{22}^2)} - \frac{1}{1-\left(\phi_{11}\phi_{22}\right)^2}\right) \\
&\quad + \rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2\left(\phi_{11}\phi_{22}\right)^h \left(\frac{1}{\left(1-\phi_{11}\phi_{22}\right)^2} - \frac{1}{1-\left(\phi_{11}\phi_{22}\right)^2}\right) \\
&\quad + \rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2\left(\phi_{11}\phi_{22}\right)^h \left(\frac{\left(\phi_{11}\phi_{22}\right)^{-h}}{\left(1-\phi_{11}\phi_{22}\right)^2} - \frac{1}{1-\left(\phi_{11}\phi_{22}\right)^2}\right) - \frac{\rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2}{\left(1-\phi_{11}\phi_{22}\right)^2} \\
&= \left(\phi_{11}\phi_{22}\right)^h \left[\frac{m_Z - \sigma_{Z,1}^2\sigma_{Z,2}^2 - 2\rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2}{1-\left(\phi_{11}\phi_{22}\right)^2} + \frac{\sigma_{Z,1}^2\sigma_{Z,2}^2}{(1-\phi_{11}^2)(1-\phi_{22}^2)}\right. \\
&\quad \left. + \frac{\rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2}{\left(1-\phi_{11}\phi_{22}\right)^2}\right]. \qquad (7.29)
\end{aligned}$$

Finally, taking $h = 0$ one obtains the variance of a random variable $Y(t)$ for each $t \in \mathbb{Z}$:

$$\mathbb{V}\text{ar}(Y(t)) = \frac{m_Z - \sigma_{Z,1}^2\sigma_{Z,2}^2 - 2\rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2}{1-\left(\phi_{11}\phi_{22}\right)^2} + \frac{\sigma_{Z,1}^2\sigma_{Z,2}^2}{(1-\phi_{11}^2)(1-\phi_{22}^2)} + \frac{\rho_Z^2\sigma_{Z,1}^2\sigma_{Z,2}^2}{\left(1-\phi_{11}\phi_{22}\right)^2}.$$

Case 3

In this case we assume $\mathbb{E}(Z_1^2Z_2^2) < \infty$ and $\mathbb{E}(Z_2^4) < \infty$. One can show that we have:

$$\phi_{11}^{(j)} = \phi_{11}^j,\ \phi_{22}^{(j)} = \phi_{22}^j,\ \phi_{21}^{(j)} = 0, \quad j = 0, 1, \ldots \qquad (7.30)$$

and the eigenvalues of the matrix Φ are equal $\nu_1 = \phi_{11}$ and $\nu_2 = \phi_{22}$. Thus, according to Eqs. (7.6) and (7.7) the following is fulfilled for $j = 1, 2, \ldots$:

$$\phi_{12}^{(j)} = \begin{cases} \frac{\phi_{22}^j - \phi_{11}^j}{\phi_{22} - \phi_{11}} \phi_{12}, \text{ if } \phi_{11} \neq \phi_{22}; \\ \\ j\phi_{11}^{j-1}\phi_{12}, \text{ if } \phi_{11} = \phi_{22}, \end{cases} \tag{7.31}$$

while for $j = 0$, $\phi_{12}^{(j)} = 0$. In order to fulfill condition (7.3) we assume that $|\phi_{11}| < 1$ and $|\phi_{22}| < 2$.

Using Eq. (7.9) one obtains:

$$\sigma_{X,1}^2 = \frac{\sigma_{Z,1}^2}{1 - \phi_{11}^2} + \sigma_{Z,2}^2 \sum_{j=0}^{\infty} \left(\phi_{12}^{(j)}\right)^2, \;\; \sigma_{X,2}^2 = \frac{\sigma_{Z,2}^2}{1 - \phi_{22}^2}. \tag{7.32}$$

Thus, we have:

$$\sigma_{X,1}^2 = \begin{cases} \frac{\sigma_{Z,1}^2}{1-\phi_{11}^2} + \frac{\sigma_{Z,2}^2 \phi_{12}^2}{(\phi_{22}-\phi_{11})^2} \left[\frac{\phi_{22}^2}{1-\phi_{22}^2} - \frac{2\phi_{11}\phi_{22}}{1-\phi_{11}\phi_{22}} + \frac{\phi_{11}^2}{1-\phi_{21}^2}\right], \text{ if } \phi_{11} \neq \phi_{22}; \\ \\ \frac{\sigma_{Z,1}^2}{1-\phi_{11}^2} + \frac{\sigma_{Z,2}^2 \phi_{12}^2 (1+\phi_{11}^2)}{(1-\phi_{11}^2)^3}, \text{ if } \phi_{11} = \phi_{22}. \end{cases} \tag{7.33}$$

Moreover from Eq. (7.11) we have:

$$\gamma_{X,1,2} = \sum_{j=1}^{\infty} \sigma_{Z,2}^2 \phi_{12}^{(j)} \phi_{22}^j.$$

Thus, we obtain the following formula for the expected value of the random variable $Y(t)$ for each $t \in \mathbb{Z}$:

$$\mathbb{E}(Y(t)) = \gamma_{X,1,2} = \begin{cases} \frac{\sigma_{Z,2}^2 \phi_{12}}{\phi_{22}-\phi_{11}} \left[\frac{\phi_{22}^2}{1-\phi_{22}^2} - \frac{\phi_{11}\phi_{22}}{1-\phi_{22}\phi_{11}}\right], \text{ if } \phi_{11} \neq \phi_{22}; \\ \\ \frac{\sigma_{Z,2}^2 \phi_{22} \phi_{12}}{(1-\phi_{11}\phi_{22})^2}, \text{ if } \phi_{11} = \phi_{22}. \end{cases} \tag{7.34}$$

To obtain the explicit formula for $\text{ACVF}_Y(h)$ we will use Eq. (7.19). For $h = 0, 1, 2 \ldots$ we have:

$$\begin{aligned}
\text{ACVF}_Y(h) &= \sum_{j,i=0}^{\infty} \sum_{m,p=-h}^{\infty} \sum_{k,l,n,r=1}^{2} \phi_{1k}^{(j)} \phi_{2l}^{(i)} \phi_{1n}^{(m+h)} \phi_{2r}^{(p+h)} \mathbb{E}\left[Z_k(t-j) Z_l(t-i) Z_n(t-m) Z_r(t-p)\right] \\
&- \gamma_{X,1,2}^2 \\
&= \sum_{j,i=0}^{\infty} \sum_{m,p=-h}^{\infty} \sum_{k,n=1}^{2} \phi_{1k}^{(j)} \phi_{22}^{(i)} \phi_{1n}^{(m+h)} \phi_{22}^{(p+h)} \mathbb{E}\left[Z_k(t-j) Z_2(t-i) Z_n(t-m) Z_2(t-p)\right] \\
&- \gamma_{X,1,2}^2 \\
&= \sum_{j,i=0}^{\infty} \sum_{m,p=-h}^{\infty} \phi_{22}^{i+p+h} \sum_{k,n=1}^{2} \phi_{1k}^{(j)} \phi_{1n}^{(m+h)} \mathbb{E}\left[Z_k(t-j) Z_2(t-i) Z_n(t-m) Z_2(t-p)\right] \\
&- \gamma_{X,1,2}^2 \\
&= \sum_{j,i=0}^{\infty} \sum_{m,p=-h}^{\infty} \phi_{22}^{i+p+h} \phi_{12}^{(j)} \phi_{12}^{(m+h)} \mathbb{E}\left[Z_2(t-j) Z_2(t-i) Z_2(t-m) Z_2(t-p)\right] \\
&+ \sum_{j,i=0}^{\infty} \sum_{m,p=-h}^{\infty} \phi_{22}^{i+p+h} \phi_{11}^{j+m+h} \mathbb{E}\left[Z_1(t-j) Z_2(t-i) Z_1(t-m) Z_2(t-p)\right] \\
&- \gamma_{X,1,2}^2.
\end{aligned}$$

Moreover, the value

$$r_{2,2}(t, j, m, i, p) = \mathbb{E}\left[Z_2(t-j) Z_2(t-i) Z_2(t-m) Z_2(t-p)\right]$$

is given by:

$$r_{2,2}(t, j, m, i, p) = \begin{cases} \kappa_Z, \text{ if } i = j = m = p; \ i, j, m, p = 0, 1, \ldots, \\ \\ \sigma_{Z,2}^4, \text{ if } i = j, \ \ m = p, \ m \neq j; \ i, j = 0, 1, \ldots, \ m, p = -h, -h+1, \ldots, \\ \\ \sigma_{Z,2}^4, \text{ if } j = m, \ \ i = p, \ m \neq i; \ i, j, m, p = 0, 1, \ldots, \\ \\ \sigma_{Z,2}^4, \text{ if } j = p, \ \ i = m, \ m \neq j; \ i, j, m, p = 0, 1, \ldots, \end{cases} \tag{7.35}$$

where $\kappa_Z = \mathbb{E}\left[Z_2^4(t)\right] < \infty$ is dependent on t. Additionally, one can show that:

$$\mathbb{E}\left[Z_1(t-j)Z_2(t-i)Z_1(t-m)Z_2(t-p)\right] = \sigma_{Z,1}^2\sigma_{Z,2}^2, \text{ if } j = m, i = p,\ i,j,p,m = 0,1,2,\ldots.$$

Thus, we have the following:

$$\begin{aligned}
\text{ACVF}_Y(h) &= \kappa_Z \sum_{j=0}^{\infty} \phi_{12}^{(j)} \phi_{22}^{2j+h} \phi_{12}^{(j+h)} + \sigma_{Z,2}^4 \sum_{j=0}^{\infty} \phi_{12}^{(j)} \phi_{12}^{(j+h)} \left[\sum_{i=0}^{\infty} \phi_{22}^{2i+h} - \phi_{22}^{2j+h} \right] \\
&+ \sigma_{Z,2}^4 \sum_{j=0}^{\infty} \phi_{22}^{j+h} \phi_{12}^{(j)} \left[\sum_{i=0}^{\infty} \phi_{22}^{i} \phi_{12}^{(i+h)} - \phi_{22}^{j} \phi_{12}^{(j+h)} \right] \\
&+ \sigma_{Z,2}^4 \sum_{j=0}^{\infty} \phi_{22}^{j+h} \phi_{12}^{(j)} \left[\sum_{m=-h}^{\infty} \phi_{22}^{m} \phi_{12}^{(m+h)} - \phi_{22}^{j} \phi_{12}^{(j+h)} \right] \\
&+ \sigma_{Z,1}^2\sigma_{Z,2}^2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \phi_{11}^{2j+h} \phi_{22}^{2i+h} - \gamma_{X,1,2}^2.
\end{aligned}$$

Let us observe that for $h = 0, 1, 2, \ldots$ the following holds:

$$\sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \phi_{11}^{2j+h} \phi_{22}^{2i+h} = \frac{\left(\phi_{11}\phi_{22}\right)^h}{(1-\phi_{11}^2)(1-\phi_{22}^2)}. \tag{7.36}$$

Now, to make the calculations simpler, let us assume that $\phi_{11} = 0$ and $\phi_{22} \neq 0$. In this case, the matrix Φ (see Eq. (7.2)) has two different eigenvalues and $\phi_{12}^{(j)} = \phi_{12}\phi_{22}^{j-1}$, $j = 1, 2 \ldots$. Clearly, $\phi_{11}^{(j)} = 0$ for $j = 1, 2, \ldots$ and for $j = 0$, $\phi_{11}^{j} = 1$. We have the following:

$$\sum_{j=0}^{\infty} \phi_{12}^{(j)} \phi_{22}^{2j+h} \phi_{12}^{(j+h)} = \phi_{12}^2 \phi_{22}^{2h} \sum_{j=1}^{\infty} \phi_{22}^{4j-2} = \frac{\phi_{12}^2 \phi_{22}^{2h} \phi_{22}^2}{1-\phi_{22}^4},\ h = 0, 1, 2, \ldots.$$

Let us first consider the case $h = 0$. We have:

$$\begin{aligned}
\sum_{j=0}^{\infty} \phi_{12}^{(j)} \phi_{12}^{(j)} \left[\sum_{i=0}^{\infty} \phi_{22}^{2i} - \phi_{22}^{2j} \right] &= \phi_{12}^2 \sum_{j=1}^{\infty} \phi_{22}^{2j-2} \left[\sum_{i=0}^{\infty} \phi_{22}^{2i} - \phi_{22}^{2j} \right] \\
&= \phi_{12}^2 \left[\frac{1}{(1-\phi_{22}^2)^2} - \frac{\phi_{22}^2}{1-\phi_{22}^4} \right]. \\
\sum_{j=0}^{\infty} \phi_{22}^{j} \phi_{12}^{(j)} \left[\sum_{i=0}^{\infty} \phi_{22}^{i} \phi_{12}^{(i)} - \phi_{22}^{j} \phi_{12}^{(j)} \right] &= \phi_{12}^2 \phi_{22}^{-2} \sum_{j=1}^{\infty} \phi_{22}^{2j} \left[\sum_{i=1}^{\infty} \phi_{22}^{2i} - \phi_{22}^{2j} \right] \\
&= \phi_{12}^2 \left[\frac{\phi_{22}^2}{(1-\phi_{22}^2)^2} - \frac{\phi_{22}^2}{1-\phi_{22}^4} \right].
\end{aligned}$$

On the other hand, for $h > 0$ we have:

$$\begin{aligned}
\sum_{j=0}^{\infty} \phi_{12}^{(j)} \phi_{12}^{(j+h)} \phi_{22}^{2h} \left[\sum_{i=0}^{\infty} \phi_{22}^{2i+h} - \phi_{22}^{2j+h} \right] &= \phi_{12}^2 \phi_{22}^{2h} \sum_{j=1}^{\infty} \phi_{22}^{2j-2} \left[\sum_{i=0}^{\infty} \phi_{22}^{2i} - \phi_{22}^{2j} \right] \\
&= \phi_{12}^2 \phi_{22}^{2h} \left[\frac{1}{(1-\phi_{22}^2)^2} - \frac{\phi_{22}^2}{1-\phi_{22}^4} \right]. \\
\sum_{j=0}^{\infty} \phi_{22}^{j+h} \phi_{12}^{(j)} \left[\sum_{i=0}^{\infty} \phi_{22}^{i} \phi_{12}^{(i+h)} - \phi_{22}^{j} \phi_{12}^{(j+h)} \right] &= \phi_{12}^2 \phi_{22}^{2h} \sum_{j=1}^{\infty} \phi_{22}^{2j-1} \left[\sum_{i=0}^{\infty} \phi_{22}^{2i-1} - \phi_{22}^{2j-1} \right] \\
&= \phi_{12}^2 \phi_{22}^{2h} \left[\frac{1}{(1-\phi_{22}^2)^2} - \frac{\phi_{22}^2}{1-\phi_{22}^4} \right]. \\
\sum_{j=0}^{\infty} \phi_{22}^{j+h} \phi_{12}^{(j)} \left[\sum_{m=-h}^{\infty} \phi_{22}^{m} \phi_{12}^{(m+h)} - \phi_{22}^{j} \phi_{12}^{(j+h)} \right] &= \phi_{12}^2 \sum_{j=1}^{\infty} \phi_{22}^{2j+h-1} \left[\sum_{m=-h+1}^{\infty} \phi_{22}^{2m+h-1} - \phi_{22}^{2j+h-1} \right] \\
&= \phi_{12}^2 \phi_{22}^{2h} \left[\frac{\phi_{22}^{-2h} \phi_{22}^2}{(1-\phi_{22}^2)^2} - \frac{\phi_{22}^2}{1-\phi_{22}^4} \right].
\end{aligned}$$

Finally, assuming $\phi_{11} = 0$ we obtain the following formula for the autocovariance function of the time series $\{Y(t)\}$

- If $h = 0$

$$\begin{aligned}
\text{ACVF}_Y(h) &= Var(Y(t)) \\
&= \frac{\kappa_Z \phi_{12}^2 \phi_{22}^2}{1-\phi_{22}^4} + \sigma_{Z,2}^4 \phi_{12}^2 \left[\frac{1+2\phi_{22}^2}{(1-\phi_{22}^2)^2} - \frac{3\phi_{22}^2}{1-\phi_{22}^4} \right] + \frac{\sigma_{Z,1}^2 \sigma_{Z,2}^2}{1-\phi_{22}^2} - \frac{\sigma_{Z,2}^4 \phi_{12}^2 \phi_{22}^2}{(1-\phi_{22}^2)^2} \\
&= \phi_{12}^2 \left[\frac{\phi_{22}^2 (\kappa_Z - 3\sigma_{Z,2}^4)}{1-\phi_{22}^4} + \frac{(1+\phi_{22}^2)\sigma_{Z,2}^4}{(1-\phi_{22}^2)^2} \right] + \frac{\sigma_{Z,1}^2 \sigma_{Z,2}^2}{1-\phi_{22}^2}. \qquad (7.37)
\end{aligned}$$

- If $h > 0$

$$\begin{aligned}
\text{ACVF}_Y(h) &= \frac{\kappa_Z \phi_{12}^2 \phi_{22}^{2h} \phi_{22}^2}{1-\phi_{22}^4} + \sigma_{Z,2}^4 \phi_{12}^2 \phi_{22}^{2h} \left[\frac{2+\phi_{22}^{-2h}\phi_{22}^2}{(1-\phi_{22}^2)^2} - \frac{3\phi_{22}^2}{1-\phi_{22}^4} \right] - \frac{\sigma_{Z,2}^4 \phi_{12}^2 \phi_{22}^2}{(1-\phi_{22}^2)^2} \\
&= \phi_{12}^2 \phi_{22}^{2h} \left[\frac{\phi_{22}^2 (\kappa_Z - 3\sigma_{Z,2}^4)}{1-\phi_{22}^4} + \frac{2\sigma_{Z,2}^4}{(1-\phi_{22}^2)^2} \right]. \qquad (7.38)
\end{aligned}$$

7.3 Simulation study

To illustrate properties of the product of the VAR(1) model components, we conduct a simulation study. To this end, we simulate bi-dimensional trajectories of the VAR(1) model $\{\mathbf{X}(t)\}$ with a bivariate Gaussian or bivariate Student's t distribution for the residual vector.

We recall that the bivariate Gaussian distributed random vector (Z_1, Z_2) has the following PDF [219]:

$$f_{Z_1,Z_2}(z_1,z_2) = \frac{\exp\left\{ -\frac{1}{2(1-\rho^2)}\left[\frac{(z_1-\mu_{Z,1})^2}{\sigma_{Z,1}^2} - 2\rho\left(\frac{z_1-\mu_{Z,1}}{\sigma_{Z,1}}\right)\left(\frac{z_2-\mu_{Z,2}}{\sigma_{Z,2}}\right) + \frac{(z_2-\mu_{Z,2})^2}{\sigma_{Z,2}^2}\right]\right\}}{2\pi\sigma_{Z,1}\sigma_{Z,2}\sqrt{1-\rho^2}}, \;\; z_1, z_2 \in \mathbb{R} \tag{7.39}$$

where $\rho \in (-1, 1)$ is the correlation coefficient between random variables Z_1 and Z_2 (denoted in the main text as ρ_Z); $\mu_{Z,1}, \mu_{Z,2} \in \mathbb{R}$ are the corresponding expected values, while $\sigma_{Z,1}^2, \sigma_{Z,2}^2 > 0$ are the corresponding variances. When $\rho = 0$, the PDF of the random vector (Z_1, Z_2) is just the product of the PDFs of the Gaussian distributed random variables.

The bivariate Student's t distribution with η degrees of freedom has PDF given by [220]:

$$f_{Z_1,Z_2}(z_1,z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}}\left[1 + \frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{\eta(1-\rho^2)}\right]^{-\frac{\eta+2}{2}}, \;\; z_1, z_2 \in \mathbb{R}. \tag{7.40}$$

The marginal random variables Z_1 and Z_2 have the one-dimensional Student's t distribution defined by the following PDF [221]:

$$f_{Z_1}(z_1) = \frac{\Gamma((\eta+1)/2)}{\sqrt{\eta\pi}\Gamma(\eta/2)}\left(1 + \frac{z_1^2}{\eta}\right)^{-\frac{\eta+1}{2}}, \; z_1 \in \mathbb{R}, \tag{7.41}$$

where $\Gamma(\cdot)$ is the gamma function, i.e. $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$ for α such that $\mathrm{Re}(\alpha) > 0$. Note that the number of degrees of freedom, η, is equal for both marginal variables. It is worth highlighting that the correlation ρ_Z between the random variables Z_1 and Z_2 is equal to the parameter ρ. However, its zero value is not equivalent to the independence of the random variables Z_1 and Z_2 is not a product of the PDFs of the marginal distributions. If Z_1 and Z_2 are independent, the PDF of the random vector is given by:

$$f_{Z_1,Z_2}(z_1,z_2) = \frac{\Gamma((\eta_{Z,1}+1)/2)\Gamma((\eta_{Z,2}+1)/2)}{\sqrt{\eta_{Z,1}\eta_{Z,2}}\pi\Gamma(\eta_{Z,1}/2)\Gamma(\eta_{Z,2}/2)}\left(1 + \frac{z_{Z,1}^2}{\eta_{Z,1}}\right)^{-\frac{\eta_{Z,1}+1}{2}}\left(1 + \frac{z_2^2}{\eta_{Z,2}}\right)^{-\frac{\eta_{Z,2}+1}{2}}, \tag{7.42}$$

where $z_1, z_2 \in \mathbb{R}$, $\eta_{Z,1} > 0, \eta_{Z,2} > 0$ are the degrees of freedom parameters of Z_1 and Z_2, respectively.

The Student's t distribution defined in (7.41) has zero mean and variance equal to $\sigma^2_{Z,1} = \frac{\eta}{\eta-2}$. It can be generalized to the Student's t location-scale distribution by applying the following transformation $Z_{(\mu,\lambda)} := \mu + \lambda Z$, where Z is Student's t distributed. It yields a three parameter (μ, λ, η) distribution, with μ being the shift parameter, $\lambda > 0$ the scale parameter and $\eta > 0$ the degrees of freedom. The variance of the Student's t location-scale random variable is equal to $\sigma^2_{Z_{(\mu,\lambda)}} = \lambda^2 \frac{\eta}{\eta-2}$.

In our simulation study the shape parameter of the Student's t distribution is set to $\eta = 5$, so it has a much heavier tail than the Gaussian one. For simplicity, we assume that the scale parameters of the Gaussian distribution for both coordinates are equal, i.e. $\sigma_{Z,1} = \sigma_{Z,2}$. Moreover, we assume that $\sigma_{Z,i} = \sqrt{\frac{\eta}{\eta-2}}, i = 1, 2$, which is equal to the standard deviation of the marginal Student's t distributions. Thus, the variances of both distributions of the residual vectors are equal.

We analyse the trajectories as well as the autocovariance functions of the product of the simulated VAR(1) vectors defined in Eq. (7.15) separately for Cases 1-3. We also derive the 5% and 95% confidence bounds (CB) for the autocovariance function using Monte Carlo calculations of the empirical ACVF with 1000 repetitions and two cases of the trajectories length, namely $n = 100$ or $n = 1000$. Note that the autocovariance at lag $h = 0$ is equal to the variance of the time series, $\text{ACVF}_Y(0) = \text{Var}(Y(t))$ for each t.

It is worth mentioning that in a case when the residual series has a bivariate Gaussian distribution, then $X_1(t)$ and $X_2(t)$ for each $t \in \mathbb{Z}$ have one-dimensional Gaussian distributions and $Y(t)$ has a variance-gamma distribution with appropriate parameters, see[222]. A detailed analysis related to the distribution of the product of Gaussian random variables was presented, for instance, by [132]. In a case where (Z_1, Z_2) have the bivariate Student's t distribution, the components of the VAR(1) model $X_1(t)$ and $X_2(t)$ are not Student's t distributed. This is related to the fact that the linear combination of Student's t distributed random variables (with different weights) is not Student's t distributed.

Case 1

In this case the time series $\{X_1(t)\}$ and $\{X_2(t)\}$ are independent. Hence, $\phi_{12} = \phi_{21} = 0$ and the residual vectors components $\{Z_1(t)\}$ and $\{Z_2(t)\}$ for each $t \in \mathbb{Z}$ have independent marginal distributions. In the Gaussian case it is simply equivalent to $\rho_Z = 0$ in Eq. (7.39). However, in the Student's t case putting $\rho_Z = 0$ in (7.40) does not lead to the independence of the marginal

distributions. Hence, the components of the residual vector are simulated as independent one-dimensional Student's t random variables, see Eq. (7.41) for the PDF formula.

Since the shape of the ACVF of the product time series depends strongly on the sign of $\phi_{11}\phi_{22}$ (see formula (7.24)), we use two parameter sets: either $\phi_{11} = 0.8$ and $\phi_{22} = 0.8$ or $\phi_{11} = 0.8$ and $\phi_{22} = -0.8$. The sample trajectories of the model components $\{X_1(t)\}$ and $\{X_2(t)\}$ as well as their product $\{Y(t)\}$ are plotted in Figures 7-1 and 7-2 for $\phi_{22} = 0.8$ and $\phi_{22} = -0.8$, respectively. Comparing both figures a clear difference can be observed in the behaviour of trajectories corresponding to $\{X_2(t)\}$ as well as $\{Y(t)\}$. Setting $\phi_{11}\phi_{22} < 0$ (see Figure 7-1) yields an antipersistent behaviour of both time series. Moreover, the distribution of the product time series $\{Y(t)\}$ has heavier tails than for $\{X_1(t)\}$ and $\{X_2(t)\}$. This effect is more pronounced for the Student's t distribution of the residuals.

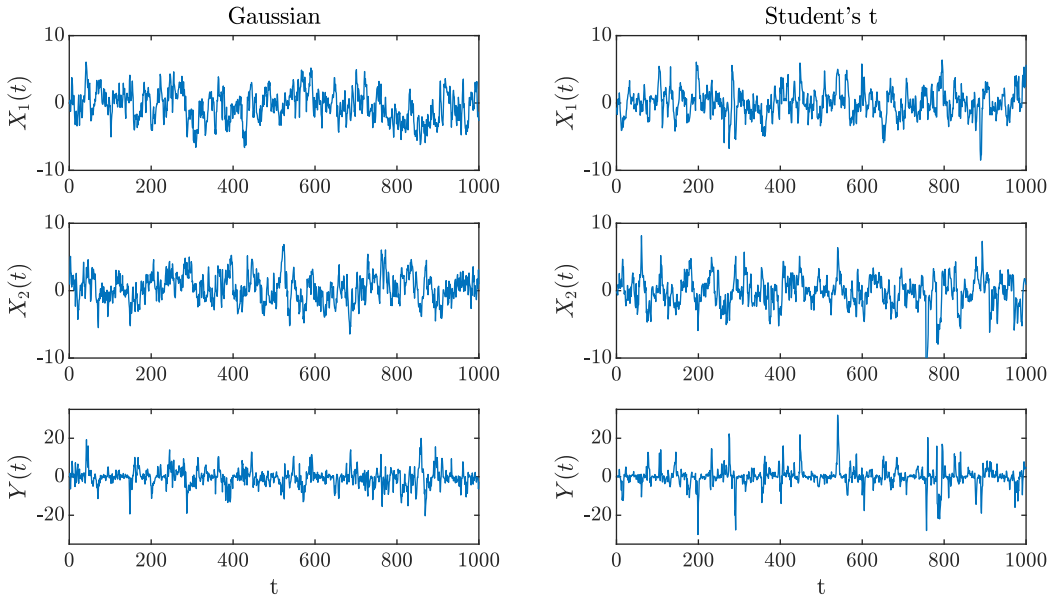


Figure 7-1: Sample trajectories of the VAR(1) model components and their product for the Gaussian (left panels) and Student's t distribution (right panels). The parameters correspond to Case 1, i.e. $\phi_{11} = 0.8$, $\phi_{22} = 0.8$, $\phi_{12} = \phi_{21} = 0$ and the residual vectors $Z_i(t), i = 1, 2$ are independent with $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

In Figure 7-3 we plot $\text{ACVF}_Y(h)$ for both sets of parameters and both distributions. The shape of the autocovariance function strongly depends on the sign of $\phi_{11}\phi_{22}$. It has a clear power decay if $\phi_{11}\phi_{22} > 0$ and antipersistent convergence to zero if $\phi_{11}\phi_{22} < 0$. The values of $\text{ACVF}_Y(h)$ are the same for both distributions, which follows directly from the formula (7.24) and the fact that the variances for both distributions are equal. However, a difference between the distributions can be observed in the widths of the intervals between CBs, i.e. the confidence intervals (CIs). Due to

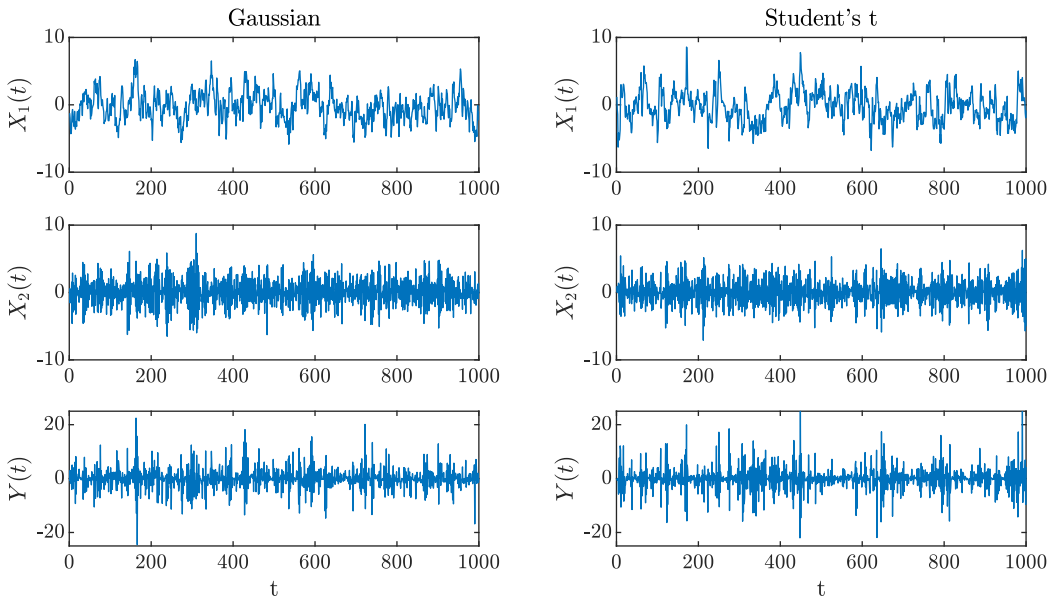


Figure 7-2: Sample trajectories of the VAR(1) model components and their product for the Gaussian (left panels) and Student's t distribution (right panels). The parameters correspond to Case 1, i.e. $\phi_{11} = 0.8$, $\phi_{22} = -0.8$, $\phi_{12} = \phi_{21} = 0$ and the residual vectors $Z_i(t), i = 1, 2$ are independent with $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

the heavier tails for the Student's t distribution wider CIs are obtained in this case. The difference is more apparent if $n = 100$, showing that the convergence of the empirical autocovariance to its theoretical value is slower than in the case of the Gaussian distribution.

Case 2

In this case, the time series $\{X_1(t)\}$ and $\{X_2(t)\}$ are dependent only through the residual vector. Hence, in the Gaussian case we set $\rho_Z \neq 0$, while in the Student's t case the bivariate specification (7.40) needs to be used, yielding dependence between Z_1 and Z_2 for any $\rho_Z \in (-1, 1)$. We start with a comparison of the behaviour of the trajectories and autocovariance functions with $\rho_Z = 0.8$ in both cases. Note that the sign of ρ_Z does not influence the value of the autocovariance function, which follows directly from the formula (7.29). On the other hand, similarly as in Case 1, the sign of $\phi_{11}\phi_{22}$ is an important factor for the ACVF behaviour. Hence, we use two parameter sets: either $\phi_{11} = 0.8$ and $\phi_{22} = 0.8$ or $\phi_{11} = 0.8$ and $\phi_{22} = -0.8$.

The sample trajectories corresponding to the time series $\{X_1(t)\}$, $\{X_2(t)\}$ and their product $\{Y(t)\}$ are plotted in Figures 7-4 and 7-5, respectively. Similarly as in Case 1, we can observe heavier tails for the Student's t distribution and a clear antipersistnecy if $\phi_{11}\phi_{22} < 0$. An interesting

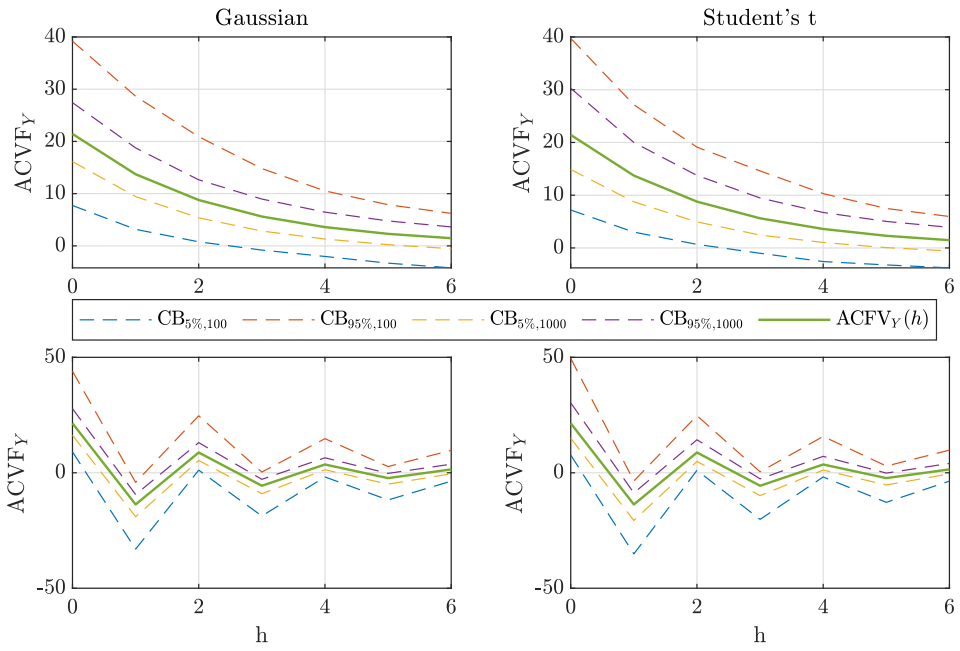


Figure 7-3: Autocovariance function of the VAR(1) components product $\text{ACVF}_Y(h)$ (solid line) and the corresponding confidence bounds $CB_{q,n}$ (dashed lines) for the Gaussian (left panels) and Student's t distribution (right panels). The confidence bounds were obtained using Monte Carlo simulations with 1000 repetitions. The sample length was set to $n = 100$ or $n = 1000$. The parameters correspond to Case 1, i.e. $\phi_{11} = 0.8$, $\phi_{22} = 0.8$, $\phi_{12} = \phi_{21} = 0$ (top panels; see Fig. 7-1 for the corresponding sample trajectories) or $\phi_{11} = 0.8$, $\phi_{22} = -0.8$, $\phi_{12} = \phi_{21} = 0$ (bottom panels; see Fig. 7-2 for the corresponding sample trajectories) and the residual vectors $Z_i(t), i = 1, 2$ are independent with $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

feature can be observed in Figure 7-4, i.e. when $\phi_{11}\phi_{22} > 0$. The strong dependence between the residuals makes the product distribution highly non-symmetric, even though the individual components do not exhibit this feature.

The empirical autocovariance function of the product time series for both parameter sets is plotted in Figure 7-6. Again, similarly as in Case 1, we can observe a power decay if $\phi_{11}\phi_{22} > 0$ and antipersistency if $\phi_{11}\phi_{22} < 0$. However, now the values of the $\text{ACVF}_Y(h)$ are different for the Gaussian and Student's t distributions. This is a consequence of different values of m_Z in Eq. (7.29). For the Gaussian distribution we have the following [132]:

$$m_Z = \mathbb{E}\left[Z_1^2 Z_2^2\right] = \sigma^2_{Z,1}\sigma^2_{Z,2} + 2\rho^2_Z\sigma^2_{Z,1}\sigma^2_{Z,2}.$$

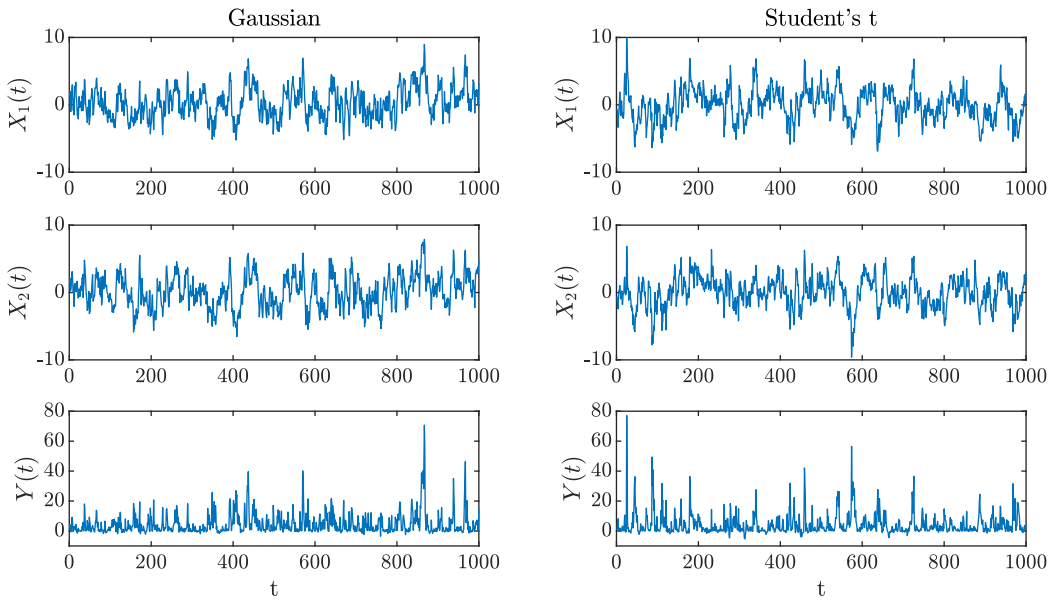


Figure 7-4: Sample trajectories of the VAR(1) model components and their product for the Gaussian (left panels) and Student's t distribution (right panels). The parameters correspond to Case 2, i.e. $\phi_{11} = 0.8$, $\phi_{22} = 0.8$, $\phi_{12} = \phi_{21} = 0$ and the residual vectors $Z_i(t), i = 1, 2$ are correlated with $\rho_Z = 0.8$. $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

Hence, the autocovariance function simplifies to:

$$\text{ACVF}_Y(h) = (\phi_{11}\phi_{22})^h \left[\frac{\sigma^2_{Z,1}\sigma^2_{Z,2}}{(1-\phi^2_{11})(1-\phi^2_{22})} + \frac{\rho^2_Z \sigma^2_{Z,1}\sigma^2_{Z,2}}{(1-\phi_{11}\phi_{22})^2} \right].$$

This is not the case for the Student's t distribution, for which in general $m_Z \neq \sigma^2_{Z,1}\sigma^2_{Z,2} + 2\rho^2_Z\sigma^2_{Z,1}\sigma^2_{Z,2}$, so the first component in the $\text{ACVF}_Y(h)$ (see formula (7.29)) is non-zero.

The difference between the distributions is also clearly visible in the widths of the confidence intervals. For the Student's t distribution much longer trajectories are needed to obtain a good convergence of the empirical autocovariance to its theoretical values.

Since for the Student's t distribution $\rho_Z = 0$ is not equivalent to independence, we also compare the $\text{ACVF}_Y(h)$ values obtained with $\rho_Z = 0$ in Case 2 with the ones obtained in Case 1 (i.e. with independent one-dimensional Student's t residuals). Recall that for the Gaussian case putting $\rho_Z = 0$ leads to independence, hence it directly corresponds to Case 1. The ACFV_Y functions for both Student's t specifications are plotted in Figure 7-7. Indeed, we can observe that the dependence between Z_1 and Z_2 changes the values of $\text{ACVF}_Y(h)$, even if $\rho_Z = 0$. This is a consequence of the m_Z value in the formula (7.29).

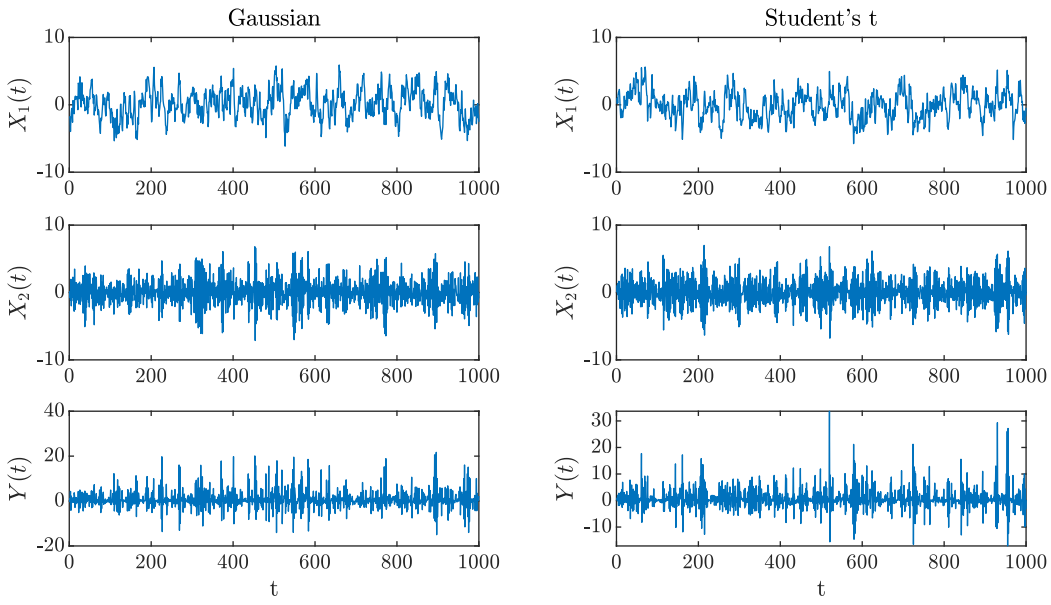


Figure 7-5: Sample trajectories of the VAR(1) model components and their product for the Gaussian (left panels) and Student's t distribution (right panels). The parameters correspond to Case 2, i.e. $\phi_{11} = 0.8$, $\phi_{22} = -0.8$, $\phi_{12} = \phi_{21} = 0$ and the residual vectors $Z_i(t), i = 1, 2$ are correlated with $\rho_Z = 0.8$. $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

Case 3

In this case, the random variables $Z_1(t)$ and $Z_2(t)$ are independent for each $t \in \mathbb{Z}$ and the relation between $\{X_1(t)\}$ and $\{X_2(t)\}$ is driven only by $\phi_{12} \neq 0$ and $\phi_{22} \neq 0$. Hence, we simulate the trajectories using the Gaussian distribution with $\rho_Z = 0$ and the independent one-dimensional Student's t distributions. Note that it follows directly from the formulas (7.37) and (7.38), that, differently than in Case 1 and Case 2, the sign of $\phi_{12}\phi_{22}$ does not influence the autocovariance values. Hence, now we use only a one-parameter set, namely $\phi_{12} = 0.8$, $\phi_{22} = 0.8$, $\phi_{11} = \phi_{21} = 0$.

The sample trajectories corresponding to time series $\{X_1(t)\}$, $\{X_2(t)\}$ and $\{Y(t)\}$ are plotted in Figure 7-8. Similarly as in Case 2, asymmetry of the product can be observed. However, now it is the effect of dependence through ϕ_{12} and ϕ_{22} coefficients and not the residuals correlation.

The autocovariance functions of the product time series together with the corresponding confidence bounds are plotted in Figure 7-9. For both distributions we can observe a power decay of $\text{ACVF}_Y(h)$, but the values are different. For the Gaussian distribution we have:

$$\kappa_Z = \mathbb{E}\left[Z_2^4\right] = 3\sigma^4_{Z,2}, \tag{7.43}$$

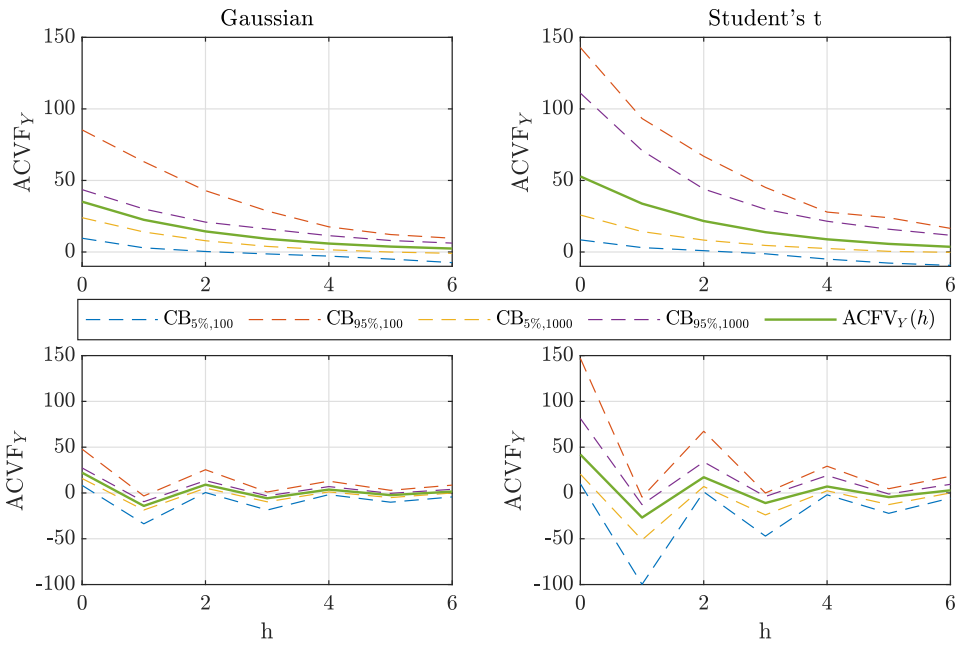


Figure 7-6: Autocovariance function of the VAR(1) components product $\text{ACVF}_Y(h)$ (solid line) and the corresponding confidence bounds $\text{CB}_{q,n}$ (dashed lines) for the Gaussian (left panels) and Student's t distribution (right panels). The confidence bounds were obtained using Monte Carlo simulations with 1000 repetitions. The sample length were set to $n = 100$ or $n = 1000$. The parameters correspond to Case 2, i.e. $\phi_{11} = 0.8$, $\phi_{22} = 0.8$, $\phi_{12} = \phi_{21} = 0$ (top panels; see Fig. 7-4 for the corresponding sample trajectories) or $\phi_{11} = 0.8$, $\phi_{22} = -0.8$, $\phi_{12} = \phi_{21} = 0$ (bottom panels; see Fig. 7-5 for the corresponding sample trajectories) and the residual vectors $Z_i(t), i = 1, 2$ are correlated with $\rho_Z = 0.8$. $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

which follows directly from the fact that the kurtosis, i.e. $\mathbb{E}[Z^4]/[\mathbb{V}\text{ar}(Z)]^2$, of the Gaussian distribution is equal to 3. This is not the case for the Student's t distribution. Hence, in the Gaussian case, $\text{ACVF}_Y(h)$ simplifies to:

$$\text{ACVF}_Y(h) = \begin{cases} \phi_{12}^2 \left[\frac{(1+\phi_{22}^2)\sigma_{Z,2}^4}{(1-\phi_{22}^2)^2} \right] + \frac{\sigma_{Z,1}^2 \sigma_{Z,2}^2}{1-\phi_{22}^2}, & \text{if} \quad h = 0, \\ \phi_{12}^2 \phi_{22}^{2h} \left[\frac{2\sigma_{Z,2}^4}{(1-\phi_{22}^2)^2} \right], & \text{if} \quad h > 0, \end{cases}$$

while for the Student's t the first component in the formulas (7.37) and (7.38) is always positive, making the $\text{ACVF}_Y(h)$ higher than in the Gaussian case. The confidence intervals are again wider for the Student's t distribution.

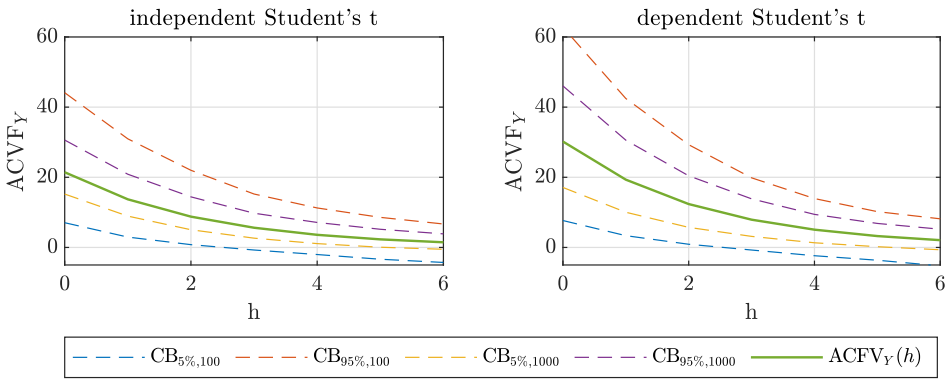


Figure 7-7: Autocovariance function of the VAR(1) components product $\text{ACVF}_Y(h)$ (solid line) and the corresponding confidence bounds $\text{CB}_{q,n}$ (dashed lines) for the bivariate Student's t distribution (right panel) with $\rho_Z = 0$ and independent Student's distributed residual vectors $Z_i(t), i = 1, 2$ (left panel). The confidence bounds were obtained using Monte Carlo simulations with 1000 repetitions. The sample length was set to $n = 100$ or $n = 1000$. The model parameters are equal $\phi_{11} = 0.8$, $\phi_{22} = 0.8$, $\phi_{12} = \phi_{21} = 0$, $\eta = 5$, source: [133].

7.4 Real data application. Modelling cost of electricity load prediction errors

Electricity prices are known to be autoregressive and highly dependent on the physical demand (or equivalently load), see e.g. [223]. The day-ahead forecasts of the load are usually published by the transmission system operators (TSO) and can be used for cost or production planning in electricity companies. However, these forecasts are burdened with prediction errors, which might cause large deviations of the actual energy cost from its predictions. The cost of these errors is equal to the product of the price and the difference between the actual and forecasted load. Hence, the methodology derived in Section 7.1 might be useful in such a context.

We apply the VAR(1) model to electricity day-ahead market data from the danish DK1 zone spanning over the time period 1.1.2016-31.12.2021, available from [224]. The time series corresponding to the first variable, denoted as $\{X_1^{\mu}(t)\}$, are the weekly means of the day-ahead electricity prices, while the second variable, corresponding to the time series $\{X_2(t)\}$, are the weekly means of the load forecast errors, i.e. the difference between the load values forecasted by the TSO and the corresponding actual values. The product of these time series $\{Y(t)\} = \{X_1^{\mu}(t)X_2(t)\}$ is the total weekly cost of the load prediction errors. The analysed time series are plotted in Figure 7-10.

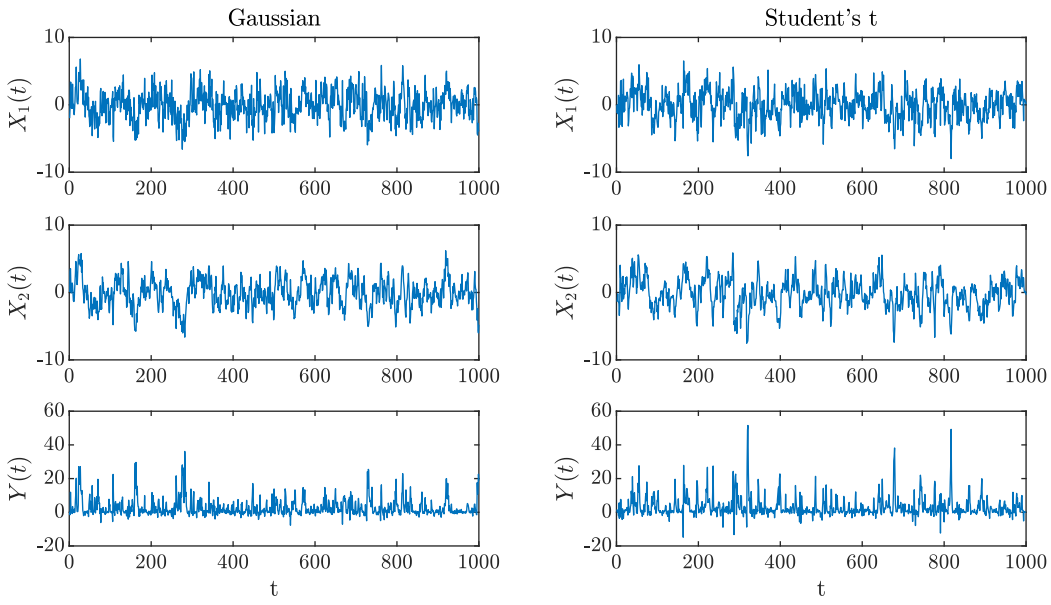


Figure 7-8: Sample trajectories of the VAR(1) model components and their product for the Gaussian (left panels) and Student's t distribution (right panels). The parameters correspond to Case 3, i.e. $\phi_{12} = 0.8$, $\phi_{22} = 0.8$, $\phi_{11} = \phi_{21} = 0$ and the residual vectors $Z_i(t), i = 1, 2$ are independent with $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

Before applying the VAR(1) model, the electricity prices are demeaned, i.e., for each t we calculate $X_1(t) = X_1^{\mu}(t) - \mu_X$, where μ_X is the mean of the prices corresponding to $\{X_1^{\mu}(t)\}$. In Figure 7-11 we plot the empirical ACVF obtained for both time series corresponding to $\{X_1(t)\}$ and $\{X_2(t)\}$, as well as the empirical cross-covariances, $\text{CCVF}_{X_1,X_2}(h) = \text{Cov}(X_1(t), X_2(t+h))$, between both components. The shapes of the obtained curves indicate a strong autoregressive effect in $\{X_1(t)\}$, a lower one in $\{X_2(t)\}$ and no such effects between the components. This observation is confirmed by the estimated matrix Φ of the coefficients of the VAR(1) model obtained using the Yule-Walker algorithm [206]:

$$\Phi = \begin{bmatrix} 0.7639 & -0.0629 \\ -0.0167 & 0.1247 \end{bmatrix}. \tag{7.44}$$

Next, we analyse the residuals obtained from the fitted VAR(1) model. The one-dimensional time series corresponding to $\{Z_1(t)\}$ and $\{Z_2(t)\}$ are plotted in Figure 7-12 together with the corresponding empirical auto- and cross-covariances. The obtained curves are close to 0 and show no time dependence in the residual series. The empirical correlation between the trajectories corresponding to $\{Z_1(t)\}$ and $\{Z_2(t)\}$ is equal to $\rho_Z = 0.0766$ and is not significantly different from 0

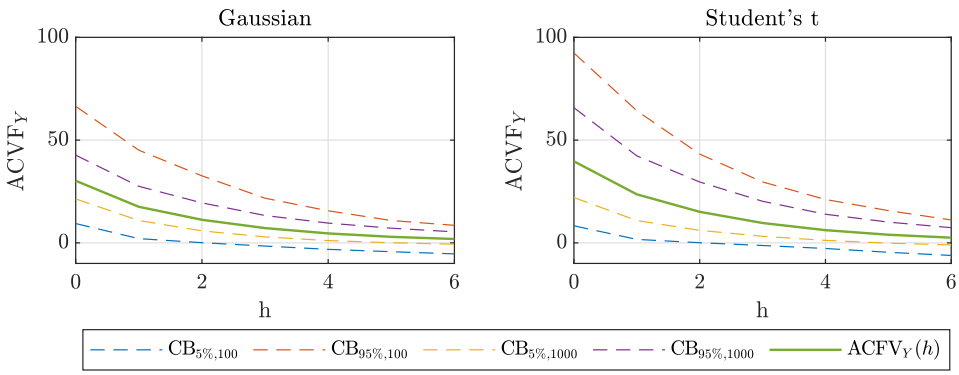


Figure 7-9: Autocovariance function of the VAR(1) components product $\mathrm{ACVF}_Y(h)$ (solid line) and the corresponding confidence bounds $\mathrm{CB}_{q,n}$ (dashed lines) for the Gaussian (left panel) and Student's t distribution (right panel). The confidence bounds were obtained using Monte Carlo simulations with 1000 repetitions. The sample length was set to $n = 100$ or $n = 1000$. The parameters correspond to Case 3, i.e. $\phi_{12} = 0.8$, $\phi_{22} = 0.8$, $\phi_{11} = \phi_{21} = 0$ (see Fig. 7-8 for the corresponding sample trajectories) and the residual vectors $Z_i(t), i = 1, 2$ are independent with $\eta = 5$ for the Student's t distribution and $\sigma^2_{Z,1} = \sigma^2_{Z,2} = \frac{\eta}{\eta-2}$ for the Gaussian one, source: [133].

according to the t test, see e.g. [225], as the p-value is equal to 0.2183. The independence of residuals' components is further confirmed by the χ^2 independence test, see e.g. [225], which yields the p-value of 0.2395.

Finally, we fit a distribution to the components of the residual series. Since the independence between time series corresponding to $\{Z_1(t)\}$ and $\{Z_2(t)\}$ for each t cannot be rejected, we analyse them separately as one-dimensional samples. In Figure 7-13 we plot the empirical probability density functions (PDFs) of sample trajectories corresponding to $\{Z_1(t)\}$ and $\{Z_2(t)\}$ together with the fitted PDFs of the Gaussian and Student's t location-scale distributions (see Appendix for details) and the corresponding quantile-quantile plots. It can be observed that the Student's t distribution yields a good fit to the empirical PDF and there is much improvement over the Gaussian one, especially in the case of load prediction errors (corresponding to $\{Z_2(t)\}$). The fit is further confirmed by the Kolmogorov-Smirnov goodness-of-fit test, see e.g.[225]. For the trajectory corresponding to the first component, the obtained p-values are equal to 0.1063 or 0.7458 for the zero mean Gaussian or Student's t location-scale distribution with $\mu = 0$, respectively. For the trajectory corresponding to $\{Z_2(t)\}$ the p-values are equal to 0.0012 and 0.5016, respectively. Hence, the Student's t distribution with the scale parameter cannot be rejected at any reasonable significance level for both variables, while the Gaussian distribution can be rejected in the case of the time series corresponding to $\{Z_2(t)\}$.

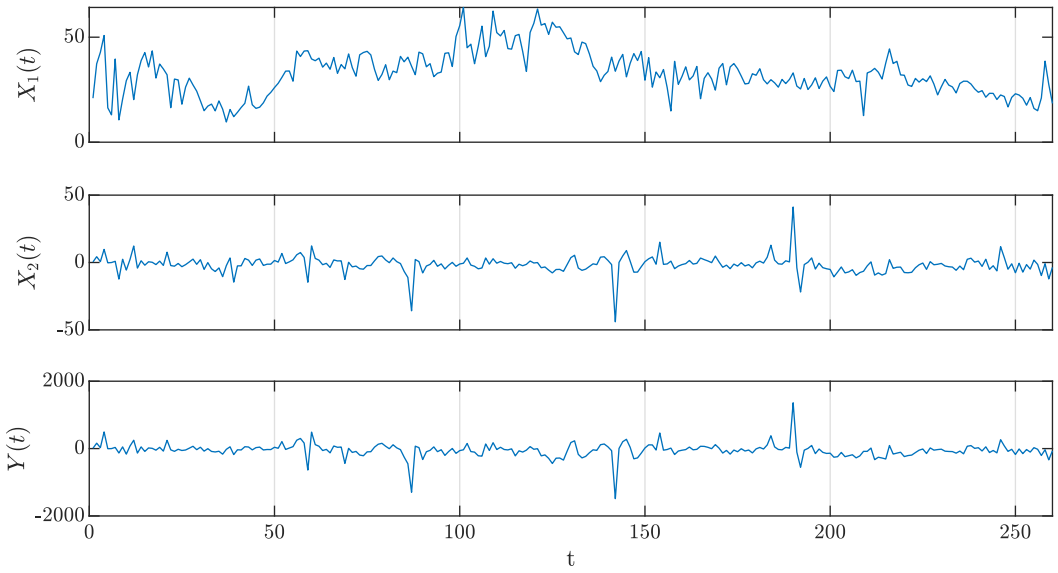


Figure 7-10: The analysed Danish electricity market data from the time period 1.1.2016-31.12.2020. Top panel: the weekly means of the day-ahead electricity prices (corresponding to time series $\{X_1^{\mu}(t)\}$). Middle panel: the weekly means of the load TSO forecast errors (corresponding to time series $\{X_2(t)\}$). Bottom panel: the product of weekly day-ahead prices and load forecast errors, ($\{Y(t)\} = \{X_1^{\mu}(t)X_2(t)\}$), i.e the total cost of these errors, source: [133].

Summarizing all of the obtained results, we conclude that a good fit is obtained for Case 1 (see Section 7.2) of the VAR(1) model with Student's t distributed residuals, i.e. $\phi_{12} = \phi_{21} = 0$ and $\{Z_1(t)\}$ and $\{Z_2(t)\}$ are independent. Estimation of the model parameters under Case 1 specification yields: $\phi_{11} = 0.7630$, $\phi_{22} = 0.1241$, while the Student's t degrees of freedom parameters $\eta_{Z,1} = 4.85$, $\eta_{Z,2} = 2.47$ and the scale parameters are equal to $\lambda_{Z,1} = 5.28$, $\lambda_{Z,2} = 3.03$.

Recall that the VAR(1) model was fitted to the demeaned prices, i.e. it was assumed that to $X_1(t) = X_1^{\mu}(t) - \mu_X$, where μ_X is the mean of the prices $\{X_1^{\mu}(t)\}$. Hence, before analyzing the final cost of the error, i.e.:

$$Y(t) = (X_1(t) + \mu_X)X_2(t), \tag{7.45}$$

there is a need to also apply the mean-shift also to the first coordinate of the fitted VAR(1) model. In Case 1 straightforward derivations lead to the following formula for the autocovariance of $\{Y(t)\}$:

$$\text{ACVF}_Y(h) = \text{ACVF}_{X_1X_2}(h) + \mu_X^2\text{ACVF}_{X_2}(h) = \frac{\sigma_{Z,1}^2\sigma_{Z,2}^2(\phi_{11}\phi_{22})^h}{(1-\phi_{11}^2)(1-\phi_{22}^2)} + \mu_{X_1}^2\phi_{22}^h\frac{\sigma_{Z,2}^2}{1-\phi_{22}^2}. \tag{7.46}$$

In Figure 7-14 we plot the derived theoretical autocovariance function for $\{Y(t)\}$, see Eq. (7.46), with the estimated parameters and the fitted Student's t residual distribution, as well as the

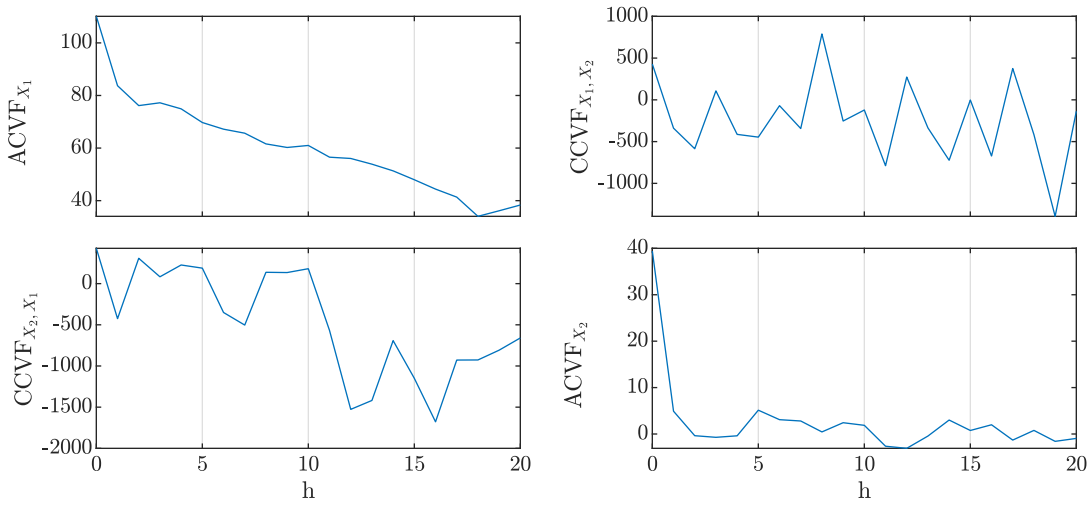


Figure 7-11: The empirical autocovariance function for time series corresponding to $\{X_1(t)\}$, ACVF$_{X_1}(h)$ (top, left panel) and $\{X_2(t)\}$, ACVF$_{X_2}(h)$) (bottom, right panel), as well as the corresponding empirical cross-covariances CCVF$_{X_1,X_2}(h)$ (top, right panel) and CCVF$_{X_2,X_1}(h)$ (bottom, right panel), source: [133].

empirical autocovariance function calculated for the product of the analysed data. Additionally, we also calculate the 5% and 95% CBs for the autocovariance function of the fitted model using Monte Carlo simulations with 1000 repetitions. As can be observed, the empirical autocovariance curve resembles the shape of the theoretical values and lies within the confidence intervals. It confirms that the fitted VAR(1) model describes well the properties of the product. Hence, the presented approach provides a well fitted model for both the prices and the load prediction errors and at the same time allows for modeling the total cost of the TSO load forecast errors, being the product of both variables. The results might be useful in cost planning for energy companies and help in a proper evaluation of the risk related to the errors of the load/demand predictions.

7.5 Discussion and summary

In this chapter, we have introduced a new times series arising as a product of the bi-dimensional VAR(1) model components and derived formulas for its main characteristics, such as the mean and the autocovariance function. Clearly, the results presented in this chapter can be further generalized for other time series models, especially for the VAR model with higher order or higher dimension. In the literature, there are also considered VAR models with heavy-tailed multidimensional distribution, e.g. stable [226, 227], thus the natural extension of the current study is the analysis of time series that is a product of two components of such models. In this case, the dependence structure can not be expressed by the means of the autocovariance function defined for finite-variance

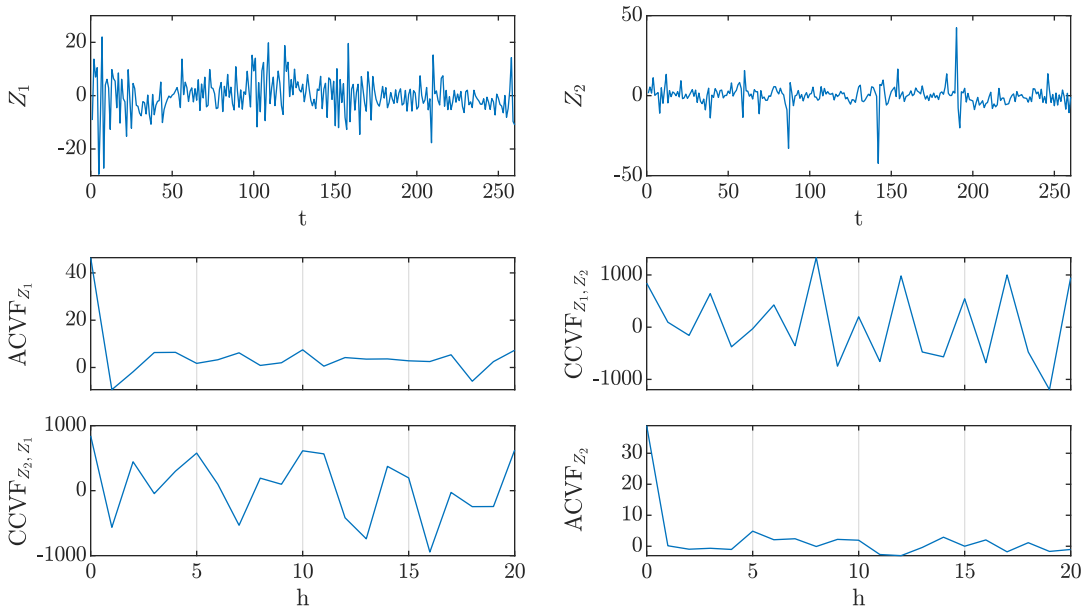


Figure 7-12: The residual series corresponding to $\{Z_1(t)\}$, $\{Z_2(t)\}$ (top right and left panels, respectively) and the empirical autocovariance function ACVF_{Z_1} (middle, left panel), ACVF_{Z_2} (bottom, right panel), as well as the empirical crosscovariances CCVF_{Z_1,Z_2} (middle, right panel) and CCVF_{Z_2,Z_1} (bottom, right panel), source: [133].

models, but by the dependence measures properly defined for the infinite-variance time series, see [226].

In the real data application section, we have conducted a case study based on the data from the Danish electricity market. We have shown that the weekly prices and load prediction errors can be modelled by a VAR(1) model, which also yields a good fit for the product time series, describing in this case the total cost of the load prediction errors. As the load forecasts are a crucial parameter for production and trade planning in electricity companies, the proper evaluation of the risk of their errors is essential for market strategies planing.

Another direction in which this topic could be analysed is the distribution of a random variable that is a product of two continuous random variables. In our study we derived formulas for the probability density functions and moments of the products of the Gaussian, log-normal, Student's t and Pareto random variables. This approach has been presented in [132]. Based on the data from continuous trading on the German energy market, we have shown a good reasonable fit of the product of log-normal and Student's t distribution to the transaction values. Since the transaction value is the final profit/cost for a trader, finding a proper density describing its distribution, which is also consistent with the prices and volumes data, can help an energy market participant in trading strategy planning.

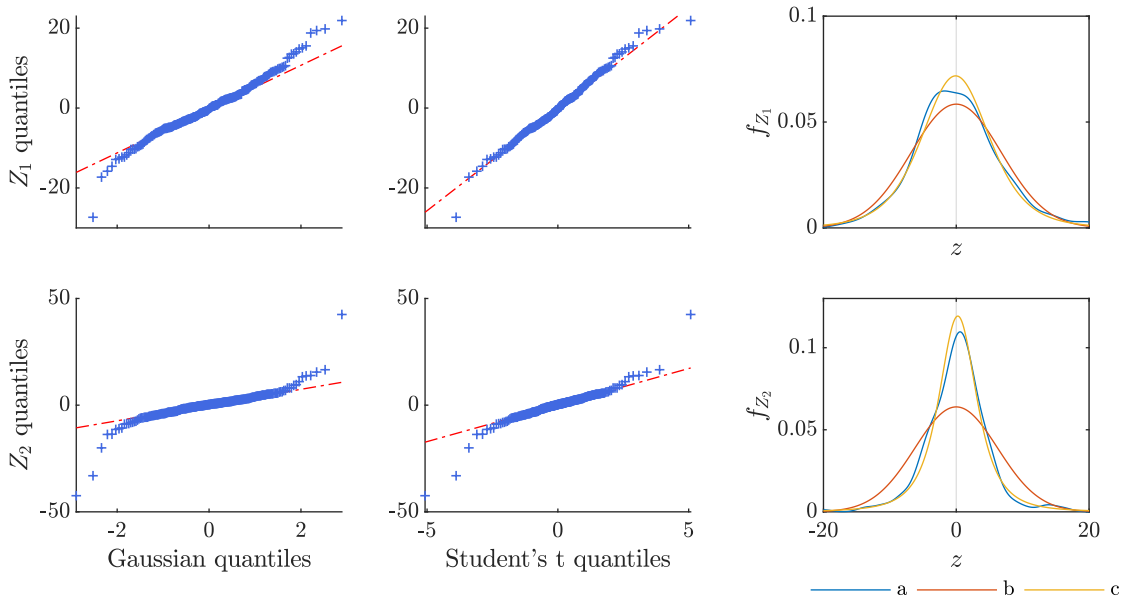


Figure 7-13: The quantile-quantile plots for the Gaussian (left panels) and Student's t location-scale distribution (middle panels) fitted to time series corresponding to $\{Z_1(t)\}$ (top panels) and $\{Z_2(t)\}$ (bottom panels). In the right panels the empirical PDFs (a, blue colour) corresponding to $\{Z_1(t)\}$, $f_{Z_1}(z)$, and $\{Z_2(t)\}$, $f_{Z_2}(z)$, together with the fitted PDFs for the Gaussian (b, red colour) and Student's t location-scale (c, yellow colour) distributions are plotted, source: [133].

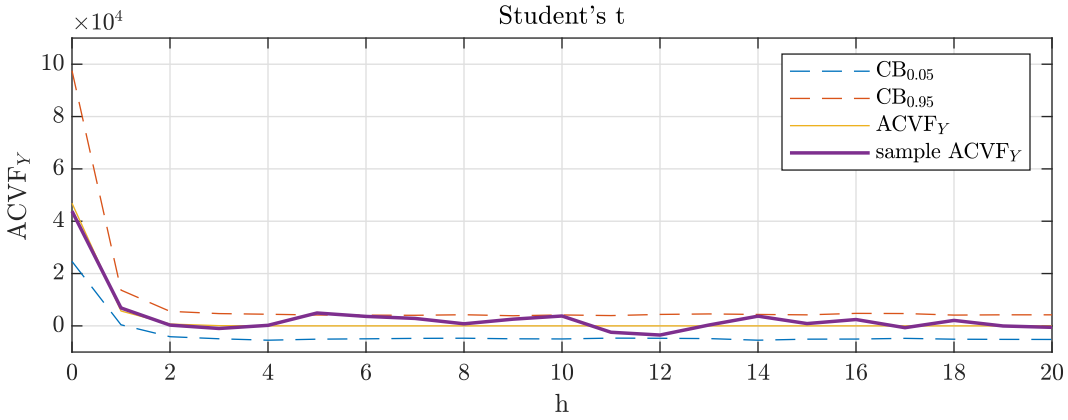


Figure 7-14: The autocovariance function of the product of prices and load prediction errors, $\mathrm{ACVF}_Y(h)$. The empirical autocovariance is plotted with a violet colour, while the theoretical values calculated for the fitted VAR(1) model, see Eq. (7.46) are plotted with a yellow colour. Additionally 5% and 95% CBs obtained from Monte Carlo simulations of the fitted model are plotted with dashed lines, source: [133].

Chapter 8

Conclusions. Potential application of discussed models for the market risk measurement process

An effective market risk management process in a industrial company consists of several stages: risk identification, the measurement of risk, determining risk appetite and, if natural methods of risk reduction are not sufficient, the use of risk mitigating instruments.

The risk measurement stage plays a special role in the market risk management process. Essentially, it quantifies the size of the company's current risk exposure and helps to answer the question of whether this exposure is optimal for the company in the context of the current market conditions and the internal situation or long-term plans of the company.

Correctly-calculated risk measures are used to estimate the impact of the company's potential actions on shaping the desired risk profile. In the risk measurement stage we can specify four successive steps [228]: Metric specification, Exposure mapping, Generating scenarios and Risk measure calculation as well as evaluation. Actually, scenario generation is exactly the risk management process stage, in which the stochastic models presented in this thesis may be applied [229]

8.1 Scenario generation - a key stage of the market risk measurement process

For calculating any risk measure, there is a requirement to generate price scenarios for market risk factors. The Monte Carlo simulation method, introduced by Stanislaw Ulam from the Lviv School

of Mathematics [230], is commonly used for sampling from the probability distribution. It can deal with often very complex and non-linear systems of equations. It assumes that the future prices of risk factors follow a predetermined predicted probability distribution.

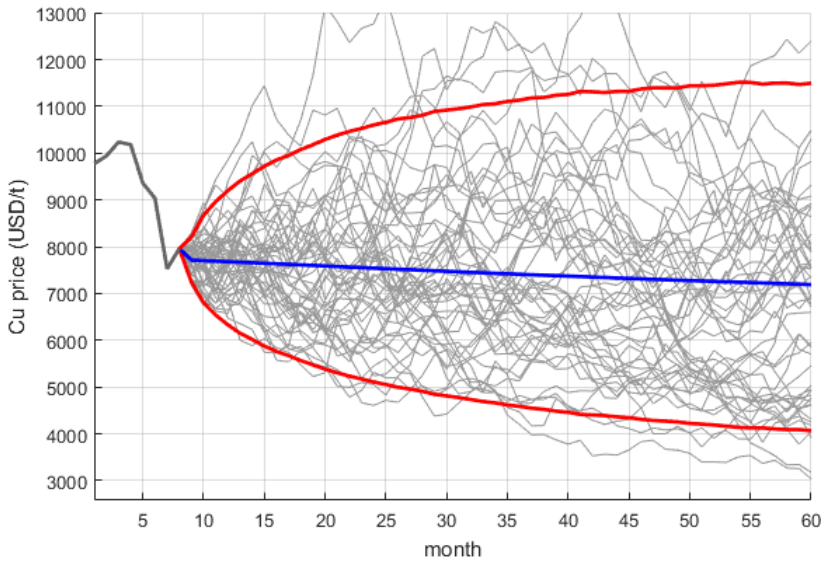


Figure 8-1: Example of copper price scenarios with mean and 5% / 95% quantiles.

Therefore it is very important to choose an appropriate model on the basis of which the distribution is created. The novel models discussed in previous chapters can be useful for improving the scenario generation process and solving the problems signaled in the Chapter 2. From a practical point of view, the areas of potential improvements indicated at the beginning of this thesis are connected to the three stages of model application, listed below.

Model specification

A model should correspond well to the actual behaviour of market data. Potential model risk may appear when for example, we assume that a stochastic process follows Gaussian distribution, whereas in fact it is fat-tailed, with a large deviation of the extreme values. We may also ignore stochastic volatility in a situation where the present observations affect the following ones. We might also misspecify relationships between variables. For instance, we might ignore correlations or get correlations wrong in VAR estimation. All of the above-mentioned examples of model risk have been addressed in this thesis.

Two novel stochastic models which could be used for modelling metals prices, introduced in Chapter 4 are based on the SGT distribution and stochastic differential equations. Additionally, in Section 4.1 we implemented and described modification of the classical Ornstein-Uhlenbeck

process by taking, instead of the constant, the time-dependent coefficients α and β. In turn, in Section 6.3, based on relation and distribution analysis, we proposed the use of a two-dimensional VAR model with $\alpha-$stable distribution. Moreover the non-homogeneity of the data is reflected in the two identified model regimes.

Model implementation

Another element which may create a problem with the functioning of the model in practice can also arise from the way models are implemented. A formal model does not and cannot provide a complete specification of model implementation in every conceivable circumstance, because there is a large number of possible instruments and markets and their varying institutional and statistical properties. Therefore, the selected model should be properly, and with great awareness, fit to the purpose of its use. This topic is developed in Section 8.2.

Model calibration

Other risk associated with model application can also result from the incorrect calibration of an otherwise good model. Parameters might be estimated with error, not kept up-to-date or estimated over inappropriate sample periods. Incorrect calibration can lead to major losses if the models are then used to price traded instruments. We indicated this problem at the beginning of this thesis, focusing especially on choosing an appropriate length of historical data for stochastic modelling. In Chapter 5 we performed an analysis based on the generalized (time-dependent) Vasicek model for the currency exchange rate data description. We demonstrated that, in the long term, the averaging of different models can give better and more stable results in modelling exchange rates.

8.2 Risk measurement applications using stochastic processes

As described in previous sections, stochastic modelling is very important part of the market risk measurement process. Obviously it can be used for many real business applications related to market the risk management process, and specifically the measurement process. Based on the business practice of KGHM, we can indicate the following market risk measurement areas in which stochastic modelling is used:

- Stress test analysis
- Sensitivity analysis

- Profitability Risk, Cash-Flow Risk and Liquidity Risk assessment
- Credit Risk exposure evaluation
- Hedging optimisation

Sensitivity analysis

Sensitivity analysis is a tool used to decrease the uncertainty of an outcome by changing the value of variables that affect it within a preset range. For corporates exposed to significant market risk, it is usually focused on finding dependence and the influence of changing prices of metals, energy prices and exchange rates among other important factors on a company's financial situation. While this often has a one-dimensional character, in business practice, adding an additional dimension is also very informative.

Modelling mining royalty tax payments in Poland is a good example of a multidimensional problem, as the tax formula is a derivative of the copper price in PLN. A model which could be used for conducting sensitivity analysis of tax payments to the copper price in PLN was described in more detail in Chapter 6 and in [131].

Stress test analysis

An extended and more extreme version of the sensitivity analysis is the analysis of stress test situations for the company. This effectively helps to evaluate the risk factor levels at which a company's key financial ratios are not acceptable from the company's point of view (break even points). In the prices simulation stage, it is extremely important to have comprehensive knowledge and a firm understanding of the stochastic models used. The fat tail distributions or homogeneous volatility assumption used can improve the quality of modelling in this area.

Stress test analysis can be used for many different purposes and be based on one-dimensional or multidimensional calculations. One of the applications can be assessing liquidity risk, for which stochastic models with heavy-tails could be used, as well as non-Gaussian distributions or addressing the higher probability of realization of extremely negative price scenarios. These kinds of models are analysed in Chapter 4 and in [129, 130].

Profitability Risk, Cash Flow Risk and Liquidity Risk assessment

Stochastic modelling is moreover used to calculate risk measures reflecting the financial situation of a company. In this part risk measures are presented based on the financial indicators used by manufacturing companies.

It is important to highlight that these measures are based on the Value-at-Risk concept proposed by Philippe Jorion [169], which a can be defined as the maximum value by which potential financial indicator may drop in relation to the assumed level (for example budget), as a result of (normal) market risk factor changes in a given reporting period and with a given probability. In the course of development of the Value-at-Risk measure a wide range of use at-Risk-based measures supporting the methodology of market risk management in manufacturing companies was created, such us CorporateMetrics, which also addressed using measures for a long-term forecasting horison [1, 228].

Below are presented examples of the measures used:

Earnings at Risk (EaR) - specifies by how much profit may decrease in relation to the assumed level (budget), as a result of (normal) changes in prices of metals, energy prices and exchange rates, for a given reporting period and with a given probability.

Fig. 8-2 presents graphically the concept of the EaR measure in relation to the planned budget.

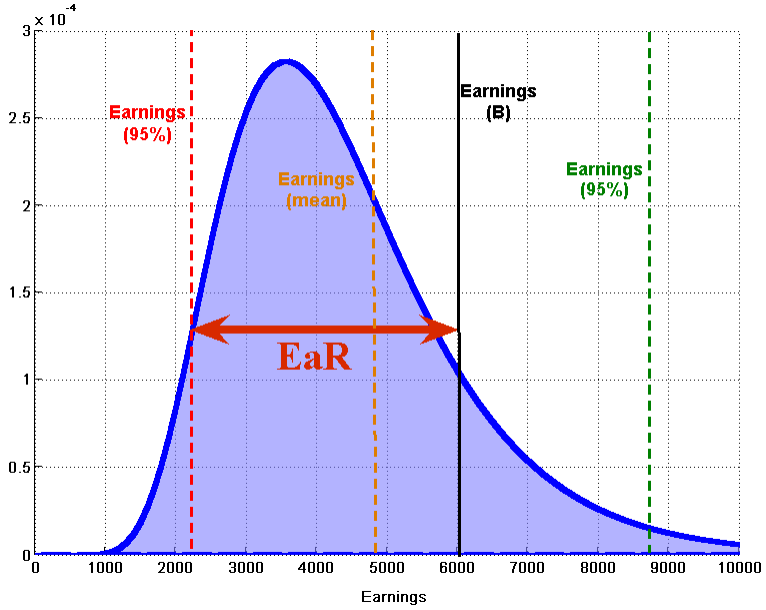


Figure 8-2: Earnings-at-Risk measure - an example.

Cash Flow at Risk (CFaR) - financial cash flows exposed to risk - defines how much financial flows may deteriorate in relation to the mean level, as a result of (normal) changes in raw material prices and exchange rates, for a given horizon and with a given probability.

Net Debt/EBITDA (Earnings Before Interest, Taxes, Depreciation and Amortization) - one of the most common financial covenant used in credit agreements. Breaking it may result in making a loan immediately payable.

Generally elements of all the models discussed in previous chapters can be considered for use in calculating risk measures. Nevertheless these measures often require simulating multiple risk factors and exposure maps are very complex. Therefore, before the implementation of more advanced models, the entire process should be analysed in terms of creating added value for the company, taking into account the system's overall complexity and costs.

Credit Risk exposure evaluation

Another area where stochastic modeling can be used is credit risk management. Derivatives used for managing market risk are usually traded with financial institutions, therefore the potential positive mark-to-market for the company with each individual institution in fact creates credit risk exposure. Classical measures like the transaction volume or portfolio mark-to-market doesn't reflect true credit exposure as they only address current market conditions, whereas the exposure can change and potentially increase, depending on market risk factors values. Therefore stochastic modelling / scenario generation is used here to estimate, with a given probability, the percentile of price levels, real potential credit risk exposure and to calculate the appropriate risk measure: Potential Receivables-at-Risk (PRaR). Such a measure can be for example calculated for a portfolio of exchange rates derivatives and the model CKLS described in Chapter 5 and in [127] can be considered here to be applied.

Hedging optimisation

When risk level is not acceptable, companies should firstly use natural methods of excess risk mitigation (for example netting, costs in the same currency as revenues, borrowing in a specific currency fit to exposure, purchasing and selling QP matching). If it is not possible or sufficient, hedging activity is often applied as a solution. Effectively this means using financial instruments (derivatives) in order to help model risk exposure in the manner accepted by the company. In fact, hedging instruments are used to manage risk and protect the financial measures described in previous sections.

Using stochastic modelling and Monte Carlo simulations in hedging activity and derivatives portfolio management is of fundamental importance as many financial instruments have an unsymmetrical payout profile, which requires advanced methods of computing to fully understand risk dynamics. Simulation is primarily used to:

- Determine the level of hedging, necessary to achieve the company's goals.
- Enable the assessment of the effect of a hedging strategy against its execution.

- Compare of strategies using different hedging instruments.

According to results presented in this thesis, based on different data length for model estimation and calibration, one can formulate varied conclusions. Taking this into account, the averaging of calibration windows, which is described in Chapter 5 and in [20], can be a valuable solution here.

8.3 Conclusions

Stochastic modelling is a key stage in the market measurement process and is used for generating price scenarios necessary to calculate risk measures. Choosing the right stochastic model, along with a comprehensive understanding of the assumptions on which is based, is extremely important.

Presented examples of stochastic modelling application presented in this chapter in the market risk measurement process have been based on the experience and the example of KGHM, a company active in the mining business, which is exposed to metals prices, energy prices and exchange rates fluctuations. However these examples can be generalised for any industrial company with a substantial market risk profile.

Applying models that take into account the price properties described in the Chapter 2 as non-Gaussian distribution, time-dependent coefficients, the changing of price regimes as well as the multidimensional approach, may help to refine the forecasting of market risk prices and thus improve the results of the market measurement process. The analysis, and the results presented in this thesis, demonstrate the usefulness of the introduced models in real applications.

Bibliography

[1] A. Y. Lee, *CorporateMetrics™ Technical Document*. RiskMetrics Group, 1999.

[2] M. Panella, F. Barcellona, and R. D'Ecclesia, "Forecasting energy commodity prices using neural networks," *Advances in Decision Sciences*, vol. 2012, p. 26, 2012.

[3] G. P. Herrera, M. Constantino, B. M. Tabak, H. Pistori, J.-J. Su, and A. Naranpanawa, "Long-term forecast of energy commodities price using machine learning," *Energy*, vol. 179, pp. 214–221, 2019.

[4] G. P. Herrera, M. Constantino, B. M. Tabak, H. Pistori, J.-J. Su, and A. Naranpanawa, "Data on forecasting energy prices using machine learning," *Data in Brief*, vol. 25, p. 104122, 2019.

[5] C. Liu, Z. Hu, Y. Li, and S. Liu, "Forecasting copper prices by decision tree learning," *Resources Policy*, vol. 52, pp. 427–434, 2017.

[6] G. Dooley and H. Lenihan, "An assessment of time series methods in metal price forecasting," *Resources Policy*, vol. 30, no. 3, pp. 208–217, 2005.

[7] C. Watkins and M. McAleer, "Econometric modelling of non-ferrous metal prices," *Journal of Economic Surveys*, vol. 18, no. 5, pp. 651–701, 2004.

[8] D. B. Reynolds, "The mineral economy: how prices and costs can falsely signal decreasing scarcity," *Ecological Economics*, vol. 31, no. 1, pp. 155–166, 1999.

[9] P. K. Achireko and G. Ansong, "Stochastic model of mineral prices incorporating neural network and regression analysis," *Mining Technology*, vol. 109, no. 1, pp. 49–54, 2000.

[10] G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[11] P. Du, J. Wang, Z. Guo, and W. Yang, "Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting," *Energy Conversion and Management*, vol. 150, pp. 90–107, 2017.

[12] C. T. Cortez, S. Saydam, J. Coulton, and C. Sammut, "Alternative techniques for forecasting mineral commodity prices," *International Journal of Mining Science and Technology*, vol. 28, no. 2, pp. 309–322, 2018.

[13] A. Rossen, "What are metal prices like? Co-movement, price cycles and long-run trends," *Resources Policy*, vol. 45, pp. 255–276, 2015.

[14] J. Lee, J. A. List, and M. C. Strazicich, "Non-renewable resource prices: Deterministic or stochastic trends?," *Journal of Environmental Economics and Management*, vol. 51, no. 3, pp. 354–370, 2006.

[15] M. C. Roberts, "Duration and characteristics of metal price cycles," *Resources Policy*, vol. 34, no. 3, pp. 87–102, 2009.

[16] M. A. Haque, E. Topal, and E. Lilford, "Iron ore prices and the value of the australian dollar," *Mining Technology*, vol. 124, no. 2, pp. 107–120, 2015.

[17] W. Walls, "Econometric analysis of the market for natural gas futures," *Energy Journal*, vol. 16, no. 1, pp. 71–83, 1995.

[18] W. Dechert, *Chaos Theory in Economics: Methods, Models and Evidence*. Cheltenham, UK: An Elgar Reference Collection, 1996.

[19] J. Miller and S. Ni, "Long-term oil price forecasts: A new perspective on oil and the macroeconomy," *Macroeconomic Dynamics*, vol. 15, no. S3, pp. 396–415, 2011.

[20] T. Serafin, A. Michalak, Ł. Bielak, and A. Wyłomańska, "Averaged-calibration-length prediction for currency exchange rates by a time-dependent Vasicek model," *Theoretical Economics Letters*, vol. 10, pp. 579–599, 2020.

[21] R. Alquist and L. Kilian, "What do we learn from the price of crude oil futures?," *Journal of Applied Econometrics*, vol. 25, no. 4, pp. 539–573, 2010.

[22] R. Alquist, L. Kilian, and R. J. Vigfusson, "Chapter 8 - Forecasting the price of oil," in *Handbook of Economic Forecasting* (G. Elliott and A. Timmermann, eds.), vol. 2, pp. 427–507, Elsevier, 2013.

[23] F. M. Baldursson, "Modelling the price of industrial commodities," *Economic Modelling*, vol. 16, no. 3, pp. 331–353, 1999.

[24] B. Dong, X. Li, and B. Lin, "Forecasting long-run coal price in China: A shifting trend time-series approach," *Review of Development Economics*, vol. 14, no. 3, pp. 499–519, 2010.

[25] J. Janczura, S. Orzeł, and A. Wyłomańska, "Subordinated α-stable Ornstein–Uhlenbeck process as a tool for financial data description," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 23–24, pp. 4379–4387, 2011.

[26] W. C. Labys, *Modeling and forecasting primary commodity prices*. Routledge, 2017.

[27] Y. Ru and H. J. Ren, "Application of ARMA model in forecasting aluminum price," in *Mechanical Engineering and Green Manufacturing II*, vol. 155 of *Applied Mechanics and Materials*, pp. 66–71, Trans Tech Publications Ltd, 2012.

[28] K. Gangopadhyay, A. Jangir, and R. Sensarma, "Forecasting the price of gold: An error correction approach," *IIMB Management Review*, vol. 28, no. 1, pp. 6–12, 2016.

[29] Y. Chen, K. He, and C. Zhang, "A novel grey wave forecasting method for predicting metal prices," *Resources Policy*, vol. 49, pp. 323–331, 2016.

[30] T. Kriechbaumer, A. Angus, D. Parsons, and M. Rivas Casado, "An improved wavelet–ARIMA approach for forecasting metal prices," *Resources Policy*, vol. 39, pp. 32–41, 2014.

[31] J. Nowicka-Zagrajek and A. Wyłomańska, "Measures of dependence for stable AR(1) models with time-varying coefficients," *Stochastic Models*, vol. 24, no. 1, pp. 58–70, 2008.

[32] M. Kulshreshtha and J. K. Parikh, "Modelling demand for coal in India: vector autoregressive models with cointegrated variables," *Energy*, vol. 25, no. 2, pp. 149–168, 2000.

[33] G. Uhlenbeck, "On the theory of the Brownian motion," *Phys. Rev.*, vol. 36, p. 823, 1930.

[34] O. Vasicek, "An equilibrium characterization of the term structure," *Journal of Financial Economics*, vol. 5, no. 2, pp. 177 – 188, 1977.

[35] C. T. Cortez, S. Saydam, and J. Coulton, "Using entropy to assess dynamic behaviour of long-term copper price," *Resources Policy*, vol. 66, 2020.

[36] R. J. Sweeney, "Mean reversion in G-10 nominal exchange rates," *The Journal of Financial and Quantitative Analysis*, vol. 41, no. 3, pp. 685–708, 2006.

[37] R. J. Sweeney and P. Jorion, "Mean reversion in real exchange rates: evidence and implications for forecasting," *Journal of International Money and Finance*, vol. 15, no. 4, pp. 535–550, 1996.

[38] E. Tully and B. M. Lucey, "A power GARCH examination of the gold market," *Research in International Business and Finance*, vol. 21, no. 2, pp. 316 – 325, 2007.

[39] J. Obuchowski and A. Wyłomańska, "The Ornstein-Uhlenbeck process with non-Gaussian structure," *Acta Physica Polonica B*, vol. 44, no. 5, pp. 11232–1136, 2013.

[40] P. J. Brockwell, "Recent results in the theory and applications of CARMA processes.," *Annals of the Institute of Statistical Mathematics*, vol. 66(4), pp. 647–685, 2014.

[41] P. J. Brockwell, "Lévy-driven CARMA processes.," *Annals of the Institute of Statistical Mathematics*, vol. 53(1), pp. 113–124, 2001.

[42] P. J. Brockwell, R. A. Davis, and Y. Yang, "Estimation for non-negative Lévy-driven CARMA processes.," *Journal of Business and Economic Statistics*, vol. 29(2), pp. 250–259, 2011.

[43] A. Wyłomańska, "Measures of dependence for Ornstein–Uhlenbeck process with tempered stable distribution.," *Acta Physica Polonica B*, vol. 42, no. 10, 2011.

[44] P. Theodossiou, "Financial data and the Skewed Generalized t-Distribution," *Management Science*, vol. 44, pp. 1650–1661, 12 1998.

[45] A. BenSaïda and S. Slim, "Highly flexible distributions to fit multiple frequency financial returns," *Physica A: Statistical Mechanics and its Applications*, vol. 442, pp. 203–213, 2016.

[46] S. Slim, Y. Koubaa, and A. Bensaida, "Value-at-Risk under Lévy GARCH models: Evidence from global stock markets," *Journal of International Financial Markets, Institutions and Money*, vol. 46, pp. 30–53, 2017.

[47] C. Hansen, J. McDonald, and P. Theodossiou, "Some flexible parametric models for partially adaptive estimators of econometric models," *Economics E-Journal*, vol. 1, no. 7, pp. 1–20, 2007.

[48] J. B. McDonald, R. A. Michelfelder, and P. Theodossiou, "Robust estimation with flexible parametric distributions: estimation of utility stock betas," *Quantitative Finance*, vol. 10, no. 4, pp. 375–387, 2010.

[49] A. M. Gambaro, "A discussion on non-gaussian price processes for energy and commodity operations," *Production and Operations Management*, vol. 30, pp. 47–67, 2021.

[50] A. Raknerud and Øivind Skare, "Multivariate stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck processes: A quasi-likelihood approach," *Statistics Norway*, vol. 614, 2010.

[51] T. S. Y. Ho and S.-B. Lee, "Term structure movements and pricing interest rate contingent claims," *The Journal of Finance*, vol. 41, no. 5, pp. 1011–1029, 1986.

[52] J. Hull and A. White, "Pricing interest-rate- derivative securities.," *Review of Financial Studies*, vol. 3, no. 4, pp. 573 – 592, 1990.

[53] F. Black, E. Derman, and W. Toy, "A one-factor model of interest rates and its application to treasury bond options," *Financial Analysts Journal*, vol. 46, no. 1, pp. 33–39, 1990.

[54] F. Black and P. Karasinski, "Bond and option pricing when short rates are lognormal," *Financial Analysts Journal*, vol. 47, p. 52, Jul 1991.

[55] J. Fan, J. Jiang, C. Zhang, and Z. Zhou, "Time-dependent diffusion models for term structure dynamics," *Statistica Sinica*, vol. 13, pp. 965–992, 10 2003.

[56] Y. Y. Su, H. J. Cui, and K. C. Li, "Parameter estimation of varying coefficients structural EV model with time series," *Acta Mathematica Sinica, English Series*, vol. 33, pp. 607–619, May 2017.

[57] H. Cui, "Estimation in partial linear ev models with replicated observations," *Science in China Series A: Mathematics*, vol. 47, p. 144, Jan 2004.

[58] C. Sophocleous, J. Hara, and P. Leach, "A model of stochastic volatility with time-dependent parameters," *Journal of Computational and Applied Mathematics*, vol. 235, 05 2011.

[59] P. Reichert and J. Mieleitner, "Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters," *Water Resources Research*, vol. 45, no. 10, 2009.

[60] J. Beckmann, A. Belke, and M. Kühl, "The dollar-euro exchange rate and macroeconomic fundamentals: a time-varying coefficient approach," *Review of World Economics*, vol. 147, p. 11–40, 2010.

[61] H. A. Dwyer and T. Petersen, "Time-dependent global energy modeling," *Journal of Applied Meteorology*, vol. 12, p. 36–42, 1973.

[62] J. Janczura and R. Weron, "Efficient estimation of markov regime-switching models: An application to electricity spot prices," *AStA Advances in Statistical Analysis*, vol. 96, 07 2011.

[63] C. J. Kim, J. Piger, and R. Startz, "Estimation of markov regime-switching regression models with endogenous switching," *Journal of Econometrics*, vol. 143, no. 2, pp. 263 – 273, 2008.

[64] K. Salhi, M. Deaconu, A. Lejay, N. Champagnat, and N. Navet, "Regime switching model for financial data: Empirical risk analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 461, pp. 148 – 157, 2016.

[65] J. Hamilton, *Regime-Switching Models*. Palgrave McMillan Ltd., 01 2008.

[66] J. Cai, "A markov model of switching-regime arch," *Journal of Business & Economic Statistics*, vol. 12, no. 3, pp. 309–316, 1994.

[67] N. Haldrup and M. Ørregaard Nielsen, "A regime switching long memory model for electricity prices," *Journal of Econometrics*, vol. 135, no. 1, pp. 349 – 376, 2006.

[68] A. Alizadeh, N. Nomikos, and P. Pouliasis, "A markov regime switching approach for hedging energy commodities," *Journal of Banking & Finance*, vol. 32, pp. 1970–1983, 09 2008.

[69] M. Naeem, A. K. Tiwari, S. Mubashra, and M. Shahbaz, "Modeling volatility of precious metals markets by using regime-switching GARCH models," *Resources Policy*, vol. 64, 2019.

[70] M. Azizi, A. Nunian, S. M. Zahari, and S. R. Shariff, "Modelling foreign exchange rates: a comparison between markov-switching and markov-switching garch," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, pp. 917–923, 2020.

[71] S. Kim and D. Y. Yang, "International monetary transmission and exchange rate regimes: Floaters vs. non-floaters in east asia," *Asian Development Bank Institute Working Papers*,

vol. 181, 2009.

[72] S. Johansen, "Modelling of cointegration in the vector autoregressive model," *Economic Modelling*, vol. 17, no. 3, pp. 359–373, 2000.

[73] P. R. Hansen, "Structural changes in the cointegrated vector autoregressive model," *Journal of Econometrics*, vol. 114, no. 2, pp. 261–295, 2003.

[74] E. Zivot and J. Wang, "Vector autoregressive models for multivariate time series," in *Modeling Financial Time Series with S-PLUS*, pp. 385–429, New York: Springer, 2006.

[75] S. Ankargren, M. Unosson, and Y. Yang, "A flexible mixed-frequency vector autoregression with a steady-state prior," *Journal of Time Series Econometrics*, vol. 12, no. 2, p. 20180034, 2020.

[76] H. Lütkepohl, "Comparison of criteria for estimating the order of a vector autoregressive process," *Journal of Time Series Analysis*, vol. 6, no. 1, pp. 35–52, 1985.

[77] P. Saikkonen and H. Lütkepohl, "Trend adjustment prior to testing for the cointegrating rank of a vector autoregressive process," *Journal of Time Series Analysis*, vol. 21, no. 4, pp. 435–456, 2000.

[78] S. Žiković, R. Weron, and I. Žiković, "Evaluating the performance of var models in energy markets," *Stochastic Models, Statistics, And Their Applications*, pp. 479–487, 2015.

[79] T.-R. Liu, M. E.Gerlow, and S. H.Irwin, "The performance of alternative var models in forecasting exchange rates," *International Journal of Forecasting*, vol. 10, no. 3, pp. 479–487, 2015.

[80] N. Apergis, "Domestic and eurocurrency yields: Any exchange rate link? evidence from a var model," *Journal of Policy Modeling*, vol. 19, no. 1, pp. 41–49, 1997.

[81] J. P.Jääskelä and D. Jennings, "Monetary policy and the exchange rate: Evaluation of var models," *Journal of International Money and Finance*, vol. 30, no. 7, pp. 1358–1374, 1994.

[82] W. E. Wecker, "A note on the time series which is the product of two stationary time series," *Stochastic Processes and their Applications*, vol. 8, no. 2, pp. 153–157, 1978.

[83] H. White and C. W. Granger, "Consideration of trends in time series," *Journal of Time Series Econometrics*, vol. 3, no. 1, 2011.

[84] P. D. Tella and C. Geiss, "Product and moment formulas for iterated stochastic integrals (associated with Lévy processes)," *Stochastics*, vol. 92, no. 6, pp. 969–1004, 2020.

[85] Y.-J. Lee and H.-H. Shih, "The product formula of multiple Lévy-Itô integrals," *Bulletin-Institute Of Mathematics Academia Sinica*, vol. 32, no. 2, pp. 71–96, 2004.

[86] F. Russo and P. Vallois, "Product of two multiple stochastic integrals with respect to a normal martingale," *Stochastic processes and their applications*, vol. 73, no. 1, pp. 47–68, 1998.

[87] S. Nadarajah, "Exact distribution of the product of two or more logistic random variables," *Methodology and Computing in Applied Probability*, vol. 11, pp. 651–660, 2008.

[88] S. Nadarajah, "Some algebra for Pearson type vii random variables," *Bulletin of The Korean Mathematical Society*, vol. 45, pp. 339–353, 2008.

[89] S. Nadarajah, "On the product xy for some elliptically symmetric distributions," *Statistics & Probability Letters*, vol. 75, pp. 67–75, 2005.

[90] M. Garg, A. Sharma, and P. Manohar, "The distribution of the product of two independent generalized trapezoidal random variables," *Communications in Statistics - Theory and Methods*, vol. 45, pp. 6369 – 6384, 2016.

[91] M. S. M. Ahsanullah, B. M. Golam Kibria, *Normal and Student's t Distributions and Their Applications*. Netherlands: Atlantis Press, 2014.

[92] S. Nadarajah and D. K. Dey, "On the product and ratio of t random variables," *Applied Mathematics Letters*, vol. 19, no. 1, pp. 45–55, 2006.

[93] A. Seijas-Macías and A. Oliveira, "An approach to distribution of the product of two normal variables," *Discussiones Mathematicae Probability and Statistics*, vol. 32, pp. 87–99, 2012.

[94] S. Nadarajah, "The product t density distribution arising from the product of two Student's t PDFs," *Statistical Papers*, vol. 50, pp. 605–615, 2009.

[95] Y. Li, Q. He, and R. S. Blum, "On the product of two correlated complex Gaussian random variables," *IEEE Signal Processing Letters*, vol. 27, pp. 16–20, 2020.

[96] M. S. M. Ahsanullah, B. M. Golam Kibria, *Normal and Student's t Distributions and Their Applications*. Atlantis Press, 2014.

[97] S. Nadarajah and D. K. Dey, "On the product and ratio of t random variables," *Applied Mathematics Letters*, vol. 19, no. 1, pp. 45–55, 2006.

[98] A. Seijas-Macías and A. Oliveira, "An approach to distribution of the product of two normal variables," *Discussiones Mathematicae Probability and Statistics*, vol. 32, pp. 87–99, 2012.

[99] Y. Li, Q. He, and R. S. Blum, "On the product of two correlated complex Gaussian random variables," *IEEE Signal Processing Letters*, vol. 27, pp. 16–20, 2020.

[100] H. J. Malik and R. Trudel, "Probability density function of the product and quotient of two correlated exponential random variables," *Canadian Mathematical Bulletin*, vol. 29, no. 4, p. 413–418, 1986.

[101] H. Homei, "The stochastic linear combination of dirichlet distributions," *Communications in Statistics - Theory and Methods*, vol. 50, pp. 2354 – 2359, 2019.

[102] J. Tang and A. Gupta, "On the distribution of the product of independent beta random variables," *Statistics & Probability Letters*, vol. 2, pp. 165–168, 1984.

[103] R. Bhargava and C. Khatri, "The distribution of product of independent beta random variables with application to multivariate analysis," *Annals of the Institute of Statistical Mathematics*, vol. 33, pp. 287–296, 1981.

[104] S. Nadarajah, "On the product of generalized Pareto random variables," *Applied Economics Letters*, vol. 15, pp. 253 – 259, 2008.

[105] T. Pham-Gia and N. Turkkan, "The product and quotient of general beta distributions," *Statistical Papers*, vol. 43, no. 4, pp. 537–550, 2002.

[106] S. Nadarajah and S. Kotz, "A note on the product of normal and laplace random variables," *Brazilian Journal of Probability and Statistics*, vol. 19, no. 1, pp. 33–38, 2005.

[107] S. Nadarajah and S. Kotz, "On the linear combination, product and ratio of normal and Laplace random variables," *J. Frankl. Inst.*, vol. 348, pp. 810–822, 2011.

[108] S. Nadarajah and S. Kotz, "On the product and ratio of Gamma and Weibull random variables," *Econometric Theory*, vol. 22, no. 2, pp. 338–344, 2006.

[109] S. Nadarajah and S. Kotz, "On the product and ratio of gamma and beta random variables," *Allgemeines Statistisches Archiv*, vol. 89, pp. 435–449, 2005.

[110] S. Nadarajah, "Sum, product and ratio of Pareto and gamma variables," *Journal of Statistical Computation and Simulation*, vol. 80, pp. 1071 – 1082, 2010.

[111] T. Pham-Gia and N. Turkkan, "The product and quotient of general beta distributions," *Statistical Papers*, vol. 43, no. 4, pp. 537–550, 2002.

[112] J. Galambos and I. Simonelli, *Products of Random Variables: Applications to Problems of Physics and to Arithmetical Functions (1st ed.)*. Boca Raton: CRC Press, 2004.

[113] H. Podolski, "The distribution of a product of n independent random variables with generalized gamma distribution," *Demonstratio Mathematica*, vol. 4, pp. 119 – 124, 1972.

[114] P. S. Wilson and R. Toumi, "A fundamental probability distribution for heavy rainfall," *Geophysical Research Letters*, vol. 32, no. 14, 2005.

[115] H. K. Cigizoglu and M. Bayazit, "A generalized seasonal model for flow duration curve," *Hydrological Processes*, vol. 14, no. 6, pp. 1053–1067, 2000.

[116] S. Ly, K.-H. Pho, S. Ly, and W.-K. Wong, "Determining distribution for the product of random variables by using copulas," *Risks*, vol. 7, no. 1, 2019.

[117] J. Salo, H. El-Sallabi, and P. Vainikainen, "The distribution of the product of independent Rayleigh random variables," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 2, pp. 639–643, 2006.

[118] Y. Yang and Y. Wang, "Tail behavior of the product of two dependent random variables with applications to risk theory," *Extremes*, vol. 16, pp. 55–74, 2013.

[119] N. Bhargav, C. R. N. da Silva, Y. J. Chun, E. J. Leonardo, S. L. Cotton, and M. D. Yacoub, "On the product of two $\kappa - \mu$ random variables and its application to double and composite fading channels," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2457–2470, 2018.

[120] S. Nadarajah and S. Kotz, "On the product and ratio of gamma and Weibull random variables," *Econometric Theory*, vol. 22, no. 2, pp. 338–344, 2006.

[121] P. Theodossiou, "Financial data and the skewed generalized t distribution," *Management Science*, vol. 44, pp. 1650–1661, 12 1998.

[122] K. Akdogan, S. Baser, M. G. Chadwick, D. Ertug, T. Hulagu, S. Kosem, F. Ogunc, U. Ozmen, and N. Tekatli, "Short-term inflation forecasting models for Turkey and a forecast combination analysis," pp. 312–325.

[123] M. H. Pesaran and A. Timmermann, "Selection of estimation window in the presence of breaks.," *Journal of Econometrics*, vol. 137, no. 6, pp. 134–161, 2007.

[124] M. H. Pesaran and A. Pick, "Forecast combination across estimation windows.," *Journal of Business & Economic Statistics*, vol. 29, no. 2, pp. 307–318, 2011.

[125] G. Marcjasz, T. Serafin, and R. Weron, "Selection of calibration windows for day-ahead electricity price forecasting.," *Energies*, vol. 11, no. 9, pp. 307–318, 2018.

[126] G. Marcjasz, T. Serafin, and R. Weron, "A note on averaging day-ahead electricity price forecasts across calibration windows.," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 321–323, 2018.

[127] G. Sikora, A. Michalak, Ł. Bielak, P. Miśta, and A. Wyłomańska, "Stochastic modeling of currency exchange rates with novel validation techniques," *Physica A: Statistical Mechanics and its Applications*, vol. 523, pp. 1202 – 1215, 2019.

[128] Ł. Bielak, P. Miśta, A. Michalak, and A. Wyłomańska, "The application of correlation models for the analysis of market risk factors in KGHM Capital Group," *Proceedings in Earth and Geosciences vol. 3, Mining Goes Digital, Mueller et al*, vol. 523, pp. 38–46, 2019.

[129] D. Szarek, Ł. Bielak, and A. Wyłomańska, "Long-term prediction of the metals' prices using non-Gaussian time-inhomogeneous stochastic process," *Physica A: Statistical Mechanics and its Applications*, vol. 555, p. 124659, 2020.

[130] D. Szarek, Ł. Bielak, and A. Wyłomańska, "Non-Gaussian regime-switching model in application to the commodity price description," *Chaari F.et al. (eds) Nonstationary Systems: Theory and Applications. WNSTA 2021. Applied Condition Monitoring, Springer*, vol. 18, pp. 108–126, 2021.

[131] Ł. Bielak, A. Grzesiek, J. Janczura, and A. Wyłomańska, "Market risk factors analysis for an international mining company. Multi-dimensional heavy-tailed-based modelling," *Resources Policy*, vol. 74, p. 102308, 2021.

[132] J. Adamska, Ł. Bielak, J. Janczura, and A. Wyłomańska, "From multi- to univariate: A product random variable with an application to electricity market transactions: Pareto and student's t-distribution case.," *Mathematics*, vol. 10, no. 18, p. 3371, 2022.

[133] J. Janczura, A. Puć, Ł. Bielak, and A. Wyłomańska, "Dependence structure for the product of bi-dimensional finite-variance VAR(1) model components. an application to the cost of electricity load prediction errors.," submitted 2021.

[134] C. Hansen, J. B. McDonald, and W. K. Newey, "Instrumental variables estimation with flexible distributions," *Journal of Business & Economic Statistics*, vol. 28, p. 13, 01 2010.

[135] A. Janicki and A. Weron, *Simulation and Chaotic Behavior of Stable Stochastic Processes*. New York: Marcel Dekker, 1994.

[136] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer, 2017.

[137] J. Marsden and A. Weinstein, *Calculus II*. Springer, 1985.

[138] A. Saleh, K. M. Ehsanes, M. Arashi, and S. M. M. Tabatabaey, "Statistical inference for models with multivariate t-distributed errors," *John Wiley & Sons*, pp. 133–170, 2014.

[139] R. Cont, "Encyclopedia of quantitative finance," *John Wiley & Sons*, vol. IV, pp. 1807–1811, 2010.

[140] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.

[141] T. Jaditz and L. A. Riddick, "Time-series near-neighbor regression.," *Studies in Nonlinear Dynamics & Econometrics*, vol. 4, no. 1, pp. 35 – 44, 2000.

[142] G. Elliott, T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," *Econometrica*, vol. 64, no. 4, pp. 813–836, 1996.

[143] F. Pozzi, T. Di Matteo, and T. Aste, "Exponential smoothing weighted correlations," *The European Physical Journal B*, vol. 85, no. 6, p. 175, 2012.

[144] T. Breusch and A. Pagan, "The lagrange multiplier test and its applications to model specifications in econometrics.," *Review of Economic Studies*, vol. 47, no. 1, p. 239, 1980.

[145] P. Tan and C. Drossos, "Invariance properties of maximum likelihood estimators," *Mathematics Magazine*, vol. 48, pp. 37–41, 01 1975.

[146] D. F. Shanno, "Conditioning of quasi-newton methods for function minimization," *Mathematics of Computation*, vol. 24, no. 111, pp. 647–656, 1970.

[147] L. Fox and D. F. Mayers, *Numerical Solution of Ordinary Differential Equations*. Chapman and Hall, 1987.

[148] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *Journal of the American Statistical Association*, vol. 69, no. 347, pp. 730–737, 1974.

[149] P. Brandimarte, "Handbook in monte carlo simulation: Applications in financial engineering, risk management, and economics," *John Wiley & Sons*, 2014.

[150] G. E. P. Box and D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 26, no. 2, pp. 211–252, 1964.

[151] F. Black and M. Scholes, "The pricing of options and corporate liabilities.," *Journal of Political Economy*, vol. 81, no. 3, p. 637, 1973.

[152] A. Charnes, E. L. Frome, and P. L. Yu, "The equivalence of generalized least squares and maximum likelihood estimates in the exponential family," *Journal of the American Statistical Association*, vol. 71, no. 353, pp. 169–171, 1976.

[153] H. A. L. Kiers, "Weighted least squares fitting using ordinary least squares algorithms," *Psychometrika*, vol. 62, no. 2, pp. 251–266, 1997.

[154] P. J. Huber, "Robust Estimation of a Location Parameter," *The Annals of Mathematical Statistics*, vol. 35, pp. 73–101, Mar 1964.

[155] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 298–311, 1997.

[156] C. Graham, *Markov Chains: Analytic and Monte Carlo Computations*. Wiley, 2014.

[157] I. L. M. Walter Zucchini, *Hidden Markov Models for Individual Time Series*. Chapman and Hall/CRC, 2009.

[158] J. Hamilton, *Time series analysis*. Princeton University Press, 1994.

[159] Y. Ephraim and N. Merhav, "Hidden markov processes," *IEEE Transactions on Information Theory*, vol. 48, pp. 1518–1569, June 2002.

[160] C. J. Kim, "Dynamic Linear Model with Markov-Switching," *Journal of Econometrics*, vol. 60, pp. 1–22, 02 1991.

[161] J. D. Hamilton, "Analysis of time series subject to changes in regime," *Journal of Econometrics*, vol. 45, no. 1, pp. 39 – 70, 1990.

[162] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, vol. 7, pp. 308–313, 01 1965.

[163] M. B. Wilk and R. Gnanadesikan, "Probability plotting methods for the analysis of data," *Biometrika*, vol. 55, no. 1, pp. 1–17, 1968.

[164] P. R. Krugman, "Target zones and exchange rate dynamics," *The Quarterly Journal of Economics*, vol. 106, no. 3, pp. 669–682, 1991.

[165] L. Svensson, "Why exchange rate bands?: Monetary independence in spite of fixed exchange rates," *Journal of Monetary Economics*, no. 1, pp. 157–199, 1994.

[166] A. Rose, "Are exchange rates macroeconomic phenomena?," *Economic Review*, no. 1, pp. 19–30, 1994.

[167] R. M. Myrvin Anthony, "On the mean-reverting properties of target zone exchange rates: Some evidence from the erm," *European Economic Review*, no. 8, pp. 1493–1523, 1994.

[168] K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders, "An empirical comparison of alternative models of the short-term interest rate," *The Journal of Finance*, vol. 47, no. 3, pp. 1209–1227, 1992.

[169] P. Jorion, *Value at Risk, The new benchmark for managing financial risk*. Mcgraw Hill Book Co, 2006.

[170] D. Tangman, N. Thakoor, K. Dookhitram, and M. Bhuruth, "Fast approximations of bond option prices under CKLS models," *Finance Research Letters*, vol. 8, no. 4, pp. 206–212, 2011.

[171] K. Nowman and G. Sorwar, "Derivative prices from interest rate models: results for Canada, Hong Kong, and United States," *International Review of Financial Analysis*, vol. 14 (4), p. 428–438, 2005.

[172] L. Chen, "Stochastic Mean and Stochastic Volatility – A Three-Factor Model of the Term Structure of Interest Rates and Its Application to the Pricing of Interest Rate Derivatives," *Financial Markets, Institutions & Instruments*, vol. 5, pp. 1–88, 1996.

[173] D. Tunaru, "Gaussian estimation and forecasting of the UK yield curve with multi-factor continuous-time models," *International Review of Financial Analysis*, vol. 52, p. 119–129, 2017.

[174] R. C. Merton, "Theory of rational option pricing," *The Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141–183, 1973.

[175] J. C. Cox, J. E. Ingersoll, and S. A. Ross, "A theory of the term structure of interest rates," *Econometrica*, vol. 53, no. 2, pp. 385–407, 1985.

[176] L. U. Dothan, "On the term structure of interest rates," *Journal of Financial Economics*, vol. 6, no. 1, pp. 59–69, 1978.

[177] M. J. Brennan and E. S. Schwartz, "An equilibrium model of bond pricing and a test of market efficiency," *Journal of Financial and quantitative analysis*, vol. 17, no. 3, pp. 301–329, 1982.

[178] J. C. Cox, J. E. Ingersoll Jr, and S. A. Ross, "An analysis of variable rate loan contracts," *The Journal of Finance*, vol. 35, no. 2, pp. 389–403, 1980.

[179] J. Cox, "Notes on option pricing I: Constant elasticity of variance diffusions," *Working paper, Stanford University, Graduate School of Business*, 1975.

[180] C. Christiansen, "Mean reversion in US and international short rates," *The North American Journal of Economics and Finance*, vol. 21, no. 3, pp. 286–296, 2010.

[181] J. R. Röman, *Analytical Finance: Volume II: The Mathematics of Interest Rate Derivatives, Markets, Risk and Valuation*. Springer, 2017.

[182] L. P. Hansen, "Large sample properties of generalized method of moments estimators," *Econometrica: Journal of the Econometric Society*, pp. 1029–1054, 1982.

[183] J. Hwang and Y. Sun, "Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework," *Journal of Econometrics*, vol. 207, no. 2, pp. 381–405, 2018.

[184] L. P. Hansen, "Method of Moments and Generalized Method of Moments," pp. 294–301, 2015.

[185] C. Li and W. Jiang, "On oracle property and asymptotic validity of Bayesian generalized method of moments," *Journal of Multivariate Analysis*, vol. 145, pp. 132–147, 2016.

[186] R. Liu and T. Lux, "Generalized Method of Moment estimation of multivariate multifractal models," *Economic Modelling*, vol. 67, pp. 136–148, 2017.

[187] M. Fan, S. Shao, and L. Yang, "Combining global Malmquist–Luenberger index and generalized method of moments to investigate industrial total factor CO2 emission performance: a case of Shanghai (China)," *Energy Policy*, vol. 79, pp. 189–201, 2015.

[188] S. Akbar, J. Poletti-Hughes, R. El-Faitouri, and S. Z. A. Shah, "More on the relationship between corporate governance and firm performance in the UK: Evidence from the application of generalized method of moments estimation," *Research in International Business and Finance*, vol. 38, pp. 417–429, 2016.

[189] S. K. Park, S. K. Ahn, and S. Cho, "Generalized method of moments estimation for cointegrated vector autoregressive models," *Computational Statistics & Data Analysis*, vol. 55, no. 9, pp. 2605–2618, 2011.

[190] G. Casale, "A generalized method of moments for closed queueing networks," *Performance Evaluation*, vol. 68, no. 2, pp. 180–200, 2011.

[191] X. Song and J. Cheng, "Joint estimation of the lognormal-Rician atmospheric turbulence model by the generalized method of moments," *Optics Communications*, vol. 285, no. 24, pp. 4727–4732, 2012.

[192] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. Berlin: Springer, 1992.

[193] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion. Physical review," vol. 36(5), 823, 1930.

[194] M. Musiela and M. Rutkowski, "Martingale methods in financial modelling: theory and applications," *Springer*, 1998.

[195] M. H. Pesaran and A. Pick, "Forecast combination across estimation windows," *Journal of Business and Economic Statistics*, vol. 29, no. 2, pp. 307–318, 2011.

[196] O. J. Dunn and V. A. Clark, *Basic statistics: a primer for the biomedical sciences*. John Wiley & Sons, 2009.

[197] M. Kendall and J. D. Gibbons, *Rank Correlation Methods*. Charles Griffin Book Series (5th ed.), Oxford: Oxford University Press, 1990.

[198] W. W. Daniel, *Kendall's tau*. Applied Nonparametric Statistics (2nd ed.), Boston: PWS-Kent, 1990.

[199] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.

[200] I. Kharisudin, D. Rosadi, Abdurakhman, and S. Suhartono, "The asymptotic property of the sample generalized codifference function of stable MA(1)," *Far East Journal of Mathematical Sciences*, vol. 99, pp. 1297–1308, 2016.

[201] A. Wyłomańska, A. Chechkin, J. Gajda, and I. M. Sokolov, "Codifference as a practical tool to measure interdependence," *Physica A: Statistical Mechanics and its Applications*, vol. 421, pp. 412–429, 2015.

[202] X. Ma and C. L. Nikias, "Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2669–2687, 1996.

[203] G. Samorodnitsky and M. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman and Hall, 1994.

[204] G. E. P. Box and G. M. Jenkins, *Introduction to time series and forecasting (2nd ed.)*. San Francisco: Holden-Day, 1976.

[205] M. Shelton Peiris and A. Thavaneswaran, "Multivariate stable ARMA processes with time dependent coefficients," *Metrika*, vol. 54, pp. 131–138, 2001.

[206] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*. Springer, 2016.

[207] A. Grzesiek, M. Mrozińska, P. Giri, S. Sundar, and A. Wyłomańska, "The covariation-based Yule-Walker method for multidimensional autoregressive time series with $\alpha-$stable distributed noise," *International Journal of Advances in Engineering Sciences and Applied Mathematics*, vol. 13, p. 394–414, 2021.

[208] C. M. Gallagher, "A method for fitting stable autoregressive models using the autocovariation function," *Statistics & Probability Letter*, vol. 53, pp. 381–390, 2001.

[209] P. Kruczek, A. Wyłomańska, M. Teuerle, and J. Gajda, "The modified Yule-Walker method for alpha-stable time series models," *Physica A*, vol. 469, pp. 588—-603, 2017.

[210] C. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*. Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control, Wiley, New York, 1995.

[211] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[212] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. R. Stat. Soc. B*, vol. 39, no. 1, pp. 1–38, 1977.

[213] M. Karson, "Handbook of methods of applied statistics. Volume I: Techniques of computation descriptive methods, and statistical inference. Volume II: Planning of surveys and experiments. I. M. Chakravarti, R. G. Laha and J. Roy, New York, John Wiley; 1967, $9.00.," *Journal of the American Statistical Association*, vol. 63, no. 323, pp. 1047–1049, 1968.

[214] R. Srinivasan, "On the Kuiper test for normality with mean and variance unknown," *Statistica Neerlandica*, vol. 25, no. 3, pp. 153–157, 1971.

[215] T. W. Anderson, "On the distribution of the two-sample Cramer-von Mises criterion," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1148–1159, 1962.

[216] T. W. Anderson and D. A. Darling, "A test of goodness of fit," *Journal of the American Statistical Association*, vol. 49, no. 268, pp. 765–769, 1954.

[217] I. A. Koutrouvelis, "Regression-type estimation of the parameters of stable laws," *Journal of the American Statistical Association*, vol. 75, no. 372, pp. 918–928, 1980.

[218] J. Wodecki, P. Kruczek, A. Bartkowiak, R. Zimroz, and A. Wyłomańska, "Novel method of informative frequency band selection for vibration signal using nonnegative matrix factorization of spectrogram matrix," *Mechanical Systems and Signal Processing*, vol. 130, pp. 585–596, 2019.

[219] G. G. Roussas, "Joint and conditional p.d.f.'s, conditional expectation and variance, moment generating function, covariance, and correlation coefficient," *In: An Introduction to Probability and Statistical Inference*, pp. 135–186, 2015.

[220] C. D. Lai and N. Balakrishnan, *Continuous Bivariate Distributions*. Springer New York, 2009.

[221] W. G. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 30, no. 2, p. 178–191, 1934.

[222] L. A. Aroian, S. V. Taneja, and L. W. Cornwell, "Mathematical forms of the distribution of the product of two normal variables," *Communications in Statistics - Theory and Methods*, vol. 7, no. 2, pp. 165–172, 1978.

[223] R. Weron, *Modeling and forecasting electricity loads and prices: a statistical approach*. Wiley Finance Series, Chichester: John Wiley & Sons, 2006.

[224] ENTSO-E, "European association for the cooperation of transmission system operators (TSOs) for electricity," 2021. https://transparency.entsoe.eu/, accessed: 2021-11-09.

[225] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures (5th ed.)*. Boca Raton: Chapman and Hall/CRC, 2011.

[226] A. Grzesiek, M. Teuerle, and A. Wyłomańska, "Cross-codifference for bidimensional VAR(1) time series with infinite variance," *Communications in Statistics - Simulation and Computation*, pp. 1–26, , 2019.

[227] A. Grzesiek, P. Giri, S. Sundar, and A. Wyłomańska, "Measures of cross-dependence for bidimensional periodic AR(1) model with alpha-stable distribution," *Journal of Time Series Analysis*, vol. 41, no. 6, pp. 785–807, 2020.

[228] J. M. Jongwoo Kim, Allan M. Malz, *LongRun Technical Document*. RiskMetrics Group, 1999.

[229] K. Dowd, *An Introduction to Market Risk Measurement*. John Wiley Sons, Ltd, 2002.

[230] R. E. Marks, *An Introduction to Market Risk Measurement*. The Palgrave Encyclopedia of Strategic Management, 2014.