IIT, Madras

Application of non-Gaussian multidimensional autoregressive model for climate data prediction

Ewa Broszkiewicz-Suwaj[1] · **Agnieszka Wyłomańska**[2]

Abstract From the point of view of agriculture, ecology, or environmental engineering, the capability of forecasting meteorological variables in the long and short term is crucial. Short-term forecasts enabling the planning of field work in agriculture, management of mass events, or tourism are important, while long-term forecasts related to advancing climate change are also very interesting. In the literature, there are known many approaches that can be used to forecast climate time series. The most common is based on the statistical modelling of the corresponding data, and the prediction is made on the fitted model. There are known one-dimensional approaches, where single variables are modeled separately; however, in the last decade, there appears a new trend which assumes the importance of the relationship between different time series. This is the approach considered in this paper. We propose to examine the climate data (temperature and precipitation) using the multidimensional vector autoregressive model (VAR). However, because in the time series we observe non-Gaussian behaviour, the classical VAR model can not be applied and the multidimensional Gaussian noise is replaced by the α−stable one. This model was previously analyzed by the authors in the context of financial data description where also non-Gaussian characteristics are observed. The main goal of this paper is to answer the question whether there are reasons to go from the Gaussian model to the generalized models, like α−stable based. The second purpose is to link total precipitation data with temperature time series. In the classical approach, precipitation was treated as a variable not correlated with temperature, which, as we will show in the paper, is inconsistent with reality. We hope the presented in this paper results open new areas of interest related to climate data modelling and prediction.

Keywords Non-Gaussian distribution · Multidimensional model · Estimation · Prediction · Climate data

✉ Ewa Broszkiewicz-Suwaj
  ewa.broszkiewicz-suwaj@upwr.edu.pl

  Agnieszka Wyłomańska
  agnieszka.wylomanska@pwr.edu.pl

1 Department of Applied Mathematics, Wrocław University of
   Environmental and Life Sciences, Norwida 25,
   50-375 Wrocław, Poland

2 Faculty of Pure and Applied Mathematics, Hugo Steinhaus
   Center, Wrocław University of Science and Technology,
   50-370 Wrocław, Poland

1 Introduction

From the point of view of agriculture, ecology, or environmental engineering, the capability of forecasting meteorological variables in the long and short term is very useful. On the one hand, short-term forecasts enabling the planning of field work in agriculture, management of mass events or tourism are important; on the other hand, long-term forecasts related to the advancing climate change are also very interesting. These changes depend on many global factors, the most important of which is the amount of greenhouse gases in the atmosphere. According to the analyzes, greenhouse gases (mainly carbon dioxide) are responsible for the temperature increase observed in many places around the world. In this case, depending on the RCP (Representative Concentration Pathways) scenario [1, 2], the content of these gases in the atmosphere is forecast using climatic indicators up to the year 2100. The

ability to model climate variables can be used for forecasting as well as for generating synthetic data in any time horizon and in any resolution using weather generators. Meteorological variables in the short-term horizon are usually described using models that take into account a large number of atmospheric factors [3, 4]. There are numerous models based on various types of statistical analysis [3–5], and from the point of view of new technologies, models based on artificial intelligence are becoming more and more popular [6]. From the point of view of generating meteorological data [7], the most popular models that are in use are models that in the case of precipitation, the model total precipitation by means of the first-order Markov chain to determine the occurrence of wet/dry days, and then for the amount of precipitation, the multidimensional two-parameter gamma distribution is used [2, 8, 9]. For other variables such as daily minimum temperature, maximum temperature, solar radiation, and wind speed, multivariate autoregression models are usually used [10]. In the first generators, it was assumed that the distributions of the described variables could be described by Gaussian distribution. However, it was soon observed that this assumption was not correct. In [11], the multivariate closed skew-normal distributions [12] were used for modeling of the residual skewness observed in climate data.

Following this path, one should ask whether the use of autoregressive Gaussian models to estimate model parameters is correct and whether there are reasons to go from the Gaussian model to the generalized models. Answering this question is the first goal of the paper. The second purpose is to link total precipitation data with temperature time series. In the first generator, precipitation was treated as a variable not correlated with temperature, which, as we will show in the paper, is inconsistent with reality.

To achieve the defined goals, in this paper, we propose a multidimensional approach for climate data description. To take into consideration the relation between the examined climate data (i.e., temperature and total precipitation), we apply a multidimensional vector autoregressive (VAR) model. It should be emphasized that this time series can be used to describe the dependence between analyzed variables, and at the same time, it takes into account the time dependence for single components. The VAR model is considered as the classical multidimensional time series [13]. However, because the analyzed data exhibit non-Gaussian behavior, the classical VAR model used in this paper, is modified to take into consideration the specific characteristics of the data. Thus, we propose to replace the Gaussian distributed innovations in the classical model by a more general class, namely $\alpha-$stable distributed variables. This class of distributions was introduced in 20's [14, 15]. The stable probability laws are important in

probability theory. According to the Generalized Central Limit Theorem, the stable laws attract distributions of sums of random variables with a diverging variance. It is a generalization of the Central Limit Theorem, which states that the Gaussian law attracts distributions with finite variance. The $\alpha-$stable distribution is considered as the generalization of the Gaussian one. Although the first application of this distribution appeared in the work of Mandelbrot [16], where the financial time series were analyzed, the $\alpha-$stable distributions and processes have found various applications, including economy [17–23], physics [24–27], signal processing [28–31], computer science [32–35], geology and geophysics [36–38], biology [39–42], and many other fields. The $\alpha-$stable distributions are also considered for climate data modelling, see, e.g., [43–46]. We also refer the readers to interesting bibliography positions [47, 48]. The VAR time series based on the $\alpha-$stable distribution was considered, for instance, in [49–52]. This model takes into consideration the relationship between the multidimensional data, and by using the $\alpha-$stable distribution instead of the Gaussian one, it allows describing the data with the heavy-tailed behaviour.

The paper is organized as follows. In Sect. 2, we present the used model. Next, in Sect. 3, we describe the analyzed climate data indicating the relationship between variables. The relation between temperature and precipitation data is a starting point to apply the VAR model. The results of the fitting are presented in Sect. 4. In this section, we also discuss the prediction results and compare them with the predictions obtained in the Klimada project. The last section concludes the paper.

2 The $\alpha-$stable vector autoregressive model

In this section, we present the model under consideration. Because the general methodology was introduced in the literature, therefore, we remind only the basic information referring to the appropriate bibliography positions.

The model examined in this paper is called the vector autoregressive time series (VAR model) with the α-stable distribution. This time series can be treated as a generalization of the Gaussian VAR model, [13].

A m-dimensional time series $\{\mathbf{X}(t)\} = \{(X_1(t),\ldots,X_m(t))^T\}$ is called a VAR model with the α-stable distribution, and if for each $t \in \mathbb{Z}$, it fulfills the following equation

$$\mathbf{X}(t) - \Theta_1\mathbf{X}(t-1) - \ldots - \Theta_p\mathbf{X}(t-p) = \mathbf{Z}(t). \qquad (1)$$

In Eq. (1), the sequence $\{\mathbf{Z}(t)\} = \{(Z_1(t),\ldots,Z_m(t))^T\}$ constitutes a sample of independent m-dimensional vectors with independent components having one-dimensional

α−stable distribution. More precisely, for each $t \in \mathbb{Z}$ $Z_j(t)$ ($j = 1, 2, \cdots, m$) has α−stable distribution and $Z_j(t)$ and $Z_i(t)$ are independent for $i \neq j$. Moreover, $\Theta_1, \ldots, \Theta_p$ are $m \times m$ matrices with time-constant coefficients. In our paper, we consider the special case of the model (1) and assume $p = 1$. We remind the one-dimensional random variable Z has α−stable distribution if its characteristic function is given by [53, 54]

$$\mathbb{E}[\exp\{i\theta Z\}] = \exp\{-\sigma^\alpha |\theta|^\alpha (1 + i\beta w(t, \theta)) + i\mu\theta\}, \quad (2)$$

where

$$w(\theta, \alpha) = \begin{cases} -sign(\theta) \tan\left(\frac{\pi\alpha}{2}\right) & \text{if } \alpha \neq 1, \\ \frac{2}{\pi} sign(\theta) \ln|\theta| & \text{if } \alpha = 1, \end{cases} \quad (3)$$

and $sign(\cdot)$ denotes a sign function. The parameter $0 < \alpha \leq 2$ is the stability index which is responsible for the rate of convergence of the distribution tail. More precisely, for $\alpha \neq 2$, the probability tail $P(Z > z) \sim z^{-\alpha}$ and the corresponding second moment are infinite. Thus, the smaller α, the convergence is slower, and in the consequence, the α−stable random variables take extreme values more likely than it in the Gaussian case. The α−stable distribution for $\alpha \neq 2$ belongs to the so-called heavy-tailed class of distributions. On the other hand, for $\alpha = 2$, it reduces to the Gaussian distribution, and in this sense, it can be considered as its generalization. The model defined as in Eq. (1) can be considered as the generalisation of the VAR model with Gaussian innovations. The other parameters of α−stable distribution are: the scale parameter $\sigma > 0$, the skewness parameter $-1 \leq \beta \leq 1$, and the shift parameter $\mu \in \mathbb{R}$. For more details about the α−stable distribution, we refer to [53, 55–59].

It should be emphasized that for $m = 1$, the model defined in Eq. (1) is the one-dimensional autoregressive time series (AR model) with α−stable innovations, [60]. For more details about the AR model and multidimensional VAR model with α−stable innovations, we refer the readers to the bibliography positions [49–51, 60–62].

In the Gaussian VAR model, one of the classical methods used for the estimation of the parameters is the Yule-Walker approach that utilizes the auto-covariance function [13]. For the VAR model with non-Gaussian α−stable distribution, this classical measure of dependence is infinite, and thus, in the literature, there are considered different approaches to the parameters' estimation. One of these is the modified Yule-Walker approach that utilizes the so-called auto-covariation function, and one of the dependency measures properly defined for the α−stable model. This method is described in detail in [49, 63], and in this paper, it is used for multi-dimensional

VAR models as well as for one-dimensional AR time series with α−stable innovations. See also [28, 64, 65].

3 Climate data description

The analyzed data are the daily measurements of the maximum temperature (Tmax), minimum temperature (Tmin), average temperature (Tavg), and total precipitation (P) for the city of Wrocław from 1961–2020. The data come from the Institute of Meteorology and Water Management (IMGW) [66]. Each variable is described by a time series of length 21915 data.

Figure 1 shows the graphs of all analyzed data. As can be easily seen, all data exhibit seasonal behavior. Therefore, they were divided into individual months, and thus, 12 time series for each variable were obtained. Figures 2, 3, , 4 and 5 show the behavior of the analyzed variables in the years 1961–2020 for individual months.

For the temperature data, an upward trend characterizing the data in individual months is visible; therefore, before the time series modelling, the long-term trend, based on a moving average averaging data from the moving 10 years, was removed. As it is known from various studies [67], this trend is related to long-term climate changes leading to global warming.

The first step of our study was to analyze the structure of the relationship between the individual time series. For this purpose, the behavior of correlation coefficients for 5-year windows sliding by one day taking into account all combinations of time series was investigated. The period of five years has been taken as long enough to detect a correlation while not being affected by the long-term trend. For the purposes of research, the Pearson, Spearman, and Kendall correlation coefficients were used, see [68]. The Person correlation coefficient is essential to study the linear relationship between two variables; however, its estimatior is sensitive to outliers. Therefore, this measure is useful, especially for the Gaussian (or light-tailed) distributed variables [69]. The Spearman correlation coefficient (called Spearman rank correlation coefficient) measures a monotonic relationship between variables. Its estimator is insensitive to large observations and thus in cases when the analyzed variables are heavy-tailed [70, 71]. The Kendall correlation coefficient (called also Kendall rank correlation coefficient) indicates not only the strength, but also the direction of the dependence. Similarly to the Spearman correlation coefficient, it is resistant to outliers and is used especially for non-Gaussian distributed data [72].

In Fig. 6, there are presented the exemplary correlation coefficients for the sliding 5-year windows for July data

Fig. 1 The daily measurements of maximum temperature, minimum temperature, average temperature, and total precipitation for Wrocław from the period 1961–2020
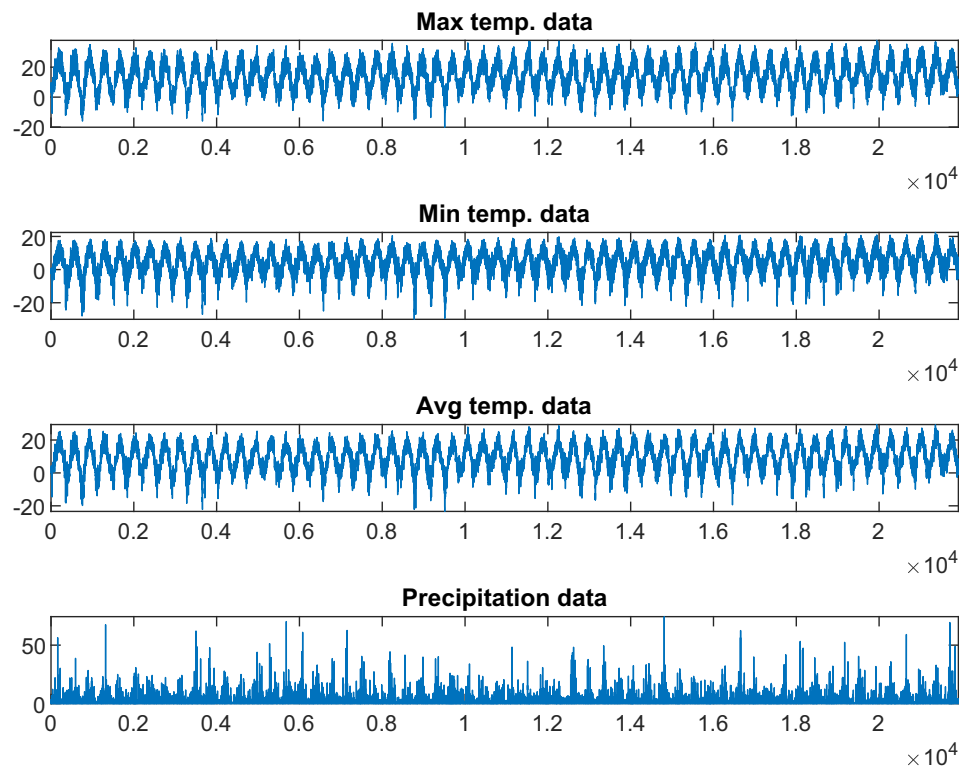

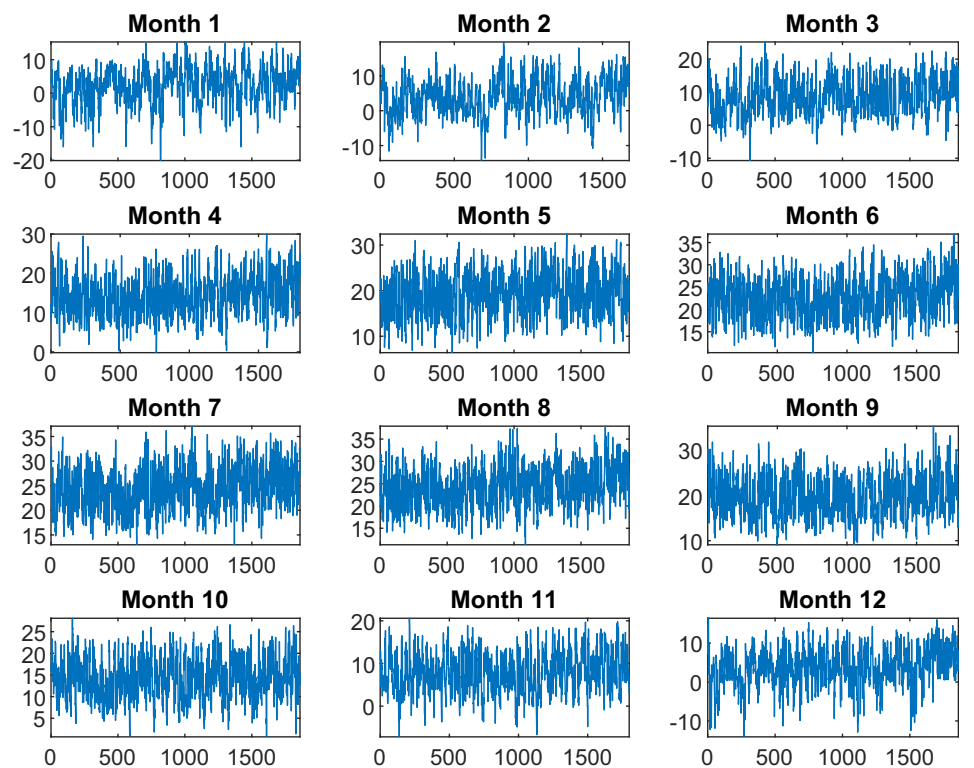**Fig. 2** Daily maximum temperature for each month in year for the period 1961–2020

Fig. 3 Daily minimum
temperature for each month in
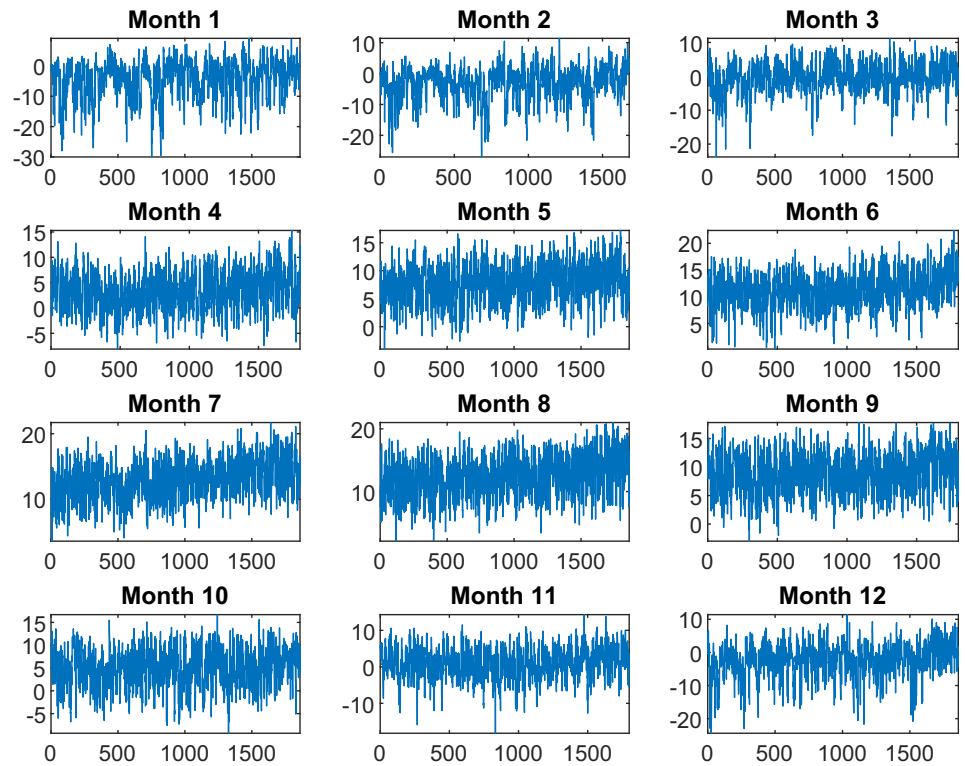year for the period 1961–2020

Fig. 4 Daily mean temperature
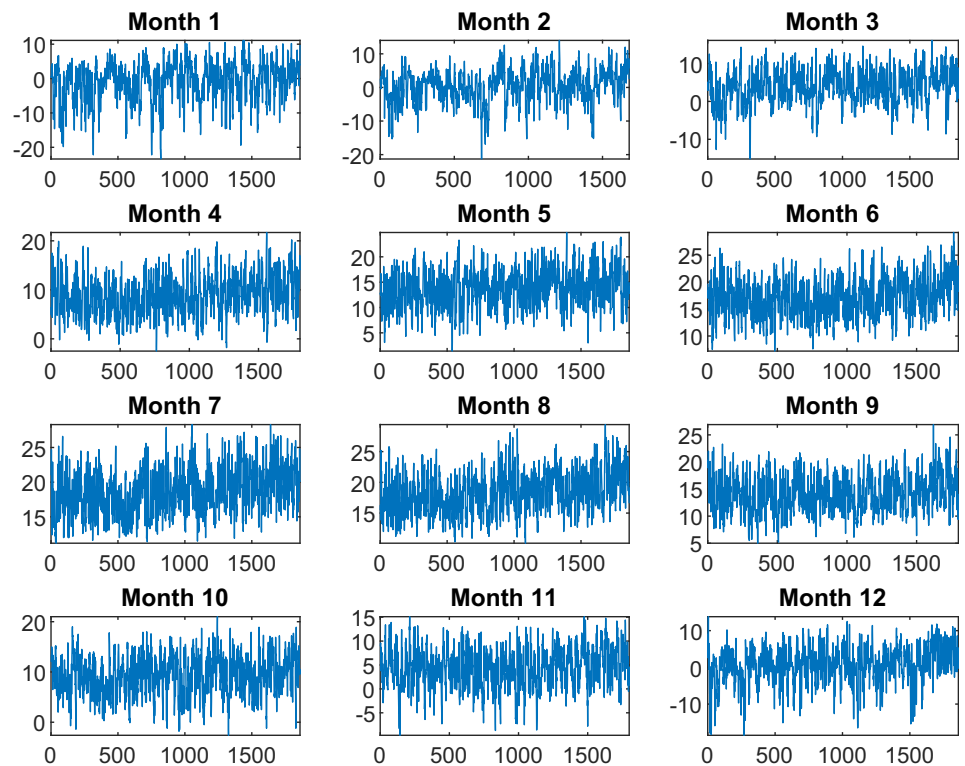for each month in year for the
period 1961–2020

Fig. 5 Daily precipitation sum
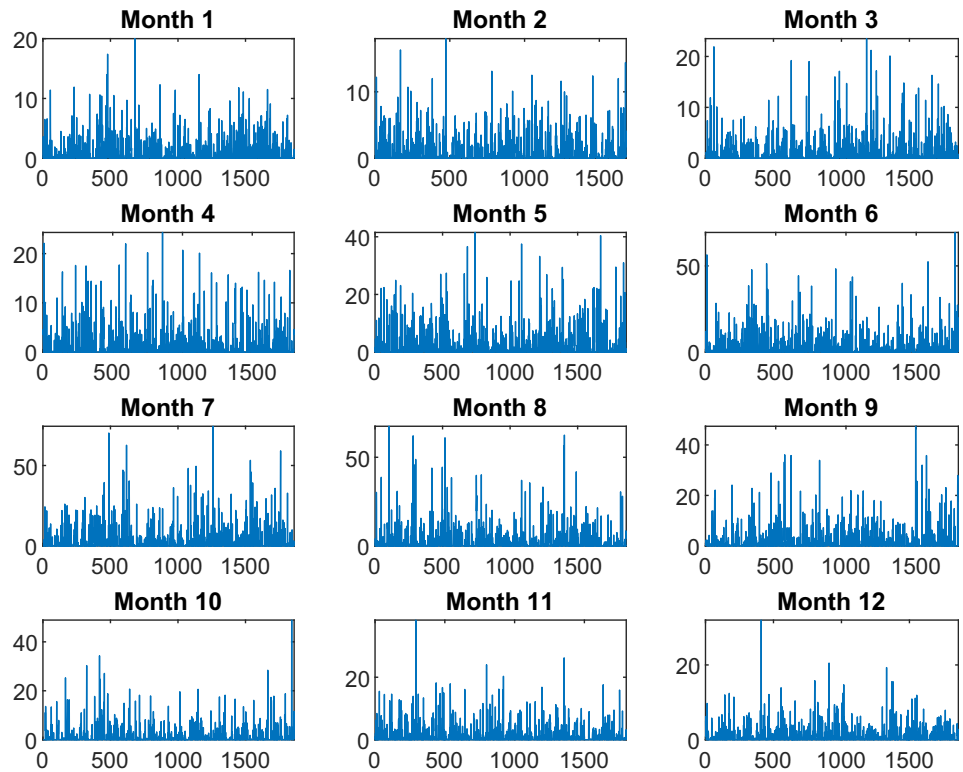for each month in year for the
period 1961–2020

Fig. 6 The analyzed correlation
coefficients for sliding five-
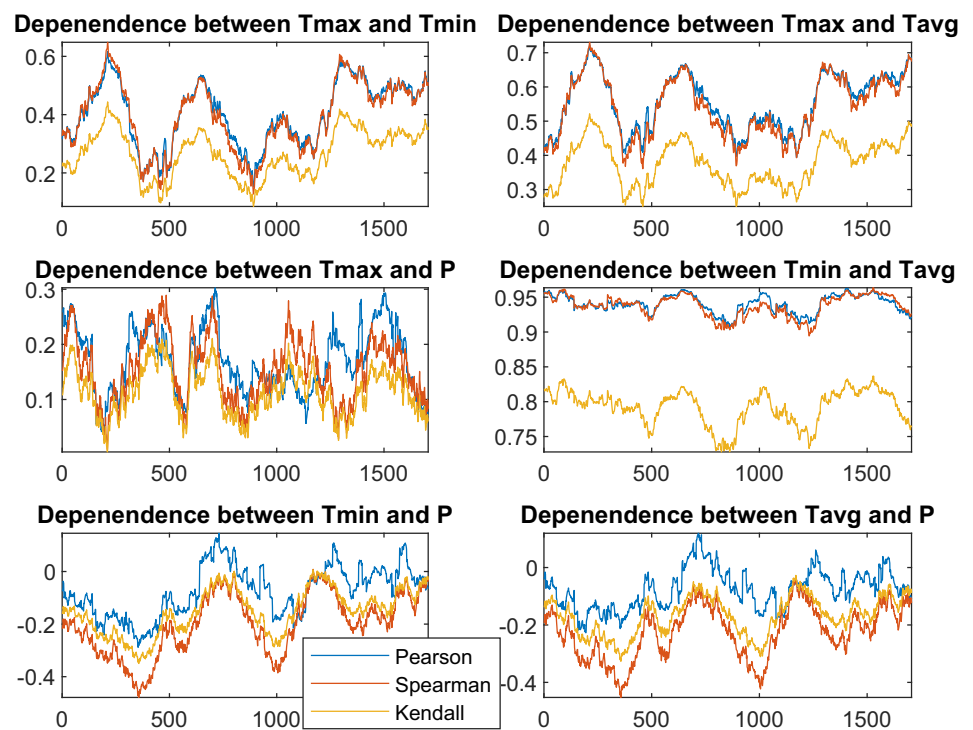years windows in July for each
data combination

Fig. 7 The analyzed correlation coefficients for sliding five-years windows in December for each data combination
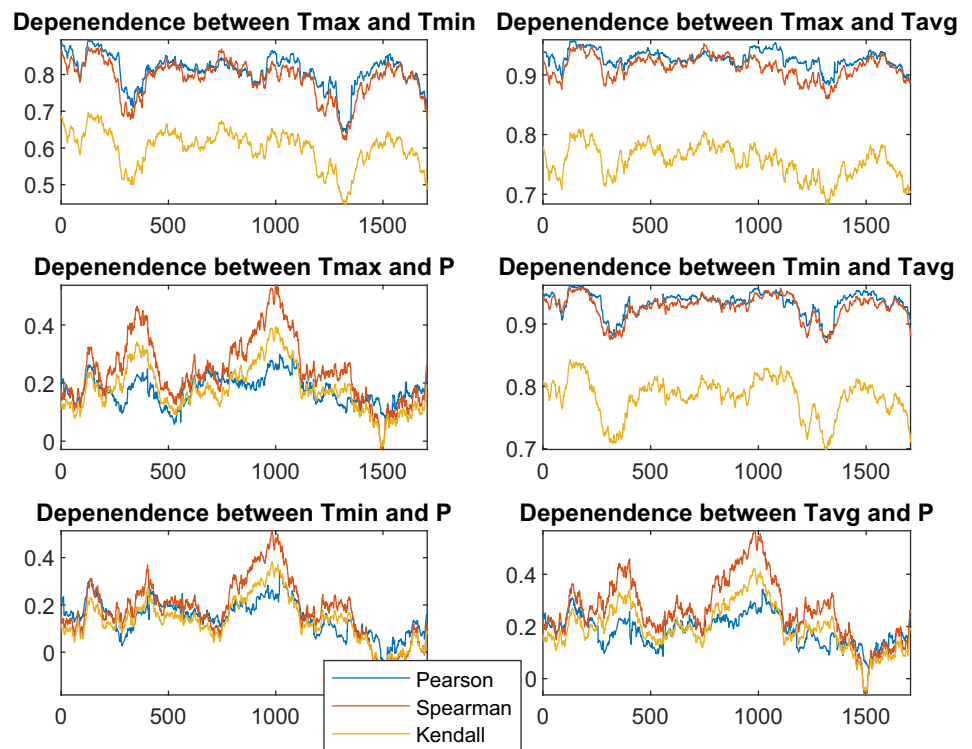

Table 1 Results of the normality and stability test statistics of non-Gaussian VAR(1) residuals for daily maximum temperatures

Month	H0:Residuals are α-stable distributed					H0:Residuals are Gaussian distributed				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
1	1.29	1.67	0.20	0.14	1.45	2.26	4.21	1.86	1.79	12.19
2	1.03	1.45	0.26	0.10	2.78	1.98	3.73	1.32	1.25	8.83
3	1.55	2.32	0.53	0.36	Inf	2.23	3.40	0.85	0.84	6.91
4	1.13	1.90	0.22	0.22	1.49	0.99	1.76	0.16	0.15	1.65
5	1.13	2.01	0.27	0.27	1.63	1.27	2.20	0.31	0.28	2.06
6	1.38	2.05	0.35	0.32	1.75	1.42	2.34	0.39	0.36	2.16
7	1.32	2.34	0.34	0.32	1.85	1.27	2.27	0.31	0.30	1.80
8	0.58	0.91	0.06	0.04	0.40	1.31	1.99	0.34	0.25	2.52
9	1.34	2.04	0.45	0.21	3.14	0.99	1.96	0.31	0.19	3.05
10	1.53	1.94	0.47	0.26	2.33	1.39	2.07	0.42	0.29	2.75
11	1.10	1.58	0.29	0.17	1.41	3.04	5.04	2.97	2.33	18.25
12	0.95	1.63	0.17	0.14	1.80	2.57	4.50	2.42	2.16	Inf

and in Fig. 7 for December. The coefficients were calculated for all combinations of the considered data.

It can be clearly seen that the Spearman and Pearson coefficients coincide for the temperature data. The Kendall coefficient is usually lower due to the way it is calculated. This indicates a linear relationship between the data. The situation is different for the relationship between the temperature and precipitation data. In this case, we observe large discrepancies between the coefficients, which indicates a nonlinear relationship and probably non-Gaussian behavior. An important observation is the fact that all considered data are significantly correlated with each other, which justifies the consideration of multivariate data modeling using VAR-type models. On the other hand, the

nonlinear relationship between the data also indicates that considering a model other than Gaussian may be justified.

4 Model fitting

As a set of four time series was analyzed and the correlation coefficients between all data are significant, an attempt was made to fit the 4-dimensional VAR(1) model, with α-stable innovations, described in the previous section, to the data under consideration. The VAR(1) model was used as the optimal according to multiple literature positions analysing this type of data [2, 8, 9].

Due to the fact that in the case of the daily sum of precipitation, the VAR model based on maximum, minimum, and average temperatures cannot reflect the exact values and the methodology of forecasting precipitation has a different nature [2, 8, 9] (which does not change the fact that precipitation has a significant impact on daily temperatures), only the estimation of the model and forecasts for maximum and minimum temperature is considered. The two remaining variables will be treated as auxiliary variables that have a significant impact on the quality of the forecast

As the relationships between 4-dimensional data are not always linear, the residuals distribution of the fitted non-Gaussian 4-dimensional VAR(1) model was investigated by testing whether these distributions are Gaussian or α-stable and fitting the $\alpha-$stable distribution. In the Tables 1 and 3, we present the values of the five statistics used for testing Gaussian and/or α-stable distributions of the residuals, namely Kolmogorov-Smirnov test (T1), Kuiper test (T2), Watson test (T3), Cramer-von Mises test (T4), and Anderson-Darling test (T5), see [73]. In Tables 2 and 4, we can see the parameters of fitted to residuals series the α-stable distribution. The results are presented for the minimum and maximum temperature.

In the case of daily maximum temperature, the residuals distributions of the fitted 4-dimensional non-Gaussian VAR (1) model are not Gaussian for all considered months. The results of the stability test statistics show that the residuals distribution is much closer to α-stable distribution which indicates a justified use of a non-Gaussian model. In Table 2, we can observe that the residuals for maximum temperature, especially for winter months, have heavier tails than in Gaussian distribution. Additionally, for most of the analysed months, the residuals β parameter is significantly positive, which indicates the right skewness of this distribution. Thus, in the case of maximum temperatures, we will observe the extreme values more often, with particular emphasis on the values with high maximum temperature.

Table 2 Fitted α-stable distribution parameters for residuals of daily maximum temperatures

Month	α	σ	β	μ
1	1.79	1.65	0.85	0.23
2	1.80	1.75	1.00	0.53
3	1.88	2.10	1.00	0.48
4	1.97	2.23	− 0.14	0.47
5	1.97	2.06	− 1.00	0.45
6	1.97	2.04	− 1.00	0.42
7	1.99	1.94	− 1.00	0.39
8	1.94	2.02	1.00	0.50
9	1.97	2.12	1.00	0.50
10	1.96	2.10	1.00	0.34
11	1.78	1.67	1.00	0.30
12	1.77	1.62	0.91	0.22

Table 3 Results of the normality and stability test statistics of non-Gaussian VAR(1) residuals for daily minimum temperatures

Month	H0:Residuals are α-stable distributed					H0:Residuals are Gaussian distributed				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
1	1.26	1.73	0.34	0.16	Inf	3.08	5.54	3.28	2.86	19.82
2	1.43	1.71	0.38	0.19	Inf	2.65	4.81	2.75	2.32	17.45
3	0.83	1.09	0.13	0.05	0.76	1.79	3.03	1.23	0.91	8.76
4	0.83	1.38	0.12	0.12	0.81	0.89	1.32	0.11	0.09	1.24
5	0.77	1.23	0.12	0.09	0.80	0.69	1.16	0.10	0.07	0.96
6	1.36	2.13	0.49	0.26	2.69	1.31	2.32	0.38	0.25	2.59
7	1.13	1.61	0.18	0.09	1.34	0.95	1.50	0.11	0.07	1.09
8	0.72	1.39	0.10	0.09	0.67	1.03	1.81	0.16	0.15	1.30
9	0.83	1.66	0.16	0.16	1.11	0.77	1.51	0.14	0.14	1.11
10	0.97	1.68	0.19	0.17	1.41	1.26	1.82	0.18	0.18	2.25
11	1.07	1.71	0.27	0.24	1.52	1.42	2.70	0.57	0.55	4.09
12	2.05	2.54	1.02	0.47	Inf	2.34	4.20	1.58	1.42	11.41

Table 4 Fitted α-stable distribution parameters for residuals of daily minimum temperatures

Month	α	σ	β	μ
1	1.74	1.82	− 1.00	− 0.38
2	1.75	1.75	− 1.00	− 0.32
3	1.89	1.80	− 1.00	− 0.41
4	1.96	1.75	− 1.00	− 0.35
5	1.98	1.64	− 1.00	− 0.41
6	1.98	1.53	− 1.00	− 0.33
7	1.99	1.44	− 1.00	− 0.30
8	1.95	1.49	− 0.62	− 0.41
9	1.99	1.78	1.00	− 0.38
10	1.94	1.82	0.15	− 0.29
11	1.90	1.81	− 0.41	− 0.23
12	1.83	1.81	− 1.00	− 0.42

Table 5 MAE of one day ahead prediction for daily maximum temperatures in case of non-Gaussian, 4-dimensional VAR(1) model and one-dimensional AR(1) model

Month	MAE 4-dimensional VAR(1)	MAE AR(1)
1	1.80	1.97
2	2.16	2.19
3	2.40	2.47
4	2.75	2.91
5	2.31	2.54
6	2.41	2.61
7	2.06	2.31
8	2.19	2.34
9	2.21	2.30
10	2.26	2.23
11	1.87	1.97
12	1.91	2.20

In the case of minimum temperature, the lack of Gaussianity of the residuals distributions is not as visible as in the case of maximum temperature. In some months, statistics for normality are lower than for stability. However, generally we can assume that the residuals distribution is closer to α- stable than to Gaussian. As in the case of maximum temperatures, the fitted α factor of α-stable distribution indicates heavier tails in winter months, and a negative β for almost all months indicates that special attention should be paid to the more frequent low values of minimum temperature.

4.1 Prediction

Based on the fitted model, the one-step prediction was made (one day ahead) for the minimum temperature and for the maximum temperature. The mean absolute error (MAE) measure, which is appropriate for non-Gaussian considerations, was used to measure the forecast error [74]

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|t_{pred} - t_{obs}|. \tag{4}$$

where t_{pred} means the predicted and t_{obs} the observed value. The mean absolute forecast error was determined for the last ten years, i.e., for the period 2010–2020. The forecast errors obtained in this way were compared with the one-dimensional model of autoregression of the first-order AR(1) with α-stable innovations for both minimum and maximum temperature separately. The forecast quality results are presented in Tables 5 and 6.

The results presented in Tables 5 and 6 show that the multivariate model has an advantage over the one-dimensional model in the case of forecasting the daily maximum and minimum temperature values. However, the quality of

the forecast for the minimum temperature is much better. The mean absolute forecast error for the minimum temperature for all months is about 0.5 degrees Celsius lower than for the univariate model, while for the maximum temperature, the difference is between 0.15 and 0.3. However, from the prognostic point of view, considering multivariate models, in which the correlation between the series under consideration is taken into account, is absolutely justified. In addition, it is also justified to use models with a nonlinear correlation structure and no assumption of Gaussianity.

Due to the fact that in the case of minimum temperature, the quality of the forecast is much better, and an attempt was made to estimate the future minimum temperature values for the year 2090, assuming the RCP 8.5 scenario [1, 75]. In the RCP 8.5 scenario, it is assumed that the warming of the climate related to the content of greenhouse gases will continue as before. The determined values were presented with confidence intervals and compared with the forecast values in the Klimada project [76].

The calculations were made for 10, 000 trajectories simulated on the basis of the estimated 4-dimensional VAR(1) model using the Monte Carlo methodology for each considered month. Figure 8 presents a chart showing the scenario for the year 2090. In the figure, we can see the monthly average minimum temperature obtained in the Klimada project and using the 4-dimensional VAR(1) model and the 99% and 95% confidence intervals from the VAR model.

It is worth noting that the forecast based on a simple VAR model, in the case of the forecast for 2090 of the average monthly minimum temperature, assuming the RCP 8.5 scenario, gives results not much different from those from the Klimada project, where the model used is

Table 6 MAE of one day ahead prediction for daily minimum temperatures in case of non-Gaussian, 4-dimensional VAR(1) model and one-dimensional AR(1) model

Month	MAE 4-dimensional VAR(1)	MAE AR(1)
1	2.00	2.68
2	2.27	2.79
3	2.13	2.46
4	1.95	2.42
5	1.91	2.47
6	1.72	2.19
7	1.60	1.98
8	1.64	2.33
9	1.91	2.43
10	2.02	2.57
11	1.93	2.54
12	2.23	2.83
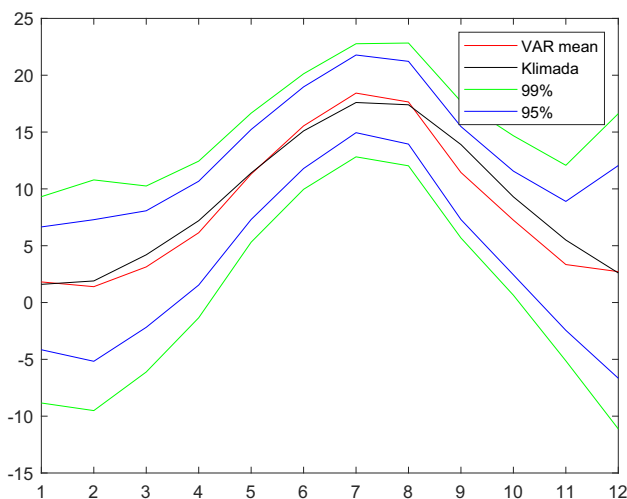
Fig. 8 The monthly average minimum temperature assuming RCP 8.5 scenario for 2090 according to 4-dimensional VAR(1) model and Klimada project plus 99% and 95% confidence intervals from the VAR model

certainly much more sophisticated. What is essential is the fact that the confidence intervals are shifted downwards against the mean (the distance between the lower limit of the confidence interval and the mean is in most cases greater than the distance between the mean and the upper limit of the interval). This asymmetry is related to the fact that non-Gaussian residuals are used for projection (thicker tails and left asymmetry). This affects the quality of the projection and the further information that the projection carries with it. As for the distribution of the structure of minimum temperatures, it is also worth noting that despite the asymmetry of the confidence intervals and a slightly lower average minimum temperature in the winter months, the percentage of days with a minimum temperature below

$0°C$ in the VAR model is 10%, while in the Klimada project, it is about 12.6%, which indicates a certain narrowing of the confidence intervals. Projection using the VAR model shows that in the RCP 8.5 scenario, we can expect, for example, daily minimum temperatures in December even around $16°C$, which clearly indicates the already observed upward trend.

5 Conclusions

In this paper, we have analyzed multidimensional climate data related to temperature (minimum, maximum, and averaged) and total precipitation. We have demonstrated the examined time series are related, and the structure of their relation changes over time. Moreover, we analysis of three different correlation coefficients clearly indicates their non-Gaussian behavior. The preliminary research was the starting point for the proposition of the stochastic model that shares similar properties as the data. More precisely, we have proposed the multidimensional VAR time series with non-Gaussian $\alpha-$stable innovations. This model was previously used by the authors for financial data modelling. We have shown that the proposed model more precisely reflects the nature of the data, and thus, it is justified to describe the time series by using non-Gaussian systems. Moreover, finally we have demonstrated that the model-based prediction corresponds to the RCP 8.5 scenario-based results. The proposed approach is universal and can be used in a more general case, e.g., to describe climate data corresponding to more variables. One of the main goals was to prove that the non-Gaussian multidimensional model is adequate to the analyzed data. This objective of this paper is fulfilled, and the obtained results clearly confirm this.

References

1. van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque,

J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J., Rose, S.K.: The representative concentration pathways: an overview. Climatic Change, **109**(5), (2011)

2. Kuchar, L., Broszkiewicz-Suwaj, E., Iwanski, S., Jelonek, L.: Comparison of daily flows simulated for the year 2060 on the kaczawa river for various scenarios of climate change by simple time series analysis. E3S Web of Conferences, **100**, (2019)

3. Boi, P.: A statistical method for forecasting extreme daily temperatures using ecmwf 2-m temperatures and ground station measurements. Meteorol. Appl. **11**, 245–251 (2004)

4. Hengl, T., Heuvelink, G., Perčec-Tadići, M., Pebesma, E.: Spatio-temporal prediction of daily temperatures using time-series of modis lst images. Theoret. Appl. Climatol. **107**(1), 265–277 (2012)

5. Murat, M., Malinowska, I., Gos, M., Krzyszczak, J.: Forecasting daily meteorological time series using arima and regression models. Int. Agrophys. **32**, 253–264 (2018)

6. Tran, T.T.K., Bateni, S.M., Ki, S.J., Vosoughifar, H.: A review of neural networks for air temperature forecasting. Water **13**, 1294 (2021)

7. Allard, D., Ailliot, P., Monbet, V., Naveau, P.: Stochastic weather generators: An overview of weather type models. J. de la Société Française de Statistique **156**, 101–113 (2015)

8. Katz, R.: Precipitation as a chain-dependant process. J. Appl. Meteorol. **16**, 671–676 (1977)

9. Richardson, C.: Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res. **17**(1), 182–190 (1981)

10. Parlange, M., Katz, R.: An extended version of the richardson model forsimulating daily weather variables. J. Appl. Meteorol. **39**, 610–622 (2000)

11. Flecher, C., Naveau, P., Allard, D., Brisson, N.: A stochastic daily weather generator for skewed data. Water Resour. Res. **46**(7), 7519 (2010)

12. Gupta, A., Gonzalez-Farias, G., Dominguez-Molina, J.: A multivariate skew-normal distribution. J. Multivar. Anal. **89**(1), 181–190 (2004)

13. Brockwell, P.J., Davis, R.A.: Introduction to Time Series and Forecasting. Springer, Berlin (2016)

14. Lévy, P.: Calcul des Probabilites. Gauthier-Villars, Paris (1925)

15. Lévy, P.: Théorie des erreurs. la loi de gauss et les lois exceptionnelles. Bull. Soc. Math. France **52**, 49–85 (1924)

16. Mandelbrot, B.: The Pareto-Lévy Law and the distribution of income. Int. Econ. Rev. **1**(2), 79–106 (1960)

17. Janczura, J., Orzeł, S., Wyłomańska, A.: Subordinated α-stable Ornstein–Uhlenbeck process as a tool for financial data description. Physica A **390**(23–24), 4379–4387 (2011)

18. McCulloch, J.H.: 13 Financial applications of stable distributions In: Statistical Methods in Finance, ser. Handbook of Statistics. Elsevier, Vol. 14, pp. 393–425 (1996)

19. Jin, H.J.: Heavy-tailed behavior of commodity price distribution and optimal hedging demand. J. Risk Insur. **74**(4), 863–881 (2007)

20. Hall, J.A., Brorsen, W., Irwin, S.H.: The distribution of futures prices: A test of the stable paretian and mixture of normals hypotheses. J. Financ. Quant. Anal. **24**(1), 105–116 (1989)

21. Cornew, R.W., Town, D.E., Crowson, L.D.: Stable distributions, futures prices, and the measurement of trading performance. J. Futur. Mark. **4**(4), 531–557 (1984)

22. Kateregga, M., Mataramvura, S., Taylor, D.: Parameter estimation for stable distributions with application to commodity futures log-returns. Cogent Econ. Finance **5**(1), 1318813 (2017)

23. Xu, W., Wu, C., Dong, Y., Xiao, W.: Modeling chinese stock returns with stable distribution. Math. Comput. Model. **54**(1), 610–617 (2011)

24. Bendler, J.: Lévy (stable) probability densities and mechanical relaxation in solid polymers. J. Stat. Phys. **36**, 625–637 (1984)

25. Weron, A., Weron, K.: Stable measures and processes in statistical physics. In: Probability in Banach Spaces V. Lecture Notes in Mathematics. Springer, Vol 1153, pp. 440–452 (1985)

26. Lan, B.L.: Bohm's quantum-force time series: Stable distribution, flat power spectrum, and implication. Phys. Rev. A **63**(4), 042105 (2001)

27. Kincses, D.: Shape analysis of HBT correlations at STAR. Phys. Part. Nucl. **51**, 267–269 (2020)

28. Nikias, C., Shao, M.: Signal processing with alpha-stable distributions and applications, ser. Signal Processing, Learning, Communications and Control. Wiley, New York, Adaptive and Cognitive Dynamic Systems (1995)

29. Stuck, B.W.: An historical overview of stable probability distributions in signal processing. In: 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100), vol. 6, pp. 3795–3797 (2000)

30. Jaoua, N., Duflos, E., Vanheeghe, P., Clavier, L., Septier, F.: Joint estimation of state and noise parameters in a linear dynamic system with impulsive measurement noise: Application to OFDM systems. Digital Signal Process. **35**, 21–36 (2014)

31. Yang, Xueshi, Petropulu, A.P.: Long-range dependent alpha-stable impulsive noise in a Poisson field of interferers. In: Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing (Cat. No.01TH8563), pp. 54–57 (2001)

32. Misra, N., Kuruoglu, E.E.: Stable graphical models. J. Mach. Learn. Res. **17**(168), 1–36 (2016)

33. Indyk, P.: Stable distributions, pseudorandom generators, embeddings, and data stream computation. J. ACM **53**(3), 307–323 (2006)

34. Simmross-Wattenberg, F., Asensio-Perez, J.I., Casaseca-de-la Higuera, P., Martin-Fernandez, M., Dimitriadis, I.A., Alberola-Lopez, C.: Anomaly detection in network traffic based on statistical inference and α-stable modeling. IEEE Trans. Dependable Secure Comput. **8**(4), 494–509 (2011)

35. Obuchowicz, A.K., Smołka, M.: Application of α-stable mutation in a hierarchic evolutionary inverse solver. J. Comput. Sci. **17**(P1), 261–269 (2016)

36. Painter, S., Beresford, G., Paterson, L.: On the distribution of seismic reflection coefficients and seismic amplitudes. Geophysics **60**(4), 1187–1194 (1995)

37. Lavallée, D., Archuleta, R.. J.: Stochastic modeling of slip spatial complexities for the:: Imperial Valley, California, earthquake. Geophys. Res. Lett. **30**(5), 2003 (1979)

38. Chave, A.D.: Magnetotelluric data, stable distributions and impropriety: an existential combination. Geophys. J. Int. **198**(1), 622–636 (2014)

39. Salas-Gonzalez, D., Kuruoglu, E.E., Ruiz, D.P.: Modelling and assessing differential gene expression using the alpha-stable distribution. Int. J. Biostat. **5**(1), 1–16 (2009)

40. Crato, N., Linhares, R.R., Lopes, S.R.: α-stable laws for noncoding regions in DNA sequences. J. Appl. Stat. **38**(2), 261–271 (2011)

41. Lan, B.L., Toda, M.: Fluctuations of healthy and unhealthy heartbeat intervals. Europhys. Lett. **102**(1), 18002 (2013)

42. Lan, B.L., Chandran, P.: Distribution of animal population fluctuations. Physica A **390**(7), 1289–1294 (2011)

43. Boucharel, J., Dewitte, B., Garel, B., du Penhoat, Y.: Enso's non-stationary and non-gaussian character: the role of climate shifts. Nonlinear Process. Geophys. **16**(4), 453–473 (2009)

44. Joelson, M., Golder, J., Beltrame, P., Néel, M.-C., Di Pietro, L.: On fractal nature of groundwater level fluctuations due to rainfall process. Chaos, Solitons Fractals **82**, 103–115 (2016)

45. Ditlevsen, P.: Observation of α-stable noise induced millennial climate changes from an ice-core record. Geophys. Res. Lett. **26**(10), 1441–1444 (1999)

46. Ditlevsen, P.D.: Anomalous jumping in a double-well potential. Phys. Rev. E **60**, 172–179 (1999)

47. Shao, M., Nikias, C.L.: Signal processing with fractional lower order moments: Stable processes and their application. Proc. IEEE **81**, 986–1010 (1993)

48. Nolan, J.P., Panorska, A.K.: Data analysis for heavy tailed multivariate samples. Commun. Stat. Stochastic Models **13**(4), 687–702 (1997)

49. Grzesiek, A., Sundar, S., Wyłomańska, A.: Fractional lower order covariance-based estimator for bidimensional AR(1) model with stable distribution. Int. J. Adv. Eng. Sci. Appl. Math. **11**, 217–229 (2019)

50. Grzesiek, A., Giri, P., Sundar, S., Wyłomańska, A.: Measures of cross-dependence for bidimensional periodic AR(1) model with alpha-stable distribution. J. Time Ser. Anal. **41**(6), 785–807 (2020)

51. Grzesiek, A., Teuerle, M., Wyłomańska, A.: Cross-codifference for bidimensional VAR(1) time series with infinite variance. Communications in Statistics - Simulation and Computation, pp. 1–26 (2019). https://doi.org/10.1080/03610918.2019.1670840

52. Bielak, L., Grzesiek, A., Janczura, J., Wyłomańska, A.: Market risk factors analysis for an international mining company. multi-dimensional heavy-tailed-based modelling. (2021). arxiv, vol. arXiv:2107.07142

53. Samorodnitsky, G., Taqqu, M.: Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. Chapman and Hall, London (1994)

54. Nolan, J.P.: Multivariate Stable Distributions: Approximation, Estimation, Simulation and Identification, pp. 509–525. Birkhauser Boston Inc., Cambridge (1998)

55. Press, S.: Multivariate stable distributions. J. Multivar. Anal. **2**(4), 444–462 (1972)

56. Paulauskas, V.: Some remarks on multivariate stable distributions. J. Multivar. Anal. **6**(3), 356–368 (1976)

57. Weron, A.: Stable processes and measures a survey. In: Szynal, D., Weron, A. (eds.) Probability Theory on Vector Spaces III, pp. 306–364. Springer, Berlin (1984)

58. Zolotarev, V.: One-dimensional Stable Distributions, ser. Translations of mathematical monographs. American Mathematical Society (1986)

59. Janicki, A., Weron, A.: Simulation and Chaotic Behavior of α-stable Stochastic Processes. Chapman& Hall/CRC Pure and Applied Mathematics, Taylor & Francis, UK (1993)

60. Nowicka-Zagrajek, J., Wyłomańska, A.: Measures of dependence for stable AR(1) models with time-varying coefficients. Stoch. Model. **24**(1), 58–70 (2008)

61. Nowicka, J.: Asymptotic behavior of the covariation and the codifference for ARMA models with stable innovations. Commun. Stat. Stochastic Models **13**(4), 673–685 (1997)

62. Nowicka-Zagrajek, J., Wylomanska, A.: The dependence structure for parma models with alpha-stable innovations. Acta Physica Polonica **37**(1), 3071–3081 (2006)

63. Grzesiek, A., Mrozińska, M., Giri, P., Sundar, S., Wyłomańska, A.: The covariation-based Yule-Walker method for multidimensional autoregressive time series with α−stable distributed noise. Submitted, pp. 1–33 (2021)

64. Hu, Y., Long, H.: Least squares estimator for Ornstein–Uhlenbeck processes driven by α-stable motions. Stochastic Processes Appl. **119**(8), 2465–2480 (2009)

65. Kruczek, P., Wyłomańska, A., Teuerle, M., Gajda, J.: The modified Yule-Walker method for alpha-stable time series models. Phys. A **469**, 588–603 (2017)

66. https://danepubliczne.imgw.pl

67. https://climate.nasa.gov/resources/global-warming-vs-climate-change

68. Pozzi, F., Matteo, T.D., Aste, T.: Exponential smoothing weighted correlations. The European Physical Journal B

69. Dunn, O.J., Clark, V.A.: Basic Statistics: A Primer for the Biomedical Sciences. Wiley, New York (2009)

70. Kendall, M., Gibbons, J.D.: Rank Correlation Methods. Charles Griffin Book Series, 5th edn. Oxford University Press, Oxford (1990)

71. Daniel, W.W.: Kendall's tau. Applied Nonparametric Statistics, 2nd edn. PWS-Kent, Boston (1990)

72. Kendall, M.G.: A new measure of rank correlation. Biometrika **30**(1/2), 81–93 (1938)

73. Pavel Čížek, R.W., Haerdle, W.: Statistical Tools for Finance and Insurance. Spinger (2005)

74. Willmott, C.J., Matsuura, K.: Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. Climate Res. **30**, 79–82 (2005)

75. Riahi, K., Rao, S., Krey, V., Cheolhung Cho, V.C., Fischer, G., Kindermann, G., Nakicenovic, N., Rafaj, P.: Rcp 8.5—a scenario of comparatively high greenhouse gas emissions. Climatic Change **109**(33), 79–82 (2011)

76. https://klimada2.ios.gov.pl/klimat-scenariusze/