

Quantitative Characteristics of Correlations of Meteorological Data

R. RAK^a AND S. BWANAKARE^b

^aFaculty of Mathematics and Natural Sciences, University of Rzeszów, Rzeszów, Poland

^bUniversity of Information Technology and Management (WSIZ), Rzeszów, Poland

This paper presents the quantitative characteristics of correlations (and cross-correlations) of plant main eco-factors i.e. the ground and over-ground temperature, the wind speed, and the humidity. The study is based upon hourly data statistical observations collected in the region of Lublin, in Poland for the period 2001.05.07–2009.04.10. This paper indicates that plant growth conditions constitute an emergent response to the above direct eco-factors. Then, the dynamics properties of each eco-factor is first analyzed alone for its multifractal structure. We apply the multifractal detrended correlation analysis and multifractal detrended cross-correlation analysis. We show that the widest multifractal spectrum is for over-ground temperature and the strongest power-law cross-correlations exist between ground and over-ground temperature. Next, an impulse response analysis is carried out to measure dynamical inter causalities within all the considered variables. As far as cross-impact between different eco-variables is concerned, one observes that the wind speed, the ground temperature and the air humidity dynamics are the most influenced, in terms of memory length time, by external temperature.

DOI: [10.12693/APhysPolA.129.922](https://doi.org/10.12693/APhysPolA.129.922)

PACS/topics: 89.75.-k, 89.75.Da, 89.65.Gh, 02.70.Rr

1. Multifractal analysis

1.1. Methodology

In recent years investigation of complex systems with regard to their fractal properties has become one of the elementary methods of such systems analysis [1]. Multifractal structures were identified in systems from various areas such as physics [2–5], biology [6–8], chemistry [9, 10], economics [11–13] and even music [14–19]. One of the most popular methods of the multifractal analysis is multifractal detrended correlation (MFDFA) [20–22] and cross-correlation analysis (MFCCA) [23, 24]. In this paper, we systematically study the multifractal characteristics of meteorological data.

The MFDFA procedure was proposed by Kantelhardt et al. in [20] and is one of the most frequently applied algorithms of calculating the multifractal spectra. Its popularity owes to, among other, the simplicity of its implementation and the reliability of obtained results in the case of non-stationary time series. The generalization of the MFDFA method for two time series is MFCCA. This procedure was proposed by Oświęcimka et al. in [23].

Let consider two time series $x(i)_{i=1,\dots,N}$ and $y(i)_{i=1,\dots,N}$ divided into $2M_s$ separate boxes v of length s . In each box v a local trend of an integrated signal is approximated by an m -th degree polynomial $Q^{(m)}$. Next, according to formula, this trend is subtracted

$$X_v(s, i) = \sum_{j=1}^i x(vs + j) - Q_{X,s,v}^{(m)}(j),$$

$$Y_v(s, i) = \sum_{j=1}^i y(vs + j) - Q_{Y,s,v}^{(m)}(j).$$

Variance of the detrended data is calculated inside each segment

$$f_{XY}^2(s, v) = \frac{1}{s} \sum_{j=1}^s X_v(s, i) Y_v(s, i).$$

Finally, the q -th-order fluctuation function is expressed by the following formulae:

$$F_{XY}^q(s) = \frac{1}{2M_s} \sum_{v=0}^{2M_s-1} \text{sgn} [f_{XY}^2(s, v)] |f_{XY}^2(s, v)|^{q/2},$$

and

$$F_{qXY}(s) = \text{sgn} [F_{XY}^q(s)] |F_{XY}^q(s)|^{1/q}.$$

Moreover, for $q = 2$ the MFCCA procedure consistently transforms itself into the basic detrended cross-correlation analysis (DCCA) [25].

A (multi-) fractal character of the cross-correlations is related to a power-law scaling

$$F_{XY}^q(s) \propto s^{\delta(q)} \text{ and } F_{qXY}(s) \propto s^{\lambda(q)},$$

where $\delta(q)$, q and $\lambda(q)$ is an exponent that quantitatively characterizes fractal properties. When $\lambda(q) = \text{const}$ we observe monofractal cross-correlation. Otherwise, the analyzed data have a fractal nature of the correlation. Of course, the estimated (multi-) fractal cross-correlations are related to the fractal properties of the respective signals $x(i)$ and $y(i)$. Therefore, it is often used the average of the generalized Hurst exponents

$$h_{xy}(q) = \frac{h_x(q) + h_y(q)}{2},$$

where $h_x(q)$ and $h_y(q)$ refer to fractal properties of individual signals and for $q = 2$ correspond to the Hurst exponent H .

For MF DFA method, when $X_v(s, i) = Y_v(s, i)$, $F_{qXX}(s) \equiv F_q s^{h(q)}$, where $h(q)$ denotes the generalized Hurst exponent. For a monofractal signal, $h(q)$ is independent of q and equals the Hurst exponent $h(q) = H$. On the other hand, for a multifractal time series, $h(q)$ is decreasing function of q and the ordinary Hurst exponent is obtained for $q = 2$. The multifractal spectrum can be calculated by means of the following relation:

$$f(\alpha) = q[\alpha - h(q)] + 1,$$

where α denotes the strength of a singularity spectrum and $f(\alpha)$ is the fractal dimension of a points set with particular α . For a multifractal time series, the shape of the singularity spectrum is similar to a wide inverted parabola. The left and right wing of the parabola refers to the positive and negative values of q , respectively. The maximum of the spectrum is located at $\alpha(q = 0)$. For a monofractal signal, $f(\alpha)$ aims to the one point. The richness of the multifractal is evaluated by the width of its spectrum

$$\Delta\alpha = \alpha_{\max} - \alpha_{\min}.$$

1.2. Results

We start our study with an analysis of a time series represented by the one 1-hour meteorological data. T1 — it is the air temperature changes at a height of 2 m, T2 — it is the ground temperature changes. We are also considering changes of “wind speed” (W) and “humidity” (H) at the height 2 m. Examples of the time series can be seen in Fig. 1.

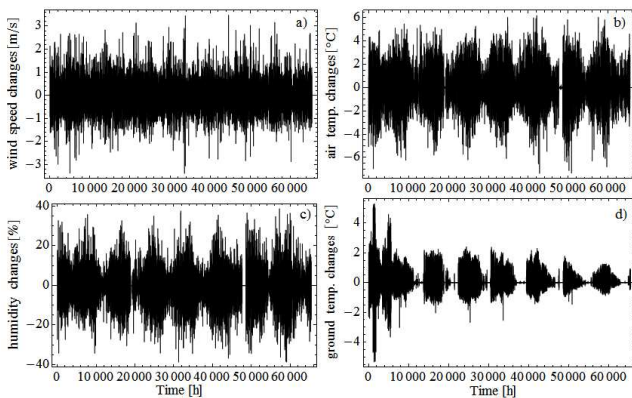


Fig. 1. Time series of meteorological data during the period of 2001.05.07–2009.04.10. Hourly changes of: (a) wind speed (W), (b) air temperature (T1), (c) humidity (H), (d) ground temperature (T2).

In all subsequent analyses we use the normalized data, i.e. data with a mean equal to zero and a standard deviation equal to 1. In order to determine the range of q values, which we will use for the multifractal analysis, we study cumulative probability distributions. The results are shown in Fig. 2. It is clear that the distributions, for all of the analyzed signals, have a power-law type tails ($x^{-\beta}$). The most extreme slope tails are observed

for temperature T1 and T2 (β value is respectively 6 and 3). Interestingly, $\beta = 3$ and $\beta = 4$ is a typical value observed among others in the high frequency financial data [11, 21, 26].

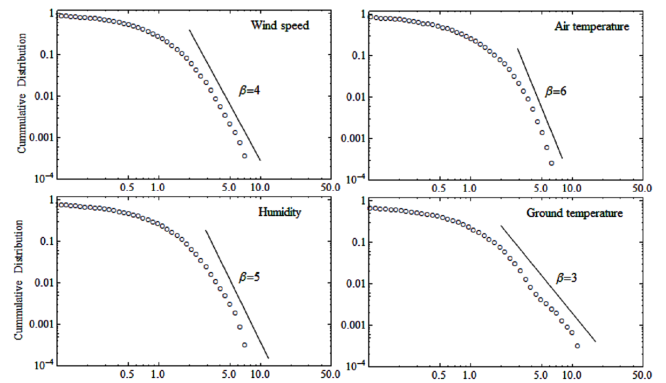


Fig. 2. Log–log plot of the cumulative distributions of absolute normalized changes of wind speed, air temperature, humidity, and ground temperature.

In order to avoid divergent moments due to fat tails in the distribution of the fluctuations, we restrict q to $\langle -4, 4 \rangle$ throughout this paper. Moreover, based on our own experience, as optimal we use a polynomial of order $m = 2$. In Fig. 3, we depict empirical singularity spectra $f(\alpha)$ for all considered data.

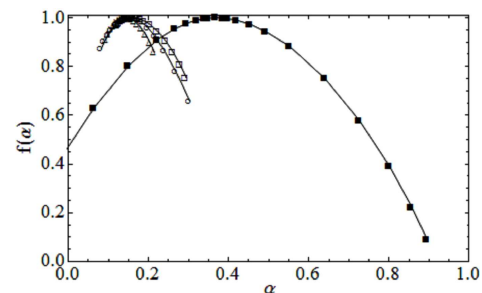


Fig. 3. The empirical singularity spectra $f(\alpha)$ for normalized changes of wind speed (\circ), air temperature (\triangle), humidity (\square) and ground temperature (\blacksquare).

It is clear that the singularity spectra estimated for each signals differ from each other. In all cases, we obtained relatively broad spectra which confirm that the analyzed meteorological time series are multifractal. The spectrum is widest for the case of T2 and its maximum is shifted to the right. Such a large width for T2 testifies to the fact that fluctuations of the temperature at a height of 2 m are much more multifractal than the ground temperature. The Hurst exponents (located in the vicinity of the maximum) estimated for wind speed, T1, H indicate stronger antipersistence than for T2.

To detect possible relationship between the analyzed meteorological data we study the multifractal cross-correlation. We consider the following cases of cross-correlation: W–H, T1–T2, W–T2, H–T2, W–T, H–T1. The obtained results are shown in Fig. 4.

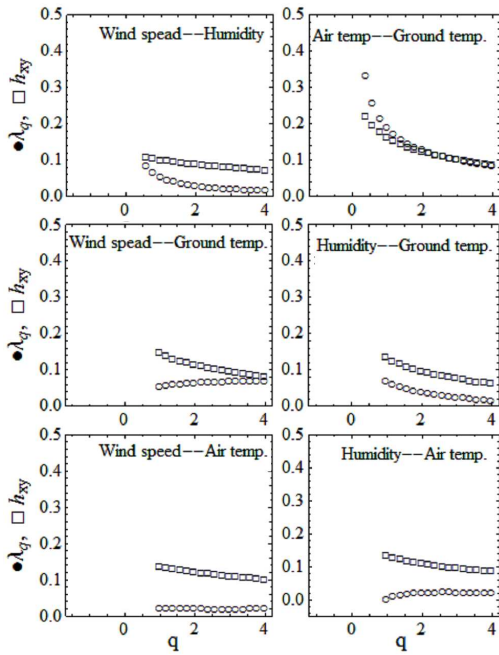


Fig. 4. The multifractal cross-correlation scaling exponents $\lambda(q)$ (\circ) and the average generalized Hurst exponents $h_{xy}(q)$ (\square) for normalized changes of W, T1, H, and T2.

There are shown the values obtained for $\lambda(q)$ (circles) and $h_{xy}(q)$ (squares). Due to the fact that in all cases $F_{qXY}(s)$ is well-defined for all scales only for q approximately greater than 0 (power law is fulfilled), while it is wildly unstable for $q \leq 0$ ($F_{qXY}(s)$ fluctuates from positive to negative values), results for $\lambda(q)$ and $h_{xy}(q)$ are shown only in the range of q (approximately) greater than 0. Of course, the spectrum of values of $h_{xy}(q)$ exist for all q but here they are presented only for those q for which the value $\lambda(q)$ exists. This means that the meteorological signals are fractal cross-correlated, but these correlations are restricted mainly to large fluctuations. This means that these data are uncorrelated noise at the level of small fluctuations. This type of effect has been also observed for financial data [24, 27]. Analyzing all pairs shown in Fig. 4 it can be seen that the strongest correlations are observed between air temperature (T2) and ground temperature (T1) — $\lambda(q)$ and $h_{xy}(q)$ almost overlap.

2. Eco-factors statistical interaction and impulse response analysis

2.1. The model

The last paragraph revealed different complexity levels through multi-fractal and cross-correlation analysis. Now we propose a next, complementary step of the ecosystem variable analysis. We point out the dynamic relationships for each of these variables thanks to the variance decomposition to discriminate the amount of information each variable contributes to the other variables in their

dynamic states. Sims* in [28] proposed a statistical tool called vector autoregression (VAR) to differentiate shocks that come about as a result of other shocks and those that occur independently through an impulse response (IR) function analysis (IRFA) [29, 30]. More generally, an IR refers to the reaction of any dynamic system in response to some external change. Thus, while this technique is largely applied in econometrics, this article extends its application to natural sciences. Traditionally, to carry out such an analysis, a VAR is built and followed by an IRFA through a variance decomposition.

Following [31–34] a discrete VAR model can be written as follows:

$$BX_t = \Gamma_0 D_t + \Gamma_1 X_{t-1} + \Gamma_2 X_{t-2} + \dots + \Gamma_k X_{t-k} + \xi_t, \quad (1)$$

or in a reduced form

$$X_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 X_{t-1} + B^{-1}\Gamma_2 X_{t-2} + \dots + B^{-1}\Gamma_k X_{t-k} + B^{-1}\xi_t,$$

where X_t is the vector of the current observations on the ecosystem variables of the model $X_t = [X_{1t}, X_{2t}, \dots, X_{nt}]$, D_t — the vector of exogenous variables of the system, Γ_j — the matrix of parameters related to the variables (exogenous and endogenous lagged variables), B — the matrix of parameters related to the endogenous present time variables, ξ_t — random term.

In our case n represents the four endogenous variables (the same as in paragraph I): T1, T2, W, and H. To allow for a better visualization of the IR mechanism, we rewrite (1) in vector moving average reduced form (MA):

$$X_t = \mu + \sum_{i=0}^{\infty} \varphi_i \xi_{t-i}, \quad (2)$$

where $\varphi_i = A_1^i B^{-1}$ and $A_i = B^{-1}\Gamma_i$ ($i = 1, \dots, k$). The value μ stands for a vector of average values of model variables. The time index i explains the lag degree [34] of variable autocorrelation. The coefficients φ_i is used to generate the impact of random shocks (impulse) on X_t .

In the case of the presently analyzed model, the reduced form can be presented as follows:

$$\begin{bmatrix} W \\ T1 \\ T2 \\ H \end{bmatrix} = \begin{bmatrix} \bar{W} \\ \bar{T1} \\ \bar{T2} \\ \bar{H} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \varphi_{11}(i) & \varphi_{12}(i) & \varphi_{13}(i) & \varphi_{14}(i) \\ \varphi_{21}(i) & \varphi_{22}(i) & \varphi_{23}(i) & \varphi_{24}(i) \\ \varphi_{31}(i) & \varphi_{32}(i) & \varphi_{33}(i) & \varphi_{34}(i) \\ \varphi_{41}(i) & \varphi_{42}(i) & \varphi_{43}(i) & \varphi_{44}(i) \end{bmatrix} \begin{bmatrix} \xi_{Wt-i} \\ \xi_{T1t-i} \\ \xi_{T2t-i} \\ \xi_{Ht-i} \end{bmatrix}. \quad (3)$$

Thus, elements of $\varphi_{jk}(0)$ stand for the direct impact

*The 2011 Nobel Prize winner in Economics.

of a shock. For instance $\varphi_{13}(0)$ explains the contemporary impact of one standard error unit change of the temperature T2 on the current W. Likewise $\varphi_{24}(1)$ estimates the last period impact of H on the current temperature T1. In empirical researches, it can be sometimes interesting to estimate the total fluctuations of a variable over a laps of time as a result of an external or an autogenerated shock. For example, a 200 h cumulative shock impact of T1 on W_t ($\varphi_{WT1(200)}$) can be written as follows:
$$\sum_{i=1}^{200} \varphi_{12}(i).$$

2.2. Outputs

Figure 5 shows the IR within all eco-variables presented in Eq. (3). The figure self-contains 16 (4×4) individual impact response figures. If we observe these figures by column, the first column relates the impact of a standard error unit of wind speed W_t impulse on each eco-variable, the rest of columns display the same impact respectively by T1, H, T2.

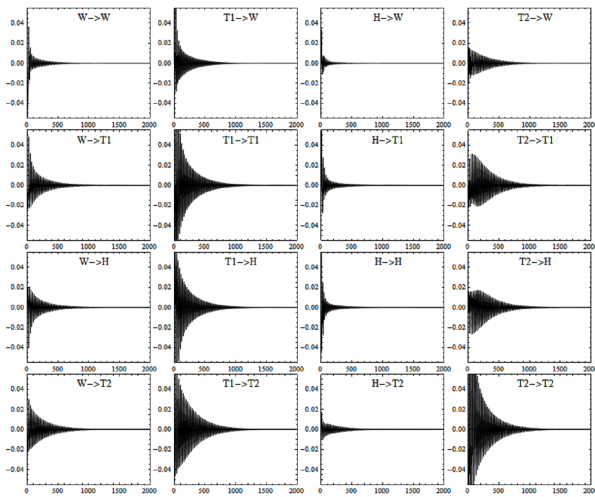


Fig. 5. The IRFA displaying the dynamic relationships between T1, T2, W, and H variables. The vertical axis displays the magnitude of a one standard error IR. The horizontal axis presents the time over which the IR signal evolves.

Each of the 16 figures presents the vertical axis and the horizontal axis. The former displays a response magnitude of the impacted variable by one unit standard error impulse from the targeting variable. The maximum IR fluctuation scale has been set on an interval length of $[-0.05, +0.05]$. Thus, for a one hundred percent unit standard error impulse the IR may display an oscillation between -5% and $+5\%$.

Next, these oscillations around zero decline over a time period (horizontal axis) arbitrarily limited to 2000 h in this study. This scale seems sufficient enough for a comparison of the IR generated by the four variables. As we can observe, figures on the diagonal describe a self-impact of each eco-variable. Then, as shown by the same Fig. 5,

a standard error unit impulse of W on itself is relatively more quickly amortized. H seems to show a similar dynamics. This suggests that the W_t or H auto-generated dynamics are relatively less lasting on the given time scale once compared to T1 and T2. Their impact on these same variables (T1 and T2) is relatively limited, too. This is not true in the case of both T1 and T2, where memory on their own self past is much longer lasting. Furthermore, as displayed in the second and the forth column, the IR generated by both temperatures present larger IR (at the beginning sometime larger than 5% with respect to a 100% of initial impulse) and more lasting impact. Seen at the considered scale, the IR of each of both T1 and T2 seems to last several hundred hours, as far as the cross-impact between the different eco-variables is H in comparison with the three remaining variables. In spite of the fact we have used the Cholecki triangular decomposition[†] model — thus excluding the contemporary mutual impact — we continue to observe an asymmetrical IRs (around zero) in the case of some eco-variables. In particular, the both temperatures impacts on each other display a higher level of asymmetry, suggesting the stronger complexity of their structures already pointed out in Sect. 1 devoted to multifractal analysis.

3. Summary and conclusions

In the present contribution, we focused our attention on a relationship between the four eco-factor variables (the ground and air temperature, the wind speed and the humidity) considered individually and collectively as a dynamic, complex system. The analysis of the probability distributions revealed relatively fat tails for all considered time series — fattest for changes of ground temperature.

An interesting observation done in the work is that all meteorological data are antipersistent — maxima of multifractal spectra are shifted to the left of $\alpha = 0.5$. Furthermore, the maximum for the ground temperature is the widest and shifted relatively closer to 0.5. Thus, the above results suggest that dynamics of changes of the ground temperature is the most complex from among all meteorological data considered herein.

Using the MFCCA method we showed that the ground and over-ground (air) temperature are most correlated and the correlations exist only for the large temperature fluctuations.

Next, the outputs from the IRFA, tend to confirm the above multifractal analysis results as shown in Fig. 5. The ground and the air temperature seem to be very correlated and their impact on other series is the biggest, which confirms their (possible) complex character.

[†]This suggests recursive shocks among the random errors of the model.

Acknowledgments

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge (University of Rzeszów).

References

- [1] J. Kwapien, S. Drozd, *Phys. Rep.* **515**, 115 (2012).
- [2] J.F. Muzy, E. Bacry, R. Baile, P. Poggi, *Europhys. Lett.* **82**, 60007 (2008).
- [3] P. Oświęcimka, J. Kwapien, S. Drozd, *Phys. Rev. E* **74**, 016103 (2006).
- [4] P. Oświęcimka, J. Kwapien, S. Drozd, R. Rak, *Acta Phys. Pol. B* **36**, 2447 (2005).
- [5] A.R. Subramaniam, L.A. Gruzberg, A.W.W. Ludwig, *Phys. Rev. B* **78**, 245105 (2008).
- [6] P.Ch. Ivanov, A.N. Amaral, A.L. Goldberger, Sh. Havlin, M.G. Rosenblum, Z.R. Struzik, H.E. Stanley, *Nature* **399**, 461 (1999).
- [7] D. Makowiec, A. Dudkowska, R. Galaska, A. Rynkiewicz, *Physica A* **388**, 3486 (2009).
- [8] A. Rosas, E. Nogueira Jr., J.F. Fontanari, *Phys. Rev. E* **66**, 061906 (2002).
- [9] H.E. Stanley, P. Meakin, *Nature* **335**, 405 (1988).
- [10] V.V. Udovichenko, P.E. Strizhak, *Theoret. Exp. Chem.* **38**, 259 (2002).
- [11] S. Drozd, J. Kwapien, P. Oświęcimka, R. Rak, *New J. Phys.* **12**, 105003 (2010).
- [12] P. Oświęcimka, J. Kwapien, S. Drozd, A.Z. Gorski, R. Rak, *Acta Phys. Pol. A* **114**, 547 (2008).
- [13] W.X. Zhou, *Europhys. Lett.* **88**, 28004 (2009).
- [14] R.F. Voss, J. Clarke, *Nature* **258**, 317 (1975).
- [15] K.J. Hsu, A.J. Hsu, *Proc. Natl. Acad. Sci. USA* **87**, 938 (1990).
- [16] G.R. Jafari, P. Pedram, L. Hedayatifar, *J. Stat. Mech.* **2007**, P04012 (2007).
- [17] Z.-Y. Su, T. Wu, *Physica D* **221**, 188 (2006).
- [18] P. Oświęcimka, J. Kwapien, I. Celińska, S. Drozd, R. Rak, [arXiv:1106.2902v1](https://arxiv.org/abs/1106.2902v1).
- [19] Z. Wu, N.E. Huang, S.R. Long, C.-K. Peng, *Proc. Natl. Acad. Sci. USA* **104**, 18889 (2007).
- [20] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, H.E. Stanley, *Physica A* **316**, 87 (2002).
- [21] S. Drozd, J. Kwapien, P. Oświęcimka, R. Rak, *Europhys. Lett.* **88**, 60003 (2009).
- [22] D. Makowiec, A. Fuliński, *Acta Phys. Pol. B* **41**, 1025 (2010).
- [23] P. Oświęcimka, S. Drozd, M. Forczek, S. Jadach, J. Kwapien, *Phys. Rev. E* **89**, 023305 (2014).
- [24] R. Rak, S. Drozd, J. Kwapien, P. Oświęcimka, *Europhys. Lett.* **112**, 48001 (2015).
- [25] B. Podobnik, H.E. Stanley, *Phys. Rev. Lett.* **100**, 084102 (2008).
- [26] R. Rak, S. Drozd, J. Kwapien, *Physica A* **374**, 315 (2007).
- [27] B. Podobnik, D. Horvatic, A.M. Petersen, H.E. Stanley, *Proc. Natl. Acad. Sci. USA* **106**, 22079 (2009).
- [28] C. Sims, *Econometrica* **48**, 1 (1980).
- [29] G. Koop, M.H. Pesaran, S.M. Potter, *J. Econom.* **74**, 119 (1996).
- [30] H. Hashem Pesaran, S. Yongcheol, *Econom. Lett.* **58**, 17 (1997).
- [31] D. Asteriou, S.G. Hall, *Applied Econometrics*, 2nd ed., Palgrave MacMillan, London 2011, p. 319.
- [32] W. Enders, *Applied Econometric Time Series*, 3rd ed., Wiley, New York 2010, p. 272.
- [33] C.W.J. Granger, *Econometrica* **37**, 424 (1969).
- [34] R.S. Hacker, A. Hatemi-J, *J. Appl. Statist.* **35**, 601 (2008).