Acta Sci. Pol.
Formatio Circumiectus 17 (3) 2018, 101–117
ENVIRONMENTAL PROCESSES

www.formatiocircumiectus.actapol.net/pl/ ISSN 1644-0765 DOI: http://dx.doi.org/10.15576/ASP.FC/2018.17.3.101

ORIGINAL PAPER Accepted: 31.08.2018

ANALYSIS OF GRANGER CAUSALITY BETWEEN DAILY AND MONTHLY TEMPERATURES OF WATER AND AIR, AS ILLUSTRATED WITH THE EXAMPLE OF NOTEĆ RIVER

Renata Graf✉

Department of Hydrology and Water Management, Institute of Physical Geography and Environmental, Adam Mickiewicz University in Poznań, ul. Bogumiła Krygowskiego 10, 61-680 Poznań

ABSTRACT

The study presents the results obtained following the analysis of the so-called Granger causality between daily and monthly temperatures of air and water for the period 1987–2013 carried out for the Noteć river and its two main tributaries: Drawa and Gwda. Granger causality relates to a situation where the data concerning past values of one time series provide important information helping to predict values of another series not included in the information about its past values. The analysis was based on the Granger causality test (of the first order). A causality relationship was established for daily temperature series both for the air-water and water-air directions of influence, which means that forecasting the pattern of river water temperatures from changes to air temperatures can yield better results when done based on data from the previous day. The model forecasting daily water temperature in the Noteć river on the basis of water and air temperatures from the previous day explained 0.07–0.27% of unique variance more than the model that used only water temperature from the previous day. The model forecasting the daily air temperature based on air and water temperatures from the previous day explained 0.3–0.79% of the variance more than the model, which uses only the air temperature from the previous day. For monthly series of water and air temperatures, different configurations of correlations in terms of Granger causality were established: one-way in water-air direction or no correlation, which may result from the river water thermal regime being disturbed by the local impact of anthropogenic factors. In addition, the analysed effect of Granger causality between series of random fluctuations of both temperature models confirmed that causal dependencies occur in both directions. The identification of causal relationships in terms of Granger causality confirms the influence of one data series on the evolution on another data series, and it defines the application potential of study results for the purpose of forecasting the changeability of thermal parameters of river waters. The obtained results may help improve the quality of forecasting changes in water thermal conditions, which is important for managing their environmental condition properly.

Keywords: water temperature, air temperature, Granger causality, unique variance, forecasting, Noteć river

INTRODUCTION

In connection with observed changes in air temperature and their impact on the temperature of streamflow waters, the need to collect and continuously record data on the variability of thermal parameters and the correlations between them is increasing. These correlations are confirmed in various spatial scales and at different levels of time resolution of the data, taking into account the nature of temperature fluctuations, from short-term to long-term (Sahoo et al. 2009, Toffolon and Piccolroaz 2015). The nature and strength

✉e-mail: rengraf@amu.edu.pl

of the correlation between water and air temperature measurement series and the synchronism of their fluctuations are analysed (Łaszewski 2014, Wiejaczka 2007). Statistical methods, which are used in the analysis of river water temperature, change from simple linear regression equations to more complex parametric and non-parametric methods (Letcher et al. 2016). To predict changes in river water temperature in the changing climatic conditions, both physical and data-based models are used. Physical models, however, require a large amount of information, often unavailable, while data-based models often overlook physical processes (Piotrowski and Napiórkowski 2018). Deterministic models use energy balance in this respect (Benyahya et al. 2007), while regression and stochastic models are based on air temperature data (Hilderbrand et al. 2014). These models are characterized by good predictability of water temperature changes in the aspect of climate change (Neumann et al. 2003, Lagergaard Pedersen and Sand-Jensen 2007). The majority of thermal energy exchange occurs at the air-water level, while the smaller dimension is at the level of riverbed-water (Caissie 2006). This aspect is also reflected in models developed as a function of heat transfer from groundwater and exchange in the hyporheic zone. The temperature of river waters is modified by the feeding type of the river and by environmental factors, including anthropogenic factors related to the way the catchment is used, and the manner and degree of the river valley development, as confirmed by a number of studies (Webb and Nobilis 2007, Wiejaczka 2007, Łaszewski 2014, Graf 2015).

In the analysis of the mutual interdependence of statistical features of water and air temperatures, a multi-level regression modelling is used in the form of a hierarchical linear model. The basis for its development is the division of measurement series into fixed and random effects, combined with a multi-level data structure including inter-level interaction and a specific approach to the components of variance (Letcher et al. 2016). Geostatistical models are applied in spatial analysis of river systems (Rushworth et al. 2015); also used is the concept of thermal equilibrium (Caissie et al. 2005) and regional conditions related to latitude (Gardner et al. 2003). An interesting approach in studying the relationship between river water and air temperatures is a hybrid model that combines the physical aspect of the parameters with their stochastic calibration. In many methodological approaches, forecasting the temperature of streamflow waters is carried out on the basis of various meteorological and hydrological variables, whilst data-controlled Artificial Neural Networks are also increasingly used (DeWeber and Wagner 2014, Napiórkowski et al. 2014). The selection of variables remains a significantproblem; and another is the selection of those hydrological and meteorological variables available from the Global Circulation Models, which would be the most important as input data for the artificial neural networks model. Piotrowski et al. (2015), among others, conducted the testing of various types of neural networks in this respect, and obtained the results of their research into two natural catchments (mountain and lowland) located in the temperate climate zone. In the opinion of the aforementioned authors, the best results are obtained when the mean, maximum and minimum daily air temperatures from the previous days are used as input data together with the current river runoff. Because neural networks do not have a physical background, they require a large number of parameters in order to be calibrated, which is usually done with the use of gradient algorithms; however, nowadays discussions about the suitability of metaheuristics for this type of research are on-going (Piotrowski et al. 2014, 2016).

The models used in recent decades in climatic and hydrological research are constantly being developed and expanded with new data. Data-based simulations directly model phenomena, while they also learn some input data sets. Application of information theory and methods of detecting causal relationships between various factors makes it easier to acquire still new information about the degree of their correlation at different intervals of time. The most commonly used data-driven approach to identifying cause-and-effect relationships consists of Granger causality modelling (Granger 1980), which refers to a situation in which past data of one time series provides important information for predicting the value of the other series that is not included in the information about its past values. Granger causality is a statistical approach used, especially in econometrics, to identify the presence of a linear causal interaction between time series of data based on the forecasting theory According to the concept of causality, in the time series the cause precedes the effect, and the series that is the cause contains information about the variable being the result that was not included in other series (Granger 1980, Syczewska 2014).

The paper presents the results of the causal relationship analysis, in the sense of Granger causality, between the daily and monthly temperatures of river water and of air, illustrated with the example of the Noteć river and its tributaries. The aim of the study was to verify whether the correlation between temperatures has a one- or two-way dimension, and whether it occurs between air-water and water-air, as well as to determine the conditions of this correlation. In the case of time series of water and air temperatures, the detection of causative patterns in variance, while using the causality analysis, in the sense of Granger causality, may serve to improve the quality of forecasts of these quantities.

STUDY AREA, SOURCE MATERIAL AND RESEARCH METHODOLOGY

Studies of the relationship between river water and air temperatures were carried out for Noteć, the right tributary of the Warta River, and its two tributaries: Gwda and Drawa (see: Fig. 1). According to the physico-geographical division of Poland (Kondracki 2008) the

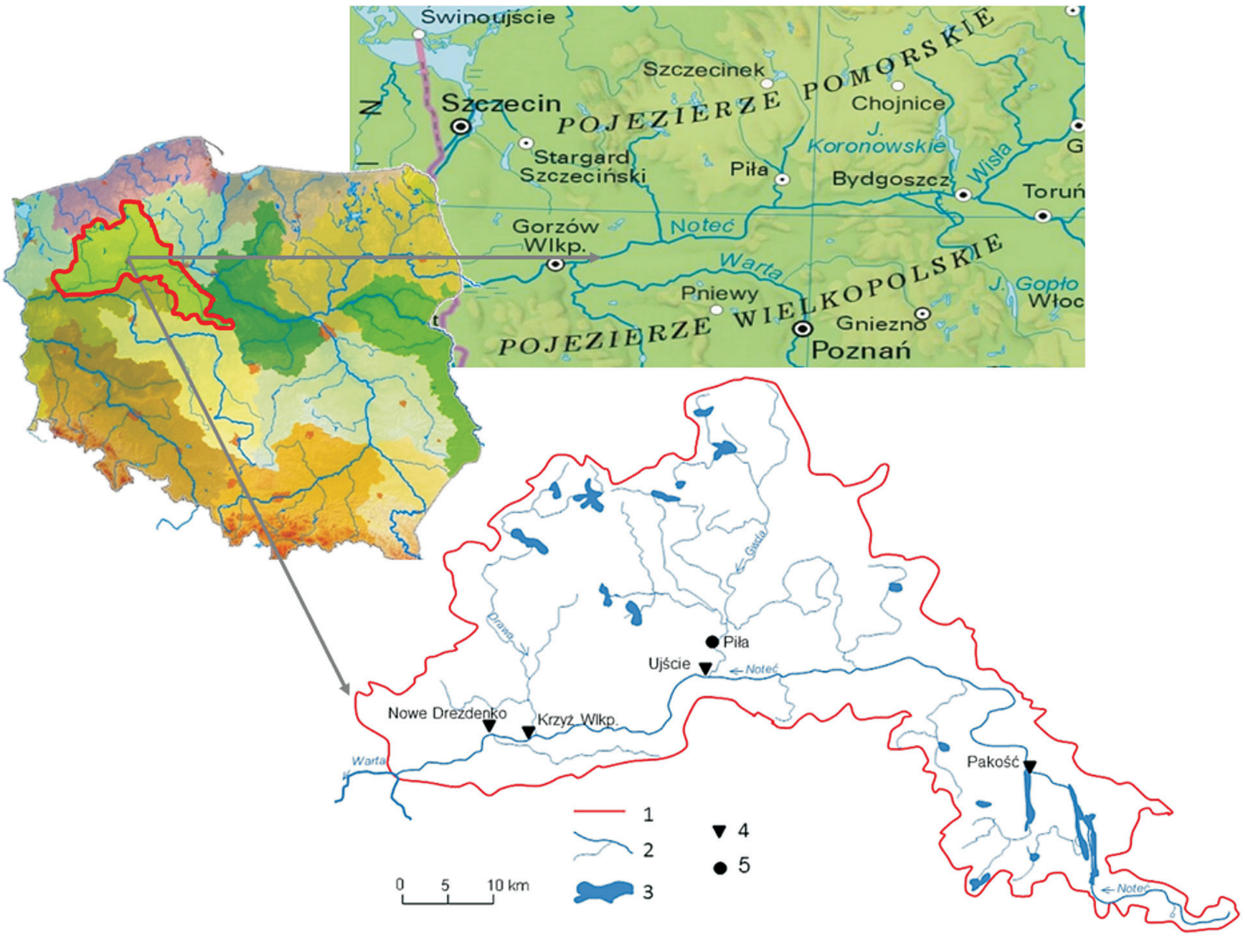


Fig. 1. Location of the Noteć river catchment within the regional layout of Poland, and the distribution of river water temperature and air temperature measurement stations within the studied area: 1) catchment boundary, 2) river, 3) lake, 4) water temperature measurement station, 5) weather station

Noteć river catchment is located in the macroregions of Pojezierze Wielkopolskie (lake district) and Pradolina Toruńsko-Eberswaldzka (valley), through which the river flows in its the main riverbed, and the Pojezierze Południowopomorskie (lake district), through which the right-bank tributaries of the Noteć river are flowing. The research area is located within the range of two climatic regions: Wielkopolski Zachodni (Western Wielkopolska) and Wielkopolski Wschodni (Eastern Wielkopolska), while its northern part includes the regions of: Pomorski Zachodni (Western Pomeranian) and Pomorski Wschodni (Eastern Pomeranian) (Woś 2010). The average annual air temperature in the two Wielkopolska regions is 8.3°C (Western) and 8.0°C (Eastern), while in the Pomeranian regions, 8.1°C (Western) and 7.5°C (Eastern), respectively.

The Noteć river and its tributaries are characterized by a snow-and-rain supply regime. The largest flows usually occur in the months of March and April. The average flow of the spring month is 130–180% of the average annual flow, which allows Noteć to be included among rivers with a medium-developed nival regime (Borowicz 2016). Low flows occur in the months from July to September. The studied rivers are characterized by the longest, average duration of low flows (from 75 to 97 days), among which the summer and autumn low flows predominate. The average annual unit runoff for the Noteć river catchment for the period of 1971–2010 (according to the IMGW-PIB data) in the Pakość section (upper course of the river) is 3.5 $dm^3s^{-1}km^{-2}$, whereas in the Nowe Drezdenko section (lower course of the river), it is 4.6 $dm^3s^{-1}km^{-2}$. Noteć river and its main tributaries, Drawa and Gwda, are characterized by the smallest variability of daily flows, which is influenced by, among other things, instability of the snow cover, precipitation of low intensity, and high infiltration capacity of the sub-surface of the catchment. Along the course of the river and its tributaries, there are numerous lakes (see: Fig. 1), which cover about 4% of the catchment area (Borowicz 2016), whereas Noteć river itself is regulated by a sluice system, and channelled in its middle course. The river functions as a migration corridor for many species of fish, whereas its valley, due to its significant natural qualities, is protected by inclusion under Natura 2000 areas: "Dolina Noteci" and "Dolina Środkowa Noteci i Kanału Bydgoskiego".

The analysis uses a homogeneous data series obtained from the Institute of Meteorology and Water Management – National Research Institute IMGW-PIB (*Instytut Meteorologii i Gospodarki Wodnej – Państwowy Instytut Badawczy IMGW-PIB*) from 1987–2013 regarding daily and monthly water temperatures from three measuring stations located along the Noteć river (see: Fig. 1) at Pakość, Ujście and Nowe Drezdenko, as well as on Gwda river (measuring station at Piła) and Drawa river (measuring station at Drawiny) (see: Fig. 2). Additionally, data regarding water temperatures in the Drawa river were taken into account, measured at the Drawsko Pomorskie station, from a shorter observation period of 1987–1991. Data regarding daily and monthly air temperatures in the analysed period were obtained from the weather station in Piła (see: Fig. 2). Calculations and statistical analyses were conducted in the R computational environment (GNU R Package), version 3.3.2. (R Core Team 2015). The results of the significance tests applied were evaluated against the assumed level of $\alpha = 0.05$. If another level of significance was adopted in particular cases, then such information was additionally included in the presentation of calculation results.

Analyses of the relationship between the diurnal and the monthly temperatures of water in the Noteć river and its tributaries versus temperatures were preceded by the assessment of the degree of normality of the data series distributions, which was carried out using frequency histograms and quantile-quantile plots. At the level of the multi-year period of 1987–2013, due to the large number of observations (9.860 daily measurements), the temperature distribution was evaluated graphically, based on the histogram and quantile-quantile plot, which determines the relationship between the quantile value expected based on the normality of distribution versus the actual, empirical quantiles. If the data is in line with the graph plotted, it means that their distribution is approximately normal. Quantile distributions, provided that we possess a significant number of data, make it possible to assess the degree of normality of data series distributions, and to solve issues related to their symmetry (Graf 2018).

With the aid of the Augmented Dickey-Fuller Test ADF (Dickey and Fuller 1981) we have characterizes a series of daily temperature data in terms of station-

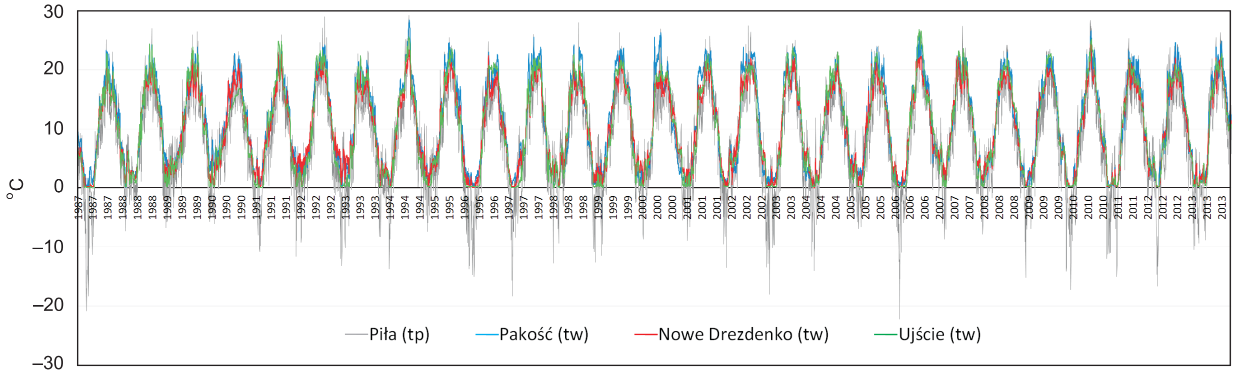


Fig. 2. Daily air temperatures measured in Piła weather station (Ta) and water temperatures of the Noteć river measured in Pakość, Nowe Drezdenko and Ujście [°C] in the period 1987–2013

arity. A detailed description and results of the test are presented in the work by Graf (2018). The studies of the stationarity of the data series were the procedure necessary in order to select the method for analysing the strength of the correlation between the analysed variables. In the ADF test, the stationarity of the data series was determined when the p-value of the test was lower than the adopted level of significance ($\alpha = 0.01$), which meant rejecting the hypothesis of the so-called unit root, and the acceptance of the stationarity of the time series. The study uses the ADF test, which determines significance levels based on the significance table, therefore, the determined levels were approximate. Decomposition of daily water and air temperature time series into additive components was also carried out, namely into: trend component, periodic component, and random (irregular) fluctuations. For this purpose, the so-called loess technique was used (i.e. locally weighted scatterplot smoothing), which is also called local-weighted regression, or weighted locally polynomial regression. The method was described in detail by Cleveland et al. (1990).

In the study of causal relationships between the time series of water and air temperatures, the first-order Granger causality test was used (Granger 1980), which refers to stationary processes (Detto et al. 2012). The Granger causality test is used in order to identify linear causal interactions between time series of data (see: Fig. 3) based on the theory of prediction, and it concerns causality in the mean or in variance, and often in risk. The phenomenon is considered causal if it meets two conditions: the cause precedes the effect, and the cause contains information about the effect that is not available in the group of other variables (Granger 2012). According to the assumptions of causality, it was assumed that the variable Y affects the variable X, if:

$$P\ (Xt{+}1 \in A|X,Y) \neq P\ (Xt{+}1 \in A|X) \quad \text{for a certain A event,} \tag{1}$$

X – historical time series X,
X, Y – historical time series X and Y,
$Xt + 1$ – value of the variable at the moment in time $t+1$

The prognostic models developed in the present work were based on explaining the percentage changes of the so-called unique variance, denoting the percentage of variance that is explained by a given factor over what is explained by the historical values of a given variable. For this purpose, in the study of cause-and-effect relationships between time series of water and air temperatures, it was determined how much the coefficient of determination (R^2) of both models increases, taking into account the past values of the second time series. The Granger causality test was used to verify whether the model predicting the variable value, that is the water temperature at the moment of e, on a given day, based on the water temperature from the previous day (Model 1) is significantly worse, in other words, whether it explains less

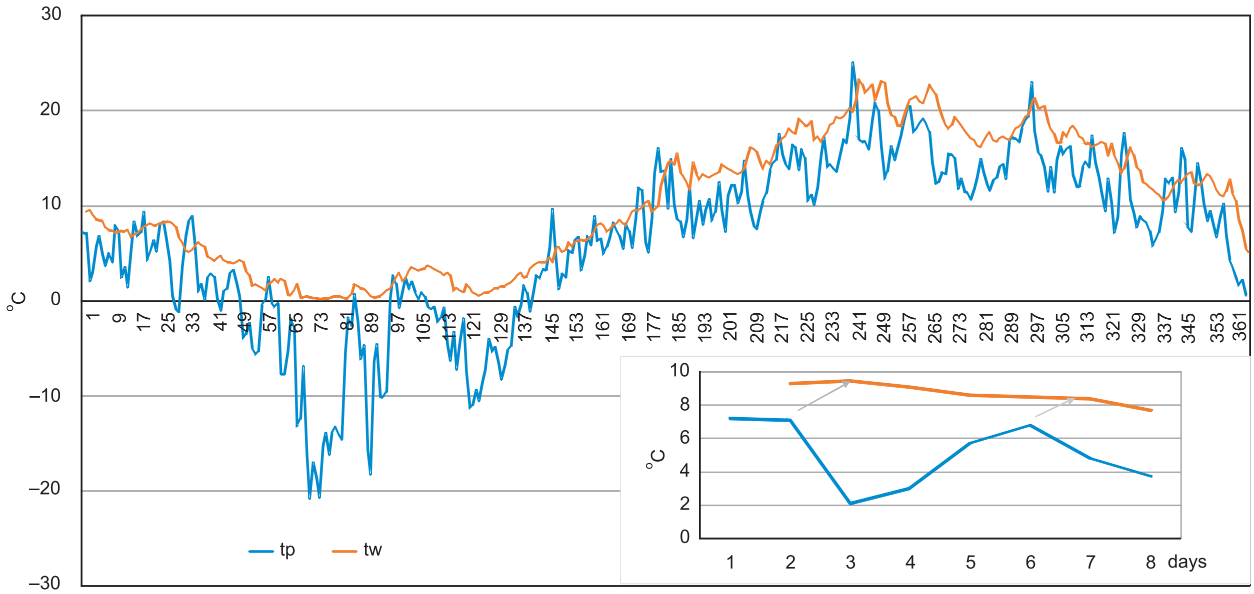


Fig. 3. The analysis of the Granger causality relationship between daily temperatures of water (Noteć river at Pakość) and of air (at Piła station) as illustrated with the example of the data series from the year 1987

variability of the predicted variable than the model which explains it additionally based on the air temperature from the previous day (Model 2). In such a system, if Model 2 is better suited to the data, there is a predictive causality in the Granger sense, which is symbolically described as:

Model 1. $W\_t = W\_t - 1 + e$ (2)

Model 2. $W\_t = W\_t - 1 + P\_t - 1 + e$ (3)

where:

- W – water temperature,
- P – air temperature,
- t – moment in time,
- e – random error component.

The significance of the difference between the models was determined using F-test statistics, and the analysis of variance. The analysis was additionally enhanced by examining causality, in the meaning of Granger causality, between the series of residues of both temperature models, that is, the part of each measurement that can be considered as random fluctuations. Having established the changes in determination coefficients (R^2) of models based on random observation series fluctuations, taking into account the past values of the second time series, the strength of cause-and-effect dependency between them was indicated.

RESULTS

The analysed distributions of a time series of daily water temperatures in the Noteć river and its tributaries (1987–2013) were for the most part relatively symmetrical and close to normal distribution, which was confirmed by quantile distributions and histograms of distributions (see: Fig. 4). Skewness has also been identified to some extent (Graf 2018). Biodimodality occurred only in the case of annual distributions, which is an obvious consequence of the seasonality associated with the occurrence of cold and warm half-year. Distributions of air temperature (measured at Piła station) were also predominantly close to normal, with a slight left skewness associated with a slight over-representation of lower air temperatures. At the level of individual months, distributions of time series of river water temperatures were also estimated as normal, or at least relatively symmetrical (see: Fig. 5). This confirms the fact that the monthly temperatures of river waters were relatively stable throughout the period under consideration.

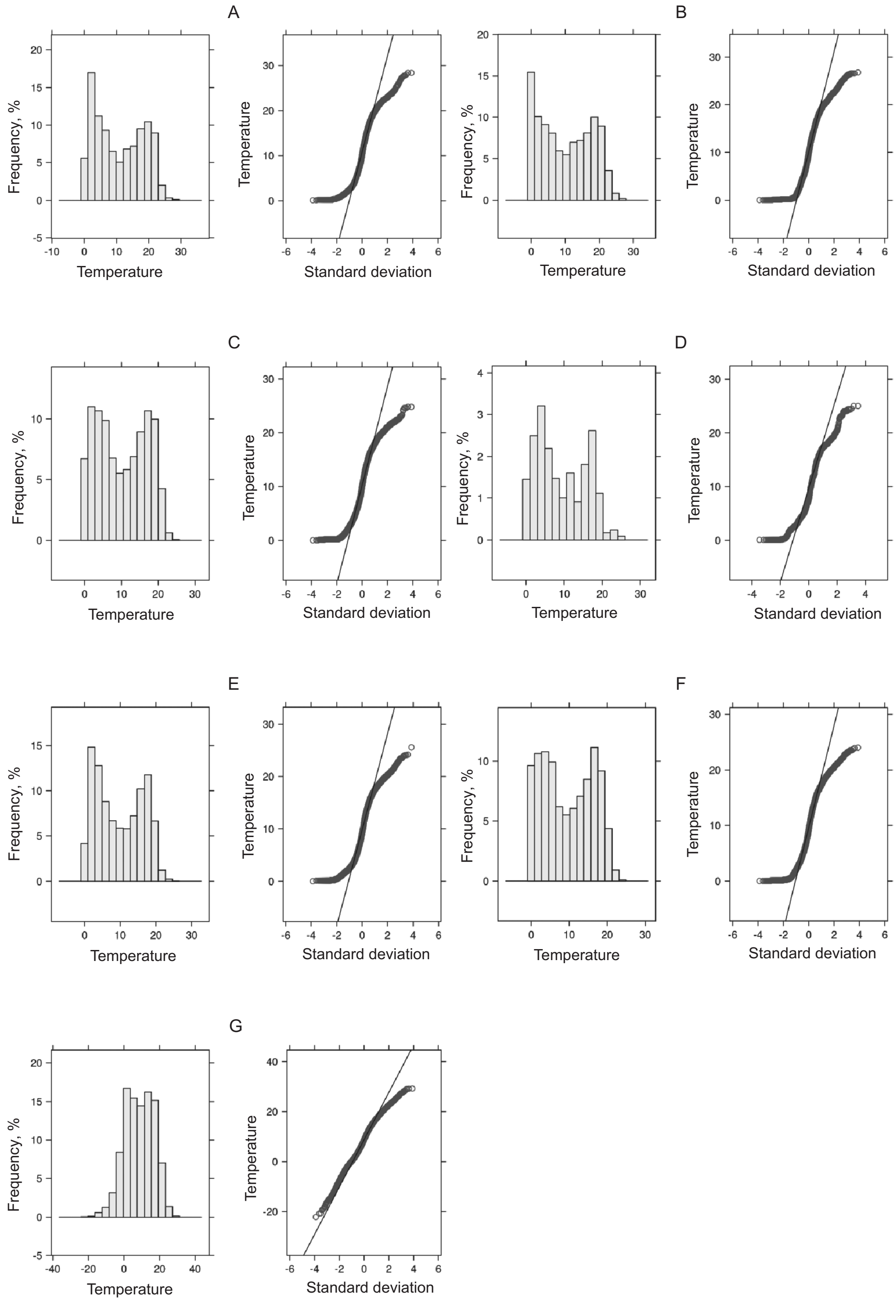
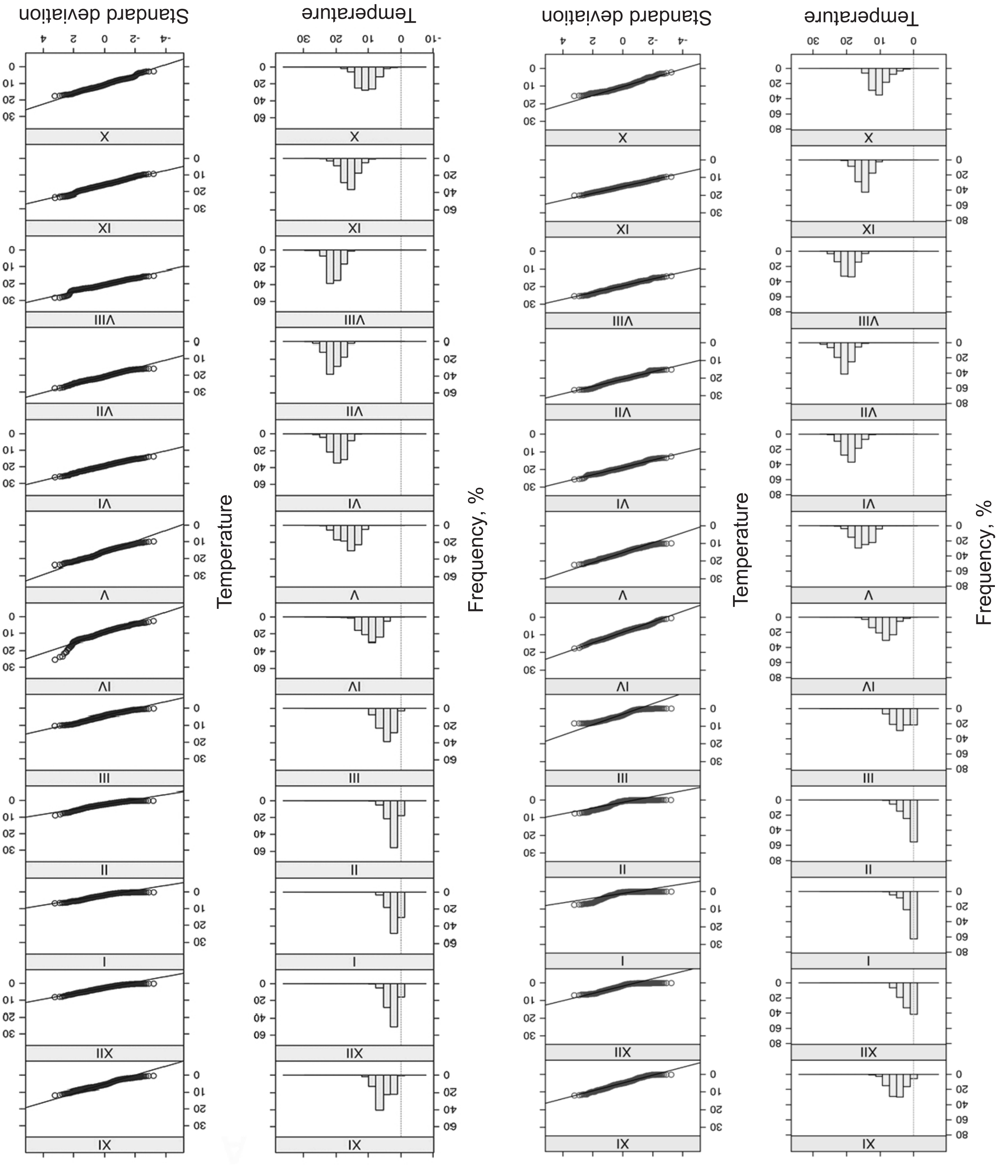


Fig. 4. Frequency histograms and quantile plots of the distribution of daily water temperature series for the Noteć river and its tributaries, and of air temperature (1987–2013): A) Noteć – Pakość, B) Noteć – Ujście, C) Noteć – Nowe Drezdenko, D) Drawa – Drawsko Pomorskie (1987–1991), E) Drawa – Drawiny, F) Gwda – Piła, G) Piła – weather station (based on Graf 2018)


Standard deviation
Temperature
Temperature
Frequency, %
Standard deviation
Temperature
Temperature
Frequency, %
X
IX
VIII
VII
VI
V
IV
III
II
I
XII
XI

A B

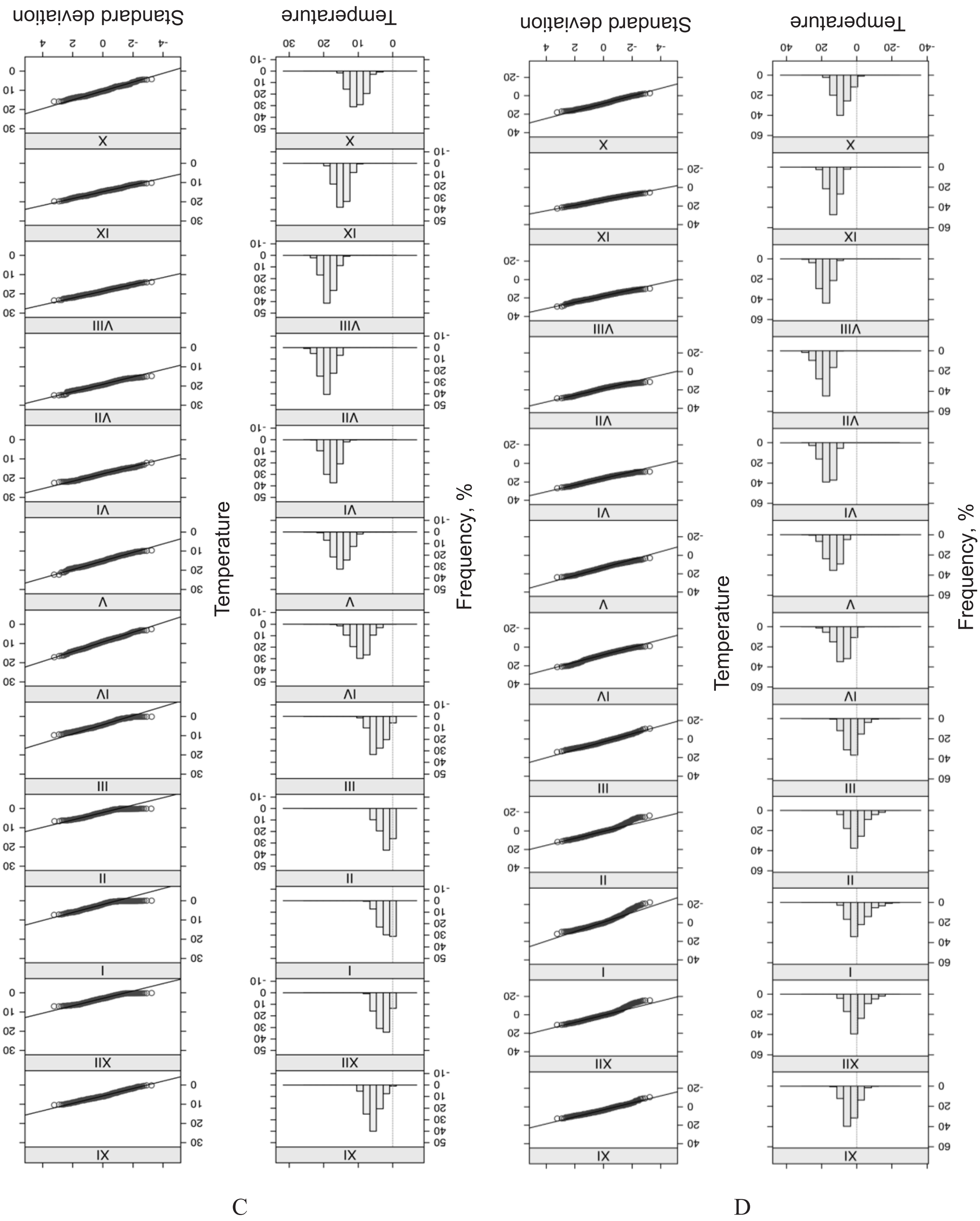


Fig. 5. Frequency histograms and quantile plots of the distribution of monthly water temperature series for the Noteć river, and of air temperature (1987–2013): A) Pakość, B) Ujście, C) Nowe Drezdenko, D) Piła – weather station (based on Graf 2018)

For Noteć river at Pakość (in the upper course), the distribution for April was clearly more right-skewed than distributions recorded for other months. Distribution of time series of river water temperatures in Ujście (in the middle course) showed a distinctive right skewness between December and March, while for the lower section of Noteć river (at Nowe Drezdenko), the water temperature distributions for January, February and March displayed a not quite accurate (right)-skewness (see: Fig. 5). Time series, although still dominated by low temperatures, were characterized by above average frequency of warmer days. For the tributary of Noteć river, the Drawa river in Drawiny, water temperature distributions at the level of individual months were approximately normal, or at least relatively symmetrical. On Gwda river in Piła, the monthly temperature distribution of water, apart from the distributions for December, January, February and March, showed approximately normal type of distribution. The aforementioned distributions were characterized by fairly clear right skewness. The monthly distributions of air temperature in Piła, apart from the distributions for December, January and February, which showed slight left skewness, were also approximately normal (see: Fig. 5). Both analysed series of temperatures – of river water and of air – were stationary time series.

In the tested series of daily temperature, the dependencies of the so-called Granger causality were defined in two directions. The cause and effect relationship was observed between the series on both the water-air and air-water axes (see: Table 1). Between the series of monthly temperatures of water and of air, various configurations of causal relationships, in the sense of Granger causality, or lack thereof, have been found. For Drawa river and Drawsko Pomorskie, the causality relationship in the Granger sense was noted between the data only on the water-air axis,

Table 1. Correlation between daily and monthly series of river water and air temperatures in terms of Granger causality (1987–2013)

River – Profile	Cause-and-effect relationship at the level of significance p < 0.001			
	Daily (24-hour) series		Monthly series	
	Water–Air	Air–Water	Air–Water	Air–Water
Noteć – Pakość	F(1, 9859) = 2901	F(1, 9859) = 363.9	F(1, 321) = 9.667; p = 0.00205	F(1, 321) = 1.803; p = 0.18 No correlation
Noteć – Ujście	F(1, 9859) = 2476	F(1, 9859) = 456.1	F(1, 321) = 18.39 p = 2.39e-05	F(1, 321) = 0.949 p = 0.331 No correlation
Noteć – Nowe Drezdenko	F(1, 9859) = 3685	F(1, 9859) = 399.2	F(1, 321) =13.21; p = 0.000324	F(1, 321) = 0.463; p = 0.497 No correlation
Drawa – Drawiny	F(1, 9859) = 1823	F(1, 9859) = 423.2	F(1, 321) = 29.43; p < 0.001	F(1,321) = 1.51; p = 0.220
*Drawa – Drawsko Pomorskie	F(1, 2005) = 53.91	F(1, 2005) = 166.2	F(1, 62) = 6.70; p = 0.012,	F(1, 62) = 0.002; p = 0.968
Gwda – Piła	F(1, 9859) = 5169	F(1, 9859) = 373.2	F(1, 321) = 3.571; p = 0.0597 **No correlation	F(1, 321) = 3.638; p = 0.0574 No correlation

F – test statistics, numbers in parentheses – degrees of freedom of F distribution, value on the right side of the sign "=" denotes the values of test statistics F.
* Drawa – Drawsko Pomorskie: lag order = 12 years (1987–1991), for other measurement stations: lag order = 26 (1987–2013), α = 0,01
** No correlation

while the air-water axis appeared statistically insignificant (see: Table 1). In the case of Noteć river, (at Pakość, Ujście, and Nowe Drezdenko stations), these correlations were recorded only on the water-air axis, which indicated the occurrence of a one-sided causal relationship, according to which the average air temperature to some extent influenced and determined the water temperature in the next month. In the case of Gwda river in Piła, no correlation was confirmed, which may be the effect of disturbing the features of the thermal regime of river waters through the local impact of anthropogenic factors (operation of hydropower plants).

When assessing the strength of the detected correlations between the water and air temperature measurement series, it was established by how much the coefficient of determination (R^2) of both models increases, taking into account the past values of the second time series (see: Table 2). The model predicting the daily temperature of the waters in Noteć river based on water and air temperatures from the previous day explained 0.07–0.27% of the unique variance more than the model that used only the water temperature of the previous day. The unique variance in this context meant the percentage of variance that is explained by a given factor over what is explained by the historical values of a given variable. In turn, the model predicting the daily air temperature based on air and water temperatures from the previous day explained 0.3–0.79% of the variance more than the model, which uses only the air temperature of the previous day (see: Table 2). The model predicting the average monthly temperature of the waters in Noteć river and its tributaries based on water and air temperatures from the previous month explained 0.87–2.75% of the unique variance more than the model that used only water temperature.

The analysis of Granger causality was additionally expanded by examining the relationship between the series of random fluctuations of both temperature models. Random fluctuations in daily temperature series are one of their features, apart from the long-term trend and periodic values (see: Fig. 6).

The decomposition of time series into additive components indicated the occurrence of a negligible trend, responsible for 0.4–0.6% of temperature variability in the analysed period. Apart from the trend, there was a very strong periodic component (explaining around 92.6–93.4% of the variability), whereas random fluctuations accounted for about 6.0–7.5% of the temperature variance.

The investigated effect of Granger causality (of the first degree) between the daily time series of random fluctuations of both models showed that cause-and-effect dependencies take place in both directions. The strength of cause-and-effect correlations was exam-

Table 2. Prediction models for water temperatures in the Noteć river and its tributaries, and air temperatures, taking into account the past values of the second time series (tp–1, tw–1).

River – Profile	Increase in the unique variance percentage explained by an additional factor [%]		
	Daily (24–hour) series		Monthly series
	Water temperature prediction model taking into account (tw–1)*	Air temperature prediction model taking into account (tp–1)	Water temperature prediction model taking into account (tw–1)**
Noteć – Pakość	0.14	0.30	1.58
Noteć – Ujście	0.07	0.38	0.87
Noteć – Nowe Drezdenko	0.19	0.33	1.13
Drawa – Drawiny	0.10	0.40	2.51
Drawa – Drawsko Pomorskie	0.01	0.79	2.75
Gwda – Piła	0.27	0.31	No correlation

* (tw–1) – water temperature from the previous day
** (tp–1) – air temperature from the previous day

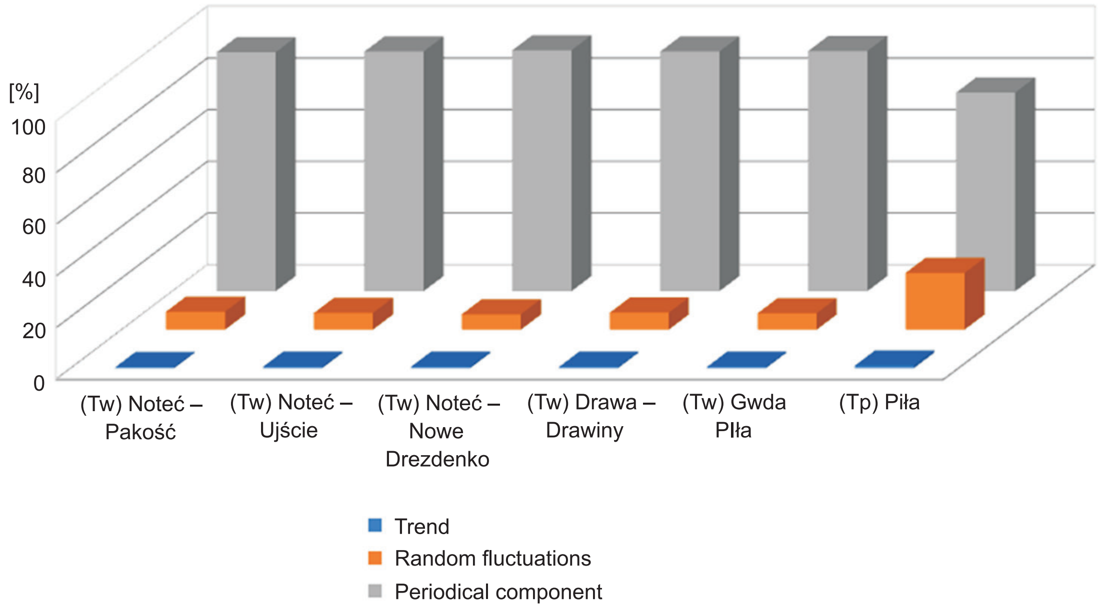


Fig. 6. Decomposition of a daily time series characteristics of water and air temperature into their additive components: trend component, random fluctuations, and periodic component (1987–2013)

ined by analysing changes in the coefficients of determination (R^2) of models taking into account the past values of the second time series (see: Fig. 7–8).

In the case of random fluctuations in air temperatures, the increase in R^2 was only 0.13% in relation to 0.02% of the variance explained by the less effective model in the case of Drawa river in Drawiny, and over 33% in relation to R^2 equal to zero in the case of Gwda river in Piła (see: Fig. 7). The obtained value determines the increment (%) of the variance of random fluctuations in air temperatures, which was explained by the model taking into account not only random fluctuations in air temperature from the previous day, but also random fluctuations in water temperature from the previous day. In the case of slight differences in R^2, this points to the fact that random fluctuations in air temperatures are relatively independent, whereas the effect that the temperature fluctuation of water have on them, although statistically significant, is so weak that it is virtually negligible. In the case of water temperature fluctuations, the largest difference in R^2 amounted to 9.98% (in Drawa river at Drawiny) versus R^2 equal to zero (see: Fig. 8). This means that in this case, air temperature fluctuations have an important, and not insignificant impact (as it is explaining almost 10% of the variability) on what random fluctuations in water temperature will be on the next day. In the other prediction models for wa-

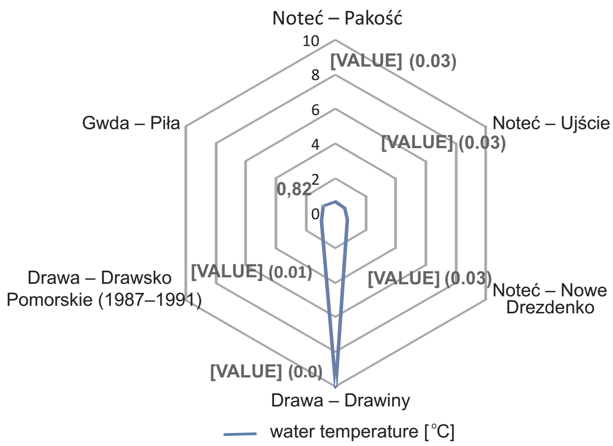


Fig. 7. Increase in the coefficient of determination (R^2) in models predicting random fluctuations of water temperatures, taking into account the past values of the second time series (tp-1) [%]

* Given in brackets are R^2 values for the model, taking account of only the random fluctuation of water temperature from the previous day

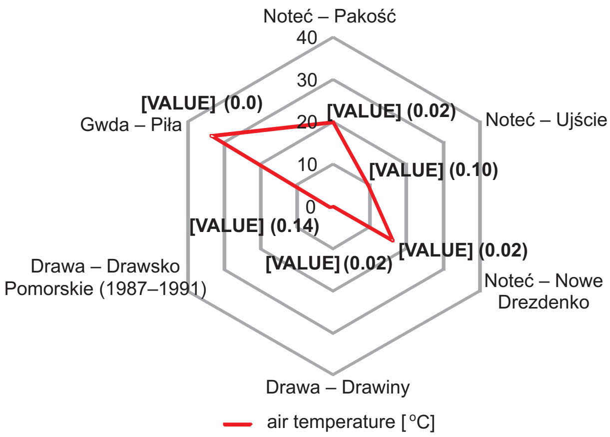


Fig. 8. Increase in the coefficient of determination (R^2) in models predicting random fluctuations of air temperatures, taking into account the past values of the second time series (tp-1) [%]
* Given in brackets are R^2 values for the model, taking account of only the random fluctuation of air temperature from the previous day.

ter temperature in the Noteć river and its tributaries, obtaining low R^2 while taking into account the past values of the second time series confirms that in this case, the temperature fluctuations of water are relatively independent of each other, and fluctuations of air temperatures have no significant effect on them.

Between monthly fluctuations, no cause-effect relationship has been discovered, which points to a different dynamics of water and air temperature changes on a monthly scale compared to the daily scale. On the monthly data level, there is a strong synchronous correlation between monthly average water and air temperatures, while on the scale of daily data, the synchronous correlation is practically negligible, but there are clear cause-and-effect dependencies, according to which fluctuations in water temperature (or air temperature) affect air temperature fluctuations (or water temperature fluctuations) in the following days.

DISCUSSION

The structure and functions of river systems, and the properties of their waters, are described by means of various data sets and quantitative characteristics, which constitute a response to a series of interactions of climatic and environmental factors, often including also anthropogenic factors. The existence of a correlation between the fluctuations in the temperature of water in the river and air temperature has been confirmed in various climatic and regional conditions (Caissie 2006, Webb and Nobilis 2007, Wiejaczka 2007, Piotrowski et al. 2015, Piotrowski and Napiórkowski 2018). In the case of water in the Noteć river water and its tributaries, as well as air temperatures, the Granger causality relationship was determined in two directions of impact: air – water, and water – air (see: Table 1). This allows us to assume that periodic and more frequent fluctuations in the air temperature in the studied area bring about specific changes in water temperature dynamics of the watercourses. The thermal regime of rivers depends primarily on the surface heat exchange with the atmosphere in the processes of radiation, evaporation and convection (Caissie 2006) as well as on the turbulent mixing of water at different temperatures, for instance, inflows, sewage, or thermal pollution. The natural thermal regime is maintained if the water temperature in the rivers is not disturbed, and equilibrium is established between the temperature of the water and that of the environment. It is assumed that molecular diffusion, heat exchange with the substrate, and conductivity in the mass of water play a lesser role in shaping the characteristics of the thermal regime.

In the aspect of Granger causality, it is important to state that water temperature is to some extent causally determined by the air temperature of the previous day, whereas air temperature is to some extent causally determined by the water temperature of the previous day. These correlations occur mainly during freezing and evaporation processes that involve huge amounts of energy. The latent heat of water phase transitions is taken or transferred from or to the environment, that is, the air in the first place. These processes cause significant heating or cooling of the air over the watercourse and the immediate surroundings. This correlation is particularly apparent over water reservoirs. In the case of Noteć river and its tributaries, however, both these correlations are weak.

Detection of the effect of the so-called first degree Granger causality between the daily and, to a lesser extent, monthly, water temperatures of the examined rivers versus air temperature (see: Table 1) confirms the impact of one time series of data on the evolution of the second time series of data. In the analysed case, the data concerning the past of the time series of air temperatures provides important information for the prediction of the time series of river water temperatures, which is not included in the information about its past values and vice versa. This means that forecasting the temperature of waters of the examined rivers from the level of changes in air temperature may give better results if the information from the previous day is used.

The study of the causal relationship in the time series of water temperature in the Noteć river and its tributaries versus air temperature also provided important information on the level of correlation of random fluctuations of both time series for the same days. Determination, at the examined level, of a statistically significant correlation of random fluctuations enables an insight into the characteristics of short-term correlations between water and air temperatures. Random events and values in the measurement series that are not typical for the thermal regime of river waters appear as a result of various types of disturbances therein. The impact of atmospheric factors in the changes of the water temperature dynamics in Noteć river and its tributaries is additionally influenced by the impact of regional factors related to hydrological conditions and the structure of the water balance as well as the effect of local factors. In the case of the Noteć river catchment, factors that have a local impact on the thermal regime of its waters include: the morphological character of the river valley (namely, the flat, wide Pradolina Toruńsko-Eberswaldzka), springs flowing out of the edges and terraces of river valleys, and the proximity of peat bogs (Borowicz 2016). Extensive wetlands are found in the lower and middle reaches of the Noteć river (see: Fig. 1). Local factors, modifying water temperatures of the river, often refer to anthropogenic changes recorded in the riverbed and in the river valley, and the inflow of pollutants that may disturb river continuum, that is, the continuity of processes and gradient nature of zones in the river (Graf 2015, 2018).

The concept of causality, in the meaning of Granger causality, is not limited to searching for a variable whose presence in the model can improve the accuracy of forecasts. According to the concept of causality, in the time series the cause precedes the effect, and the series that causes it contains information about a variable that was not included in other series (Granger 1980, Syczewska 2014). In the case of established correlations, significant results of Granger causality tests can only be an artefact of a large number of the analysed data. On the other hand, it should be emphasized that narrowing the interpretation concerns linear causality (on the water temperature-air temperature and air temperature-water temperature axes), while in a wider scope it is about the dependencies of Data Generating Process (so-called DGP), which assumes that generating information about a given process consists in acquiring that information from other processes or systems (Detto et al. 2012).

In testing Granger causality between the series of water temperatures of the Noteć river and its tributaries versus air temperature, several directions of correlations were indicated: the existence of two-directional causality in the data in the case of daily measurement series; one-way dependence or lack of dependence in the series of monthly data; and lack of confirmation of the relevance of the correlation. The lack of correlation between variables X and Y is usually determined when both are caused by another, third variable. In the studied case, correlation effects are more visible at the level of higher frequency data, that is, the daily data. Data at the monthly or longer aggregation level may "mask" any correlation effects. In addition, Granger causality relates to stationary data that has been verified for a series of daily and monthly water and air temperatures. Instead, the series of annual data of water temperatures in the Noteć river displayed the features of non-stationarity (Graf 2018). The effect of the frequency of the measurement series on the detection of their causal relationships was interpreted, among others, in studies of graphic patterns for the hydrological time series (Jangyodsuk et al. 2014). According to Stern and Kaufmann (2014), Granger causality between time series of data demonstrates the probability of such causality or lack of such causality with greater force than simple correlation. It is argued that a statistical test of Granger causality between the time series of data proves useful. These conclusions have also been confirmed in the context of environmental research into climate change (Wang et al. 2004, Mosedale and Stephenson 2006, Granger 2012). Extracting relevant and confirmed conclusions about the existence and direction of causal dependence on the basis of the Granger causality test is possible only if we possess the knowledge of the mechanism combining the two time series (Maziarz 2015).

CONCLUSIONS

Studies of the correlation between river water temperatures, recorded in selected measurement profiles of the Noteć river and its tributaries, and air temperatures, applying the causality relationship in the meaning of Granger causality, confirmed their interdependence in the analysed period. In the case of daily (24-hour) data series, cause-and-effect relationships were determined on the "water-air" axis, as well as "air-water" axis, while in the case of the monthly data series, there were also cases of one-sided causality or lack thereof. This means that forecasting the course of temperatures of river waters based on the level of air temperature changes can render better results if the information from the previous day is used.

In the case of time series of temperatures, the detection of causal patterns in variance, using the causality analysis in the meaning of Granger causality, may serve to improve the quality of forecasts of the studied variables. It was demonstrated that the past data on the time series of air temperatures provides important information for the prediction of the time series values related to the waters of Noteć river and its tributaries, which is not included in information about its past values and vice versa. It was also found that a weak but statistically significant correlation of random fluctuations exists for the same days of both series, that is, those temperature changes that do not result from long-term trend, historical values, or periodicity. Therefore, the occurrence of short-term correlations between water and air temperatures has been confirmed. The influence of regional and local factors, which are most often associated with anthropogenic pressure in the catchment and river valley, is overlapping the influence of atmospheric factors pertaining to water temperature changes.

Causal relationship in the Granger sense between water temperature in the river and air temperature means time precedence between temperatures, which confirms the applicability of test results in forecasting the variability of thermal parameters of river waters. The obtained results constitute a database for reference indications of thermal parameters, which are significant in the prediction of climate change, and in studying the thermal regime of the river. They can also contribute to improving the quality of predictions of water temperature changes in the Noteć river and its tributaries, which is important for the proper management of their environmental and ecological status. The established correlations between thermal parameters of water and air also determine the level of risk for the functioning of water ecosystems represented by habitats and species susceptible to changes in river water temperature. This problem is particularly important in the case of rivers with important ecological functions, for instance, including that of migration corridors for diversified species of fish, which is also one of the functions of the Noteć river.

REFERENCES

Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T.B.M., Bobée, B. (2007). A review of statistical water temperature models. Canadian Water Resources Journal, 32, 79–192.

Borowicz, J. (2016). Charakterystyka zjawisk lodowych na Noteci i jej prawobrzeżnych dopływach w latach 1982–2011, Archiwum Prac Dyplomowych, Poznań: UAM.

Caissie, D. (2006). The thermal regime of rivers: a review. Freshwater Biology, 51, 1389–1406.

Caissie, D., Satish, M.G., El-Jabi, N. (2005). Predicting river water temperatures using the equilibrium temperature concept with application on the Miramichi River catchments (New Brunswick, Canada). Hydrological Processes, 19, 2137–2159.

Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics, 6, 3–73.

Detto, M., Molini, A., Katul, G., Stoy, P., Palmroth, S., Baldocchi, D. (2012). Causality and persistence in ecological systems: A nonparametric spectral Granger causality approach. Am. Nat., 179, 524–535.

Dickey, D.A., Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49, 1057–1072.

Gardner, B., Sullivan, P.J., Lembo, A.J. (2003). Predicting stream temperatures: geostatistical model comparison using alternative distance metrics. Canadian Journal of Fisheries and Aquatic Sciences, 60, 344–351.

Graf, R. (2015). Zmiany termiki wód Warty w profilu łączącym pradolinny i przełomowy odcinek doliny (Nowa Wieś Podgórna-Śrem-Poznań). In: Absalon D., Matysik M., Ruman M. (ed.), Nowoczesne metody i rozwiązania w hydrologii i gospodarce wodnej. Katowice: Komisja Hydrologiczna PTG, PTG Oddział Katowice, 177–194.

Graf, R. (2018). Distribution Properties of a Measurement Series of River Water Temperature at Different Time Resolution Levels (Based on the Example of the Lowland River Noteć, Poland). Water, 10, 203.

Granger, C.W.J. (1980). Testing for Causality. A personal Viewpoint, Journal of Economic Dynamic and Control, 2, 4.

Granger, C.W.J. (2012). Forecasting, [in:] U. Maki (ed.), Philosophy of Economics. Amsterdam: Elsevier.

Hilderbrand, R.H., Kashiwagi, M.T., Prochaska, A.P. (2014). Regional and local scale modeling of stream temperatures and spatio-temporal variation in thermal sensitivities. Environmental Management, 54, 14–22.

Jangyodsuk, P., Seo, D-J., Gao, J. (2014). Causal Graph Discovery For Hydrological Time Series Knowledge Discovery. CUNY Academic Works. http://academicworks.cuny.edu/cc_conf_hic/430.

Kondracki, J. (2008). Geografia regionalna Polski. Warszawa: PWN.

Lagergaard Pedersen, N., Sand-Jensen, K. (2007). Temperature in lowland Danish streams: Contemporary patterns, empirical models and future scenarios, Hydrological Processes, 21, 348–358.

Łaszewski, M. (2014). Metody określania związków temperatury wody rzecznej i temperatury powietrza na przykładzie rzeki Świder. Prace Geograficzne, 136, 45–60.

Letcher, B.H., Hocking, D.J., O'Neil, K., Whiteley, A.R., Nislow, K.H., O'Donnell, M.J. (2016). A hierarchical model of daily stream temperature using air-water temperature synchronization, autocorrelation, and time lags. PeerJ – the Journal of Life and Environmental Sciences, 4. https://doi.org/10.7717/peerj.1727.

Maziarz, M. (2015). O wartości informacyjnej testów przyczynowości w sensie Grangera. Optimum. Studia ekonomiczne, 2(74), 152–170.

Mosedale, T.J., Stephenson, D.B. (2006). Granger causality of coupled climate processes: ocean feedback on the north Atlantic oscillation. Journal of Climate, 19, 1182–1194.

Napiórkowski, M.J., Piotrowski, A.P., Napiórkowski, J.J. (2014). Stream temperature forecasting by means of ensemble of neural networks: Importance of input variables and ensemble size. In: A. J. Schleiss, G. de Cesare, M. J. Franca, M. Pfister (Eds), River Flow 2014. London: Taylor & Francis Group.

Neumann, D.W., Rajagopalan, B., Zagona, E.A. (2003). Regression Model for Daily Maximum Stream Temperature, Journal of Environmental Engineering, 7, 667–674.

Piotrowski, A.P., Napiórkowski, J.J. (2018). Performance of the air2stream model that relates air and stream water temperatures depends on the calibration method. Journal of Hydrology, 561, 395–412.

Piotrowski, A.P., Napiórkowski, M.J., Napiórkowski, J.J., Osuch, M. (2015). Comparing various artificial neural network types for water temperature prediction in rivers. Journal of Hydrology, 529, 302–315.

Piotrowski, A.P., Osuch, M., Napiórkowski, M.J., Rowiński, P.M., Napiórkowski, J.J. (2014). Comparing large number of metaheuristics for artificial neural networks training to predict water temperature in a natural river. Computers & Geosciences, 64, 136–151.

Piotrowski, A.P., Napiórkowski, M.J., Kalinowska, M., Napiórkowski, J.J., Osuch, M. (2016). Are Evolutionary Algorithms Effective in Calibrating Different Artificial Neural Network Types for Streamwater Temperature Prediction? Water Resources Management, 30, 3, 1217–1237.

R Core Team, (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online: http://www.R-project.org (15.02.2017).

Rushworth, A.M., Peterson, E.E., Ver Hoef, J.M., Bowman A.W. (2015). Validation and comparison of geostatistical and spline models for spatial stream networks. Environmetrics, 26, 327–338.

Sahoo, G.B, Schladow, S., Gand Reuter, J.E. (2009). Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models. J. Hydrol., 378, 325–42.

Stern, D.I., Kaufmann, R.K. (2014). Anthropogenic and natural causes of climate change. Climatic Change, 122, 257–269.

Syczewska, E.M. (2014). Przyczynowość w sensie Grangera – wybrane metody. Metody Ilościowe w Badaniach Ekonomicznych 15, 4, 169–180.

Toffolon, M., Piccolroaz, S. (2015). A hybrid model for river water temperature as a function of air temperature and discharge, Environ. Res. Lett., 10 114011, DOI:10.1088/1748-9326/10/11/114011.

Wang, W., Anderson, B.T., Kaufmann, R.K., Myneni, R.B. (2004). The relation between the North Atlantic oscillation and SST's in the North Atlantic basin. Journal of Climate, 17, 4752–4759.

Webb, B.W., Nobilis, F. (2007). Long-term changes in river temperature and the influence of climatic and hydrological factors. Hydrological Sciences, 52, 74–85.

Wiejaczka, Ł. (2007). Relacje pomiędzy temperaturą wody w rzece a temperaturą powietrza (na przykładzie rzeki Ropy), Folia Geogr., Geograph.-Physica., 37–38, 95–105.

Woś, A. (2010). Klimat Polski w drugiej połowie XX wieku. Poznań: Wydawnictwo Naukowe UAM.

ANALIZA ZALEŻNOŚCI PRZYCZYNOWYCH W SENSIE GRANGERA MIĘDZY TEMPERATURĄ WODY RZECZNEJ I POWIETRZA NA PRZYKŁADZIE RZEKI NOTEĆ

ABSTRAKT

W opracowaniu przedstawiono wyniki analizy zależności tzw. przyczynowości Grangera między dobowymi i miesięcznymi temperaturami wody i powietrza dla okresu 1987–2013, przeprowadzonej na przykładzie rzeki Noteć i jej dwóch zasadniczych dopływów: Drawy i Gwdy. Przyczynowość Grangera odnosi się do sytuacji, w której dane, dotyczące przeszłości jednego szeregu czasowego, dostarczają istotnej informacji do przewidywania wartości drugiego szeregu, która nie jest zawarta w informacji dotyczącej jego przeszłych wartości. W analizie zastosowano test przyczynowości Grangera pierwszego rzędu. Dla serii dobowych temperatur stwierdzono relację przyczynowości zarówno na linii powietrze–woda, jak również woda–powietrze, co oznacza, że prognozowanie przebiegu temperatury wód rzecznych z poziomu zmian temperatury powietrza może dać lepsze efekty, jeżeli wykorzystane zostaną informacje z dnia poprzedniego. Model przewidujący dobową temperaturę wód Noteci na podstawie temperatur wody i powietrza z dnia poprzedniego wyjaśnił o 0,07–0,27% unikalnej wariancji więcej niż model, który wykorzystywał wyłącznie temperaturę wody z dnia poprzedniego. Model przewidujący dobową temperaturę powietrza na podstawie temperatur powietrza i wody z poprzedniego dnia wyjaśnił o 0,3–0,79% wariancji więcej niż model, który wykorzystuje wyłącznie temperaturę powietrza z dnia poprzedniego. Dla serii miesięcznych temperatur wody i powietrza stwierdzono różne konfiguracje zależności w sensie Grangera, jednostronne, na linii woda–powietrze lub jej brak, co może stanowić efekt zaburzenia cech reżimu termicznego wód rzecznych przez lokalne oddziaływanie czynników antropogenicznych. Dodatkowo zbadany efekt zależności przyczynowości Grangera między szeregami fluktuacji losowych obu modeli temperatur potwierdził, że relacje przyczynowe zachodzą w obu kierunkach. Wskazanie efektu zależności przyczynowych w sensie Grangera stanowi potwierdzenie wpływu jednego szeregu danych na ewolucję drugiego szeregu danych i określa możliwości aplikacyjne wyników badań w prognozowaniu zmienności parametrów termicznych wód rzecznych. Uzyskane wyniki mogą przyczynić się do polepszenia jakości prognoz zmian termiki wód, co ma istotne znaczenie dla właściwego zarządzania ich stanem ekologicznym.

Słowa kluczowe: temperatura wody, temperatura powietrza, przyczynowość Grangera, unikalna wariancja, prognozowanie, rzeka Noteć