

## Zastosowanie metod korelacji do analizy dyskretnych sygnałów pomiarowych o charakterze losowym

**Streszczenie.** W części teoretycznej pracy omówiono metody korelacyjne (diagram rozrzutu, nożyce korelacji, współczynnik korelacji liniowej Pearsona), które należy stosować podczas określania zależności pomiędzy dyskretnymi sygnałami pomiarowymi o charakterze losowym. W części praktycznej pracy zaprezentowano przykładową analizę korelacyjną (metodami graficznymi i analitycznymi) rzeczywistych wyników pomiarów (temperatury, wilgotności względnej oraz natężenie promieniowania słonecznego). Na podstawie uzyskanych wyników sformułowano wnioski dotyczące zależności pomiędzy rozpatrywanymi sygnałami, reprezentującymi wyżej wymienione wielkości fizyczne.

**Abstract.** The theoretical part of the work discusses the correlation methods (scatter diagram, correlation scissors, Pearson's linear correlation coefficient), which should be used when determining the relationship between discrete random measurement signals. The practical part of the work presents an example of a correlation analysis (using graphical and analytical methods) of the actual measurement results (temperature, relative humidity and solar radiation intensity). On the basis of the obtained results, conclusions were formulated regarding the relationship between the analyzed signals, representing the above-mentioned physical quantities. (**Application of correlation methods for analysis discrete measurement signals of a random nature**)

**Słowa kluczowe:** analiza korelacyjna, diagram rozrzutu, nożyce korelacji, współczynnik korelacji liniowej Pearsona

**Keywords:** correlation analysis, scatter diagram, correlation scissors, Pearson's linear correlation coefficient.

### Wstęp

Klimat to charakterystyczny dla danego obszaru zespół zjawisk i procesów atmosferycznych, które kształtują się pod wpływem czynników geograficznych i fizycznych. Określa się go na podstawie wieloletnich pomiarów i obserwacji [1]. Wszechstronne badanie i poznanie klimatu jest niezbędne z naukowego punktu widzenia, ponieważ wpływa on na wiele aspektów życia ludzkiego. Ze względu na zachodzące zmiany klimatu, m.in. pod wpływem działalności człowieka, konieczne jest stałe monitorowanie czynników stanowiących o klimacie (temperatury powietrza, wilgotności względnej powietrza oraz natężenia promieniowania słonecznego), badanie skutków ich wpływu na środowisko i człowieka oraz zależności między nimi.

Jednym ze strategicznych miast Polski jest Płock, ponieważ znajduje się w jego granicach kombinat rafineryjno-petrochemiczny PKN ORLEN S.A., będący jednym z czynników klimatotwórczych w mieście i jego okolicach. Zasadnym więc stało się przeprowadzenie pomiarów temperatury powietrza, wilgotności względnej powietrza oraz natężenia promieniowania słonecznego na terenie miasta. Pomiary przeprowadzono w przedziale czasowym od grudnia 2020 roku do maja 2021 roku. Akwizycja danych następowała co 30 minut przy użyciu stacji meteorologicznej LB-719, umieszczonej na dachu Politechniki Warszawskiej Filii w Płocku, wyposażonej w czujnik temperatury i wilgotności powietrza LB-710R oraz czujnik promieniowania słonecznego LB-900/LB-901 [2].

Uzyskane zbiory danych poddano korekcji, a następnie wyznaczono ich wartości średnie, które odniesiono do wyników krajowych [3] (Tabela 1.).

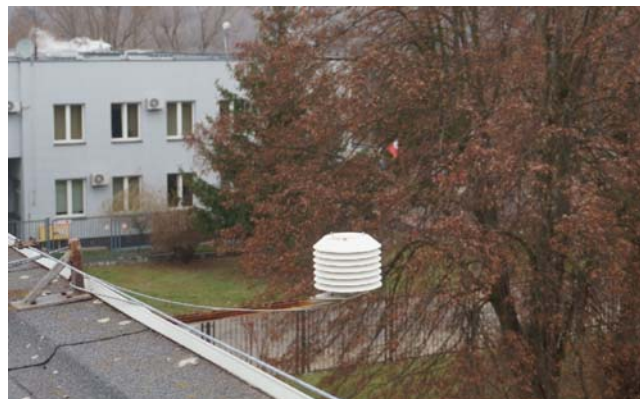
Tabela 1. Porównanie wartości średnich wielkości fizycznych charakteryzujących klimat w Płocku i Polsce

	Płock	Polska
Temperatura powietrza [°C]	3,6	2,8
Wilgotność względna powietrza [%]	82,4	80,0
Natężenie promieniowania słonecznego [W/m <sup>2</sup> ]	67,4	75,3

Rozbieżności pomiędzy wynikami średnimi dla miasta Płocka i dla Polski stały się asumptem do przeprowadzenia analizy korelacyjnej umożliwiającej określenie zależności pomiędzy temperaturą powietrza, wilgotnością względną powietrza i natężeniem promieniowania słonecznego w mieście Płocku oraz do sformułowania wniosków.

### Charakterystyka użytego systemu pomiarowego

W skład użytego systemu pomiarowego wchodziła stacja meteorologiczna LB-719, umieszczonej na dachu Politechniki Warszawskiej Filii w Płocku, który w linii prostej znajduje się w odległości 2 km od kombinatu rafineryjno-petrochemicznego PKN ORLEN S.A. oraz komputera klasy PC z dedykowanym oprogramowaniem umożliwiającym akwizycję wyników pomiarów.



Rys. 1. Klatka termohigrometru (czujnika temperatury i względnej wilgotności powietrza LB-710R) stacji meteorologicznej LB-719 zamontowanej na dachu Politechniki Warszawskiej Filii w Płocku



Rys. 2. Czujnik promieniowania słonecznego LB-900/LB-901 stacji meteorologicznej LB-719 zamontowanej na dachu Politechniki Warszawskiej Filii w Płocku

Stacja meteorologiczna LB-719 jest wyposażona w termohigrometr (czujnik temperatury i względnej wilgotności powietrza) LB-710R (Rys. 1.) oraz czujnik promieniowania słonecznego LB-900/LB-901 (Rys. 2.). Właściwości metrologiczne wyżej wymienionych czujników zawarto w Tabeli 2. i Tabeli 3.

Tabela 2. Właściwości metrologiczne termohigrometru LB-710R stacji meteorologicznej LB-719 [2]

POMIAR TEMPERATURY	
Niepewność typu B	0,1°C
Zakres pomiaru / odczytu	od -40,0°C do +85,0°C
Rozdzielczość	0,1°C
POMIAR WILGOTNOŚCI WZGLĘDNEJ	
Niepewność typu B	2,0% w zakresie od 10% do 90% 4,0% poza zakresem od 10% do 90%
Zakres pomiaru / odczytu	od 0,0% do 100,0%
Rozdzielczość	0,1%
PARAMETRY TRANSMISJI SZEREGOWEJ W STANDARDZIE RS232C	
<ul style="list-style-type: none"> <li>▪ przepływność bitowa 300 bps</li> <li>▪ 7 bitów informacyjnych, bez kontroli parzystości</li> <li>▪ 1 bit stopu pracujący w standardzie tzw. cyfrowej pętli prądowej</li> </ul>	

Tabela 3. Właściwości metrologiczne czujnika promieniowania słonecznego LB-900/LB-901 stacji meteorologicznej LB-719 [2]

POMIAR PROMIENIOWANIA SŁONECZNEGO	
Dokładność	0,05%
Rozdzielczość	0,1 W/m <sup>2</sup>
PARAMETRY TRANSMISJI SZEREGOWEJ W STANDARDZIE RS232C	
<ul style="list-style-type: none"> <li>▪ przepływność bitowa 300 bps</li> <li>▪ 7 bitów informacyjnych, bez kontroli parzystości</li> <li>▪ 1 bit stopu pracujący w standardzie tzw. cyfrowej pętli prądowej</li> </ul>	

### Analiza metrologiczna uzyskanych wyników pomiarów z wykorzystaniem teorii niepewności pomiaru

Na podstawie teorii niepewności pomiaru można stwierdzić, że każdy wynik pomiaru jest obarczony błędem (systematycznym lub przypadkowym). Oznacza to, że podczas prezentacji wyniku powinno się podać błąd, który mu towarzyszy (błąd bezwzględny i względny).

Zalecany jest przeprowadzanie bardziej szczegółowej analizy z wykorzystaniem teorii niepewności, aby określić wpływ grupy czynników na wartości wyników pomiarów. Korzystając z metody zawartej w opracowaniu [4] oraz danych zawartych w Tabeli 2. i Tabeli 3. Wykonano stosowne obliczenia, których wyniki zawarto w Tabeli 4.

Tabela 4. Wyniki obliczeń niepewności pomiaru temperatury powietrza, względnej wilgotności powietrza oraz natężenia promieniowania słonecznego w mieście Płocku

Parametr	Wyniki z uwzględnieniem rozdzielczości		
	Temperatura powietrza [°C]	Względna wilgotność powietrza [%]	Natężenie promieniowania słonecznego [W/m <sup>2</sup> ]
Standardowa niepewność typu A	0,1	0,3	3,2
Standardowa niepewność typu B	0,1	2,0	0,0
Standardowa niepewność całkowita	0,2	2,0	3,2
Standardowa niepewność rozszerzona i odpowiadający jej przedział ufności dla $t_{\alpha} = 1,2816$	0,2 (3,5; 3,7)	2,6 (82,1; 82,7)	4,1 (63,2; 71,5)
Standardowa niepewność rozszerzona i odpowiadający jej przedział ufności dla $t_{\alpha} = 1,9600$	0,3 (3,5; 3,8)	4,0 (81,9; 82,9)	6,3 (61,0; 73,7)
Standardowa niepewność rozszerzona i odpowiadający jej przedział ufności dla $t_{\alpha} = 3,0902$	0,5 (3,6; 3,8)	6,2 (81,6; 83,2)	10,0 (57,4; 77,3)

Standardowa niepewność całkowita jest sumą geometryczną standardowej niepewności typu A (reprezentująca błędy przypadkowe) oraz standardowej niepewności typu B (reprezentująca błędy instrumentalne). Porównując jej wartość ze składowymi można stwierdzić, że na standardową niepewność całkowitą:

- w przypadku pomiaru temperatury powietrza, w tym samym stopniu wpływają standardowa niepewność typu A oraz standardowa niepewność typu B,
- w przypadku pomiaru względnej wilgotności powietrza, znacząco wpływa standardowa niepewność typu B. Oznacza to, że podczas pomiarów dominują błędy instrumentalne wynikające z konstrukcji czujnika wilgotności,
- w przypadku pomiaru natężenia promieniowania słonecznego, jedynie wpływa standardowa niepewność typu A. Oznacza to, że szybkie zmiany w czasie nasłonecznienia czujnika promieniowania słonecznego, znacząco różniły się od wartości średniej.

Odnosząc wartości średnie dla Polski (Tabela 1.) do uzyskanych przedziałów ufności, można stwierdzić, że tylko w przypadku pomiaru natężenia promieniowania słonecznego, wynik średni zawiera się w przedziale ufności dla miasta Płocka i to jedynie dla współczynnika rozszerzenia rozkładu  $t_{\alpha} = 3,0902$ , który odpowiada prawdopodobieństwu bliższemu jedności, które symbolizuje zdarzenie prawie pewne. Taką wartość współczynnika rozszerzenia rozkładu stosuje się podczas analizy wyników pomiarów z obszaru zdrowia i bezpieczeństwa.

### Charakterystyka i zastosowanie wybranych graficznych metod analizy korelacyjnej dyskretnych sygnałów pomiarowych o charakterze losowy

W pomiarach próbkujących mamy do czynienia z dyskretnymi sygnałami pomiarowymi, które opisują wielkości fizyczne jedynie w ściśle określonych (dyskretnych) chwilach czasu. Tego typu sygnały uzyskuje się w wyniku operacji próbkowania (dyskretyzacji) ciągłego sygnału pomiarowego.

Zjawiska losowe, do których zalicza się klimat oraz czynniki stanowiące o nim (temperatura powietrza, wilgotność względna powietrza oraz natężenie promieniowania słonecznego), nie mogą być opisane dokładnymi zależnościami matematycznymi, ponieważ faktem jest, że wynik każdej obserwacji sygnału jest jedyny (nieodtwarzalny). Innymi słowy, wynik każdej konkretnej obserwacji (dyskretny sygnał pomiarowy o charakterze losowym) jest tylko jednym z wielu możliwych wyników, jakie mogą się zdarzyć [5].

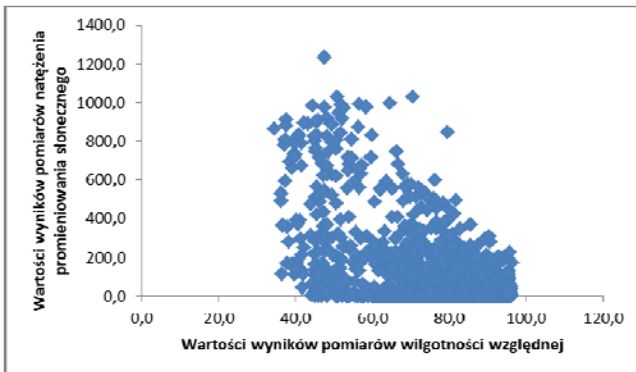
Tego typu sygnały opisuje się za pomocą parametrów statystycznych, a zależności między nimi bada się stosując graficzne metody analizy korelacyjnej [6,7]:

- diagram rozrzutu (wykres rozrzutu, wykres korelacji): zbiór punktów na płaszczyźnie, które odpowiadają uporządkowanym parom liczb  $(x, y)$ , gdzie:  $x, y$  – obserwacje niezależnych zmiennych losowych  $X, Y$ . W przypadku tej metody zakłada się, że dwuwymiarowy rozkład badanych cech zmiennych losowych  $X$  i  $Y$  jest normalny lub zbliżony do normalnego. Należy podkreślić, że diagram rozrzutu nie bada związku przyczynowo-skutkowego zachodzącego między zmiennymi losowymi, lecz związek korelacyjny. Na jego podstawie (dokonując wzrokowej oceny) można określić rodzaj zależności (korelację dodatnią, ujemną, liniową, krzywoliniową),
- nożyce korelacji (krzywe korelacji): proste w układzie współrzędnych, na podstawie których można określić stopień skorelowania zmiennych losowych  $X$  i  $Y$ . Opisują je równania:

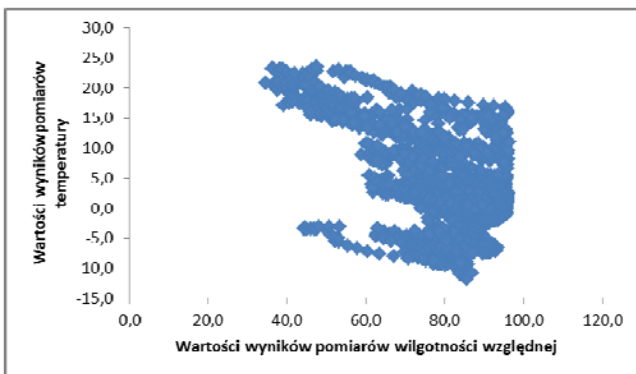
$$(1) \quad x = m_1 + r_{XY} \cdot \frac{\sigma_x}{\sigma_y} \cdot (y - m_2)$$

$$(2) \quad y = m_2 + r_{XY} \cdot \frac{\sigma_y}{\sigma_x} \cdot (x - m_1)$$

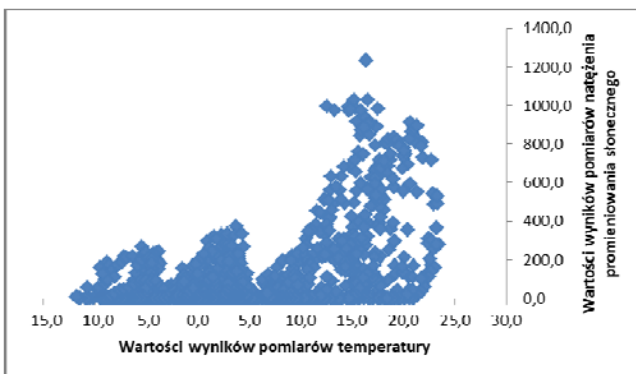
gdzie:  $m_1$  – średnia arytmetyczna z elementów zmiennej losowej  $X$ ,  $m_2$  – średnia arytmetyczna z elementów zmiennej losowej  $Y$ ,  $r_{XY}$  – współczynnik korelacji liniowej Pearsona,  $\sigma_x$  – odchylenie standardowe zmiennej losowej  $X$ ,  $\sigma_y$  – odchylenie standardowe zmiennej losowej  $Y$ . Na ich podstawie można określić rodzaj oraz siłę zależności (brak korelacji, korelację umiarkowaną bądź silną).



Rys. 3. Diagram korelacji wyników pomiarów wilgotności względnej powietrza i natężenia promieniowania słonecznego



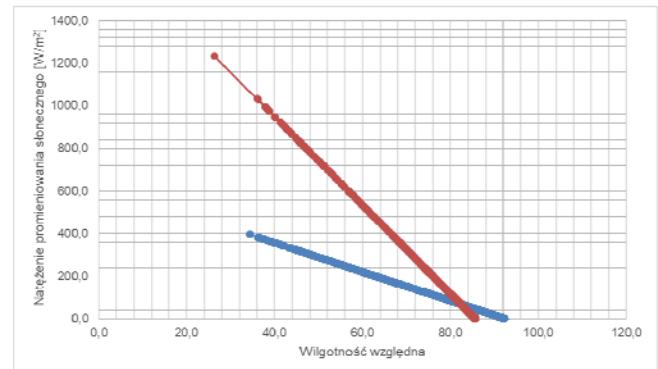
Rys. 4. Diagram korelacji wyników pomiarów wilgotności względnej powietrza i temperatury powietrza



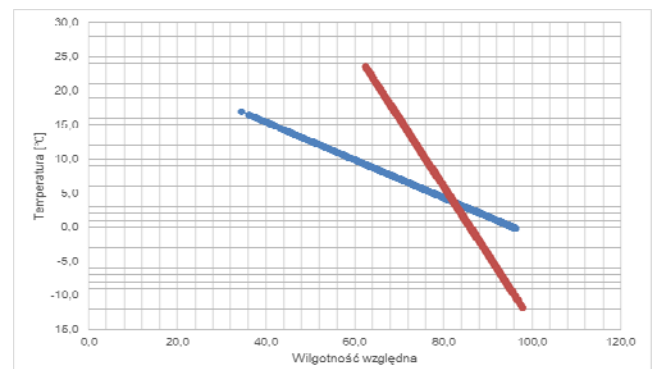
Rys. 5. Diagram korelacji wyników pomiarów temperatury powietrza i natężenia promieniowania słonecznego

Na podstawie uzyskanych wyników pomiarów temperatury powietrza, wilgotności względnej powietrza oraz natężenia promieniowania słonecznego, wyznaczono diagramy rozrzutu (Rys. 3., Rys. 4. i Rys. 5.). Następnie sformułowano wnioski dotyczące rodzaju zależności:

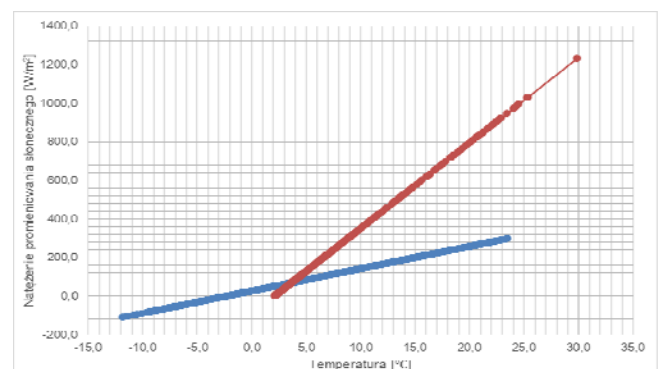
- pomiędzy wilgotnością względną powietrza i natężeniem promieniowania słonecznego (Rys. 3.), występuje korelacja ujemna, gdyż analizowane wartości wielkości fizycznych zmieniają się w przeciwnym kierunku (wzrost jednej powoduje zmniejszenie się drugiej),
- pomiędzy wilgotnością względną powietrza i temperaturą powietrza (Rys. 4.), występuje korelacja ujemna, gdyż analizowane wartości wielkości fizycznych zmieniają się w przeciwnym kierunku (wzrost jednej powoduje zmniejszenie się drugiej),
- pomiędzy temperaturą powietrza i natężeniem promieniowania słonecznego (Rys. 5.), występuje korelacja dodatnia, gdyż analizowane wartości wielkości fizycznych zmieniają się w tym samym kierunku (wzrost jednej powoduje wzrost drugiej).



Rys. 6. Nożyce korelacji wyników pomiarów wilgotności względnej powietrza i natężenia promieniowania słonecznego



Rys. 7. Nożyce korelacji wyników pomiarów wilgotności względnej powietrza i temperatury powietrza



Rys. 8. Nożyce korelacji wyników pomiarów temperatury powietrza i natężenia promieniowania słonecznego

Na podstawie uzyskanych wyników pomiarów temperatury powietrza, wilgotności względnej powietrza oraz natężenia promieniowania słonecznego, wyznaczono nożyce korelacji (Rys. 6., Rys. 7. i Rys. 8.). Następnie sformułowano wnioski dotyczące rodzaju i siły zależności:

- pomiędzy wilgotnością względną powietrza i natężeniem promieniowania słonecznego (Rys. 6.), występuje korelacja ujemna, ponieważ współczynniki kierunkowe prostych są ujemne. Dodatkowo, kąt między prostymi dąży do 45°, co wskazuje na umiarkowaną zależność pomiędzy analizowanymi wielkościami fizycznymi,
- pomiędzy wilgotnością względną powietrza i temperaturą powietrza (Rys. 7.), występuje korelacja ujemna, ponieważ współczynniki kierunkowe prostych są ujemne. Dodatkowo, kąt między prostymi dąży do 45°, co wskazuje na umiarkowaną zależność pomiędzy analizowanymi wielkościami fizycznymi,
- pomiędzy temperaturą powietrza i natężeniem promieniowania słonecznego (Rys. 8.), występuje korelacja dodatnia, ponieważ współczynniki kierunkowe prostych są dodatnie. Dodatkowo, kąt między prostymi dąży do 45°, co wskazuje na umiarkowaną zależność pomiędzy analizowanymi wielkościami fizycznymi.

### Charakterystyka i zastosowanie analitycznej metody badania korelacji dyskretnych sygnałów pomiarowych o charakterze losowy z wykorzystaniem współczynnik korelacji liniowej Pearsona

Do analitycznego badania zależności korelacyjnej często stosuje się współczynnik korelacji liniowej Pearsona, zdefiniowany jako iloraz kowariancji i iloczynu odchyłań standardowych przyjętych zmiennych losowych  $X$  i  $Y$  [7].

W praktyce, podczas korelacyjnej analizy danych (dyskretnych sygnałów pomiarowych o charakterze losowym), dla których trudno jest określić prawdopodobieństwo ich występowania, korzysta się z estymatora współczynnika korelacji liniowej Pearsona, opisanego wzorem [7]:

$$(3) \quad r_{xy} = \frac{\sum_{i=1}^n (x_i - m_1) \cdot (y_i - m_2)}{\sqrt{\sum_{i=1}^n (x_i - m_1)^2} \cdot \sqrt{\sum_{i=1}^n (y_i - m_2)^2}}$$

Współczynnik korelacji liniowej Pearsona oraz jego estymata przyjmuje wartości z przedziału  $[-1;1]$ . Jeżeli współczynnik korelacji bądź jego estymata przyjmuje wartość  $-1$  lub  $+1$ , to między zmiennymi losowymi  $X$  i  $Y$  istnieje ścisła zależność w postaci funkcji liniowej (zmienne losowe są skorelowane). Jeżeli przyjmuje wartość  $0$ , wówczas zmienne losowe  $X$  i  $Y$  nie są skorelowane (nie ma między nimi żadnej zależności). Im wartość bezwzględna współczynnika korelacji bądź jego estymata jest bliższa jedności, tym mocniejsza jest korelacja pomiędzy zmiennymi losowymi  $X$  i  $Y$ .

Tabela 5. Sposób interpretacji stopnia skorelowania zmiennych losowych na podstawie wartość współczynnika korelacji liniowej Pearsona bądź jego estymaty [8]

Wartość współczynnika korelacji liniowej Pearsona bądź jego estymaty	Sposób interpretacji stopnia skorelowania zmiennych losowych
$ r_{xy}  < 0,2$	brak związku liniowego
$0,2 \leq  r_{xy}  < 0,4$	słaba zależność
$0,4 \leq  r_{xy}  < 0,7$	umiarkowana zależność
$0,7 \leq  r_{xy}  < 0,9$	dość silna zależność
$ r_{xy}  \geq 0,9$	bardzo silna zależność

W celu ułatwienia interpretowania stopnia skorelowania zmiennych losowych sformułowano w pracy [8] progi dla wartości współczynnika korelacji liniowej Pearsona bądź jego estymaty i przypisano im stopnie zależności między zmiennymi losowymi (Tabela 5.).

Na podstawie uzyskanych wyników pomiarów temperatury powietrza, wilgotności względnej powietrza oraz natężenia promieniowania słonecznego, wyznaczono estymaty współczynnika korelacji liniowej Pearsona. Następnie sformułowano wnioski dotyczące rodzaju i siły zależności:

- pomiędzy wilgotnością względną powietrza i natężeniem promieniowania słonecznego, występuje umiarkowana korelacja ujemna ( $r_{XY} = -0,52659$ ),
- pomiędzy wilgotnością względną powietrza i temperaturą powietrza, występuje umiarkowana korelacja ujemna ( $r_{XY} = -0,57196$ ),
- pomiędzy temperaturą powietrza i natężeniem promieniowania słonecznego, występuje umiarkowana korelacja dodatnia ( $r_{XY} = 0,510331$ ).

### Wnioski

Niniejsza praca zawiera zestawienie wyników analiz statystycznych. Dotyczą one wpływu emisji pary wodnej do atmosfery, w okresie zimy i wiosny, przez kombinat rafinerijno-petrochemiczny PKN ORLEN S.A. Na ich podstawie można sformułować konkluzje dotyczące czynników stanowiących o klimacie w mieście Płocku. Są one następujące:

- średnia temperatura powietrza oraz średnia wilgotność względna powietrza są większe od średnich wartości na terenie Polski. Średnia moc promieniowania słonecznego wykazuje odwrotną zależność,
- emitowana do atmosfery para wodna wywołuje wzrost wilgotności względnej powietrza, co przyczynia się do zmniejszenia natężenia promieniowania słonecznego oraz temperatury powietrza (ujemna i umiarkowana korelacja),
- emitowana do atmosfery para wodna wywołuje wzrost temperatury powietrza, co przyczynia się do zwiększenia natężenia promieniowania słonecznego (dodatnia i umiarkowana korelacja).

Podsumowując, można stwierdzić, że ograniczenie emisji pary wodnej do atmosfery spowoduje normalizację klimatu na terenie miasta Płocka. Oznacza to, że średnie wartości temperatury powietrza, wilgotności względnej powietrza, mocy promieniowania świetlnego, będą dążyły do średnich wartości w Polsce.

**Autorzy:** dr inż. Sławomir Andrzej Torbus, prof. uczelni, Uniwersytet Kazimierza Wielkiego w Bydgoszczy, Instytut Matematyki, al. Powstańców Wielkopolskich 2, 85-090 Bydgoszcz, E-mail: slawomir.torbus@ukw.edu.pl; dr inż. Karolina Mroczyńska, Uniwersytet Kazimierza Wielkiego w Bydgoszczy, Instytut Matematyki, al. Powstańców Wielkopolskich 2, 85-090 Bydgoszcz, E-mail: karolina.mroczyńska@ukw.edu.pl; Patryk Szlichta, E-mail: patryk.szlichta@student.ukw.edu.pl;

### LITERATURA

- [1] Niedźwiedz T., Słownik meteorologiczny, Instytut Meteorologii i Gospodarki Wodnej, Warszawa (2003)
- [2] <https://www.label.pl/po/rek710r.html>
- [3] Cedro A., Walczakiewicz Sz., Podstawy meteorologii i klimatologii Polski, ZAPOL Sobczyk, Szczecin (2017)
- [4] Jaworski J. M., Wyrażanie niepewności pomiaru. Przewodnik, Główny Urząd Miar, Warszawa (1999)
- [5] Bendat J. S., Piersol A. G., Metody analizy i pomiaru sygnałów losowych, PWN, Warszawa (1976)
- [6] Sobczak M., Statystyka, PWN, Warszawa (2022)
- [7] Gerstenkorn T., Śródka T., Kombinatoryka i rachunek prawdopodobieństwa, PWN, Warszawa (1980)
- [8] Cohen J., Statistical Power Analysis for the Behavioral Sciences, Routledge, New York (1988)