

Politechnika Częstochowska
Wydział Infrastruktury i Środowiska

mgr inż. Mariusz Filak

Rozprawa doktorska

**Predykcja stężeń zanieczyszczeń powietrza
w perceptronowych modelach regresyjnych**

Promotor:
dr hab. Szymon Hoffman

Częstochowa, 2022

Podziękowania

Autor dziękuje Głównemu Inspektoratowi Ochrony Środowiska w Warszawie i Wojewódzkiemu Inspektoratowi Ochrony Środowiska w Katowicach za udostępnienie danych z regionalnego monitoringu powietrza, które wykorzystano w analizie.

Finansowanie

Badania naukowe zostały sfinansowane z subwencji statutowej Wydziału Infrastruktury i Środowiska Politechniki Częstochowskiej, przyznanej w latach 2021-2022: BS/PB-400-301/21, BS/PB-400-301/22.

Spis treści

Streszczenie	
Summary	
Wykaz stosowanych skrótów i oznaczeń	
1. WSTĘP.....	strona 1
2. CZĘŚĆ TEORETYCZNA.....	strona 3
2.1. Akty prawne regulujące funkcjonowanie monitoringu powietrza w ramach Państwowego Monitoringu Środowiska	strona 3
2.2. Wpływ zanieczyszczeń powietrza na zdrowie człowieka.....	strona 5
2.3. Sztuczne sieci neuronowe.....	strona 7
2.4. Modele jakości powietrza.....	strona 10
3. CEL I TEZY PRACY.....	strona 13
4. CZĘŚĆ BADAWCZA.....	strona 14
4.1. Charakterystyka stacji monitoringu powietrza i danych wykorzystanych w analizie.....	strona 14
4.1.1. Charakterystyka wybranych stacji monitoringu powietrza.....	strona 14
4.1.2. Charakterystyka danych wykorzystanych w analizie.....	strona 15
4.1.3. Zmienność sezonowa stężeń zanieczyszczeń powietrza.....	strona 18
4.2. Wybór typu sieci neuronowej.....	strona 22
4.3. Wybór architektury sieci MLP.....	strona 24
4.4. Metodyka badań.....	strona 28
4.4.1. Przygotowanie danych.....	strona 28
4.4.2. Metodyka modelowania.....	strona 29
4.4.3. Podział zbiorów pełnozakresowych na podzbiory.....	strona 30
4.4.4. Ocena dokładności modeli.....	strona 33
4.4.5. Epizody.....	strona 34
5. WYNIKI I ICH DYSKUSJA.....	strona 35
5.1. Modele RVS.....	strona 35
5.1.1. Wyniki modelowania stężeń O ₃	strona 35
5.1.2. Wyniki modelowania stężeń NO.....	strona 41
5.1.3. Wyniki modelowania stężeń NO ₂	strona 45
5.1.4. Wyniki modelowania stężeń SO ₂	strona 50
5.1.5. Wyniki modelowania stężeń PM ₁₀	strona 54

5.1.6. Wyniki modelowania stężeń CO.....	strona 60
5.2. Modele PVS.....	strona 62
5.2.1. Wyniki modelowania stężeń O ₃	strona 62
5.2.2. Wyniki modelowania stężeń NO.....	strona 67
5.2.3. Wyniki modelowania stężeń NO ₂	strona 72
5.2.4. Wyniki modelowania stężeń SO ₂	strona 76
5.2.5. Wyniki modelowania stężeń PM ₁₀	strona 80
5.2.6. Wyniki modelowania stężeń CO.....	strona 84
6. PODSUMOWANIE I DYSKUSJA.....	strona 87
7. WNIOSKI.....	strona 90
7.1. Główne wnioski.....	strona 90
7.2. Wnioski szczegółowe.....	strona 90
BIBLIOGRAFIA.....	strona 92
Spis zamieszczonych tabel.....	strona 104
Spis zamieszczonych rysunków.....	strona 106

Streszczenie

Jakość powietrza ocenia się na podstawie danych z monitoringu powietrza. Dane z monitoringu często nie są wystarczająco kompletne, aby przeprowadzić ocenę jakości powietrza. Do wypełnienia luk pomiarowych można wykorzystać modele predykcyjne, które umożliwiają aproksymację brakujących danych. Modele takie wykorzystują dane historyczne i zależności między mierzonymi wielkościami, w tym stężeniami zanieczyszczeń powietrza i czynnikami meteorologicznymi. Znane modele predykcyjne jakości powietrza nie są dokładne, dlatego ważne jest, aby szukać modeli, które dają niższy błąd aproksymacji. Zastosowanie sztucznych sieci neuronowych pozwala zmniejszyć błąd predykcji w porównaniu z klasycznymi metodami regresji. We wcześniejszych badaniach do aproksymacji stężeń wybranego zanieczyszczenia stosowano pojedynczy model regresji w całym zakresie stężeń. W niniejszym badaniu założono, że do predykcji można wykorzystać grupę modeli zamiast pojedynczego modelu. W tym podejściu każdy model z grupy jest dedykowany innemu podzakresowi stężeń modelowanego zanieczyszczenia. Celem analizy było sprawdzenie, czy takie podejście poprawi jakość modelowania. Do badań wykorzystano zbiór wieloletnich danych zarejestrowanych na dwóch stacjach monitoringu powietrza w województwie śląskim – Zabrze i Złoty Potok. Do stworzenia predykcyjnych modeli regresji zostały wykorzystane dane uśrednione w okresach 1-godzinnych, obejmujące stężenia podstawowych zanieczyszczeń powietrza i pomiary meteorologiczne. Błędy predykcji dla modeli podzakresowych porównano z odpowiednimi błędami obliczonymi dla jednego pełnozakresowego modelu regresji. Stwierdzono, że zastosowanie modeli podzakresowych pozwala zmniejszyć błąd modelowania stężeń podstawowych zanieczyszczeń powietrza.

Summary

Air quality is evaluated on the basis of air monitoring data. Monitoring data is often not complete enough to carry out an air quality assessment. To fill the measurement gaps, predictive models can be used, which enable the approximation of missing data. Predictive models use historical data and relationships between measured variables, including air pollutant concentrations and meteorological factors. The known predictive air quality models are not accurate, so it is important to look for models that give a lower approximation error. The use of artificial neural networks allows to reduce the prediction error compared to classical regression methods. In previous studies, a single regression model over the entire concentration range was used to approximate the concentrations of a selected pollutant. In this study, it was assumed that a group of models instead of a single model could be used for prediction. In this approach, each model from the group is dedicated to a different sub-range of the concentration of the modeled pollutant. The aim of the analysis was to check whether this approach would improve the quality of modeling. A long-term data set recorded at 2 air monitoring stations in Silesian Voivodeship (South Poland) was used in the examination. Hourly data of basic air pollutants concentrations and meteorological parameters were used to create predictive regression models. The prediction errors for the sub-range models were compared with the corresponding errors calculated for one full-range regression model. It was found that the application of sub-range models allows to reduce the modeling error of basic air pollutants.

Wykaz stosowanych skrótów i oznaczeń

Symbole zmiennych:

D – data w postaci numerycznej

G – godzina w postaci numerycznej

O₃ – średnie godzinne stężenie O₃, μg/m³

NO – średnie godzinne stężenie NO, μg/m³

NO₂ – średnie godzinne stężenie NO₂, μg/m³

SO₂ – średnie godzinne stężenie SO₂, μg/m³

CO – średnie godzinne stężenie CO, mg/m³

PM10 – średnie godzinne stężenie PM10, μg/m³

WS – średnia godzinna prędkość wiatru, m/s

T – średnia godzinna temperatura powietrza, °C

I – średnia godzinne natężenie promieniowania słonecznego, W/m²

W_{wzg} – średnia godzinna wilgotność względna, %

Rodzaj sieci neuronowej:

MLP – perceptron wielowarstwowy

RBF – sieć o radialnych funkcjach bazowych

Inne oznaczenia:

MAE - średni błąd bezwzględny

RMSE – pierwiastek z błędu średniokwadratowego

r – współczynnik korelacji Pearsona

BFGS – algorytm Broyden-Fletcher-Goldfarb-Shanno

Oznaczenia modeli:

RVS (Real Values Sorting) – modele uzyskane dla podzakresów otrzymanych w wyniku podziału całego zakresu przypadków na określoną liczbę podzakresów. Przez podziałem przypadki uszeregowano według rosnących wartości stężeń rzeczywistych wybranego zanieczyszczenia

RVS-1/1-ZAB – model pełnozakresowy utworzony dla wszystkich przypadków zarejestrowanych na stacji monitoringu powietrza w Zabrze

RVS-1/2-ZAB – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane z Zabrze

RVS-2/2-ZAB – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane z Zabrze

RVS-1/4-ZAB – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

RVS-2/4-ZAB – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

RVS-3/4-ZAB – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

RVS-4/4-ZAB – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

RVS-1/8-ZAB – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-2/8-ZAB – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-3/8-ZAB – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-4/8-ZAB – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-5/8-ZAB – model utworzony dla piątego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-6/8-ZAB – model utworzony dla szóstego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-7/8-ZAB – model utworzony dla siódmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-8/8-ZAB – model utworzony dla ósmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

RVS-1/1-ZLP – model pełnozakresowy utworzony dla wszystkich przypadków zarejestrowanych na stacji monitoringu powietrza w Żółym Potoku

RVS-1/2-ZLP – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane ze Żółtego Potoku

RVS-2/2-ZLP – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane ze Żółtego Potoku

RVS-1/4-ZLP – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

RVS-2/4-ZLP – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

RVS-3/4-ZLP – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

RVS-4/4-ZLP – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

RVS-1/8-ZLP – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

RVS-2/8-ZLP – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Złotego Potoku

RVS-3/8-ZLP – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Złotego Potoku

RVS-4/8-ZLP – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Złotego Potoku

RVS-5/8-ZLP – model utworzony dla piątego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Złotego Potoku

RVS-6/8-ZLP – model utworzony dla szóstego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Złotego Potoku

RVS-7/8-ZLP – model utworzony dla siódmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Złotego Potoku

RVS-8/8-ZLP – model utworzony dla ósmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Złotego Potoku

PVS (Predicted Values Sorting) – modele uzyskane dla podzakresów otrzymanych w wyniku podziału całego zakresu przypadków na określoną liczbę podzakresów. Przed podziałem przypadki uszeregowano według rosnących wartości stężeń predykcyjnych wybranego zanieczyszczenia.

PVS-1/1-ZAB – model pełnozakresowy utworzony dla wszystkich przypadków zarejestrowanych na stacji monitoringu powietrza w Zabrze

PVS-1/2-ZAB – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane z Zabrze

PVS-2/2-ZAB – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane z Zabrze

PVS-1/4-ZAB – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

PVS-2/4-ZAB – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

PVS-3/4-ZAB – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

PVS-4/4-ZAB – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane z Zabrze

PVS-1/8-ZAB – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

PVS-2/8-ZAB – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

PVS-3/8-ZAB – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

PVS-4/8-ZAB – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrze

PVS-5/8-ZAB – model utworzony dla piątego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrza

PVS-6/8-ZAB – model utworzony dla szóstego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrza

PVS-7/8-ZAB – model utworzony dla siódmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrza

PVS-8/8-ZAB – model utworzony dla ósmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane z Zabrza

PVS-1/1-ZLP – model pełnozakresowy utworzony dla wszystkich przypadków zarejestrowanych na stacji monitoringu powietrza w Żółtym Potoku

PVS-1/2-ZLP – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane ze Żółtego Potoku

PVS-2/2-ZLP – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na dwie części, dane ze Żółtego Potoku

PVS-1/4-ZLP – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

PVS-2/4-ZLP – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

PVS-3/4-ZLP – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

PVS-4/4-ZLP – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na cztery części, dane ze Żółtego Potoku

PVS-1/8-ZLP – model utworzony dla pierwszego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

PVS-2/8-ZLP – model utworzony dla drugiego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

PVS-3/8-ZLP – model utworzony dla trzeciego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

PVS-4/8-ZLP – model utworzony dla czwartego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

PVS-5/8-ZLP – model utworzony dla piątego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

PVS-6/8-ZLP – model utworzony dla szóstego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

PVS-7/8-ZLP – model utworzony dla siódmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

PVS-8/8-ZLP – model utworzony dla ósmego z podzakresów, uzyskanych po podziale zbioru przypadków na osiem części, dane ze Żółtego Potoku

1. WSTĘP

Jednym z najważniejszych problemów współczesnego świata jest zanieczyszczenie powietrza atmosferycznego. Ekspozycja na zanieczyszczone powietrze – szczególnie na obszarach miejskich – może prowadzić do wielu chorób, takich jak: udar, nowotwory tchawicy, oskrzeli i płuc. Jednak zanieczyszczenie powietrza oddziałuje negatywnie nie tylko na zdrowie, ale również na ekonomię, powodując skrócenie przewidywanej długości życia, wzrost wydatków związanych z leczeniem chorób układu oddechowego oraz zmniejszenie wydajności pracowników (Raport EEA, 2021). Do najbardziej szkodliwych dla zdrowia zanieczyszczeń należą cząstki aerozolu (PM10 oraz PM2,5), dwutlenek azotu (NO₂) oraz ozon (O₃). Szacuje się, że drobne cząstki pyłu zawieszonego PM2,5 skracają średnią długość życia mieszkańców Unii Europejskiej o ponad osiem miesięcy. Duże obawy wzbudza również rakotwórczy benzo(a)piren, którego niebezpiecznie wysokie poziomy obserwuje się w wielu obszarach miejskich, zwłaszcza w środkowej i wschodniej Europie (Raport EEA, 2020). Od czasu opublikowania przez Światową Organizację Zdrowia w 2005 roku wytycznych dotyczących jakości powietrza (World Health Organization, 2006), pojawiło się wiele prac naukowych dokładniej opisujących negatywny wpływ zanieczyszczeń powietrza na zdrowie przy jeszcze niższych stężeniach tych zanieczyszczeń. W związku z tym Światowa Organizacja Zdrowia rekomenduje zaostrzenie norm dotyczących dopuszczalnych poziomów stężeń zanieczyszczeń powietrza (World Health Organization, 2021). Wiele polskich miast zajmuje wysokie miejsca na listach europejskich miast z najbardziej zanieczyszczonym powietrzem (Wielgosiński i in., 2018; Sówka i in., 2019). Mając na uwadze zagrożenia związane z wpływem zanieczyszczeń na zdrowie, niezwykle ważną staje się kwestia monitoringu jakości powietrza, który ma na celu – między innymi – dostarczać informacji o przekroczeniach dopuszczalnych poziomów stężeń zanieczyszczeń, czy też stopniach realizacji programów społeczno-gospodarczych mających na celu poprawę stanu środowiska.

W Polsce monitoringiem poziomów stężeń zanieczyszczeń powietrza zajmuje się Główny Inspektorat Ochrony Środowiska i jego oddziały regionalne (GIOŚ, 2021). Poprzez system stacji monitoringu wyposażonych w automatyczne analizatory gromadzone są dane opisujące, między innymi, stężenia podstawowych zanieczyszczeń powietrza, takich jak O₃, NO, NO₂, SO₂, CO i pyłów zawieszonych PM2,5 oraz PM10. Stacje monitoringu powietrza zbierają również dane opisujące parametry

meteorologiczne takie jak temperatura powietrza, prędkość i kierunek wiatru, natężenie promieniowania słonecznego, wilgotność względna oraz ciśnienie atmosferyczne. Stacje monitoringu powietrza dokonują pomiarów w sposób ciągły, a wyniki rejestrowane są w postaci 1-godzinnych wartości średnich. Tak zarejestrowane dane są podstawą do oceny jakości powietrza na terenie reprezentowanym przez daną stację monitoringu powietrza. Niestety, dane zbierane na stacji monitoringu powietrza często nie są kompletne, a normy jakości powietrza dopuszczają możliwość wykorzystania danych do oceny jakości powietrza zwykle pod warunkiem, że ich kompletność nie jest mniejsza niż 90% (Rozporządzenie, 2020). Przyczyny luk w danych pomiarowych są różne, najczęściej są to przerwy w zasilaniu w energię elektryczną, uszkodzenia analizatorów, zakłócenia w transmisji danych, procedury serwisowe i kalibracyjne (Latini, Passerini, 2004; Hauck i in., 1999). Zbyt duża niekompletność może być przyczyną braku oceny stanu powietrza atmosferycznego na terenie reprezentowanym przez daną stację monitoringu powietrza. W takich sytuacjach nie jest spełniony podstawowy cel funkcjonowania stacji monitoringu powietrza – ocena jakości powietrza (Hoffman, Jasiński, 2009). Brakujące dane można aproksymować wykorzystując techniki modelowania oparte na eksploracji danych historycznych. Do tego celu, często wykorzystywane są modele regresyjne. Tradycyjne metody mogą być zastąpione przez metody oparte na sieciach neuronowych. Przeprowadzone w pracy badania opierają się na takich metodach aproksymacji brakujących stężeń.

2. CZĘŚĆ TEORETYCZNA

2.1. Akty prawne regulujące funkcjonowanie monitoringu powietrza w ramach Państwowego Monitoringu Środowiska

Państwowy Monitoring Środowiska został powołany Ustawą o Państwowej Inspekcji Ochrony Środowiska i zgodnie z art. 23 tej ustawy „Jest to system pomiarów, ocen oraz prognoz stanu środowiska, realizowany przez jednostki organizacyjne organów administracji państwowej i rządowej, organów gmin, jak również przez szkoły wyższe i podmioty gospodarcze” (Ustawa, 1991). Ustawa określała także cele Państwowego Monitoringu Środowiska, a wśród nich zwiększenie skuteczności działań na rzecz ochrony środowiska poprzez gromadzenie, analizowanie i udostępnianie zebranych danych opisujących stan środowiska i zmian w nim zachodzących. Początkowo, zwięźle określona charakterystyka i cele przedstawione we wspomnianej ustawie były jedynym obszarem prawnym, w którym działał system pomiarów, ocen i prognoz, dotyczących stanu środowiska. Zadania przeprowadzane w ramach systemu monitoringu środowiska sprecyzowane były w wieloletnich programach Państwowego Monitoringu Środowiska (Program PMŚ, 2006; Program PMŚ, 2009; Program PMŚ, 2012; Program PMŚ, 2015). Na skutek nowelizacji Ustawy o Inspekcji Ochrony Środowiska z 2018 roku (Ustawa, 2017) zmianie uległy uwarunkowania dotyczące realizacji zadań Państwowego Monitoringu Środowiska. Zgodnie z nowymi przepisami, zasoby i zadania Państwowego Monitoringu Środowiska realizowane do końca 2018 roku przez Wojewódzkie Inspektoraty Ochrony Środowiska zostały przeniesione do Głównego Inspektoratu Ochrony Środowiska. Na skutek tej zmiany, od 1 stycznia 2019 roku zadania określone w programach Państwowego Monitoringu Środowiska realizowane są wyłącznie przez Głównego Inspektora Ochrony Środowiska (GIOŚ). Zmiana struktury funkcjonalnej PMŚ miała na celu efektywniejsze zarządzanie wyspecjalizowaną wiedzą w zakresie badania i ocen stanu środowiska oraz skuteczniejsze informowanie społeczeństwa i władz zarządzających o stanie i zmianach zachodzących w środowisku.

Obecnie, aktualnym programem jest Strategiczny Program PMŚ na lata 2020-2025 (Program PMŚ, 2020). W ramach PMŚ funkcjonuje podsystem monitoringu jakości powietrza. Podstawą tego podsystemu jest wykonywanie pomiarów stężeń zanieczyszczeń powietrza i ich rejestrowanie. Monitoring powietrza funkcjonuje

w dwóch trybach wykonywania pomiarów: jako monitoring automatyczny i manualny. Zadania PMS w zakresie monitoringu jakości powietrza to, między innymi:

- badania i roczne oceny jakości powietrza w strefach,
- pięcioletnia ocena jakości powietrza na potrzeby ustalenia odpowiedniego sposobu dokonywania rocznych ocen jakości powietrza,
- informowanie o ryzyku wystąpienia przekroczenia lub wystąpieniu przekroczenia poziomu informowania, poziomu alarmowego, poziomu dopuszczalnego i poziomu docelowego substancji w powietrzu,
- monitoring benzo(a)pirenu i wybranych metali ciężkich w pyłe zawieszonym PM10,
- wspomaganie systemu ocen jakości powietrza metodami modelowania matematycznego.

Sposób dokonywania pomiarów i wymagania, jakie muszą spełniać pozyskane wyniki reguluje obowiązujące Rozporządzenie Ministra Klimatu i Środowiska (Rozporządzenie, 2020), które w obecnej formie jest efektem nowelizacji poprzednich rozporządzeń (Rozporządzenie, 2012; Rozporządzenie, 2018). W najnowszym rozporządzeniu wprowadzono zmiany kompetencyjne wynikające z nowej struktury organizacyjnej Inspekcji Ochrony Środowiska, które zostały wprowadzone Ustawą z dnia 20 lipca 2018 r. O zmianie ustawy o Inspekcji Ochrony Środowiska, zgodnie z którą obowiązek prowadzenia pomiarów jakości powietrza prowadzonych w ramach Państwowego Monitoringu Środowiska przeniesiony został z Wojewódzkiego Inspektora Ochrony Środowiska na Głównego Inspektora Ochrony Środowiska (Rozporządzenie, 2020). Rozporządzenia te określają minimalny procent ważnych danych uzyskanych w wyniku pomiarów ciągłych (90 % dla dwutlenku siarki, dwutlenku azotu, tlenku węgla, pyłu zawieszonego oraz 90 % dla ozonu i tlenku azotu w okresie letnim i 75 % w okresie zimowym). Udziały procentowe dotyczą pojedynczych pomiarów, uśrednionych odpowiednio do okresu uśredniania wyników pomiarów, dla którego określono poziomy dopuszczalne, poziomy docelowe lub poziomy celów długoterminowych substancji w powietrzu. Wymagania określające minimalną ilość ważnych danych nie uwzględniają utraty danych wynikających z powodu okresowego sprawdzania i kalibracji aparatury pomiarowej lub konserwacji urządzeń pomiarowych.

Rozporządzenie Ministra Klimatu i Środowiska z dnia 11 grudnia 2020 r. (Rozporządzenie, 2020) określa też metody referencyjne poboru próbek i oznaczania stężeń, między innymi, dla dwutlenku azotu i tlenków azotu (Norma PN-EN 14211:2013-02); dwutlenku siarki (Norma PN-EN 14212:2013-02); ozonu (Norma PN-EN

14625:2013-02), pyłu zawieszonego PM_{2,5} i PM₁₀ (Norma PN-EN 12341:2014-07) i tlenku węgla (Norma PN-EN 14626:2013-02). Wymienione rozporządzenia wynikają z podstawowej polskiej ustawy dotyczącej ochrony środowiska – Ustawy z 27 kwietnia 2001 r. Prawo Ochrony Środowiska (Ustawa, 2001), która określa zasady ochrony środowiska, warunki korzystania ze środowiska oraz obowiązki administracji publicznej związane z ochroną środowiska i powstały w oparciu o Dyrektywy Europejskie (Dyrektywa, 2004; Dyrektywa, 2008; Dyrektywa, 2015).

2.2. Wpływ zanieczyszczeń powietrza na zdrowie człowieka

W Polsce, z powodu niskiej świadomości i zrozumienia problemu przez społeczeństwo, kwestia negatywnego wpływu zanieczyszczeń powietrza na zdrowie człowieka jest często ignorowana. Powodem tego może być niewystarczająca wiedza na temat tego ważnego zagadnienia. Do efektywnego rozwiązania problemu zanieczyszczenia powietrza konieczne są stanowcze działania podjęte we wszystkich obszarach odpowiedzialnych za emisję zanieczyszczeń. Do obszarów tych można zaliczyć przemysł, energetykę, transport, gospodarstwa domowe, ale również ogół społeczeństwa, którego postawa i świadomość ma fundamentalne znaczenie dla powodzenia działań mających na celu ograniczenie emisji zanieczyszczeń powietrza (Jędrak i in., 2017).

Istnieje wiele publikacji naukowych opisujących badania, które udowadniają istnienie zależności między ekspozycją na pewne zanieczyszczenia powietrza a występowaniem rozmaitych defektów zdrowotnych (Brunekreef, 1997; Krewski, 2009; Yamazaki i in., 2015; Tétreault i in., 2016). Początkowo, aby dostrzec ujemny wpływ zanieczyszczeń powietrza na zdrowie, konieczne było wystąpienie zdarzeń o katastrofalnej skali. Przykładem mogą być wydarzenia, które miały miejsce w dniach 1-5 grudnia 1930 roku w dolinie Mozy (wschodnia Belgia), kiedy to wysokie stężenie zanieczyszczeń powietrza oraz niekorzystne warunki meteorologiczne spowodowały śmierć 63 osób. Śmiertelność i dolegliwości oddechowe ustały wraz z poprawą warunków pogodowych. Firket, analizując ten incydent, zasugerował, że gdyby podobne warunki wystąpiły w Londynie, liczba ofiar mogłaby wynieść nawet 3200 osób (Firket, 1936). Wspomniane zjawisko, czyli połączenie dymu i mgły dały początek słowu smog, które zaczęto używać w odniesieniu do zjawisk występujących w Londynie na początku

lat pięćdziesiątych (Wielgosiński, Czerwińska, 2020). Właśnie w Londynie miała miejsce największa katastrofa środowiskowa tego typu, kiedy to smog utrzymujący się od 5 do 9 grudnia 1952 roku, powstały z powodu gęstej mgły i wyemitowanej do atmosfery dużej ilości pyłu zawieszonego oraz dwutlenku siarki, stał się przyczyną ok. 4000 zgonów. Wyraźnie podwyższona umieralność związana z chorobami układu oddechowego i krążeniowego utrzymywała się jeszcze przez wiele tygodni (Logan, 1953). Współczesne badania wskazują, że całkowita liczba ofiar londyńskiego smogu z grudnia 1952 roku może sięgać 12 tys. osób (Bell, Davis, 2001; Hunt i in., 2003; Polivka, 2018). Innym przykładem poważnego incydentu smogowego z ubiegłego wieku mogą być wydarzenia z października 1948 roku, które miały miejsce w amerykańskim mieście Donora, do dzisiaj określane jako największa katastrofa związana z zanieczyszczeniem powietrza w historii Stanów Zjednoczonych. Wysokie stężenia zanieczyszczeń, których źródłem był miejscowy przemysł związany z produkcją stali i cynku oraz niekorzystne warunki pogodowe spowodowały śmierć około 20 osób (Vallero, 2014; Vallero, 2019). Incydent ten został szczegółowo opisany w raporcie opublikowanym w 1949 roku (Schrenk i in., 1949). W literaturze można znaleźć też dowody, które potwierdzają wpływ jakości powietrza na zdrowie w odwrotny sposób – mniejszy poziom zanieczyszczeń powoduje spadek liczby zgonów. Przykładem może być strajk w zakładzie metalurgicznym w Utah Valley, USA, który trwał od sierpnia 1986 do września 1987 roku. W tym okresie średnie stężenie pyłu zawieszonego PM10 spadło o około $15 \mu\text{g}/\text{m}^3$. Na skutek tego umieralność zmniejszyła się o 3,2% (Pope, Dockery, 2006). Spadła również liczba hospitalizacji spowodowanej chorobami układu oddechowego, spadek ten był wyraźny zwłaszcza wśród dzieci do lat pięciu (Pope, 1991). Innym przykładem może być zakazanie dystrybucji i sprzedaży paliw stałych, które wprowadził irlandzki rząd 1 września 1990 roku na terenie miasta Dublin. Jakość powietrza w Dublinie uległa pogorszeniu w latach osiemdziesiątych XX w., kiedy to zrezygnowano z oleju wykorzystywanego do celów grzewczych na korzyść paliw stałych, głównie węgla kamiennego. Po wprowadzeniu wspomnianego zakazu obrotu paliwami stałymi zaobserwowano wyraźną poprawę jakości powietrza. Odnotowano spadek stężenia zanieczyszczeń pyłowych o $35,6 \mu\text{g}/\text{m}^3$, a w efekcie spadek liczby zgonów o około 116 z powodu chorób układu oddechowego i o około 243 z powodu chorób układu krążeniowego (Clancy i in., 2002).

Wspomniane przykłady związane są głównie z wysokimi stężeniami pyłu zawieszonego, którego szkodliwość zależy od wielkości cząstek oraz jego składu chemicznego. Pył zawieszony PM10 podrażnia głównie gardło, krtani oraz tchawicę,

podczas gdy pył PM_{2,5} (cząstki o średnicy aerodynamicznej do 2,5 μm), z racji mniejszych rozmiarów, trafia poprzez pęcherzyki płucne do krwioobiegu, a stamtąd do organów wewnętrznych (Atkinson i in., 2001; Cembrzyńska i in., 2012). Z kolei skład chemiczny pyłu zawieszonego zależy od jego źródła (Rogula-Kozłowska, 2015; Wielgosiński, Czerwińska, 2020) i wpływa na właściwości pyłu, w tym gęstość, reaktywność oraz toksyczność (Chlebowska-Styś i in., 2016). Szczególną cechą pyłu zawieszonego jest jego zdolność do adsorbowania na swojej powierzchni innych zanieczyszczeń, na przykład związków z grupy wielopierścieniowych węglowodorów aromatycznych (WWA) (Lee i in., 2005; Mantis i in., 2005; Wieczorek i in., 2011; Jamhari i in., 2014; Kozielska, Rogula-Kozłowska, 2014). Niektóre z tych związków, jak na przykład benzo(a)piren, mają udowodnione działanie kancerogenne (Sims i in., 1974; Collins i in., 1991; Armstrong i in., 2004; Punethai in., 2022). Pozostałe zanieczyszczenia gazowe, których źródło nie zawsze jest pochodzenia antropogenicznego (Heaviside i in., 2021), również wywierają ujemny wpływ na zdrowie ludzi, powodując choroby układu oddechowego i krążeniowego, a w efekcie przedwczesną śmierć (Nuvolone i in., 2017; Li i in., 2020; Wang i in., 2021; De Vries, 2021; Orellano i in., 2021; Farzad i in., 2021; Khomenko i in., 2021;). Biorąc pod uwagę powyższe zagrożenia związane z emisją zanieczyszczeń powietrza, niezwykle ważna staje się kwestia kontroli jakości powietrza.

2.3. Sztuczne sieci neuronowe

Ze względu na korelacje obserwowane między stężeniami zanieczyszczeń powietrza, zasadne jest wykorzystywanie do predykcji metod eksplorujących ukrytą wiedzę w danych historycznych. Proste metody statystyczne rozwijano stopniowo wraz z postępem i upowszechnianiem technik komputerowych (Bennis i in., 1997; Nosal i in., 2000). Od kilkadziesiąt lat ulepszano koncepcję sztucznego neuronu, która po raz pierwszy pojawiła się w latach 40-tych ubiegłego wieku (McCulloch, Pitts, 1943). Przez wiele lat nie potrafiono wykorzystać pojedynczego neuronu do zadań związanych z aproksymacją. Postępem okazało się wprowadzenie przez Franka Rosenblatta w 1958 roku pojęcia perceptronu, czyli najprostszej sieci neuronowej, składającej się z jednego lub wielu niezależnych neuronów (Rosenblatt, 1958). Perceptron odznaczał się innowacyjną możliwością uczenia się. Jednak już w 1969 roku ukazała się praca, w której udowodniono ograniczenia perceptronów o liniowej funkcji aktywacji (Minsky, Papert,

1969). Wykazano w niej, że jednowarstwowa sieć zdolna jest jedynie do klasyfikacji wzorców separowalnych liniowo i stwierdzono, że dowolna wielowarstwowa sieć neuronowa może zostać zredukowana i zastąpiona przez porównywalną sieć jednowarstwową. W praktyce oznaczało to, że klasyfikacja wzorców nieseparowalnych liniowo przy pomocy sieci neuronowej z liniową funkcją aktywacji jest niemożliwa. Po opublikowaniu tych twierdzeń, nastąpił okres stagnacji w dziedzinie rozwoju sieci neuronowych. Sytuacja odwróciła się w 1986 roku, kiedy ukazała się praca „Learning representations by back propagating errors” (Rumelhart i in, 1986), która na nowo wzbudziła zainteresowanie sieciami. Po tej publikacji nastąpił renesans sieci neuronowych. Szybko pojawiła się znaczna liczba publikacji i badań wykorzystujących sieci neuronowe w rozmaitych dziedzinach nauki i techniki. Pojawiło się pojęcie konekcjonizmu, jako teorii badawczej opisującej proces przetwarzania informacji (sygnałów) przez ludzki mózg. Obecnie konekcjonizm przejawia się w postulatcie użycia sztucznych sieci neuronowych jako narzędzia odwzorowującego przepływ sygnałów w ludzkim mózgu. Tak jak neurologia dokonuje tych badań w sposób bezpośredni, konekcjonizm w teorii sztucznej inteligencji robi to w pośredni sposób, modelując procesy mózgowie w sieciach neuronowych (Drozdek, 2000). W odniesieniu właśnie do sieci neuronowych i innych technik komputerowych, zdolnych do naśladowania ludzkiej inteligencji, zaczęto używać terminu „sztuczna inteligencja”.

Elementem wyróżniającym sieć neuronową jest możliwość wykorzystywania jej jako komputerowego narzędzia informatycznego służącego do praktycznego rozwiązywania przeróżnych problemów, bez konieczności przedstawiania matematycznych formuł (Tadeusiewicz, Szaleniec, 2015). Zastosowania praktyczne sieci neuronowych mogą być różne, na przykład rozpoznawanie obiektów (Kwasigroch, Grochowski, 2018; Błaszczuk, Janowski, 2021), wykrywanie wad w zautomatyzowanych procesach produkcyjnych (Adamczak, 2021; Siljama i in., 2021), prognozowanie remontów maszyn i urządzeń (Rosienkiewicz, 2019), określanie mechanicznych właściwości skał na podstawie właściwości fizycznych (Słota-Valim, 2018), czy też modelowanie stężeń zanieczyszczeń powietrza (Hoffman, 2004; Siwek, Osowski, 2011; Wilkosz i in., 2021). Zaletą sieci neuronowych jest fakt, że podczas korzystania z sieci neuronowej nie zachodzi konieczność odwoływania się do teoretycznych założeń na temat rozwiązywanego problemu.

Najbardziej jednak charakterystyczną właściwością sieci neuronowej jest jej umiejętność uczenia się na podstawie przedstawionych przykładów oraz zdolność do

uogólniania zdobytych informacji, czyli generalizacji. Pojęcie generalizacji związane jest z procesem uczenia, który oparty jest na zbiorze uczącym, zawierającym przypadki uczące z poprawnymi rozwiązaniami. Sieć po przeprowadzeniu procesu uczenia musi rozwiązywać zadania podobne do tych, których była nauczona, jednak nie takich samych. Przenoszenie nabytej wiedzy na nowe przypadki nazywane jest właśnie generalizacją (Tadeusiewicz, Szaleniec, 2015). Aby oszacować możliwości generalizacji sieci neuronowej, należy ją poddać testowi na podstawie niezależnych danych, które nie zostały wykorzystane w procesie uczenia. W tym celu z części danych tworzy się tzw. zbiór testowy. Aby uzyskać obiektywną ocenę generalizacji sieci, ważne jest aby poszczególne zestawy danych zbioru testowego nie zawierały bardzo podobnych przypadków (Krogh, 2008). Obok zbioru uczącego i testowego, do stworzenia sieci neuronowej niezbędny jest też zbiór walidacyjny, stanowiący część zbioru uczącego i służący do przeprowadzenia okresowej walidacji, mającej na celu ocenę zdolności sieci do generalizacji. Proces walidacji przeprowadza się w celu zapobiegania wystąpienia zjawiska przeuczenia, które występuje wtedy, gdy sieć uzyskuje niski stopień generalizacji (Tadeusiewicz, Szaleniec, 2015).

Proces uczenia sieci neuronowej można podzielić na trzy etapy. W pierwszym dane wejściowe są wysyłane do tzw. warstwy wejściowej sieci. Neurony warstwy wejściowej współpracują wzajemnie, wykorzystując aktualne parametry sieci i obliczają wynik końcowy, będący odpowiedzią sieci. W drugim etapie odpowiedź sieci jest porównywana z prawidłową odpowiedzią (wartością rzeczywistą) wchodzącą w skład zbioru uczącego dla rozważanego przykładu. W tym etapie zostaje obliczony błąd modelowania na wyjściu sieci. Związane jest to z najpopularniejszą metodą uczenia sieci neuronowej, wykorzystującą algorytm wstecznej propagacji błędów (Patterson, 1996; Haykin, 1994; Fausett, 1994). Trzeci etap to korekta parametrów dla wszystkich neuronów w sieci. Przeprowadzenie korekty opiera się na obliczeniach gradientu błędów sieci. W wyniku tej korekty zmienia się wiedza sieci reprezentowana przez wymienione parametry (Tadeusiewicz, 2015).

Z punktu widzenia niniejszej pracy, najistotniejszą i najbardziej użyteczną w praktyce cechą sieci neuronowych jest ich umiejętność predykcji. Zagadnienia związane z predykcją można podzielić na dwie kategorie: klasyfikację oraz regresję. Celem klasyfikacji jest przyporządkowanie danego przypadku do jednej z określonych wcześniej klas. Przykładem może być rozpatrywanie wniosków kredytowych (wniosek może być przyjęty lub odrzucony), wykrywanie nieszczelności w sieciach sanitarnych

(wykrycie lub wykluczenie), rozpoznawanie podpisów (autentyczny lub fałszywy). W powyższych przykładach na wyjściu oczekiwana jest pojedyncza zmienna nominalna. W przypadku regresji, celem jest oszacowanie wartości zmiennej wyjściowej, gdy wartości zmiennych wejściowych są znane. Do rozwiązywania zagadnień regresyjnych mogą być wykorzystywane sieci neuronowe typu MLP (Multi-Layer Perceptron), RBF (Radial Basis Functions) oraz sieci liniowe. W odróżnieniu od problemów rozwiązywanych za pomocą klasyfikacji, na wyjściu modelu regresyjnego oczekiwana jest zmienna o charakterze numerycznym. Sieci neuronowe mogą rozwiązywać zagadnienia związane z regresją wielowymiarową i umożliwiają również badanie zależności między kilkoma zmiennymi wejściowymi (objaśniającymi), a zmienną wyjściową (objaśnianą) (Siderska, 2021). Sieci liniowe są najprostszym typem sieci neuronowej i są odpowiednie do tworzenia prostych modeli opisujących liniowe zależności. Sieci MLP składają się z neuronów o nieliniowych charakterystykach, które mogą być ułożone w kilka warstw. Sieć MLP z reguły składa się z warstwy wejściowej, warstwy ukrytej oraz warstwy wyjściowej. Sieci typu RBF odznaczają się rozbudowaną warstwą ukrytą składającą się z neuronów o gausoidalnej charakterystyce (Tadeusiewicz, 2009).

2.4. Modelowanie jakości powietrza

Zjawiska fizyczne i chemiczne zachodzące w atmosferze uchodzą za niezwykle złożone. W latach pięćdziesiątych XX wieku opracowano pierwsze modele opisujące jakość powietrza, jednak od tego czasu nastąpił znaczny postęp w dziedzinie badań nad atmosferą. Wciąż istnieje wiele zagadnień związanych z modelowaniem jakości powietrza, które – z racji skali problemu – wymagają dalszych badań (Juda-Rezler, 2010)

Modele służące do opisu jakości powietrza można podzielić na dwie grupy: deterministyczne i statystyczne. Modele deterministyczne charakteryzują się tworzeniem prognoz dla wybranego obszaru, co jest możliwe nawet w przypadku braku zarejestrowanych poziomów stężeń na tym obszarze (IMGW, 2022). W wyniku modelowania matematycznego rozprzestrzeniania się zanieczyszczeń możliwe jest przedstawienie prognozy dla konkretnego obszaru. Przykładem praktycznego wykorzystania modelu deterministycznego może być system FAPPS, stworzony na potrzeby tworzenia prognoz dla województwa małopolskiego (Hajto i in., 2012).

Drugą grupą modeli są modele statystyczne, które mogą być tworzone na podstawie danych historycznych. Różnią się one od modeli deterministycznych mniejszymi wymaganiami dotyczącymi mocy obliczeniowych oraz tym, że mogą służyć do predykcji stężeń zanieczyszczeń powietrza w określonym punkcie. Przykładami modeli statystycznych mogą być modele wykorzystujące uczenie maszynowe, czyli między innymi modele sztucznych sieci neuronowych (Pawul, 2019; Pawul, Śliwka, 2016).

Modele deterministyczne opierają się na wykorzystaniu równań ciągłości masy, ruchu i pędu, uwzględniając w ten sposób procesy związane z konwekcją i dyfuzją. Przykładami modeli deterministycznych mogą być na przykład modele Eulera lub modele smugi Gaussa (Bogacki, Oleniacz, 2004). Z kolei modele statystyczne (stochastyczne) bazują na autokorelacjach występujących między stężeniami zanieczyszczeń powietrza oraz wpływie parametrów meteorologicznych na poziom stężeń zanieczyszczeń na określonym terenie. W związku z różnicami w charakterystyce modeli jakości powietrza, traktują one zanieczyszczenia w dwóch ujęciach: przestrzennych (punkt, objętość) oraz czasowych (chwila, odcinek czasu) (Tumidajski i in., 2009).

Do technik deterministycznych można zaliczyć modelowanie numeryczne, które stosowane jest w opisie zależności między emisją zanieczyszczeń, parametrami meteorologicznymi, stężeniami tła zanieczyszczeń i innymi czynnikami. Pomiar zanieczyszczeń powietrza dostarcza istotnych, ilościowych informacji o stężeniach zanieczyszczeń w otoczeniu i depozycji, ale opisują one jakość powietrza jedynie w określonych lokalizacjach i określonym czasie, bez podawania wskazówek dotyczących identyfikacji przyczyn złej jakości powietrza. Modelowanie może dać pełniejszy, deterministyczny opis jakości powietrza na podstawie analizy czynników i przyczyn (źródła emisji, procesy meteorologiczne oraz zmiany fizyczne i chemiczne). W związku z tym mogą stanowić istotną pomoc we wdrażaniu programów mających na celu poprawę jakości powietrza (Daly, Zanetti, 2007).

Wśród modeli statystycznych, popularne w modelowaniu jakości powietrza są metody wykorzystujące metody autoregresyjne oparte na analizie szeregów czasowych (Kumar, Jain, 2009; Wang, Guo, 2009; Pohoata, Lungu, 2017; Zhang i in., 2017; Nadeem i in., 2020; Rekhi, Nagrath, Jain, 2020). Metody te polegają na badaniu liniowych zależności między zmiennymi objaśniającymi (danymi historycznymi opisującymi poziomy stężenie zanieczyszczeń powietrza), a zmienną objaśnianą (przewidywane stężenie). Do ich wad zaliczyć można wrażliwość na pojawiające się wartości

ekstremalne, które wpływają ujemnie na dokładność predykcji (Li, Ren, 2022). Od lat dziewięćdziesiątych ubiegłego wieku w modelowaniu powszechnie wykorzystuje się również sieci neuronowe (Osowski, 1996). Ze względu na możliwość rozpoznawania bardzo złożonych i nieliniowych zależności między zanieczyszczeniami powietrza, sztuczne sieci neuronowe są szeroko stosowane do zadań związanych z aproksymacją i prognozowaniem. Najczęściej wykorzystywanym do tych celów typem sieci neuronowej jest perceptron wielowarstwowy (Pawlak, Jarosławski, 2019). Oprócz łatwości użycia, sieci neuronowe często odznaczają się większą dokładnością niż tradycyjne metody statystyczne (Hadjiiski i in., 1999; Shams i in., 2020).

Badania dotyczące modelowania jakości powietrza przy wykorzystaniu sztucznych sieci neuronowych często skupiają się – między innymi – na zagadnieniach związanych z uzupełnianiem brakujących danych (Plaia, Bondi, 2006), predykcją stężeń zanieczyszczeń powietrza (Karppinen i in., 2000; Nagendra, Khare, 2005) i zmniejszeniem błędu modelowania (Hoffman, 2015; Hoffman, 2021). Propozycję modyfikacji metodyki zmierzającej do poprawy jakości modelowania zaprezentowano w niniejszej pracy.

3. CEL I TEZY PRACY

Prezentowana praca oparła się na kilku tezach, które można sformułować następująco:

- archiwalne dane pomiarowe mogą być wykorzystane do modelowania aktualnych stężeń zanieczyszczeń;
- do modelowania predykcyjnego można wykorzystać sztuczne sieci neuronowe, które stanowią wygodne i praktyczne narzędzie do modelowania, w porównaniu do tradycyjnych metod;
- istnieje możliwość zmniejszenia błędów modelowania regresyjnego, poprzez zastąpienie jednej sieci neuronowej przez kilka sieci, obejmujących różne sektory zakresu stężeń modelowanego zanieczyszczenia.

Głównym celem pracy było poprawienie dokładności predykcji stężeń zanieczyszczeń powietrza w neuronowych modelach regresyjnych poprzez zastosowanie wielu modeli predykcyjnych, utworzonych dla różnych podzakresów stężeń zanieczyszczeń powietrza.

4. CZĘŚĆ BADAWCZA

4.1. Charakterystyka stacji monitoringu powietrza i danych wykorzystanych w analizie

Do przeprowadzenia analizy wykorzystano dane pochodzące z lat 2011-2016, zarejestrowane na dwóch stacjach monitoringu powietrza. Obie stacje znajdowały się na terenie województwa śląskiego. Wybór konkretnych stacji monitoringu powietrza podyktowany był stosunkowo wysoką kompletnością danych zarejestrowanych na tych stacjach.

4.1.1. Charakterystyka wybranych stacji monitoringu powietrza

W analizie wykorzystano dane zarejestrowane na dwóch stacjach monitoringu powietrza działających w ramach systemu monitoringu jakości powietrza, którego administratorem jest obecnie Główny Inspektorat Ochrony Środowiska. Stacjom przypisano nazwy związane z ich położeniem – Zabrze oraz Złoty Potok. Położenie stacji na mapie Polski zobrazowano na rysunku 1, a ich charakterystykę w tabeli 1. Obie stacje znajdują się w województwie śląskim, stacja Zabrze – w strefie Aglomeracji Górnośląskiej, stacja Złoty Potok – w tzw. strefie śląskiej. Obie stacje są stacjami tła, jednak w związku z różnicami w typie reprezentowanego obszaru, na stacji Złoty Potok można spodziewać się niższych poziomów stężeń zanieczyszczeń pierwotnych, takich jak dwutlenek siarki, tlenki azotu, pył PM10 i tlenek węgla.



Rys. 1. Lokalizacja wybranych stacji monitoringu powietrza z zaznaczonymi największymi miastami na mapie Polski

Tab. 1. Charakterystyka wybranych stacji monitoringu powietrza (źródło: gios.gov.pl)

Stacja monitoringu	Adres stacji	Strefa	Typ stacji	Typ obszaru	Wysokość n.p.m.	Współrzędne geograficzne
Zabrze	Zabrze, ul. M.	Aglomeracja	tło	miejski	255 m	Φ 50,3165
	Skłodowskiej-Curie 34	Górnośląska				λ 18,772375
Złoty Potok	Złoty Potok, Leśniczówka Kamienna Góra	Strefa śląska	tło	pozamiejski	291 m	Φ 50,710889 λ 19,458797

4.1.2. Charakterystyka danych wykorzystanych w analizie

W analizie wykorzystano dane, które można podzielić na trzy grupy: dane opisujące poziomy stężenie zanieczyszczeń powietrza w postaci 1-godzinnych stężeń (tzw. stężenia chwilowe), dane opisujące parametry meteorologiczne oraz dane opisujące czas pomiaru w postaci daty (D) i godziny (G). W przypadku danych opisujących zanieczyszczenia powietrza były to stężenia O₃, NO, NO₂, SO₂, PM₁₀ oraz CO. Parametrami meteorologicznymi były prędkość wiatru (PW), temperatura powietrza (T), natężenie promieniowania słonecznego (I) i wilgotność względna (W_{wzg}). Na stacjach nie były rejestrowane wszystkie zmienne (brak danych dotyczących stężeń CO na stacji Złoty

Potok oraz wartości opisujących wilgotność względną na stacji Zabrze). Dostępność danych pozyskanych z wybranych stacji monitoringu powietrza przedstawiono w tabeli 2.

Tab. 2. Zestawienie danych zarejestrowanych na poszczególnych stacjach monitoringu powietrza (+ - mierzone, - - niemierzone)

Stacja monitoringu	D	G	O ₃	NO	NO ₂	SO ₂	PM10	CO	PW	T	I	W _{wzg.}
Zabrze	+	+	+	+	+	+	+	+	+	+	+	-
Złoty Potok	+	+	+	+	+	+	+	-	+	+	+	+

Podstawowe parametry statystyczne zarejestrowanych danych opisujących poziomy stężenie zanieczyszczeń powietrza zamieszczono w tabeli 3.

Tab. 3. Statystyka opisowa zbioru chwilowych stężeń monitorowanych zanieczyszczeń powietrza

Parametr	Jednostka	Zabrze						Złoty Potok					
		O ₃	NO	NO ₂	SO ₂	PM10	CO	O ₃	NO	NO ₂	SO ₂	PM10	CO
Kompletność	%	94,2	78,8	94,5	91,0	80,2	94,2	92,3%	87,1%	87,1%	88,5%	95,7%	-
Wartość minimalna	µg/m ³	0,7	0,0	1,0	0,1	0,1	0,07	1,1	0,0	0,0	0,0	1,0	-
Wartość maksymalna	µg/m ³	198,0	709,0	160,0	362,0	1187,0	9,02	212,6	50,0	82,0	116,0	298,0	-
Średnia	µg/m ³	42,3	10,2	24,4	16,8	51,0	0,60	59,5	1,1	9,6	7,4	27,8	-
Mediana	µg/m ³	37,0	3,0	20,0	10,0	33,8	0,42	56,0	0,97	7,0	4,6	22,3	-
Odchylenie standardowe	µg/m ³	32,4	24,2	16,8	19,4	61,1	0,59	31,0	1,8	7,9	8,1	20,7	-

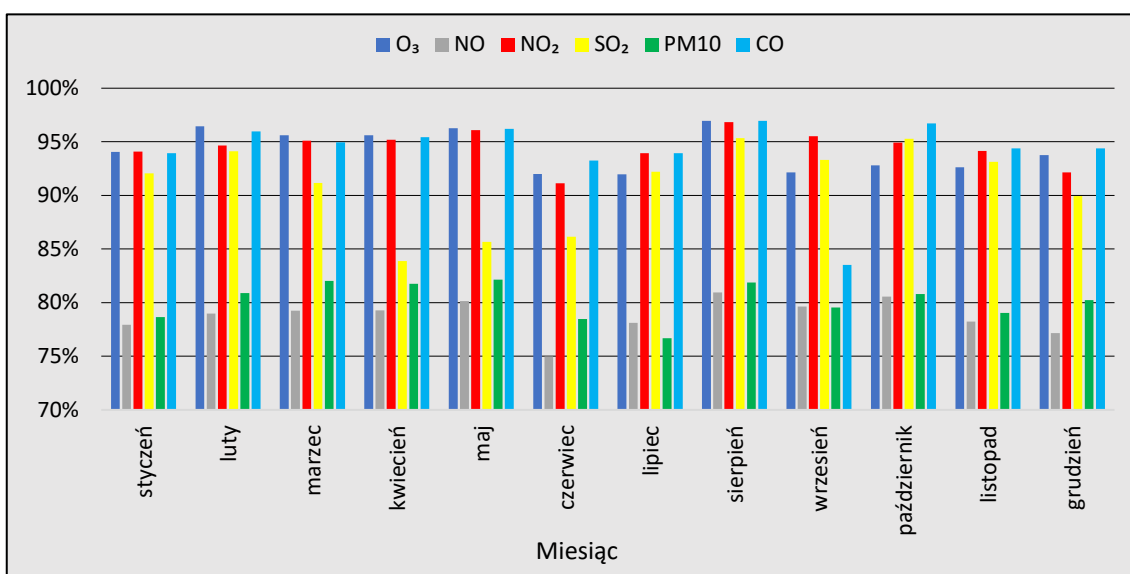
Pełna liczba przypadków stężeń chwilowych w okresie 2011-2016, z którego pochodzą dane wykorzystane w analizie, powinna wynosić 52608. W praktyce jednak taka 100-procentowa kompletność jest nieosiągalna. Powodem tego mogą być przerwy w dostawach energii elektrycznej niezbędnej do pracy oprzyrządowania stacji monitoringu powietrza, awarie aparatury pomiarowej czy systemu gromadzenia danych.

Braki w ciągłości danych mogą niekiedy trwać przez wiele dni, a nawet miesiące lub lat (Hoffman, 2013).

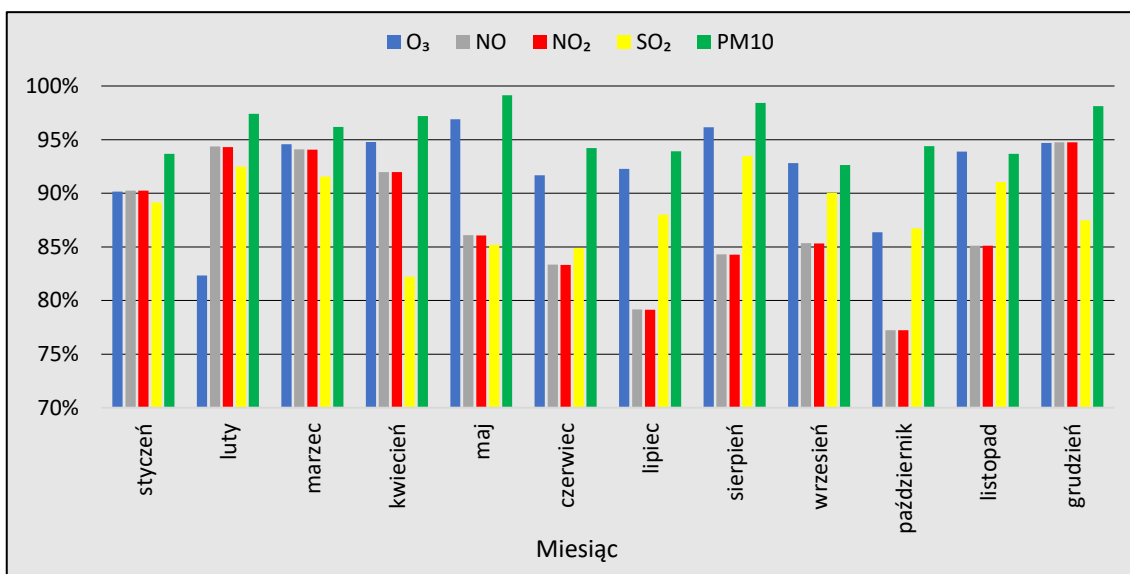
W tabeli 4 zamieszczono średnią procentową kompletność danych zarejestrowanych w okresie 2011-2016 w poszczególnych miesiącach. W przypadku większości zarejestrowanych zanieczyszczeń dane wykazują znaczną, ponad 90-procentową kompletność. Zdarzają się jednak miesiące, w których kompletność jest niższa, tzn. 75,0-81,0 % dla stężeń NO zarejestrowanych na stacji Zabrze, czy 76,7-82,1 % dla stężeń PM10 na tej samej stacji. Średnią kompletność danych w różnych miesiącach roku w rozpatrywanym okresie dla stacji Zabrze przedstawiono na rys. 2, a dla stacji Złoty Potok na rys. 3.

Tab. 4. Procentowa średnia kompletność danych zarejestrowanych w okresie 2011-2016 w poszczególnych miesiącach na stacjach Zabrze oraz Złoty Potok

Stacja monitoringu	Miesiąc	O ₃	NO	NO ₂	SO ₂	PM10	CO
Zabrze	Styczeń	94,1	77,9	94,1	92,1	78,7	93,9
	Luty	96,4	79,0	94,7	94,1	80,9	96,0
	Marzec	95,6	79,3	95,1	91,2	82,0	95,0
	Kwiecień	95,6	79,3	95,2	83,9	81,8	95,4
	Maj	96,3	80,2	96,1	85,7	82,1	96,2
	Czerwiec	92,0	75,0	91,1	86,1	78,5	93,3
	Lipiec	92,0	78,1	93,9	92,2	76,7	93,9
	Sierpień	97,0	81,0	96,8	95,3	81,9	97,0
	Wrzesień	92,2	79,6	95,5	93,3	79,6	83,5
	Październik	92,8	80,6	94,9	95,3	80,8	96,7
	Listopad	92,6	78,2	94,1	93,1	79,0	94,4
	Grudzień	93,8	77,2	92,1	89,9	80,2	94,4
Złoty Potok	Styczeń	90,2	90,2	90,2	89,1	93,7	-
	Luty	82,3	94,4	94,3	92,5	97,4	-
	Marzec	94,6	94,1	94,1	91,6	96,2	-
	Kwiecień	94,8	92,0	92,0	82,2	97,2	-
	Maj	96,9	86,1	86,1	85,2	99,1	-
	Czerwiec	91,7	83,3	83,3	84,9	94,2	-
	Lipiec	92,3	79,2	79,1	88,0	93,9	-
	Sierpień	96,1	84,3	84,3	93,5	98,4	-
	Wrzesień	92,8	85,3	85,3	90,0	92,6	-
	Październik	86,4	77,2	77,2	86,7	94,4	-
	Listopad	93,9	85,1	85,1	91,0	93,7	-
	Grudzień	94,7	94,7	94,7	87,5	98,1	-



Rys. 2. Procentowa średnia kompletność stężeń chwilowych zarejestrowanych na stacji Zabrze w okresie 2011-2016 dla poszczególnych miesięcy

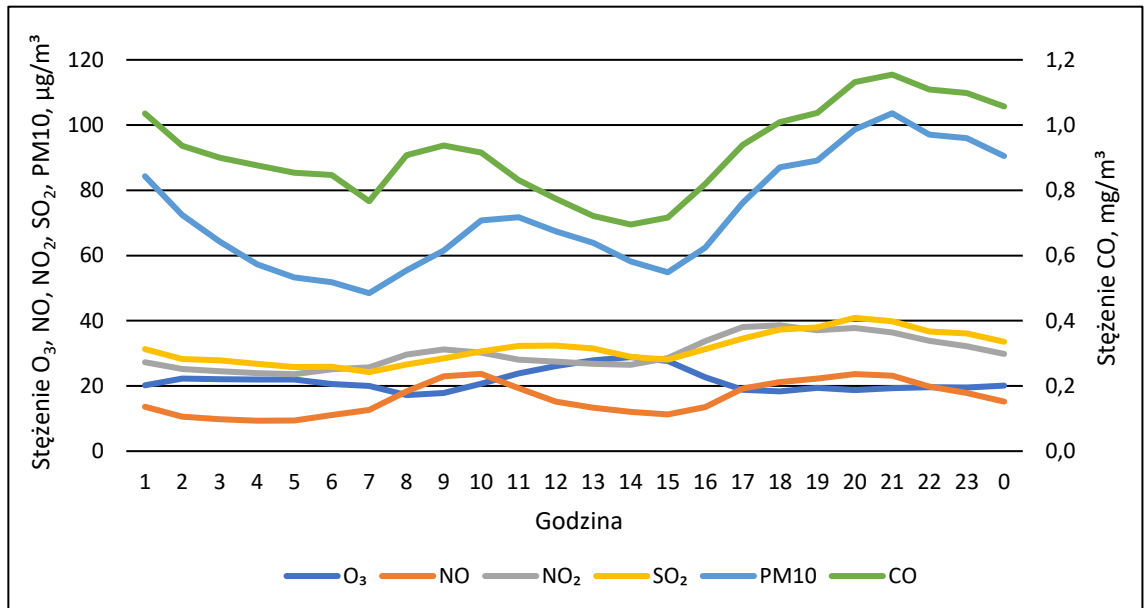


Rys. 3. Procentowa średnia kompletność stężeń chwilowych zarejestrowanych na stacji Złoty Potok w okresie 2011-2016 dla poszczególnych miesięcy

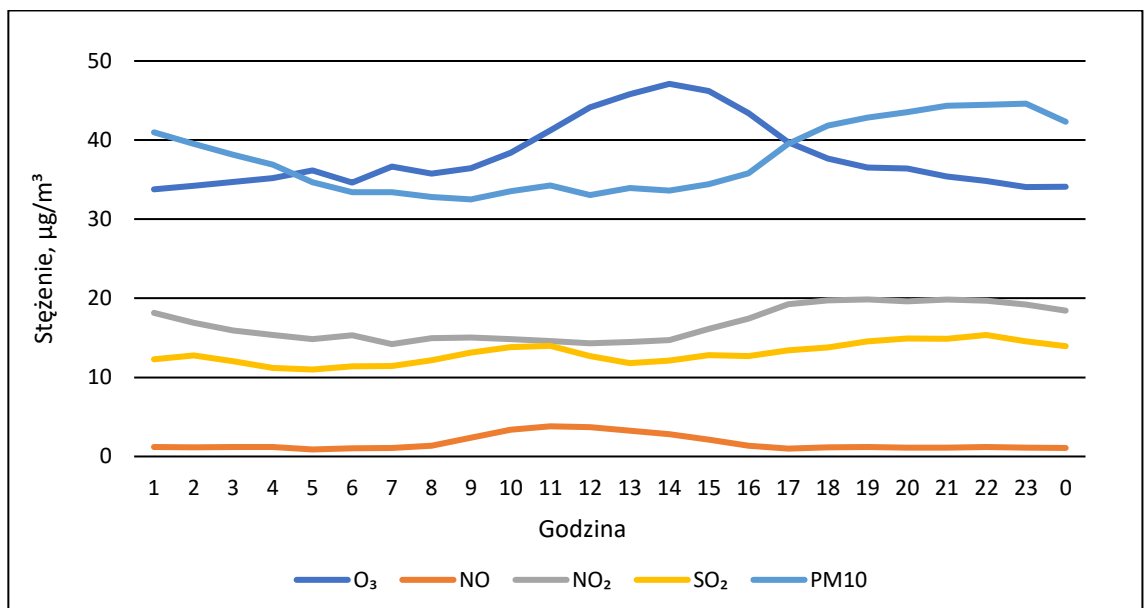
4.1.3. Zmienność sezonowa stężeń zanieczyszczeń powietrza

Surowe dane opisujące 1-godzinne stężenia zanieczyszczeń powietrza zostały przygotowane do analizy poprzez usunięcie przypadków z brakującymi danymi. Aby ocenić zmienność dobową stężeń zanieczyszczeń powietrza w zależności od sezonu w pozostawionym do analizy zbiorze danych uśredniono wartości 1-godzinne dla

wybranych miesięcy z okresu 2011-2016. Uzyskano w ten sposób średnie przebiegi stężeń chwilowych dla całego okresu pomiarowego w wybranych miesiącach. Porównano ze sobą typowe miesiące z okresu wiosenno-letniego (czerwiec i lipiec) oraz jesienno-zimowego (styczeń i grudzień). Na rys. 4 i 5 przedstawiono uśrednione dobowe przebiegi stężeń chwilowych podstawowych zanieczyszczeń powietrza dla stacji Zabrze i Złoty Potok w miesiącach styczeń i grudzień.

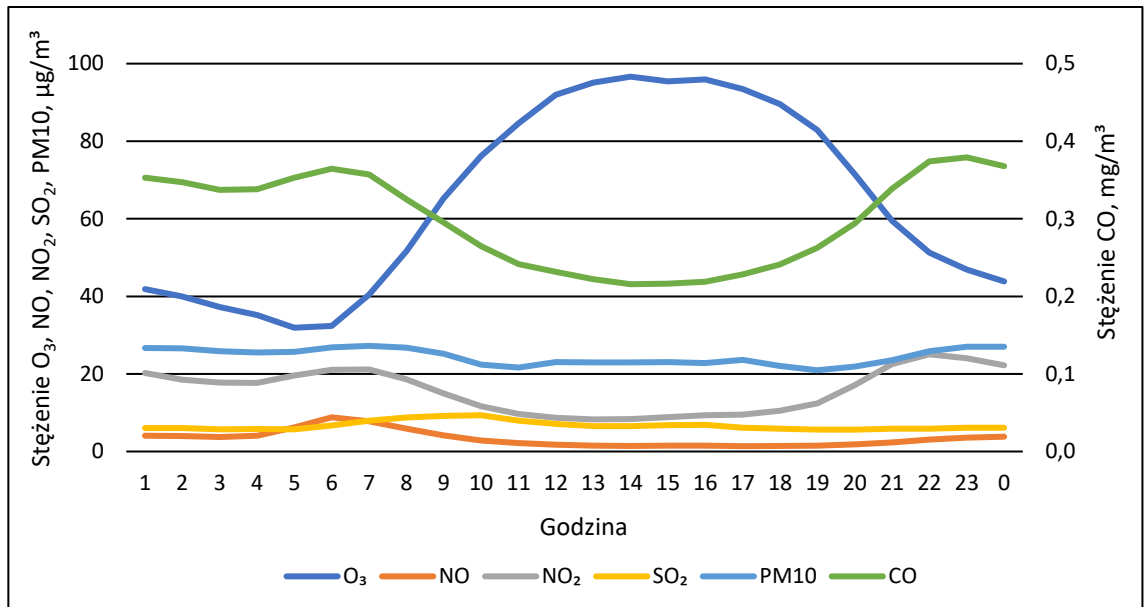


Rys. 4. Uśrednione przebiegi dobowe chwilowych stężeń zanieczyszczeń zarejestrowanych na stacji Zabrze w miesiącach styczeń i grudzień

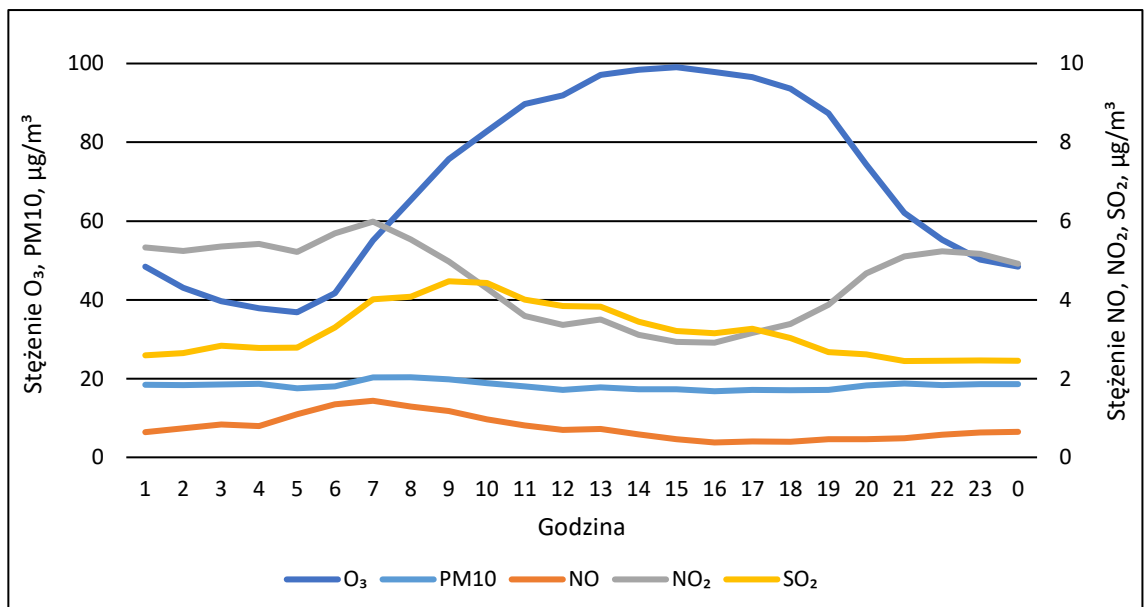


Rys. 5. Uśrednione przebiegi dobowe chwilowych stężeń zanieczyszczeń zarejestrowanych na stacji Złoty Potok w miesiącach styczeń i grudzień

Na rys. 6 i 7 przedstawiono uśrednione przebiegi stężeń zarejestrowane w miesiącach czerwiec i lipiec (okres wiosenno-letni).



Rys. 6. Uśrednione przebiegi dobowe chwilowych stężeń zanieczyszczeń zarejestrowanych na stacji Zabrze w miesiącach czerwiec i lipiec



Rys. 7. Uśrednione przebiegi dobowe chwilowych stężeń zanieczyszczeń zarejestrowanych na stacji Złoty Potok w miesiącach czerwiec i lipiec

W tabeli 5 zamieszczono minimalne i maksymalne wartości stężeń chwilowych dla obu stacji w wybranych miesiącach z badanego okresu. Rozpatrując różnice w stężeniach, wyraźnie widać różnicę w typie obszaru reprezentowanym przez konkretną stację monitoringu. Stężenia pyłu PM10 na stacji Złoty Potok osiągają maksymalną wartość znacznie niższą ($44,6 \mu\text{g}/\text{m}^3$) niż poziom stężenia zarejestrowany na stacji Zabrze ($103,6 \mu\text{g}/\text{m}^3$). W przypadku zanieczyszczeń pierwotnych wartości maksymalne zarejestrowane na stacji Złoty Potok są na ogół niższe niż wartości minimalne zarejestrowane na stacji Zabrze. Jedynym wyjątkiem są stężenia ozonu, które osiągają wyższe wartości w przypadku stacji Złoty Potok, z minimum wynoszącym $33,8 \mu\text{g}/\text{m}^3$ (względem $17,2 \mu\text{g}/\text{m}^3$ na stacji Zabrze) oraz z maksimum na poziomie $47,1 \mu\text{g}/\text{m}^3$ (względem $28,8 \mu\text{g}/\text{m}^3$ na stacji Zabrze).

W okresie wiosenno-letnim, stężenia ozonu na obu stacjach osiągają podobne wartości – minima na poziomie $31,9 \mu\text{g}/\text{m}^3$ na stacji Zabrze i $36,8 \mu\text{g}/\text{m}^3$ na stacji Złoty Potok oraz maksima wynoszące odpowiednio $96,6 \mu\text{g}/\text{m}^3$ na stacji Zabrze i $99,1 \mu\text{g}/\text{m}^3$ na stacji Złoty Potok.

Tab. 5. Minimalne i maksymalne wartości 1-godzinnych stężeń zanieczyszczeń powietrza zarejestrowane w typowych miesiącach wiosenno-letnich i jesienno-zimowych

Miesiące	Zanieczyszczenie	Zabrze		Złoty Potok	
		Wartość minimalna, $\mu\text{g}/\text{m}^3$	Wartość maksymalna, $\mu\text{g}/\text{m}^3$	Wartość minimalna, $\mu\text{g}/\text{m}^3$	Wartość maksymalna, $\mu\text{g}/\text{m}^3$
Styczeń, grudzień	O ₃	17,2	28,8	33,8	47,1
	NO	9,3	23,7	0,9	3,8
	NO ₂	23,6	38,6	14,2	19,8
	SO ₂	24,2	40,9	11,0	15,4
	PM10	48,5	103,6	32,5	44,6
	CO	0,7	1,2	-	-
Czerwiec, lipiec	O ₃	31,9	96,6	36,8	99,1
	NO	1,4	8,8	0,4	1,4
	NO ₂	8,3	25,1	2,9	6,0
	SO ₂	5,6	9,3	2,4	4,5
	PM10	20,9	27,2	16,8	20,3
	CO	0,2	0,4	-	-

4.2. Wybór typu sieci neuronowej

W pierwszym etapie badań dokonano porównania podstawowych typów sieci neuronowych: liniowej, perceptronu wielowarstwowego (MLP – Multi-Layer Perceptron) oraz sieci o radialnych funkcjach bazowych (RBF – Radial Basis Functions). Porównanie przeprowadzono w programie Statistica Neural Networks.

Liniowe sieci neuronowe są układami, które tylko w znacznym uproszczeniu odwzorowują rzeczywiste działanie biologicznych struktur. Z uwagi na prostą budowę, nadają się do budowy nieskomplikowanych modeli opisujących liniowe zależności (Tadeusiewicz, 2009). Sieci liniowe mogą być dobrym punktem odniesienia do porównania z innymi rodzajami sieci neuronowych. Zdarza się, że problem uchodzący za złożony i skomplikowany, może zostać rozwiązany przez najprostszą sieć liniową. Model liniowy jest odwzorowywany przez sieć, która nie posiada warstw ukrytych, a neurony znajdujące się w warstwie wyjściowej są w pełni liniowe, co oznacza, że są to neurony, w których pobudzenie określone jest jako liniowa kombinacja wartości wejściowych i które posiadają liniową funkcję aktywacji (StatSoft, 2022).

Sieć typu MLP jest siecią często wybieraną do rozwiązywania zadań związanych z predykcją (Osowski i in., 1996; Hippert i in., 2001; Siwek, Osowski, 2001; Rojek, Studziński, 2012; Ciechulski, Osowski, 2014; Poloczek i in., 2021). Sieć MLP składa się z neuronów o nieliniowych charakterystykach, które mogą być umieszczone w wielu warstwach i pełnić wielorakie funkcje. Na budowę perceptronu wielowarstwowego (MLP) składa się warstwa wejściowa, warstwa ukryta (lub kilka warstw ukrytych) oraz warstwa wyjściowa. Na warstwę wejściową wysyłane są sygnały wejściowe zawierające informację o zadaniu, które sieć ma rozwiązać. Warstwa ukryta to warstwa, do której nie ma bezpośredniego dostępu ani od strony wejścia, ani od strony wyjścia sieci. W warstwie ukrytej zachodzi proces przetwarzania danych wejściowych tak, aby dane uzyskane na wyjściu sieci były jak najbardziej przydatne dla potrzeb uzyskania rozwiązania postawionego zadania (Tadeusiewicz, 2009). Wybór odpowiedniej liczby warstw ukrytych i liczby neuronów wchodzących w skład poszczególnych warstw jest zadaniem trudnym i najlepszym jego rozwiązaniem jest przetestowanie wielu różnych architektur sieci (Wilkoś i in., 2021).

W sieciach typu RBF (zwanymi sieciami o radialnych funkcjach bazowych), zamiast jednej funkcji aktywacji dla wszystkich neuronów, stosowane są tak zwane

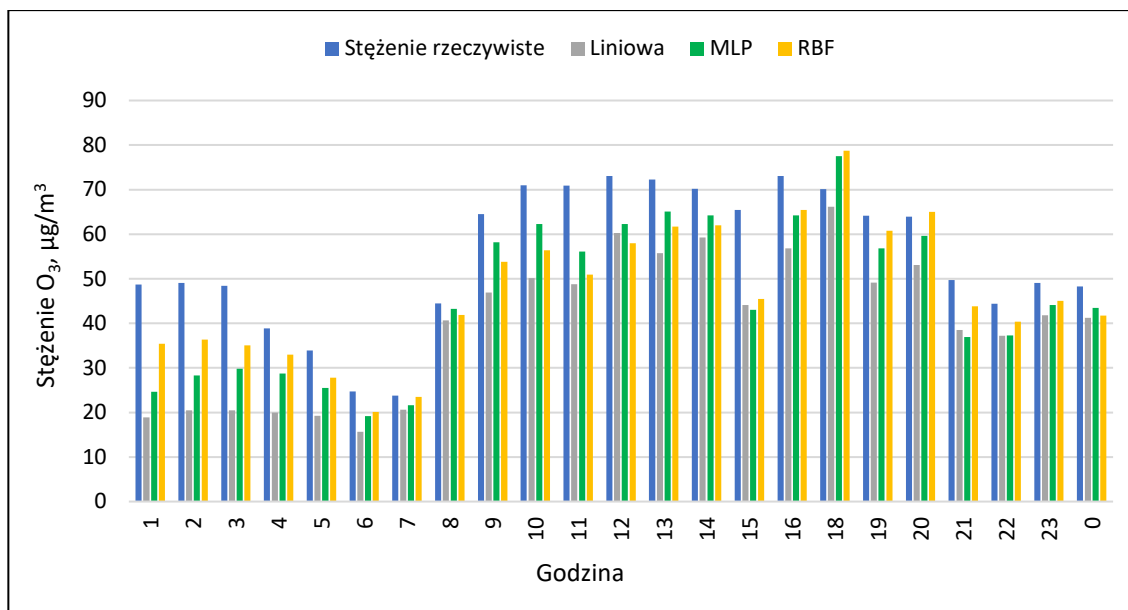
radialne funkcje bazowe, określane osobno dla każdego neuronu (Flasiński, 2011). Najczęściej sieci typu RBF służą do nieliniowej aproksymacji zmiennych numerycznych, również w zagadnieniach związanych z klasyfikacją (Bishop 1995; Nabney 2001; Pluciński 2002).

W celu wyboru najkorzystniejszego typu sieci do modelowania stężeń zanieczyszczeń powietrza, stworzono modele sieci neuronowych typu MLP, RBF oraz liniowej. Wyjściem modelu było stężenie O₃ zarejestrowane na stacji Zabrze. W przypadku każdego typu sieci wykorzystano 10 zmiennych wejściowych, w postaci pozostałych stężeń zanieczyszczeń i parametrów meteorologicznych oraz daty i godziny. Sieć liniowa oraz MLP składała się z 10 neuronów w warstwie ukrytej, sieć RBF składała się z 50 neuronów w warstwie ukrytej. Modelowano stężenia ozonu w całym okresie pomiarowym 2011-2016, a następnie obliczono wartości błędów odnosząc stężenia modelowane do stężeń rzeczywistych. W tab. 6 przedstawiono obliczone błędy dla każdego typu sieci. Dla sieci liniowej otrzymano najwyższe wartości błędów MAE i RMSE, wynoszące odpowiednio 14,1 µg/m³ i 17,8 µg/m³. Nieco lepsze wyniki uzyskano dla sieci typu RBF - wartość błędu MAE wyniosła 10,8 µg/m³, a błędu RMSE 13,9 µg/m³. Najniższe wartości błędu uzyskano w przypadku sieci MLP: wartość błędu MAE wyniosła 10,3 µg/m³, a wartość błędu RMSE 13,4 µg/m³. Na podstawie uzyskanych wyników, do dalszej analizy wybrano perceptron wielowarstwowy MLP.

Tab. 6. Wartości błędów predykcji obliczone dla każdego typu uzyskanych sieci

Typ sieci	MAE, µg/m ³	RMSE, µg/m ³
Liniowa	14,1	17,8
MLP	10,3	13,4
RBF	10,8	13,9

Na rysunku 8 zilustrowano efekty modelowania dla wybranego epizodu. Porównano dobowe przebiegi stężenia rzeczywistego oraz stężeń predykcyjnych uzyskanych za pomocą sieci: liniowej, MLP i RBF, w wybranym dniu. Wybrano dzień z pełną kompletnością zarejestrowanych danych.



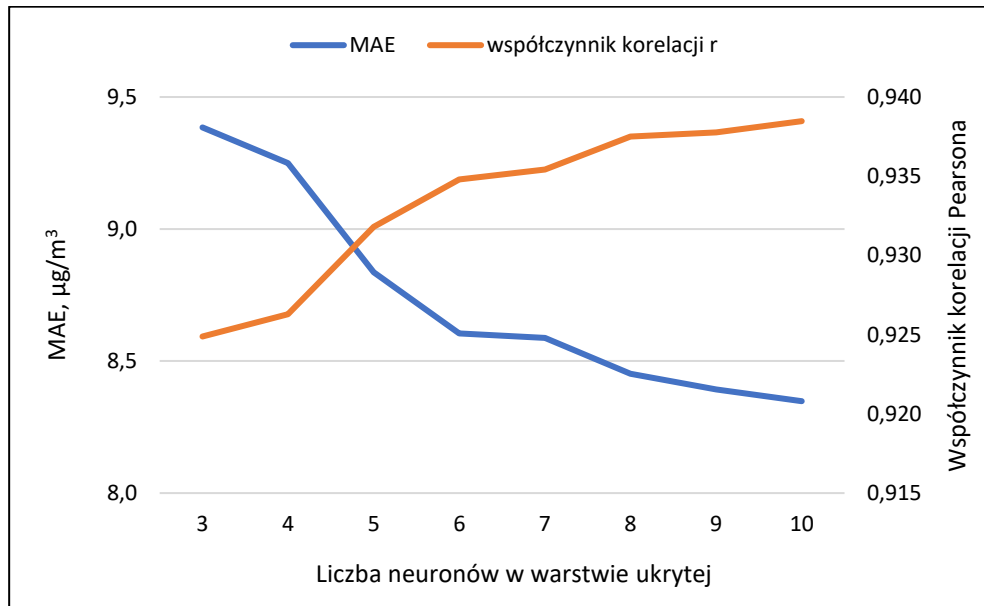
Rys. 8. Porównanie wartości rzeczywistych i predycyjnych uzyskanych przez sieć liniową, MLP oraz RBF, stężenie O₃ zarejestrowane na stacji Zabrze, 06.04.2015 r.

4.3. Wybór architektury sieci MLP

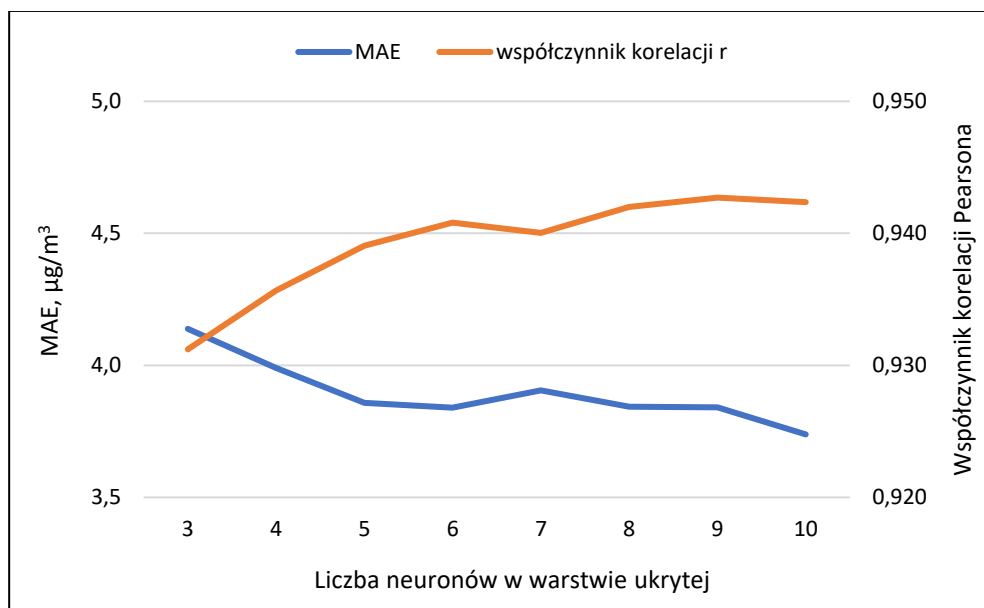
W drugim etapie badań dokonano porównania dokładności predykcji różnych architektur sieci MLP pod kątem ilości neuronów w warstwie ukrytej. Mając na uwadze, że dobry model powinien być stosunkowo prosty i dokładny, rozważono perceptrony z jedną warstwą ukrytą i ograniczoną liczbą neuronów. Dla każdego zanieczyszczenia utworzono model perceptronu wielowarstwowego, który w warstwie ukrytej posiadał od 3 do 10 neuronów. Architektura sieci o oznaczeniu MLP 10-3-1 oznacza perceptron wielowarstwowy o 10 zmiennych wejściowych (objasniających), 3 neuronach w warstwie ukrytej i 1 zmiennej wyjściowej (objasnianej). Odpowiednio oznaczenie MLP 10-10-1 oznacza perceptron wielowarstwowy z 10 zmiennymi wejściowymi, 10 neuronami w warstwie ukrytej i 1 zmienną wyjściową.

Do oceny jakości modelowania wykorzystano średni błąd bezwzględny (MAE) oraz współczynnik korelacji Pearsona (r). Na rysunkach 9-14 przedstawiono wartości obliczonego błędu MAE oraz współczynnika korelacji Pearsona, w zależności od liczby neuronów w warstwie ukrytej, odpowiednio dla różnych zanieczyszczeń powietrza. Otrzymane wyniki wskazują, że zwiększenie liczby neuronów nie zawsze przekłada się na poprawę dokładności modelowania, jednak biorąc pod uwagę wszystkie zanieczyszczenia, najniższą wartością błędu MAE i najwyższym współczynnikiem

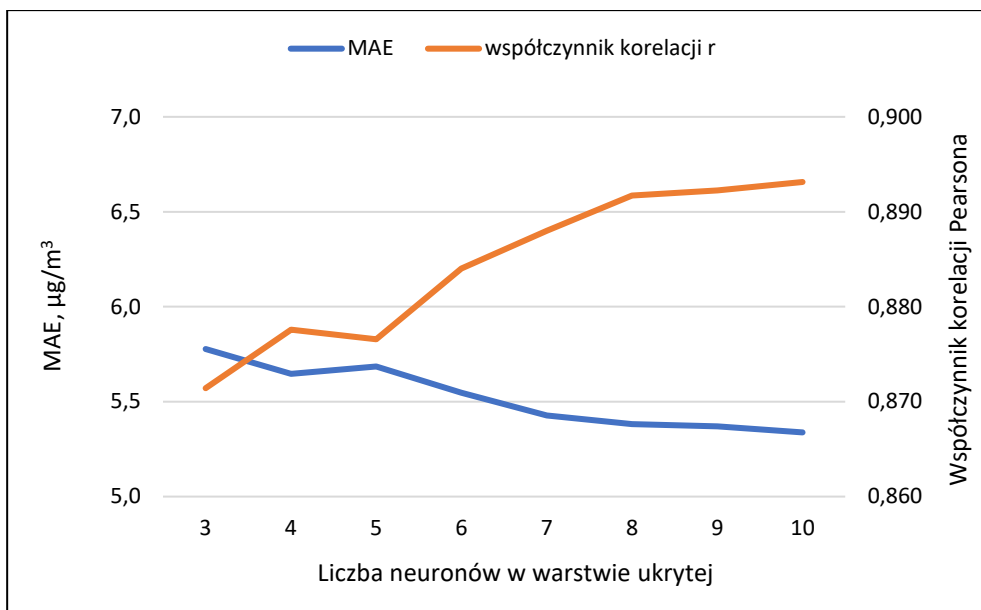
korelacji Pearsona r wykazuje się model z 10 neuronami. Na podstawie uzyskanych wyników, do dalszych badań rekomendowano architekturę o 10 neuronach w pojedynczej warstwie ukrytej. Wcześniejsze badania udowodniły też, że do aproksymacji dowolnej gładkiej funkcji wymiernej pomiędzy zmiennymi niezależnymi i zależnymi wystarczająca jest jedna warstwa ukryta (Hornik i in., 1988; Gardner, Dorling, 1998). Zbyt duża liczba neuronów nie gwarantuje większej dokładności predykcji, a sprawia, że rośnie złożoność modelu i ryzyko przeuczenia sieci (Hoffman, 2004).



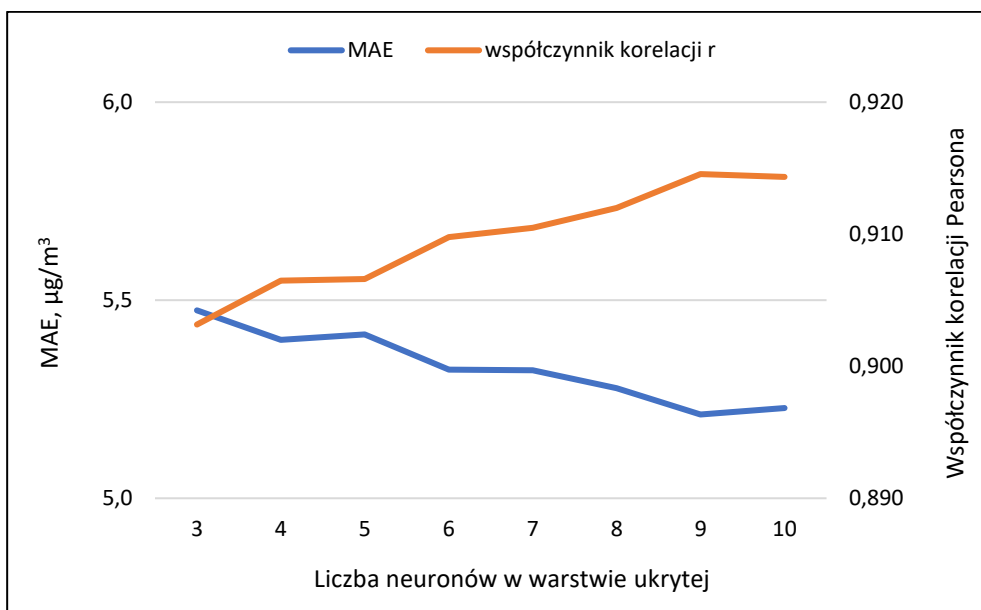
Rys. 9. Wartości błęd MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia O_3 , Zabrze 2011-2016



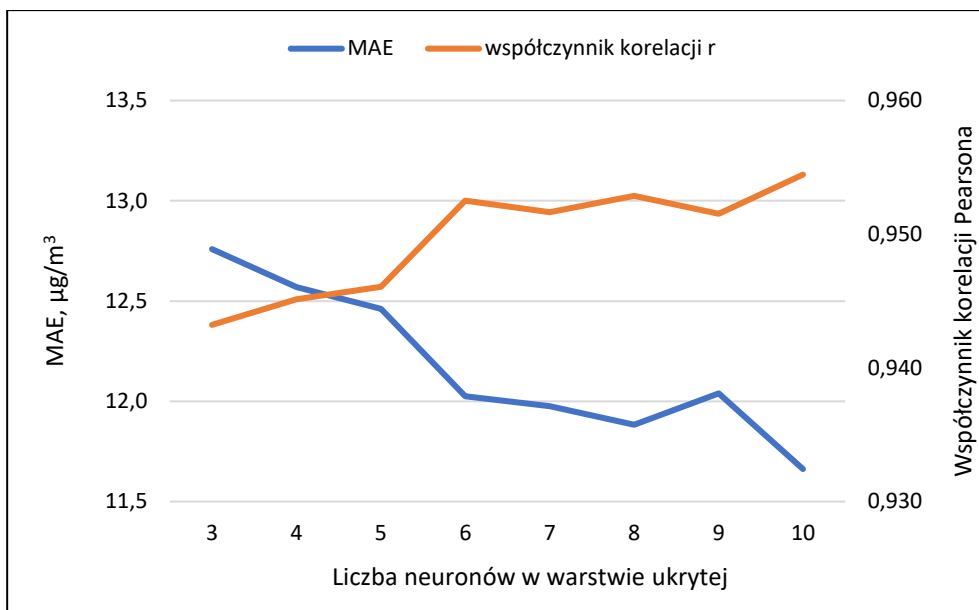
Rys. 10. Wartości błęd MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia NO , Zabrze 2011-2016



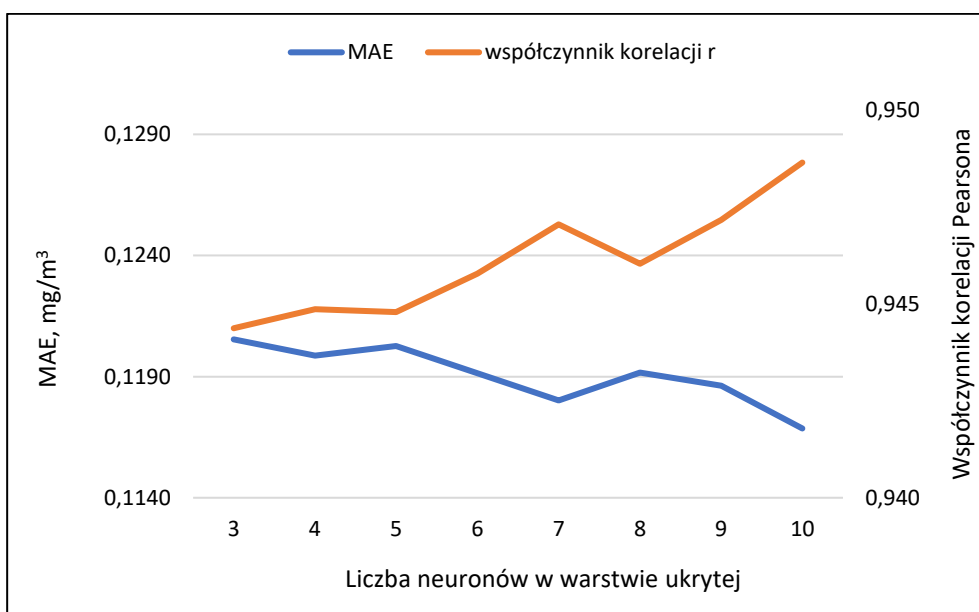
Rys. 11. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia NO₂, Zabrze 2011-2016



Rys. 12. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia SO₂, Zabrze 2011-2016



Rys. 13. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia PM10, Zabrze 2011-2016



Rys. 14. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia CO, Zabrze 2011-2016

4.4. Metodyka badań

4.4.1. Przygotowanie danych

Przed przystąpieniem do modelowania, surowe dane pochodzące z bazy danych Głównego Inspektoratu Ochrony Środowiska (Bank danych pomiarowych, 2021),

przygotowano usuwając przypadki dla godzin, w których występowały braki danych. Po usunięciu przypadków z brakującymi danymi otrzymano zbiór przypadków z kompletnymi danymi, który poddano dalszej analizie. Ten zbiór nazwano zbiorem pełnozakresowym. Zbiory pełnozakresowe zawierały 36460 i 15536 przypadków, odpowiednio dla stacji Zabrze i Złoty Potok.

Z uwagi na to, że poziom stężeń niektórych zanieczyszczeń uzależniony jest od pory roku i pory dnia, w celu zwiększenia ich przydatności w modelowaniu, zastosowano przeskalowanie wartości zmiennych opisujących datę i godzinę. W tym celu tradycyjny zapis daty i godziny zastąpiono liczbą z przedziału $<0,1>$ w sposób, który zapewni powtarzalność wartości daty dla określonej pory roku oraz godziny dla określonej pory dnia. Przyjęto, że maksymalna wartość daty równa 1,000000 przypada 31 grudnia, natomiast wartość 0,000000 przypada dla daty 2 lipca. Od 2 lipca wartość zmiennej wzrasta o 0,005494 (w latach przestępnych o 0,005479), a następnie od 31 grudnia ponownie spada o taką samą wartość, aż do poziomu 0,000000 przypisanego dla 2 lipca. Z kolei dla zmiennej opisującej porę dnia, godzinie 12:00 przypisano wartość 0,00000, rosnącą o 0,08333 dla każdej kolejnej godziny, aż do wartości 1,00000 przypisanej godzinie 24:00. Po osiągnięciu maksimum o godzinie 24:00, w kolejnych godzinach następuje skokowy spadek o 0,08333 aż do wartości 0,00000 dla godziny 12:00. W wyniku transformacji daty i godziny zwiększa się współzależność między zmiennymi, co wpływa dodatnio na ich przydatność w modelowaniu regresyjnym (Hoffman, 2013).

4.4.2. Metodyka modelowania

Podczas modelowania tworzone regresyjne modele perceptronowych sieci neuronowych, w których wyjściem było stężenie wybranego zanieczyszczenia powietrza (zmienna objaśniana), a wejściami data i godzina w postaci numerycznej, pozostałe zanieczyszczenia powietrza oraz parametry meteorologiczne (zmiennie objaśniające). W tabeli 7 wyszczególniono zmienne objaśniane i objaśniające, wykorzystywane podczas modelowania.

Tab. 7. Opis zmiennych w modelach stężeń poszczególnych zanieczyszczeń dla obu stacji monitoringu powietrza

Stacja	Zmienna objaśniana	Zmienna objaśniająca											
		G	D	O ₃	NO	NO ₂	SO ₂	CO	PM10	WS	T	I	HR
Zabrze	O ₃	+	+		+	+	+	+	+	+	+	+	-
	NO	+	+	+		+	+	+	+	+	+	+	-
	NO ₂	+	+	+	+		+	+	+	+	+	+	-
	SO ₂	+	+	+	+	+		+	+	+	+	+	-
	CO	+	+	+	+	+	+		+	+	+	+	-
	PM10	+	+	+	+	+	+	+		+	+	+	-
Złoty Potok	O ₃	+	+		+	+	+	+	+	+	+	+	+
	NO	+	+	+		+	+	+	+	+	+	+	+
	NO ₂	+	+	+	+		+	+	+	+	+	+	+
	SO ₂	+	+	+	+	+		+	+	+	+	+	+
	CO	-	-	-	-	-	-	-	-	-	-	-	-
	PM10	+	+	+	+	+	+	+	+		+	+	+

Do modelowania regresyjnego wykorzystano sztuczne sieci neuronowe o architekturze perceptronu wielowarstwowego. Perceptron wielowarstwowy składał się z 10 neuronów w pojedynczej warstwie ukrytej. Przed rozpoczęciem procesu uczenia zbiór danych został podzielony na trzy podzbiory: podzbiór uczący składający się z 70 % przypadków, podzbiór walidacyjny składający się z 15 % przypadków oraz podzbiór testowy składający się z 15 % przypadków. W procesie uczenia wykorzystano algorytm BFGS (Broyden-Fletcher-Goldfarb-Shanno). Proces uczenia ograniczono do 300 epok. Funkcją aktywacji neuronów ukrytych i neuronów wyjściowych była funkcja logistyczna. Sieć była inicjalizowana losowo, metodą gaussowską. Modelowanie przeprowadzono za pomocą programu Statistica. Każda predykcja została przeprowadzona 5-krotnie. Do raportowania wybierano najdokładniejszy z utworzonych 5 modeli.

4.4.3. Podział zbiorów pełnozakresowych na podzbiory

Zbiór pełnozakresowy był zbiorem wyjściowym, zawierającym tylko kompletne dane rzeczywiste wszystkich zmiennych. W analizie eksplorowano dwa takie zbiory: pierwszy dla stacji monitoringu Zabrze, a drugi dla stacji Złoty Potok. Przygotowując się do modelowania, uszeregowano wszystkie przypadki w zbiorze pełnozakresowym względem rosnącej wartości stężenia rzeczywistego wybranego zanieczyszczenia. W ten sposób, dla każdej stacji monitoringu uzyskano 6 różnych zbiorów pełnozakresowych,

dla każdego z zanieczyszczeń oddzielny. Każdy z tych zbiorów zawierał te same przypadki, ale ułożone w innej kolejności.

Zbiory przypadków, które zostały posortowane według rosnących stężeń rzeczywistych modelowanego zanieczyszczenia, nazwano zbiorami RVS (Real Values Sorting). Zbiory pełnozakresowe oznaczono, jako zbiory RVS-1/1. Zbiory pełnozakresowe RVS-1/1 dzielono na równoliczne podzbiory. Przyjęto trzy różne stopnie podziału: na 2, 4 i 8 podzbiorów. W wyniku podziału uzyskano:

- 2 podzbiory (RVS-1/2, RVS-2/2), po podziale na dwa podzbiory;
- 4 podzbiory (RVS-1/4, RVS-2/4, RVS-3/4, RVS-4/4), po podziale na cztery podzbiory;
- 8 podzbiorów (RVS-1/8, RVS-2/8, RVS-3/8, RVS-4/8, RVS-5/8, RVS-6/8, RVS-7/8, RVS-8/8), po podziale na osiem podzbiorów.

Tak przygotowane zbiory/podzbiory poddawano modelowaniu, w celu uzyskania stężeń predykcyjnych wybranego zanieczyszczenia. Do aproksymacji stężeń predykcyjnych wykorzystano sieci neuronowej MLP 10-10-1, według metodyki opisanej w poprzednim podrozdziale. W efekcie, dla każdego zbioru/podzbioru uzyskano oddzielny model predykcyjny, który oznaczono takim samym symbolem, jak modelowany zbiór/podzbior. Do oznaczeń zbiorów i podzbiorów dodano jeszcze nazwę stacji (ZAB albo ZP). Dzięki podziałom zbioru pełnozakresowego możliwe było sprawdzenie, jak zmienia się dokładność modelowania w podzakresach stężeń i czy modelowanie przeprowadzone w podzakresach umożliwi poprawę jakości modelowania, w stosunku do modelowania pełnozakresowego.

W sytuacji, gdy stężenia rzeczywiste modelowanego zanieczyszczenia nie są znane, modele typu RVS dla podzakresów tych stężeń nie mogą być stosowane, ponieważ nie można przeprowadzić sortowania i klasyfikacji przypadków do podzakresów stężeń rzeczywistych. Modele RVS odzwierciedlają zatem potencjalne, a nie praktyczne, możliwości poprawy jakości modelowania, poprzez segmentację procesu predykcji. Jeżeli stężenia rzeczywiste nie są znane, a istnieje potrzeba ich predykcji, to modele podzakresowe typu RVS są niedostępne. W takiej sytuacji można zaproponować inne podejście, które także umożliwi podział na podzakresy, a później modelowanie sektorowe w wyznaczonych podzakresach. Najistotniejszym etapem w takim podejściu jest wstępne modelowanie stężeń predykcyjnych wybranego zanieczyszczenia w zbiorze pełnozakresowym, rozumianym jako zbiór wszystkich przypadków zawierających kompletne dane wszystkich zmiennych objaśniających (wejść modelu) dla modelowanego zanieczyszczenia. Po wstępnym modelowaniu całego zakresu

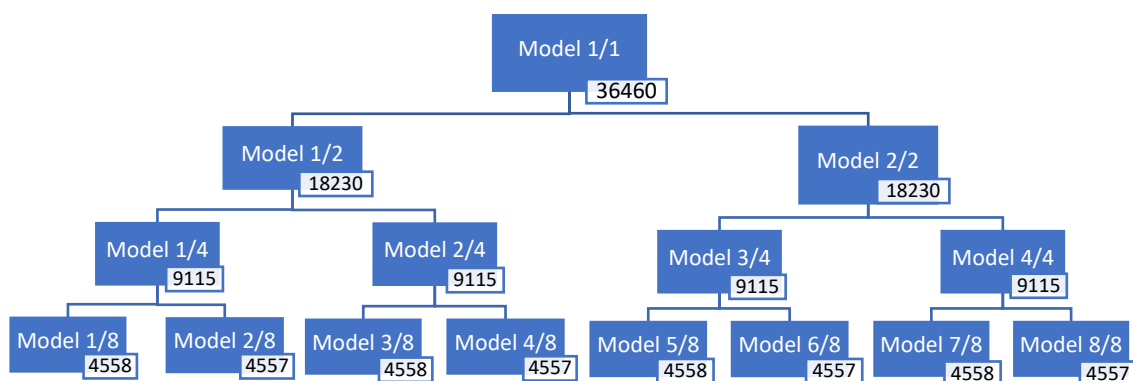
przypadków uzyskuje się stężenia predykcyjne zmiennej objaśnianej. Następnie, cały zbiór przypadków sortuje się według rosnących stężeń predykcyjnych modelowanego zanieczyszczenia. Kolejnym etapem jest podział tak posortowanego pełnozakresowego zbioru przypadków na określoną liczbę równolicznych podzakresów. W opisany sposób można wygenerować modele, które nazwano modelami PVS (Predicted Values Sorting).

Stosując takie podejście, otrzymano pełnozakresowe modele dla poszczególnych zanieczyszczeń powietrza. Modele pełnozakresowe oznaczono symbolem PVS-1/1. Zbiory pełnozakresowe PVS-1/1 podzielono na równoliczne podzbiory. Tak, jak w przypadku modeli RVS, przyjęto trzy różne stopnie podziału: na 2, 4 i 8 podzbiorów. W wyniku podziału uzyskano:

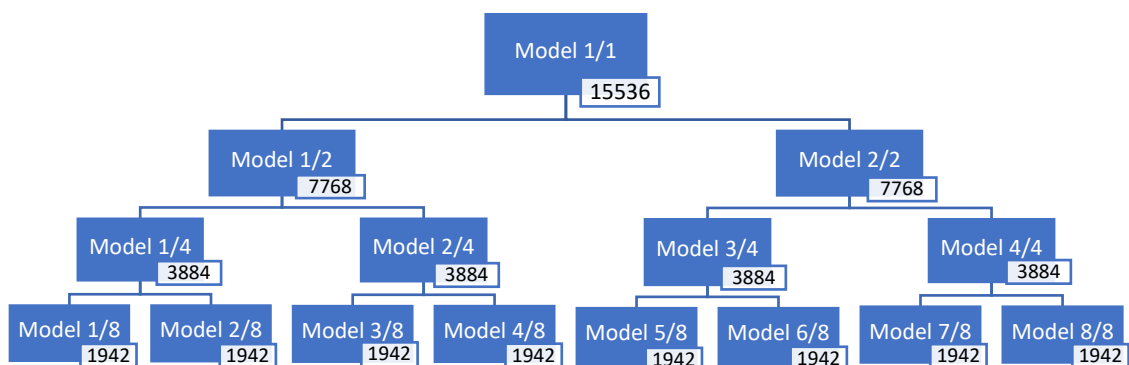
- 2 podzbiory (PVS-1/2, PVS-2/2), po podziale na dwa podzbiory;
- 4 podzbiory (PVS-1/4, PVS-2/4, PVS-3/4, PVS-4/4), po podziale na cztery podzbiory;
- 8 podzbiorów (PVS-1/8, PVS-2/8, PVS-3/8, PVS-4/8, PVS-5/8, PVS-6/8, PVS-7/8, PVS-8/8), po podziale na osiem podzbiorów.

Tak przygotowane podzbiory poddano ponownemu modelowaniu, wykorzystując zmienne objaśniające. Dla modelowanego zanieczyszczenia uzyskano nowe wartości stężeń predykcyjnych, dla każdego przypadku w analizowanym podzbiore. Do aproksymacji stężeń predykcyjnych wykorzystano regresyjne sieci neuronowe MLP 10-10-1, według metodyki opisanej w poprzednim podrozdziale. Dla każdego zbioru/podzbiory uzyskano model predykcyjny, który oznaczono takim samym symbolem, jak modelowany zbiór/podzbiór. Dokładność otrzymanych modeli podzakresowych oceniano na podstawie wartości nowych stężeń predykcyjnych.

Schematy podziału na podzakresy dla obu stacji przedstawiono na rysunkach 15 i 16, oddzielnie dla stacji Zabrze i Złoty Potok. Na schematach podano liczbę przypadków zaklasyfikowanych do poszczególnych zbiorów i podzbiorów. Schematy ilustrują podział zbiorów przypadków dla modeli obu typów RVS i PVS. Zbiory pełnozakresowe RVS-1/1 i PVS-1/1 miały te same przypadki, ale inaczej posortowane. Z powodu różnic w sortowaniu, otrzymane w wyniku podziału zbiorów pełnozakresowych podzbiory RVS i PVS miały taką samą liczbę przypadków, ale nie te same przypadki.



Rys. 15. Schemat podziału na podzakresy dla stacji Zabrze, ze wskazaniem liczby przypadków w poszczególnych podzakresach modelowania



Rys. 16. Schemat podziału na podzakresy dla stacji Złoty Potok, ze wskazaniem liczby przypadków w poszczególnych podzakresach modelowania

4.4.4. Ocena dokładności modeli

Istnieje wiele różnych miar, które mogą zostać wykorzystane do oceny jakości modelowania, jak na przykład współczynnik korelacji Pearsona, czy indeksy zgodności Willmotta, które zostały stworzone z myślą o badaniach geofizycznych, zwłaszcza w zadaniach związanych z modelowaniem poziomu stężeń zanieczyszczeń powietrza (Willmott, 1981; Willmott i in., 1985). Jednak, stosowanie tych miar, z uwagi na odnośnienie się w formułach do średniej oraz odległości od średniej, może prowadzić do nieprawidłowych interpretacji (Hoffman, 2021). Ze względu na to, do oceny dokładności uzyskanych modeli regresyjnych wykorzystano wartości błędów MAE (średni błąd bezwzględny) i RMSE (pierwiastek błędu średniokwadratowego), które obliczono na podstawie różnic między stężeniami rzeczywistymi, a stężeniami predykcyjnymi.

Średni błąd bezwzględny określa średnią wartość z badanej próby bezwzględnych różnic między wartościami przewidywanymi, a wartościami rzeczywistymi. Pierwiastek błędu średniokwadratowego oblicza się biorąc pod uwagę kwadraty różnic między wartościami rzeczywistymi, a przewidywanymi (Wilkoś i in., 2021). Wzory do obliczeń poszczególnych błędów zamieszczono poniżej:

MAE – Mean Absolute Error

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (1)$$

RMSE – Root Mean Squared Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (2)$$

, gdzie n — liczba przypadków, y — stężenia predykcyjne w zbiorze przewidywanych stężeń, x — stężenia rzeczywiste w zbiorze rzeczywistych stężeń, i — numer kolejnego przypadku.

4.4.5. Epizody

W celu zobrazowania różnic w uzyskanych stężeniach predykcyjnych, utworzono wykresy przedstawiające porównanie przebiegów dobowych stężeń rzeczywistych zarejestrowanych na danej stacji monitoringu powietrza oraz stężeń aproksymowanych modelami regresyjnymi. Wybrano jeden dzień z okresu wiosenno-letniego i jeden z okresu jesienno-zimowego. Przy wyborze konkretnego dnia kierowano się kompletnością stężeń w ciągu doby. Dla stężeń zarejestrowanych w okresie wiosenno-letnim wybrano epizody: na stacji Zabrze z dnia 01.06.2015 r., a na stacji Żłoty Potok z dnia 05.06.2016 r. Dla stężeń zarejestrowanych w okresie jesienno-zimowym wybrano epizody: na stacji Zabrze z dnia 24.12.2016 r., a na stacji Żłoty Potok z dnia 01.01.2016 r.

5. WYNIKI I ICH DYSKUSJA

Dla każdego zanieczyszczenia obliczono błędy modelowania względem rzeczywistych stężeń zanieczyszczeń: O₃, NO, NO₂, SO₂, CO, PM₁₀ dla stacji Zabrze oraz O₃, NO, NO₂, SO₂, PM₁₀ dla stacji Złoty Potok. W celu oceny dokładności modelowania obliczono dwie miary błędu: średni błąd bezwzględny MAE oraz pierwiastek błędu średniokwadratowego RMSE. Wyniki przedstawiono w dwóch oddzielnych podrozdziałach, jednym dla modeli uzyskanych w wyniku sortowania i podziału względem rosnących wartości rzeczywistych (modele RVS) i drugim dla modeli uzyskanych w wyniku sortowania i podziału względem rosnących wartości predykcyjnych (modele PVS).

5.1. Modele RVS

W tym rozdziale przedstawiono wyniki modelowania po wstępnym sortowaniu przypadków względem stężeń rzeczywistych modelowanego zanieczyszczenia. Modele podzakresowe RVS powstały w wyniku podziału zbioru pełnozakresowego przypadków na podzbiory. Podziału na podzakresy dokonano zgodnie ze schematami przedstawionymi na rys. 15 i 16.

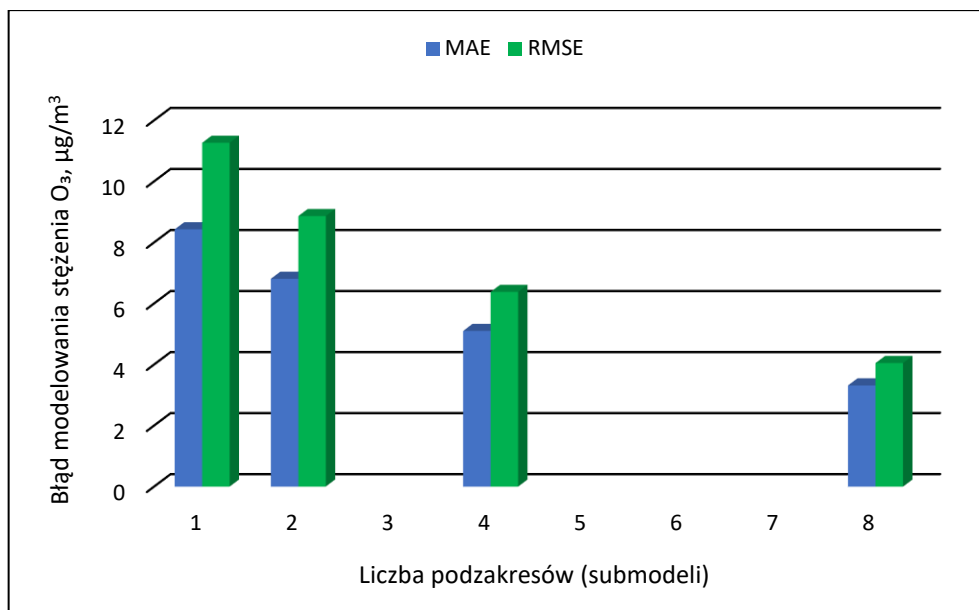
5.1.1. Wyniki modelowania stężeń O₃

W tabelach 8 i 9 zamieszczono wartości błędów aproksymacji obliczonych dla modeli RVS stężeń ozonu: pełnozakresowego i podzakresowych. W przypadku obu stacji, podział na kolejne podzakresy powodował wyraźny spadek średnich wartości błędów MAE i RMSE. Już podział na 2 podzakresy zapewnił poprawę dokładności modelowania względem modelu pełnozakresowego, objawiającą się zmniejszeniem średnich wartości błędu MAE – w przypadku stacji Zabrze z 8,43 do 6,81 $\mu\text{g}/\text{m}^3$ (zmiana o 19,2 %), a błędu RMSE z 11,26 do 8,86 $\mu\text{g}/\text{m}^3$ (zmiana o 21,3 %). Analogicznie, w przypadku stacji Złoty Potok, podział na 2 podzakresy skutkowało zmniejszeniem błędu aproksymacji z 8,38 do 6,51 $\mu\text{g}/\text{m}^3$ (zmiana o 22,3 %) oraz z 10,71 do 8,25 $\mu\text{g}/\text{m}^3$ (zmiana o 23 %), odpowiednio dla miar MAE i RMSE. Rozpatrując poprawę modelowania modeli 8-podzakresowych, zmniejszenie błędu MAE na stacji Zabrze wynosiło aż 60,7 % (z 8,43

do 3,31 $\mu\text{g}/\text{m}^3$), a na stacji Złoty Potok – 63,5 % (z 8,38 do 3,06 $\mu\text{g}/\text{m}^3$). Błąd RMSE na stacji Zabrze uległ obniżeniu z 11,26 do 4,05 $\mu\text{g}/\text{m}^3$ (zmiana o 64,1 %), a na stacji Złoty Potok – z 10,71 do 3,74 $\mu\text{g}/\text{m}^3$ (zmiana o 65,1 %). Uzyskane wyniki wskazują, że wielkość błędu uzależniona jest od poziomu stężeń w danym podzakresie, co znajduje potwierdzenie w innych pracach (Hoffman, 2015; Hoffman, 2021). Na ogół, największe wartości błędów występują przy modelach dedykowanych dla podzakresów z najwyższymi poziomami stężeń, a najniższe – przy modelach dedykowanych dla najniższych podzakresów reprezentujących najniższe poziomy stężeń. Podobne tendencje zaobserwowano w przypadku pozostałych zanieczyszczeń. Na rys. 17 i 18 zobrazowano graficznie zmiany średnich wartości błędów dla stężeń ozonu, w zależności od liczby podzakresów.

Tab. 8. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych O₃, Zabrze 2011-2016

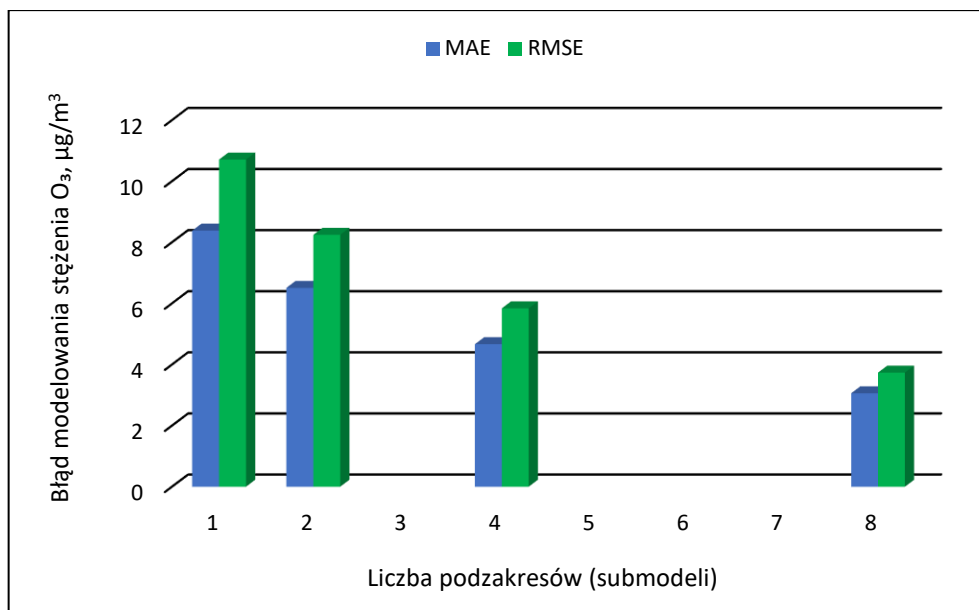
Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych O ₃ , $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	RVS-1/1-ZAB	0,7-198,0	36460	8,43	8,43	11,26	11,26
2 podzakresy	RVS-1/2-ZAB	0,7-37,0	18230	4,79	6,81	6,29	8,86
	RVS-2/2-ZAB	37-198,0	18230	8,83		11,44	
4 podzakresy	RVS-1/4-ZAB	0,7-15,0	9115	2,17	5,10	2,76	6,38
	RVS-2/4-ZAB	15,0-37,0	9115	4,59		5,55	
	RVS-3/4-ZAB	37,0-62,0	9115	5,06		6,11	
	RVS-4/4-ZAB	62,0-198,0	9115	8,58		11,09	
8 podzakresów	RVS-1/8-ZAB	0,7-6,0	4558	0,96	3,31	1,17	4,05
	RVS-2/8-ZAB	6,0-15,0	4557	1,80		2,17	
	RVS-3/8-ZAB	15,0-26,0	4558	2,59		3,04	
	RVS-4/8-ZAB	26,0-37,0	4557	2,75		3,22	
	RVS-5/8-ZAB	37,0-49,0	4558	2,77		3,25	
	RVS-6/8-ZAB	49,0-62,0	4557	3,09		3,64	
	RVS-7/8-ZAB	62,0-82,0	4558	4,24		5,08	
	RVS-8/8-ZAB	82,0-198,0	4557	8,31		10,81	



Rys. 17. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O₃, Zabrze 2011-2016

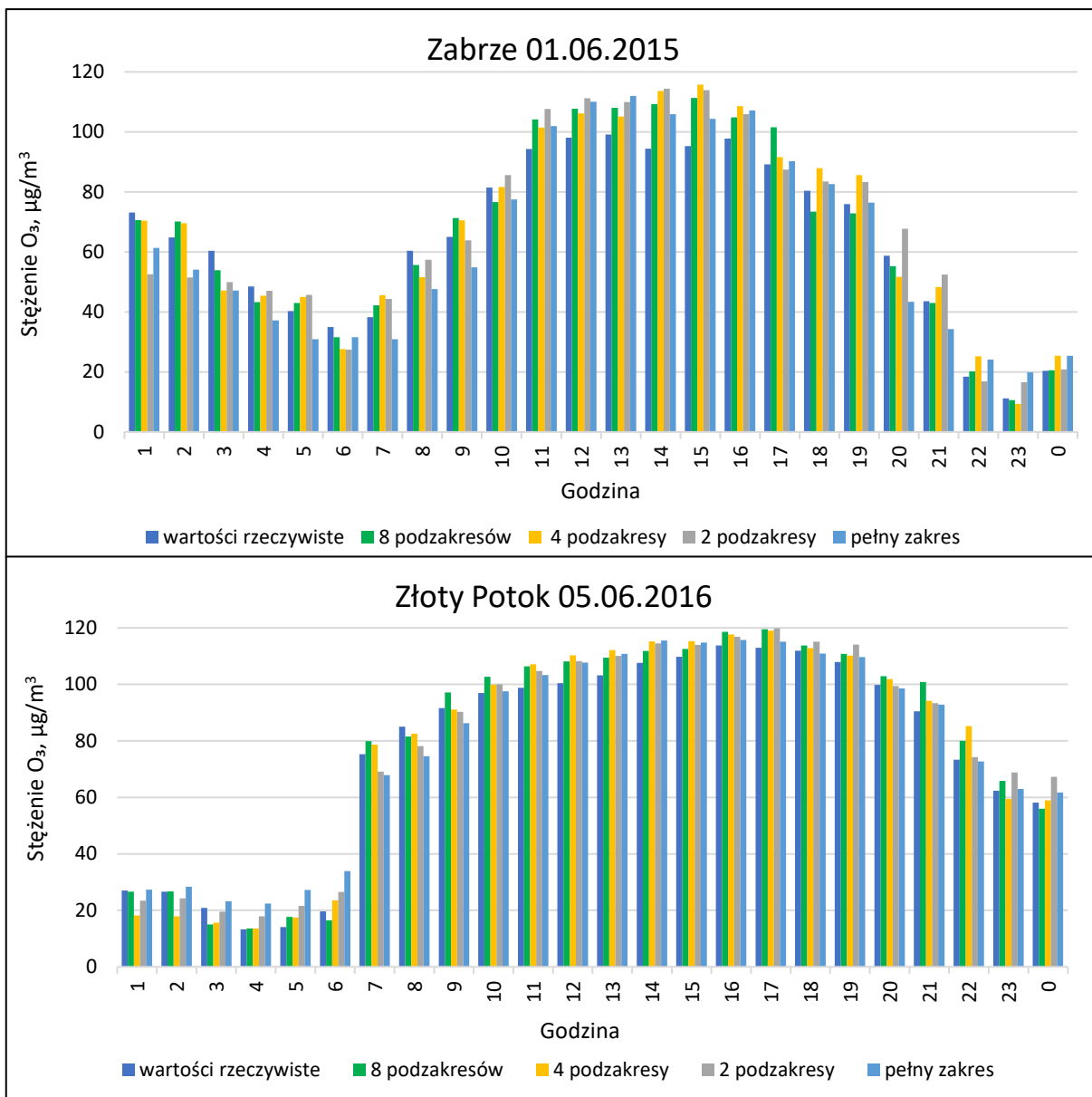
Tab. 9. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych O₃, Złoty Potok 2011-2016

Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych O ₃ , μg/m ³	Liczba przypadków	MAE, μg/m ³	Średnia wartość MAE, μg/m ³	RMSE, μg/m ³	Średnia wartość RMSE, μg/m ³
Pełny zakres	RVS-1/1-ZLP	1,1-162	15536	8,38	8,38	10,71	10,71
2 podzakresy	RVS-1/2-ZLP	1,1-51,0	7768	5,97	6,51	7,61	8,25
	RVS-2/2-ZLP	51,0-162,0	7768	7,05		8,88	
4 podzakresy	RVS-1/4-ZLP	1,1-32,3	3884	4,06	4,67	5,14	5,84
	RVS-2/4-ZLP	32,3-51,0	3884	3,85		4,67	
	RVS-3/4-ZLP	51,0-72,0	3884	4,20		5,14	
	RVS-4/4-ZLP	72,0-162,0	3884	6,56		8,41	
8 podzakresów	RVS-1/8-ZLP	1,1-21,1	1942	2,98	3,06	3,64	3,74
	RVS-2/8-ZLP	21,1-32,3	1942	2,48		2,96	
	RVS-3/8-ZLP	32,3-42,0	1942	2,27		2,66	
	RVS-4/8-ZLP	42,0-51,0	1942	2,17		2,55	
	RVS-5/8-ZLP	51,0-60,4	1942	2,28		2,64	
	RVS-6/8-ZLP	60,4-72,0	1942	2,56		3,07	
	RVS-7/8-ZLP	72,0-90,3	1942	3,68		4,47	
	RVS-8/8-ZLP	90,4-162,0	1942	6,06		7,88	

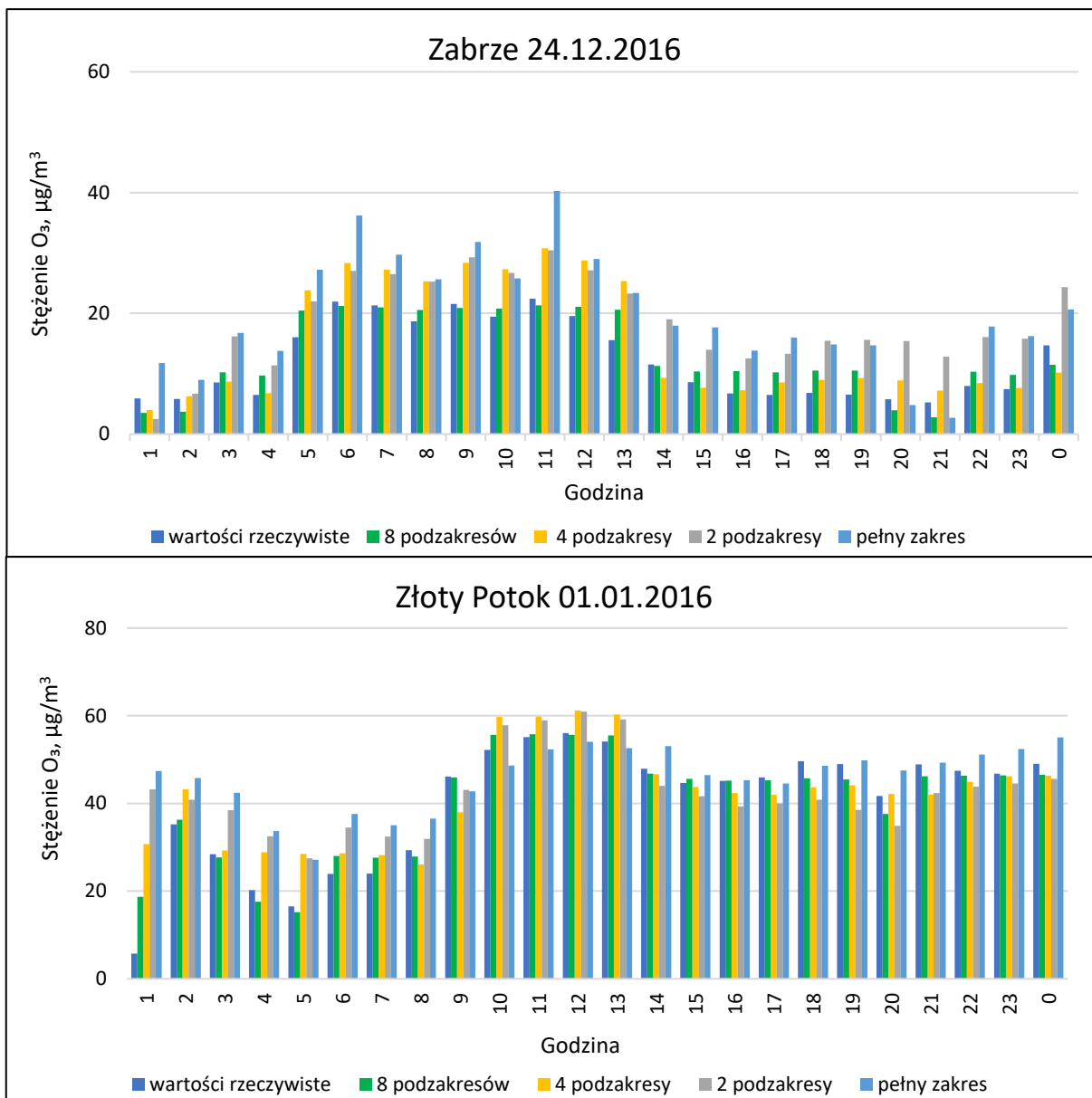


Rys. 18. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O_3 , Złoty Potok 2011-2016

Na rys. 19 i 20 przedstawiono dobowe przebiegi stężeń predykcyjnych, które zostały uzyskane za pomocą modeli RVS, w zestawieniu ze stężeniami rzeczywistymi w okresach, odpowiednio - wiosenno-letnim i jesienno-zimowym. Modele uzyskane dla epizodów w okresie wiosenno-letnim uzyskują nieco lepsze dopasowanie do stężeń rzeczywistych w porównaniu do modeli uzyskanych dla epizodów w okresie jesienno-zimowym. Szczególnie niezbyt dokładne dopasowanie do rzeczywistych stężeń można zaobserwować w okresie zimowym na stacji Zabrze. Dopasowanie modeli predykcyjnych do wartości rzeczywistych cechuje się dużą nieregularnością. Zdarza się, że nawet model teoretycznie cechujący się najniższą dokładnością ocenianą w okresie wieloletnim, w epizodzie osiąga wartości najbardziej zbliżone do rzeczywistych, jak w przypadku stacji Złoty Potok, w dniu 01.01.2016 r. w godzinach 18:00 i 19:00.



Rys. 19. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



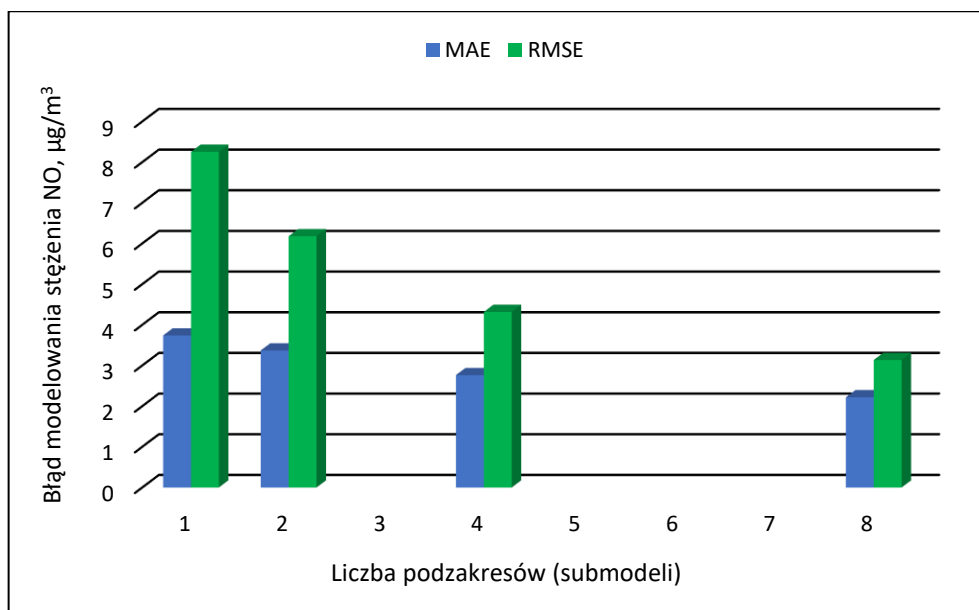
Rys. 20. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

5.1.2. Wyniki modelowania stężeń NO

W tabelach 10 i 11 przedstawiono obliczone błędy modelowania stężeń NO dla podzakresów uzyskanych w wyniku podziału na podstawie stężeń rzeczywistych O₃. Wraz ze wzrostem liczby podzakresów zmniejszały się średnie błędy modelowania na obu stacjach monitoringu powietrza. Dla modeli 2-podzakresowych dla stacji Zabrze wykazano zmniejszenie błędu MAE z 3,74 do 3,37 µg/m³, w porównaniu do modelu pełnozakresowego (zmiana o 9,9 %), a dla stacji Złoty Potok – z 0,43 do 0,31 µg/m³ (zmiana o 27,9 %) oraz zmniejszeniem błędu RMSE na stacji Zabrze z 8,26 do 6,18 µg/m³ (zmiana o 25,2 %), a na stacji Złoty Potok z 0,75 do 0,57 µg/m³ (zmiana o 23,5 %). Dla stacji Zabrze w modelach 4-podzakresowych osiągnięto poprawę dokładności modelowania, objawiającą się zmniejszeniem wartości błędu MAE z 3,74 do 2,76 µg/m³ (zmiana o 26,1 %) i błędu RMSE z 8,26 do 4,32 µg/m³ (zmiana o 47,7 %). W przypadku modeli 8-podzakresowych średnia wartość błędu MAE względem modelu pełnozakresowego obniżyła się z 3,74 do 2,22 µg/m³ (zmiana o 40,6 %), a błędu RMSE – z 8,26 do 3,14 µg/m³ (zmiana o 62 %). Analogicznie, dla stacji Złoty Potok poprawa dokładności modeli 4-podzakresowych względem modelu pełnozakresowego wynosiła 46,8 % i 45,3 % (spadek wartości błędu MAE z 0,43 do 0,23 µg/m³ oraz błędu RMSE z 0,746 do 0,408 µg/m³). W przypadku stacji Złoty Potok i modeli RVS-1/8-ZLP oraz RVS-6/8-ZLP (zakresy stężeń odpowiednio 0,0 – 0,0 µg/m³ i 1,0 – 1,0 µg/m³), nie udało się stworzyć modeli sieci neuronowych z uwagi na brak zmienności w tych zakresach stężeń. Najwyższe wartości błędów występują w przypadku modeli dedykowanych dla podzakresów z najwyższymi poziomami stężeń – na przykład wartość błędu MAE uzyskana przez model RVS-8/8-ZAB, stworzonego dla zakresu stężeń wynoszącego 19,4 – 709,0 µg/m³ wynosi 13,86 µg/m³. Dla porównania, dla modelu RVS-1/8-ZAB dedykowanego dla najniższego podzakresu stężeń w tym podzakresie (0,0 – 1,0 µg/m³) wartość obliczonego błędu MAE wynosiła 0,23 µg/m³. Z kolei dla modelu RVS-2/8-ZAB, który dedykowany był dla wąskiego zakresu stężeń wynoszącego 1,0 – 1,1 µg/m³, uzyskano wartość błędu MAE wynoszącą 0,01 µg/m³. Na rysunkach 21 i 22 przedstawiono graficznie średnie wartości błędów modelowania w zależności od liczby podzakresów (submodeli).

Tab. 10. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO, Zabrze 2011-2016

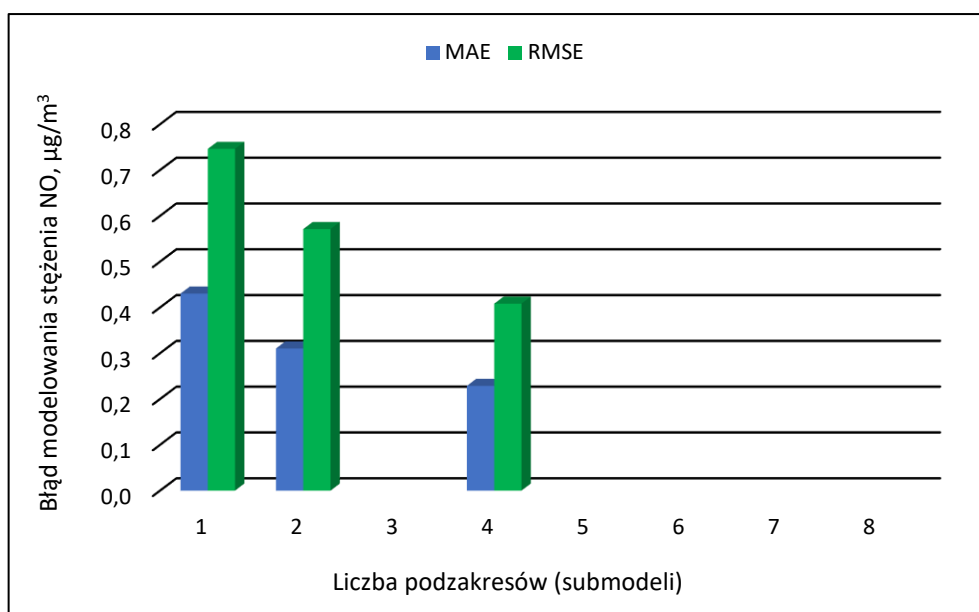
Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych NO, $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	RVS-1/1-ZAB	0,0-709,0	36460	3,74	3,74	8,26	8,26
2 podzakresy	RVS-1/2-ZAB	0,0-3,0	18230	0,52	3,37	0,66	6,18
	RVS-2/2-ZAB	3,0-709,0	18230	6,22		11,70	
4 podzakresy	RVS-1/4-ZAB	0,0-1,1	9115	0,23	2,76	0,32	4,32
	RVS-2/4-ZAB	1,1-3,0	9115	0,34		0,42	
	RVS-3/4-ZAB	3,0-8,0	9115	0,94		1,17	
	RVS-4/4-ZAB	8,0-709,0	9115	9,55		15,37	
8 podzakresów	RVS-1/8-ZAB	0,0-1,0	4558	0,23	2,22	0,32	3,14
	RVS-2/8-ZAB	1,0-1,1	4557	0,01		0,01	
	RVS-3/8-ZAB	1,1-2,0	4558	0,15		0,21	
	RVS-4/8-ZAB	2,0-3,0	4557	0,19		0,26	
	RVS-5/8-ZAB	3,0-4,7	4558	0,39		0,47	
	RVS-6/8-ZAB	4,7-8,0	4557	0,77		0,92	
	RVS-7/8-ZAB	8,0-19,3	4558	2,16		2,64	
	RVS-8/8-ZAB	19,4-709,0	4557	13,86		20,25	



Rys. 21. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO, Zabrze 2011-2016

Tab. 11. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO, Złoty Potok 2011-2016

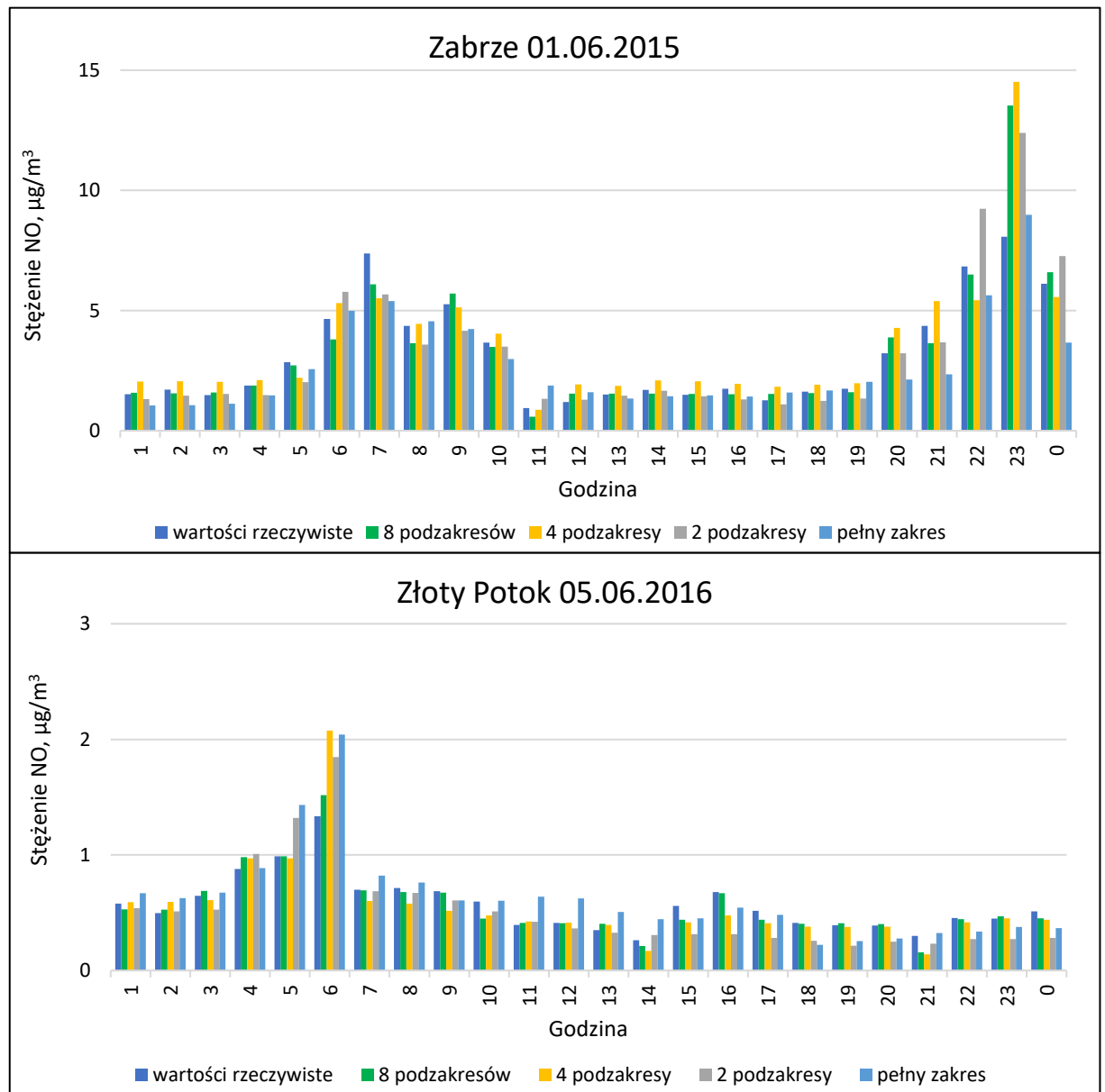
Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych NO, $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	RVS-1/1-ZLP	0,0-50,0	15536	0,43	0,43	0,75	0,75
2 podzakresy	RVS-1/2-ZLP	0,0-0,8	7768	0,18	0,31	0,23	0,57
	RVS-2/2-ZLP	0,8-50,0	7768	0,44		0,91	
4 podzakresy	RVS-1/4-ZLP	0,0-0,3	3884	0,05	0,23	0,09	0,41
	RVS-2/4-ZLP	0,3-0,8	3884	0,06		0,08	
	RVS-3/4-ZLP	0,8-1,0	3884	0,04		0,05	
	RVS-4/4-ZLP	1,0-50,0	3884	0,76		1,41	
8 podzakresów	RVS-1/8-ZLP	0,0-0,0	1942	-	-	-	-
	RVS-2/8-ZLP	0,0-0,3	1942	0,07		0,10	
	RVS-3/8-ZLP	0,3-0,6	1942	0,05		0,06	
	RVS-4/8-ZLP	0,6-0,8	1942	0,04		0,05	
	RVS-5/8-ZLP	0,8-1,0	1942	0,05		0,06	
	RVS-6/8-ZLP	1,0-1,0	1942	-		-	
	RVS-7/8-ZLP	1,0-2,0	1942	0,15		0,21	
	RVS-8/8-ZLP	2,0-50,0	1942	0,99		1,75	



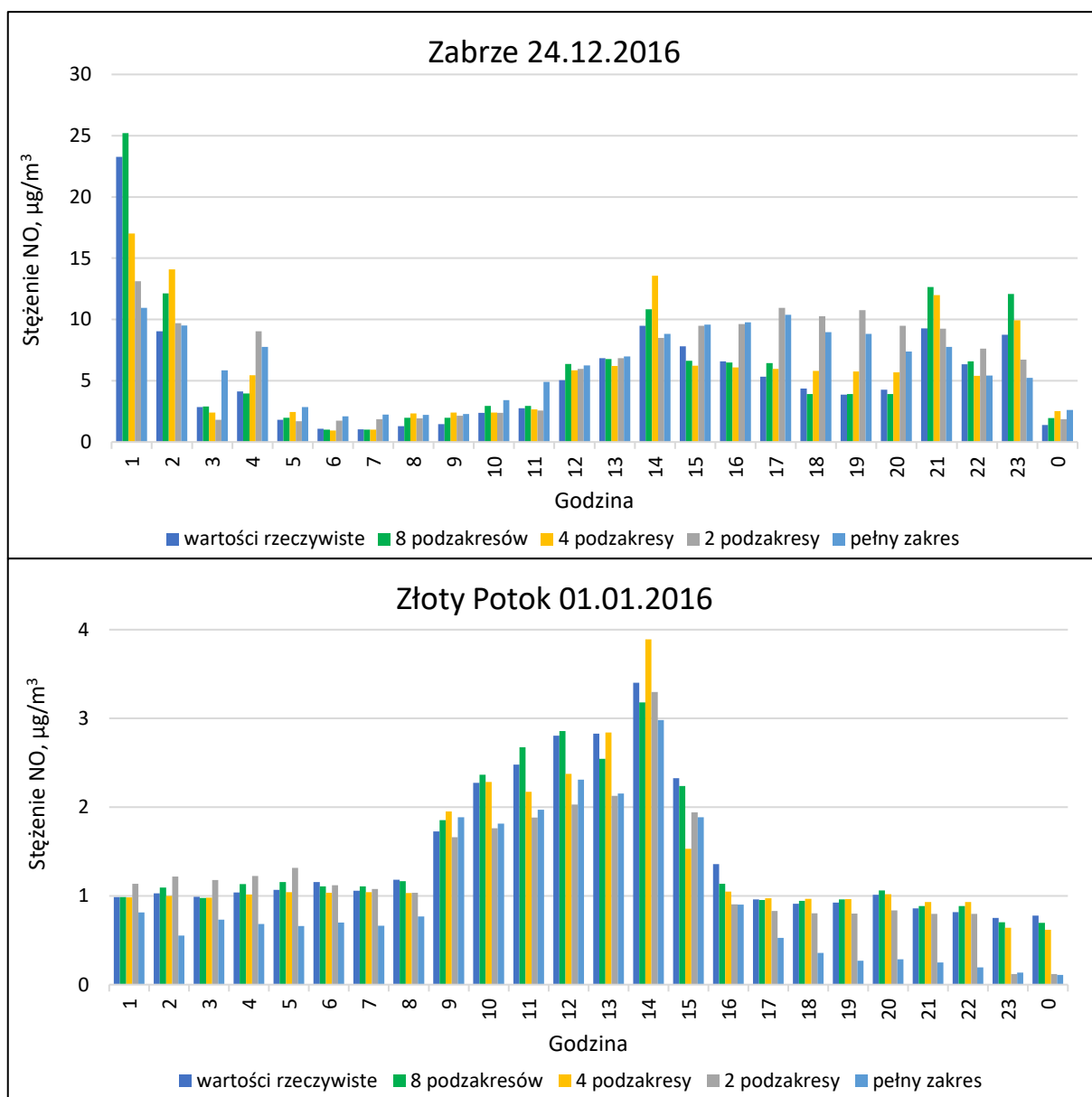
Rys. 22. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO, Złoty Potok 2011-2016

Na rys. 23 i 24 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele RVS w zestawieniu ze stężeniami rzeczywistymi, w wybranych dniach w okresie wiosenno-letnim i jesienno-zimowym. Na rys. 23 można zauważyć stosunkowo dobre wyniki uzyskane przez modele predykcyjne na stacji Zabrze, które odnotowano w godzinach 12:00 – 19:00. Pojawiły się też wyraźne niedopasowania, jak w przypadku

stężeń uzyskanych przez modele podzakresowe o godzinie 23:00. W okresie jesienno-zimowym, na stacji Złoty Potok w godzinach 1:00 – 10:00 można zaobserwować dobre dopasowania odnotowane przez modele 4- i 8-podzakresowe w porównaniu do zarejestrowanych wartości rzeczywistych.



Rys. 23. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



Rys. 24. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

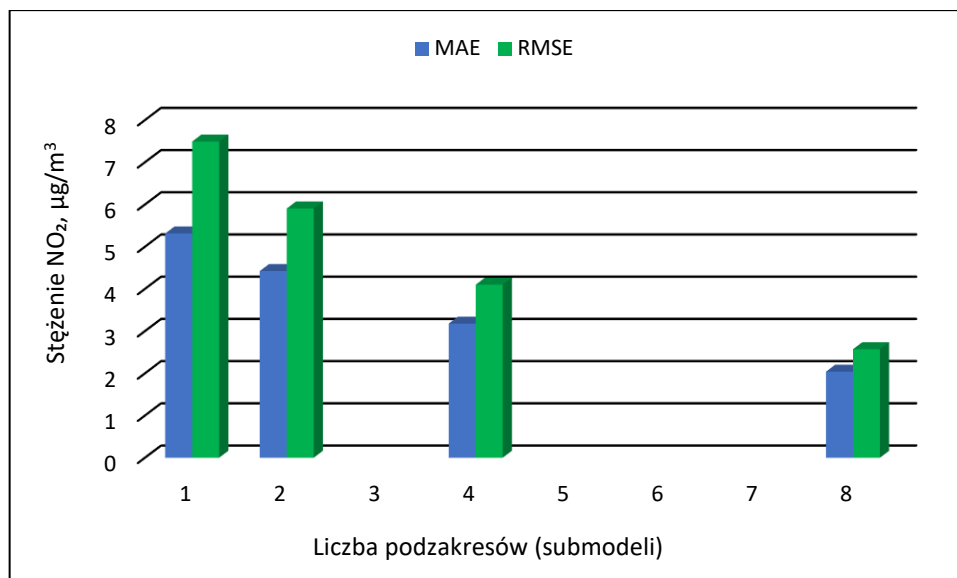
5.1.3. Wyniki modelowania stężeń NO₂

W tabelach 12 i 13 przedstawiono obliczone błędy modelowania stężeń NO₂ dla podzakresów uzyskanych w wyniku podziału na podstawie stężeń rzeczywistych. Wraz z podziałem na kolejne podzakresy stężeń, modele predykcyjne uzyskiwały znaczną poprawę dokładności modelowania. W modelach utworzonych dla dwóch podzakresów stężeń zarejestrowanych na stacji Zabrze uzyskano zmniejszenie średniej wartości błędu względem modelu pełnozakresowego: MAE z 5,31 do 4,42 µg/m³ (zmiana o 17 %) oraz błędu RMSE z 7,49 do 5,91 µg/m³ (zmiana o 21 %). Dla stacji Złoty Potok w modelach

2-podzakresowych również uzyskano poprawę dokładności modelowania w postaci zmniejszenia błędu MAE z 1,92 do 1,53 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 20 %) i błędu RMSE z 2,67 do 2,05 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 23 %). W przypadku 8-podzakresowych modeli na stacji Zabrze odnotowano zmniejszenie wartości błędu MAE względem modelu pełnozakresowego z 5,31 do 2,04 $\mu\text{g}/\text{m}^3$ (zmiana o 61,6 %), a błędu RMSE – z 7,49 do 2,57 $\mu\text{g}/\text{m}^3$ (zmiana o 65,7 %). Dla stacji Złoty Potok poprawa dokładności predykcji skutkowałą zmniejszeniem błędu MAE – z 1,92 do 0,77 $\mu\text{g}/\text{m}^3$ (zmiana o 59,7 %), a w przypadku błędu RMSE – z 2,67 do 0,97 $\mu\text{g}/\text{m}^3$ (zmiana o 63,8 %). Na rys. 25 i 26 zilustrowano wartości uzyskanych błędów w zależności od liczby podzakresów.

Tab. 12. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO_2 , Zabrze 2011-2016

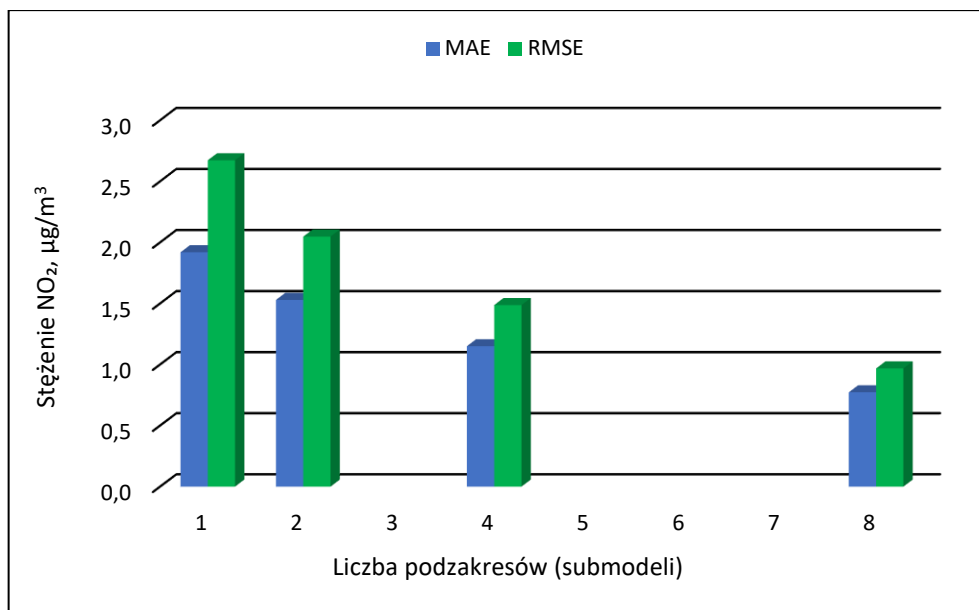
Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych NO_2 , $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	RVS-1/1-ZAB	1,2-145	36460	5,31	5,31	7,49	7,49
2 podzakresy	RVS-1/2-ZAB	1,2-20,0	18230	2,37	4,42	2,97	5,91
	RVS-2/2-ZAB	20,0-145	18230	6,47		8,84	
4 podzakresy	RVS-1/4-ZAB	1,2-11,5	9115	1,56	3,18	1,91	4,10
	RVS-2/4-ZAB	11,5-20,0	9115	1,82		2,19	
	RVS-3/4-ZAB	20,0-33,0	9115	2,73		3,27	
	RVS-4/4-ZAB	33,0-145,0	9115	6,61		9,03	
8 podzakresów	RVS-1/8-ZAB	1,2-8,0	4558	1,03	2,04	1,26	2,57
	RVS-2/8-ZAB	8,0-11,5	4557	0,86		1,00	
	RVS-3/8-ZAB	11,5-15,2	4558	0,93		1,07	
	RVS-4/8-ZAB	15,2-20,0	4557	1,12		1,32	
	RVS-5/8-ZAB	20,0-26,0	4558	1,40		1,64	
	RVS-6/8-ZAB	26,0-33,0	4557	1,68		1,97	
	RVS-7/8-ZAB	33,0-43,1	4558	2,44		2,87	
	RVS-8/8-ZAB	43,1-145,0	4557	6,84		9,44	



Rys. 25. Wartości błęd MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO₂, Zabrze 2011-2016

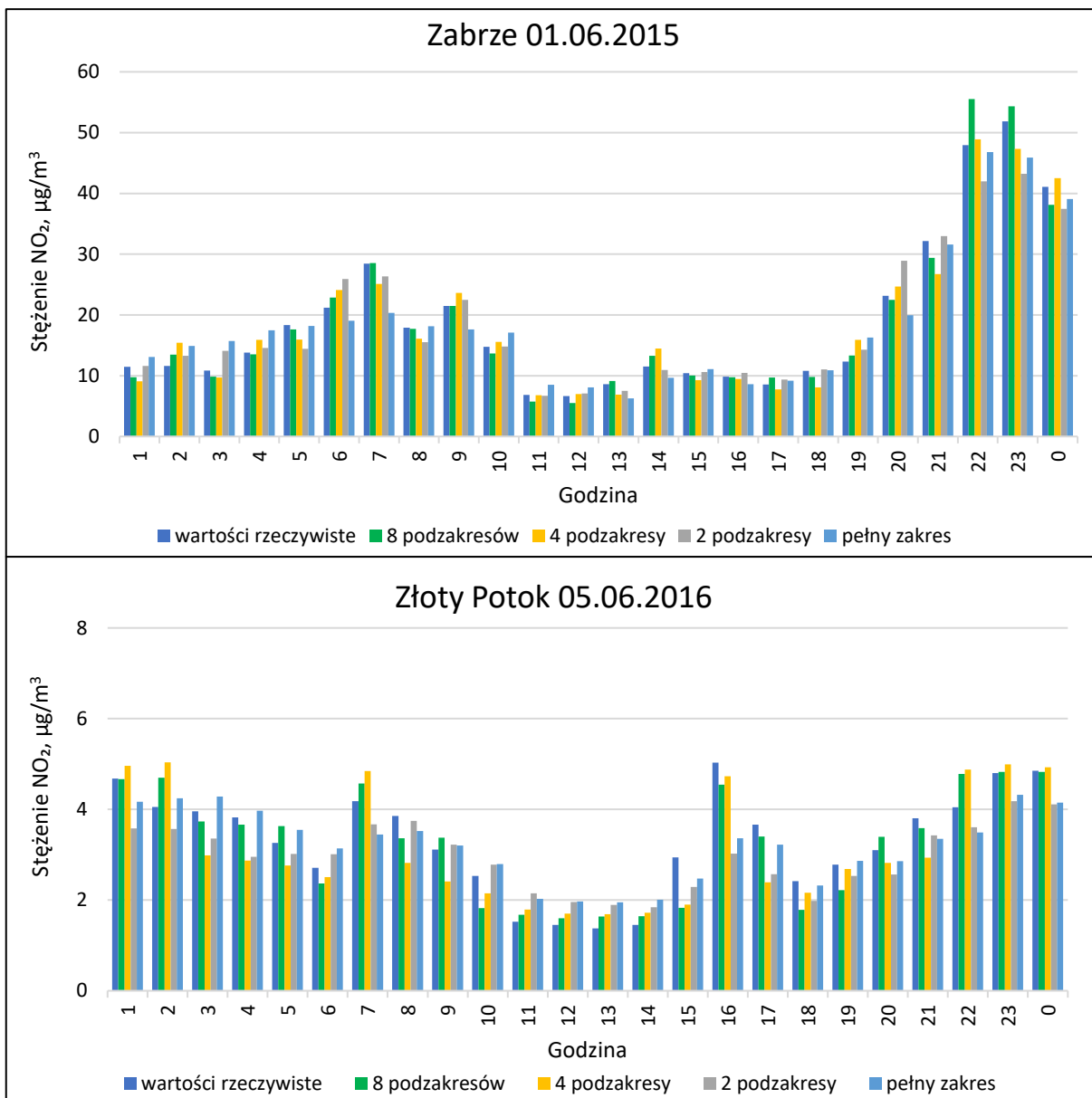
Tab. 13. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO₂, Złoty Potok 2011-2016

Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych NO ₂ , µg/m ³	Liczba przypadków	MAE, µg/m ³	Średnia wartość MAE, µg/m ³	RMSE, µg/m ³	Średnia wartość RMSE, µg/m ³
Pełny zakres	RVS-1/1-ZLP	0,4,0-60,3	15536	1,92	1,92	2,67	2,67
2 podzakresy	RVS-1/2-ZLP	0,4-7,0	7768	0,89	1,53	1,11	2,05
	RVS-2/2-ZLP	7,0-60,3	7768	2,16		2,98	
4 podzakresy	RVS-1/4-ZLP	0,4-4,0	3884	0,59	1,15	0,73	1,49
	RVS-2/4-ZLP	4,0-7,0	3884	0,59		0,71	
	RVS-3/4-ZLP	7,0-11,1	3884	0,89		1,09	
	RVS-4/4-ZLP	11,1-60,3	3884	2,53		3,43	
8 podzakresów	RVS-1/8-ZLP	0,4-3,0	1942	0,43	0,77	0,52	0,97
	RVS-2/8-ZLP	3,0-4,0	1942	0,30		0,35	
	RVS-3/8-ZLP	4,0-5,4	1942	0,28		0,35	
	RVS-4/8-ZLP	5,4-7,0	1942	0,33		0,41	
	RVS-5/8-ZLP	7,0-8,7	1942	0,44		0,50	
	RVS-6/8-ZLP	8,7-11,1	1942	0,63		0,74	
	RVS-7/8-ZLP	11,1-16,6	1942	1,11		1,33	
	RVS-8/8-ZLP	16,6-60,3	1942	2,66		3,54	

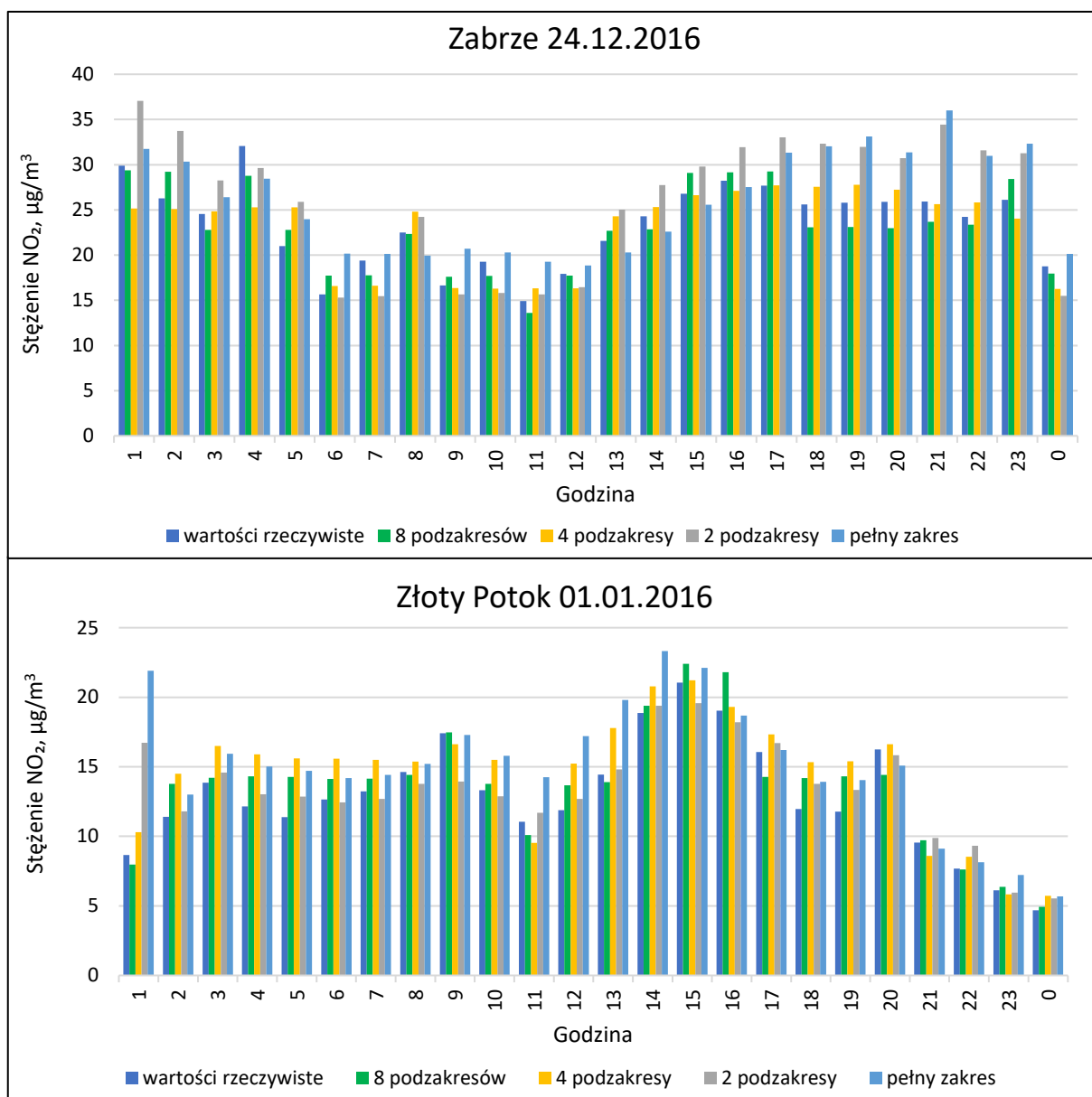


Rys. 26. Wartości błęd MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO₂, Złoty Potok 2011-2016

Na rys. 27 i 28 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele RVS w zestawieniu ze stężeniami rzeczywistymi, w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W epizodach okresu wiosenno-letniego na obu stacjach modele predykcyjne odznaczały się dobrym dopasowaniem do wartości rzeczywistych. W przypadku stacji Złoty Potok i modeli 8- i 4-podzakresowych uzyskano wartości zbliżone do wyraźnie odstającej wartości stężenia rzeczywistego, zarejestrowanej o godzinie 16:00. W epizodach okresu jesienno-zimowego wyraźne niedopasowania do wartości rzeczywistych wykazują modele 2-podzakresowe i model pełnozakresowy.



Rys. 27. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym w okresie wiosenno-letnim



Rys. 28. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

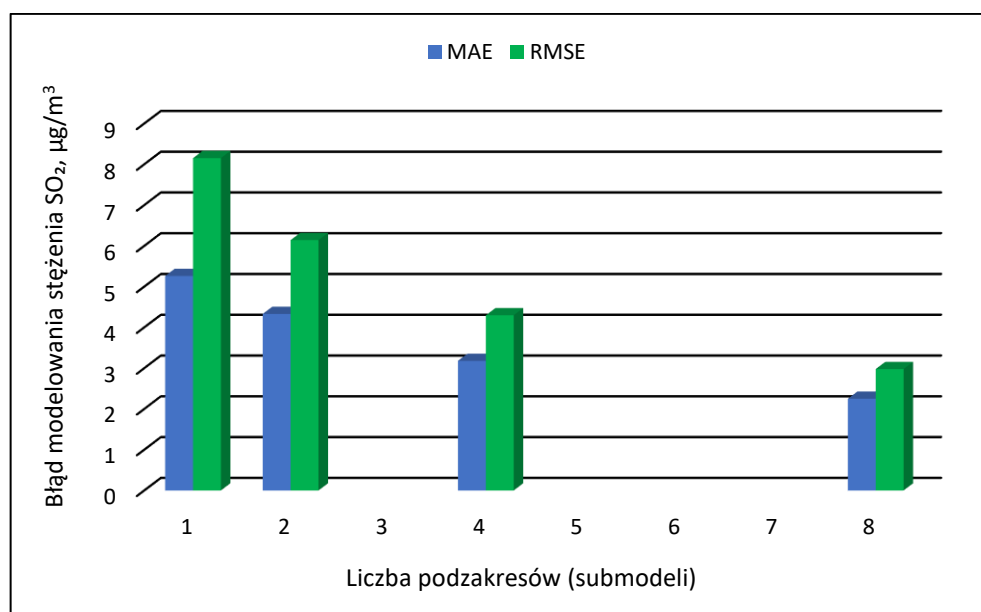
5.1.4. Wyniki modelowania stężeń SO₂

W tabelach 14 i 15 przedstawiono obliczone błędy modelowania stężeń SO₂ dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń rzeczywistych SO₂. Wraz za podziałem na kolejne podzakresy stężeń i wygenerowaniu odpowiednich submodeli, uzyskano wyraźną poprawę dokładności modelowania. Już po podziale na 2 podzakresy średnia wartość błędu MAE i RMSE na podstawie danych zarejestrowanych na stacji Zabrze spadła odpowiednio z 5,26 do 4,33 µg/m³ (zmiana

o 17,7 %) i z 8,15 do 6,14 $\mu\text{g}/\text{m}^3$ (zmiana o 24,7 %), a w przypadku stacji Złoty Potok z 1,96 do 1,66 (zmiana o 15,4 %) i z 3,19 do 2,45 $\mu\text{g}/\text{m}^3$ (zmiana o 23,2 %). W przypadku modeli 8-podzakresowych, obniżenie wartości błędu MAE na obu stacjach wynosiło około 57 %, natomiast błędu RMSE – ponad 63 %. Na rys. 29 i 30 zilustrowano wartości uzyskanych błędów w zależności od liczby podzakresów.

Tab. 14. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych SO_2 , Zabrze 2011-2016

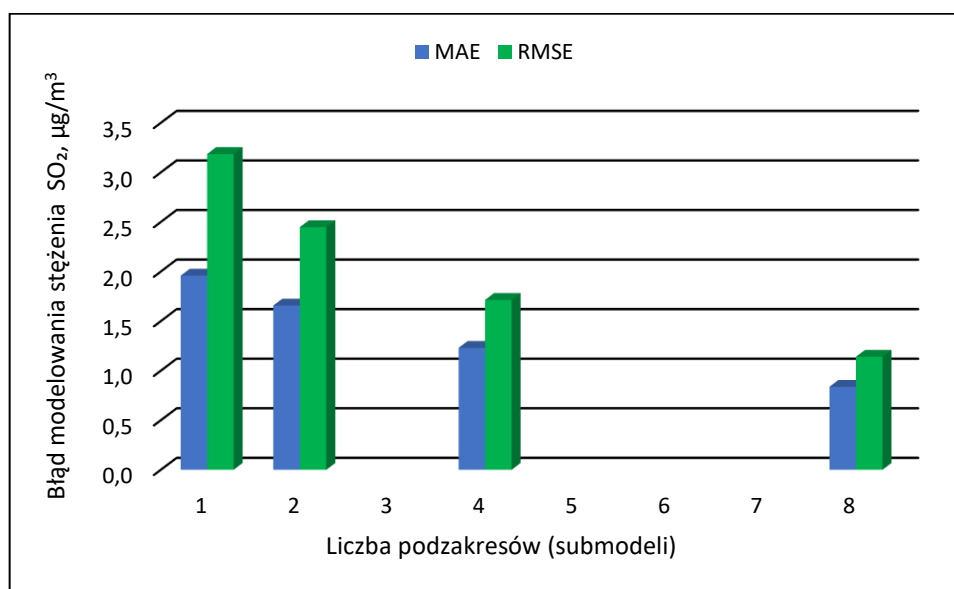
Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych SO_2 , $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	RVS-1/1-ZAB	0,1-362	36460	5,26	5,26	8,15	8,15
2 podzakresy	RVS-1/2-ZAB	0,1-10,0	18230	1,58	4,33	1,95	6,14
	RVS-2/2-ZAB	10,0-362,0	18230	7,08		10,33	
4 podzakresy	RVS-1/4-ZAB	0,1-5,0	9115	0,85	3,18	1,03	4,30
	RVS-2/4-ZAB	5,0-10	9115	1,11		1,32	
	RVS-3/4-ZAB	10,0-22,3	9115	2,66		3,21	
	RVS-4/4-ZAB	22,4-362,0	9115	8,11		11,64	
8 podzakresów	RVS-1/8-ZAB	0,1-3,2	4558	0,55	2,25	0,64	2,98
	RVS-2/8-ZAB	3,2-5,0	4557	0,41		0,48	
	RVS-3/8-ZAB	5,0-7,0	4558	0,48		0,56	
	RVS-4/8-ZAB	7,0-10	4557	0,73		0,85	
	RVS-5/8-ZAB	10,0-14,9	4558	1,16		1,36	
	RVS-6/8-ZAB	14,9-22,3	4557	1,82		2,14	
	RVS-7/8-ZAB	22,4-36,0	4558	3,02		3,58	
	RVS-8/8-ZAB	36,0-362,0	4557	9,84		14,21	



Rys. 29. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO_2 , Zabrze 2011-2016

Tab. 15. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych SO₂, Złoty Potok 2011-2016

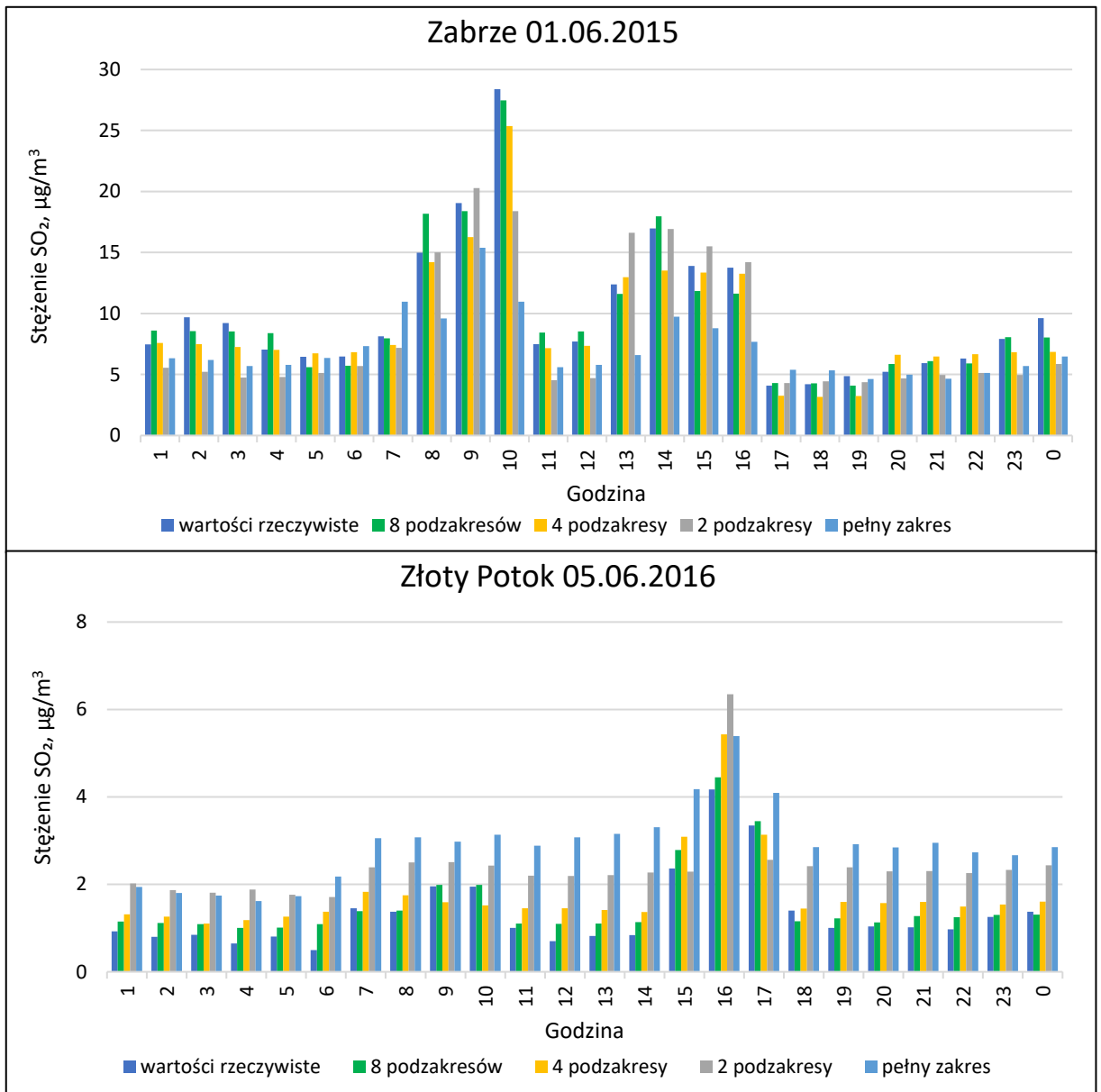
Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych SO ₂ , µg/m ³	Liczba przypadków	MAE, µg/m ³	Średnia wartość MAE, µg/m ³	RMSE, µg/m ³	Średnia wartość RMSE, µg/m ³
Pełny zakres	RVS-1/1-ZLP	0,0-85,0	15536	1,96	1,96	3,19	3,19
2 podzakresy	RVS-1/2-ZLP	0,0-4,0	7768	0,63	1,66	0,78	2,45
	RVS-2/2-ZLP	4,0-85,0	7768	2,68		4,12	
4 podzakresy	RVS-1/4-ZLP	0,0-2,1	3884	0,38	1,23	0,46	1,71
	RVS-2/4-ZLP	2,1-4,0	3884	0,36		0,46	
	RVS-3/4-ZLP	4,0-7,4	3884	0,73		0,89	
	RVS-4/4-ZLP	7,4-85,0	3884	3,45		5,05	
8 podzakresów	RVS-1/8-ZLP	0,0-1,8	1942	0,26	0,84	0,33	1,14
	RVS-2/8-ZLP	1,8-2,1	1942	0,04		0,06	
	RVS-3/8-ZLP	2,1-3,0	1942	0,19		0,25	
	RVS-4/8-ZLP	3,0-4,0	1942	0,23		0,29	
	RVS-5/8-ZLP	4,0-5,0	1942	0,34		0,39	
	RVS-6/8-ZLP	5,0-7,4	1942	0,58		0,68	
	RVS-7/8-ZLP	7,4-12,0	1942	1,01		1,22	
	RVS-8/8-ZLP	12,0-85,0	1942	4,05		5,91	



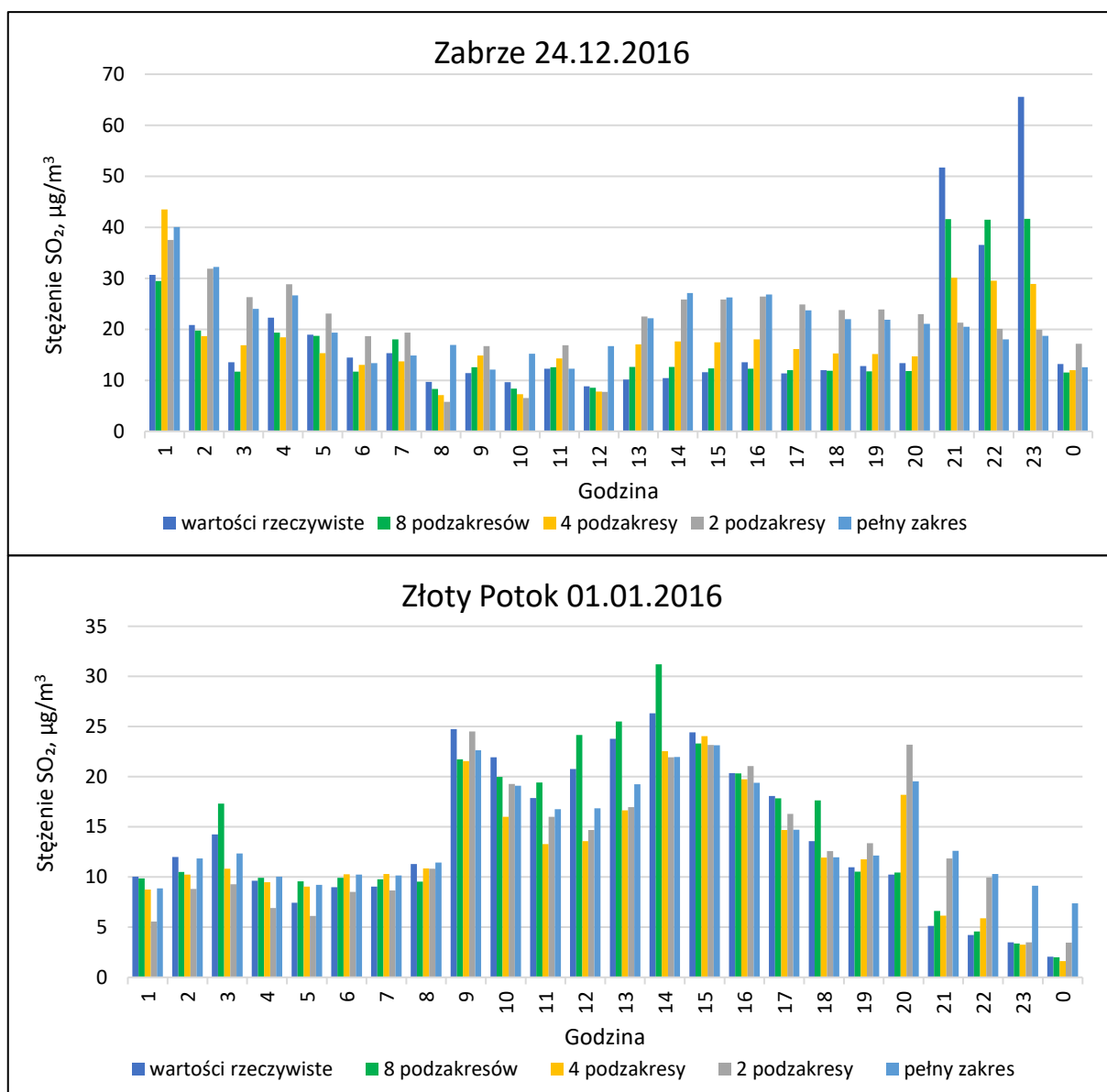
Rys. 30. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO₂, Złoty Potok 2011-2016

Na rys. 31 i 32 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele RVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W obu rozpatrywanych epizodach na obu stacjach najbardziej dopasowanymi modelami predykcyjnymi były modele 8-podzakresowe, które wygenerowały stężenia najbardziej zbliżone do wartości rzeczywistych. Modele te stosunkowo dobrze poradziły sobie również z nagłym

wzrostem wartości rzeczywistych, które odnotowano w dniu 24.12.2016 na stacji Zabrze w godzinach 21:00, 22:00 i 23:00.



Rys. 31. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



Rys. 32. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

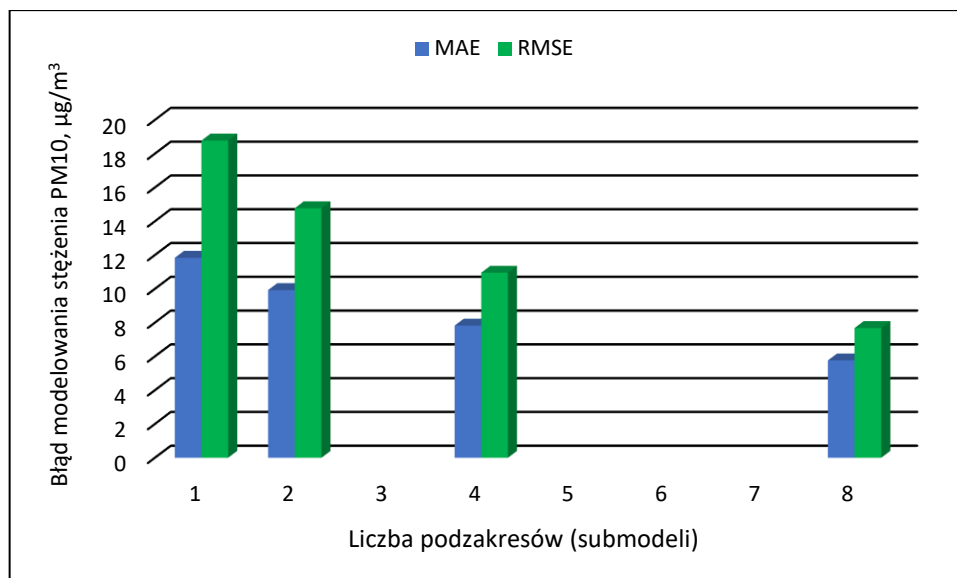
5.1.5. Wyniki modelowania stężeń PM10

W tabelach 16 i 17 przedstawiono obliczone błędy modelowania stężeń PM10 dla podzakresów uzyskanych w wyniku podziału względem stężeń rzeczywistych PM10. Podział na kolejne podzakresy skutkowało wyraźnym obniżeniem wartości błędów MAE i RMSE. Dzięki modelom 2-podzakresowym uzyskano zmniejszenie średniej wartości błędu MAE z 11,84 do 9,93 µg/m³ (zmiana o 16,1 %) w przypadku stacji Zabrze oraz z 6,88 do 5,05 µg/m³ (zmiana o 26,6 %) w przypadku stacji Złoty Potok. Zastosowanie modeli 4-podzakresowych skutkowało zmniejszeniem średniej wartości błędu MAE z 11,84 do 7,83 µg/m³ (zmiana o 33,9 %) w przypadku stacji Zabrze i z 6,88 do 3,46

$\mu\text{g}/\text{m}^3$ (zmiana o 49,6 %) w przypadku stacji Złoty Potok. W przypadku modeli 8-podzakresowych poprawa była jeszcze wyraźniejsza i na stacji Zabrze objawiła się zmniejszeniem średniej wartości błędu MAE z 11,84 do 5,76 $\mu\text{g}/\text{m}^3$ (zmiana o 51,3 %), a na stacji Złoty Potok – z 6,88 do 2,11 $\mu\text{g}/\text{m}^3$ (zmiana o 69,3 %). Na stacji Zabrze procentowe obniżenie błędu RMSE dla modeli 2-, 4- i 8-podzakresowych wynosiło, odpowiednio 21,3 % (z 18,80 do 14,77 $\mu\text{g}/\text{m}^3$), 41,7 % (z 18,80 do 10,96 $\mu\text{g}/\text{m}^3$) i 59,1 % (z 18,80 do 7,68 $\mu\text{g}/\text{m}^3$), a na stacji Złoty Potok – 28,2 % (z 9,34 do 6,70 $\mu\text{g}/\text{m}^3$), 52 % (z 9,34 do 4,48 $\mu\text{g}/\text{m}^3$) i 71,5 % (z 9,34 do 2,66 $\mu\text{g}/\text{m}^3$). Na rys. 33 i 34 zilustrowano wartości uzyskanych błędów w zależności od liczby podzakresów.

Tab. 16. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych PM10, Zabrze 2011-2016

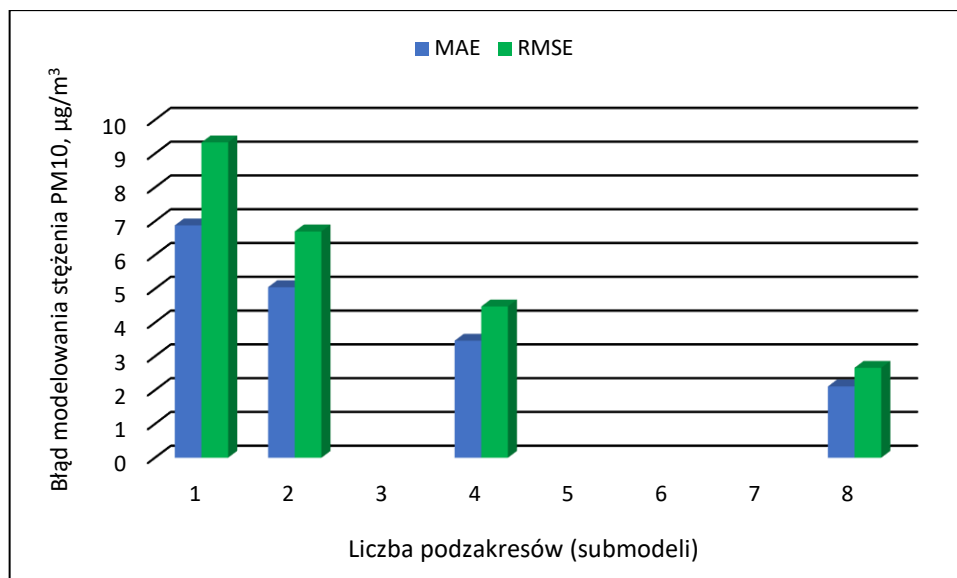
Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych PM10, $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	RVS-1/1-ZAB	0,1-1145,0	36460	11,84	11,84	18,78	18,80
2 podzakresy	RVS-1/2-ZAB	0,1-33,0	18230	4,66	9,93	5,80	14,77
	RVS-2/2-ZAB	33,0-1145,0	18230	15,20		23,75	
4 podzakresy	RVS-1/4-ZAB	0,1-20,0	9115	3,09	7,83	3,80	10,96
	RVS-2/4-ZAB	20,0-33,0	9115	2,97		3,53	
	RVS-3/4-ZAB	33,0-57,1	9115	5,02		6,06	
	RVS-4/4-ZAB	57,1-1145,0	9115	20,23		30,44	
8 podzakresów	RVS-1/8-ZAB	0,1-14,0	4558	2,26	5,76	2,78	7,68
	RVS-2/8-ZAB	14,0-20,0	4557	1,49		1,74	
	RVS-3/8-ZAB	20,0-26,0	4558	1,51		1,76	
	RVS-4/8-ZAB	26,0-33,0	4557	1,74		2,03	
	RVS-5/8-ZAB	33,0-42,5	4558	2,30		2,66	
	RVS-6/8-ZAB	42,5-57,1	4557	3,51		4,11	
	RVS-7/8-ZAB	57,1-91,1	4558	6,99		8,42	
	RVS-8/8-ZAB	91,1-1145,0	4557	26,32		37,97	



Rys. 33. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń PM10, Zabrze 2011-2016

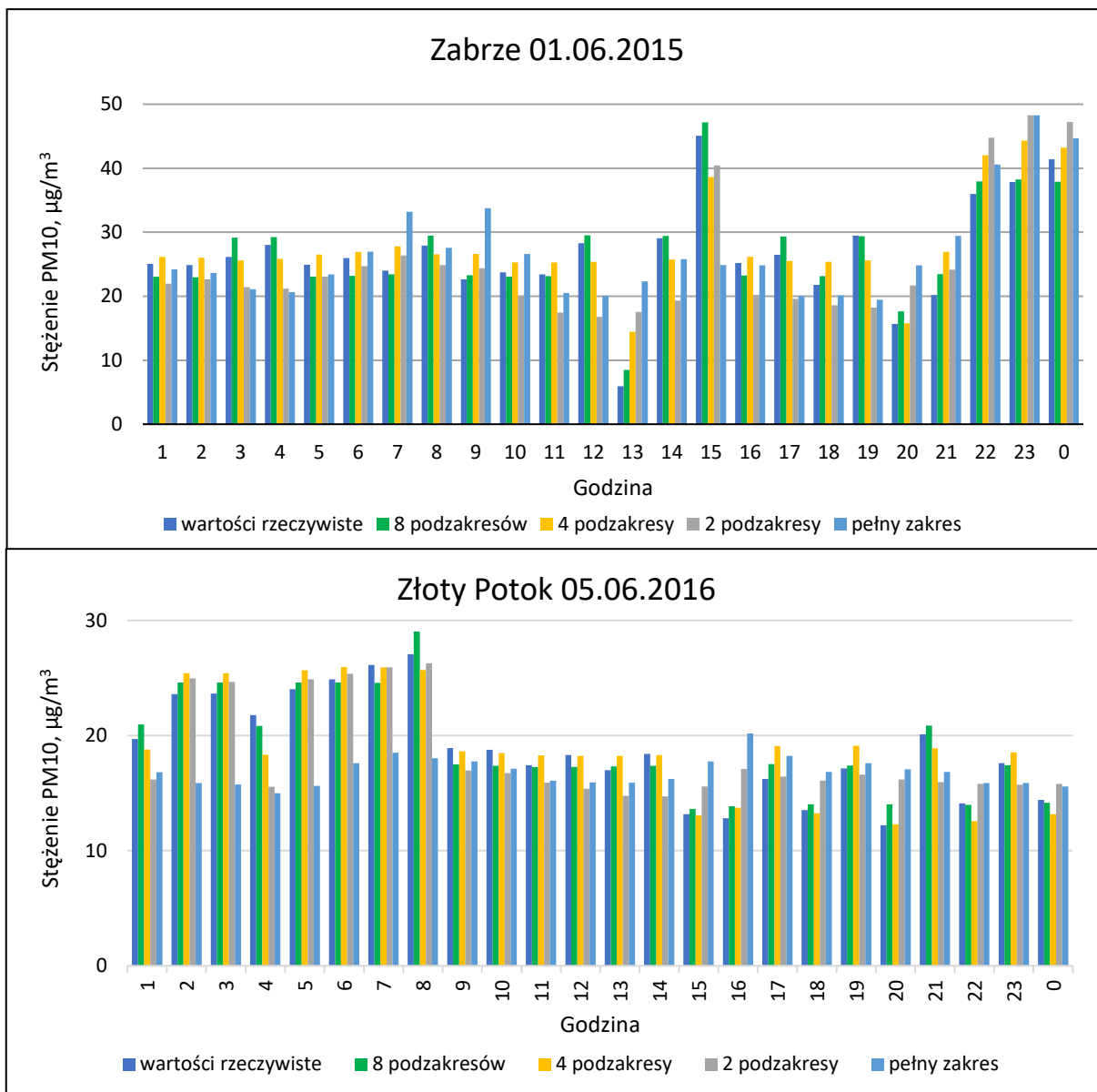
Tab. 17. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych PM10, Złoty Potok 2011-2016

Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych PM10, µg/m³	Liczba przypadków	MAE, µg/m³	Średnia wartość MAE, µg/m³	RMSE, µg/m³	Średnia wartość RMSE, µg/m³
Pełny zakres	RVS-1/1-ZLP	3,0-138,0	15536	6,88	6,88	9,34	9,34
	RVS-1/2-ZLP	3,0-22,8	7768	2,99		3,64	
2 podzakresy	RVS-2/2-ZLP	22,8-138,0	7768	7,11	5,05	9,75	6,70
	RVS-1/4-ZLP	3,0-16,0	3884	1,88		2,31	
4 podzakresy	RVS-2/4-ZLP	16,0-22,8	3884	1,62	3,46	1,90	4,48
	RVS-3/4-ZLP	22,8-32,9	3884	2,29		2,71	
	RVS-4/4-ZLP	32,9-138,0	3884	8,07		11,00	
8 podzakresów	RVS-1/8-ZLP	3,0-12,1	1942	1,43	2,11	1,76	2,66
	RVS-2/8-ZLP	12,1-16,0	1942	0,84		0,99	
	RVS-3/8-ZLP	16,0-19,0	1942	0,84		0,97	
	RVS-4/8-ZLP	19,0-22,8	1942	0,89		1,03	
	RVS-5/8-ZLP	22,8-27,0	1942	1,04		1,21	
	RVS-6/8-ZLP	27,0-32,9	1942	1,46		1,69	
	RVS-7/8-ZLP	32,9-43,0	1942	2,50		2,94	
	RVS-8/8-ZLP	43,0-138,0	1942	7,88		10,71	

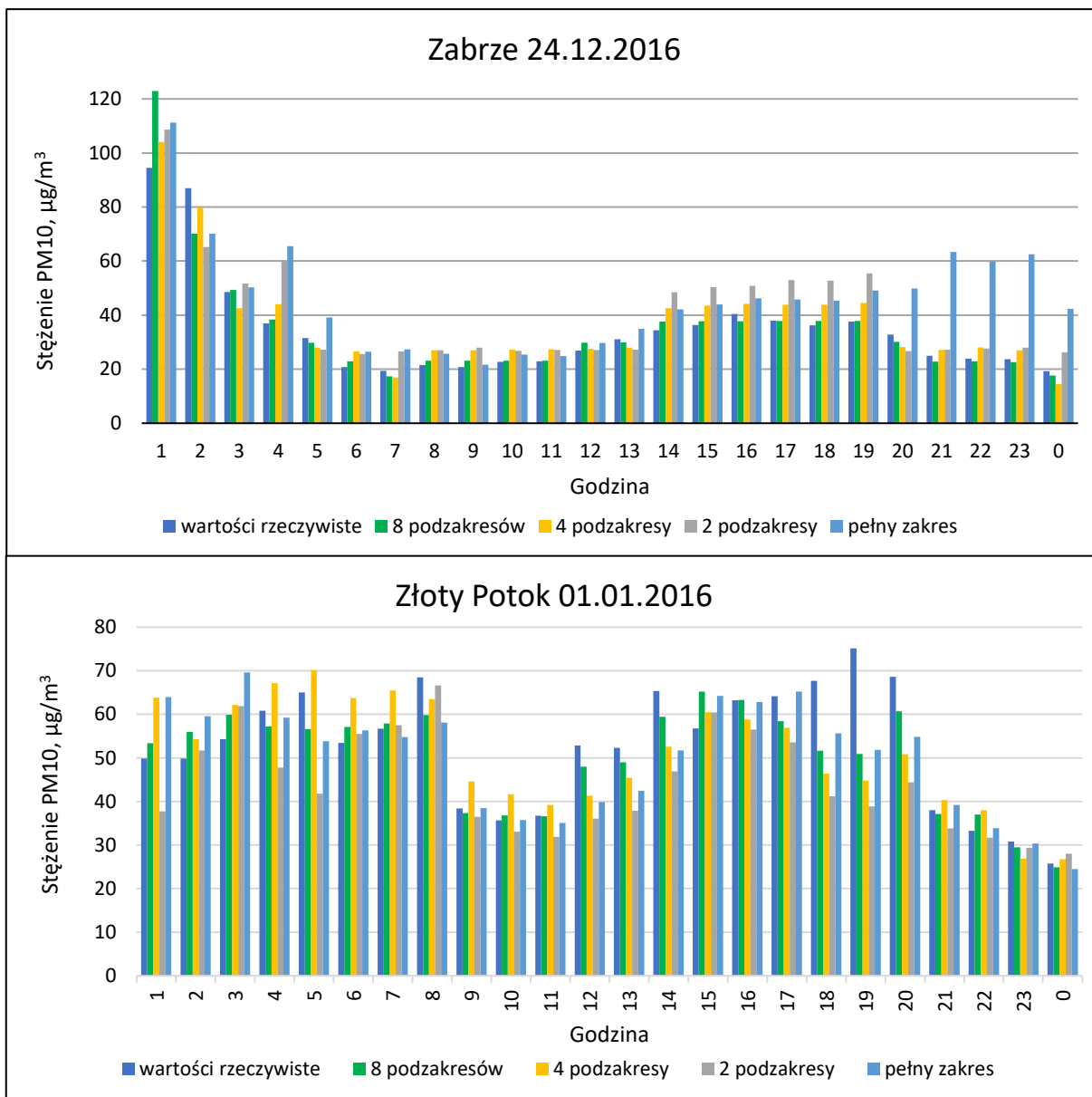


Rys. 34. Wartości błęd MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO_2 , Złoty Potok 2011-2016

Na rys. 35 i 36 zilustrowano dobowe przebiegi stężeń predykcyjnych, uzyskanych za pomocą modeli RVS, w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W epizodach trudno wskazać model najbardziej dopasowany do wartości rzeczywistych. Wyraźne niedopasowanie odnotowuje model pełnozakresowy w dniu 24.12.2016 na stacji Zabrze w godzinach 21:00-0:00.



Rys. 35. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



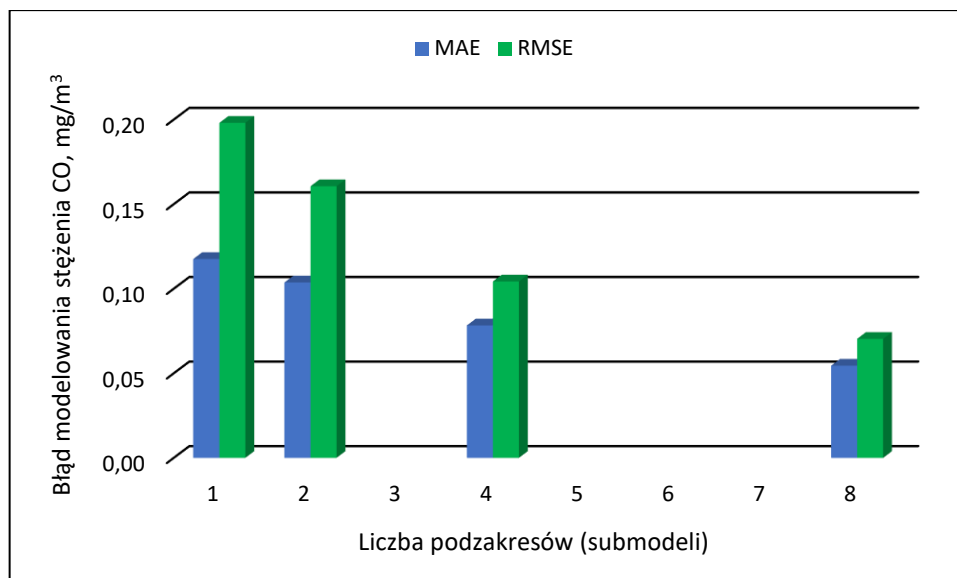
Rys. 36. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

5.1.6. Wyniki modelowania stężeń CO

W tabelach 18 i 19 przedstawiono obliczone błędy modelowania stężeń CO dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń rzeczywistych CO. Przedstawione wyniki dotyczą wyłącznie stacji Zabrze, na stacji Złoty Potok stężenia CO nie były rejestrowane. Wraz z podziałem na kolejne podzakresy stężeń i wprowadzaniem coraz większej liczby submodeli, obserwowano wyraźną poprawę dokładności modelowania. Już dzięki zastosowaniu modeli predykcyjnych utworzonych na podstawie danych podzielonych na dwa podzakresy uzyskano obniżenie średniej wartości błędu MAE z 0,118 do 0,104 mg/m³ (zmiana o 11,8 %) i błędu RMSE z 0,198 do 0,161 mg/m³ (zmiana o 18,9 %), w stosunku do modelu utworzonego na podstawie pełnego zakresu stężeń. W przypadku modeli 4-podzakresowych odnotowano wartości błędu MAE niższe o 33,4 % (spadek z 0,118 do 0,078 mg/m³), a błędu RMSE o 47,3 % (spadek z 0,198 do 0,104 mg/m³). W przypadku modeli 8-podzakresowych obniżenie średniej wartości błędu MAE wynosiło 53,6 % (z 0,118 do 0,055 mg/m³), a błędu RMSE o 64,5 % (z 0,198 do 0,070 mg/m³). Na rys. 37 zilustrowano wartości uzyskanych błędów w zależności od liczby podzakresów.

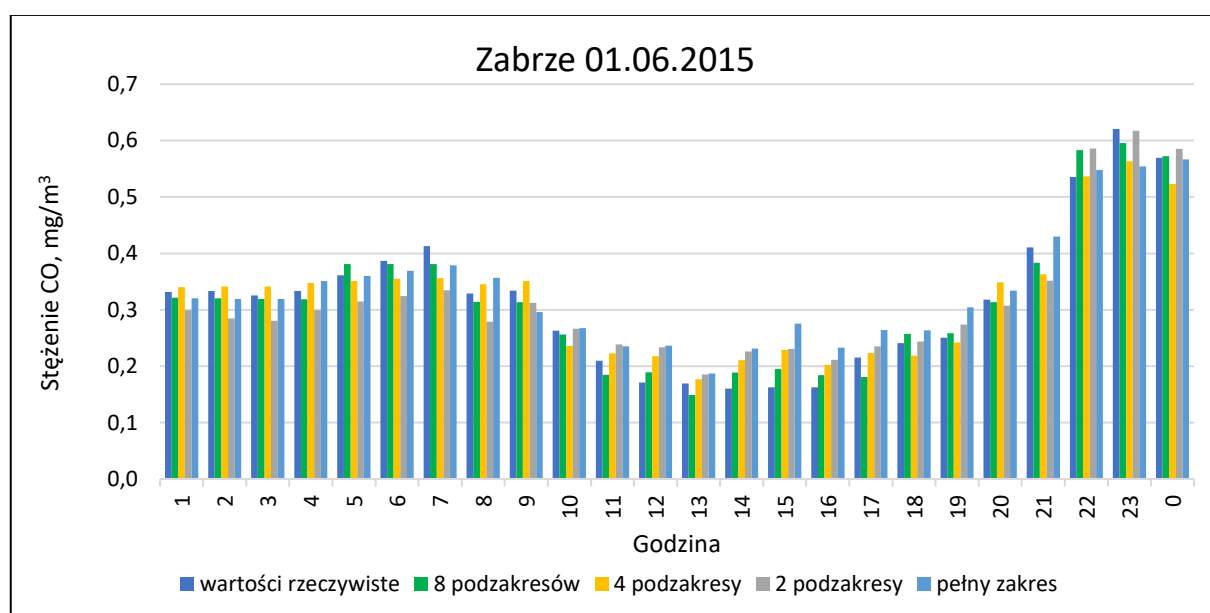
Tab. 18. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych CO, Zabrze 2011-2016

Liczba podzakresów	Nazwa modelu	Zakres stężeń rzeczywistych CO, mg/m ³	Liczba przypadków	MAE, mg/m ³	Średnia wartość MAE, mg/m ³	RMSE, mg/m ³	Średnia wartość RMSE, mg/m ³
Pełny zakres	RVS-1/1-ZAB	0,1-9,0	36460	0,118	0,118	0,198	0,198
2 podzakresy	RVS-1/2-ZAB	0,1-0,4	18230	0,047	0,104	0,059	0,161
	RVS-2/2-ZAB	0,4-9,0	18230	0,160		0,262	
4 podzakresy	RVS-1/4-ZAB	0,1-0,3	9115	0,034	0,078	0,042	0,104
	RVS-2/4-ZAB	0,3-0,4	9115	0,030		0,035	
	RVS-3/4-ZAB	0,4-0,7	9115	0,055		0,066	
	RVS-4/4-ZAB	0,7-9,0	9115	0,196		0,274	
8 podzakresów	RVS-1/8-ZAB	0,1-0,2	4558	0,027	0,055	0,033	0,070
	RVS-2/8-ZAB	0,2-0,3	4557	0,014		0,016	
	RVS-3/8-ZAB	0,3-0,3	4558	0,014		0,017	
	RVS-4/8-ZAB	0,3-0,4	4557	0,018		0,021	
	RVS-5/8-ZAB	0,4-0,5	4558	0,025		0,030	
	RVS-6/8-ZAB	0,5-0,7	4557	0,038		0,045	
	RVS-7/8-ZAB	0,7-1,1	4558	0,077		0,093	
	RVS-8/8-ZAB	1,1-9,0	4557	0,223		0,309	

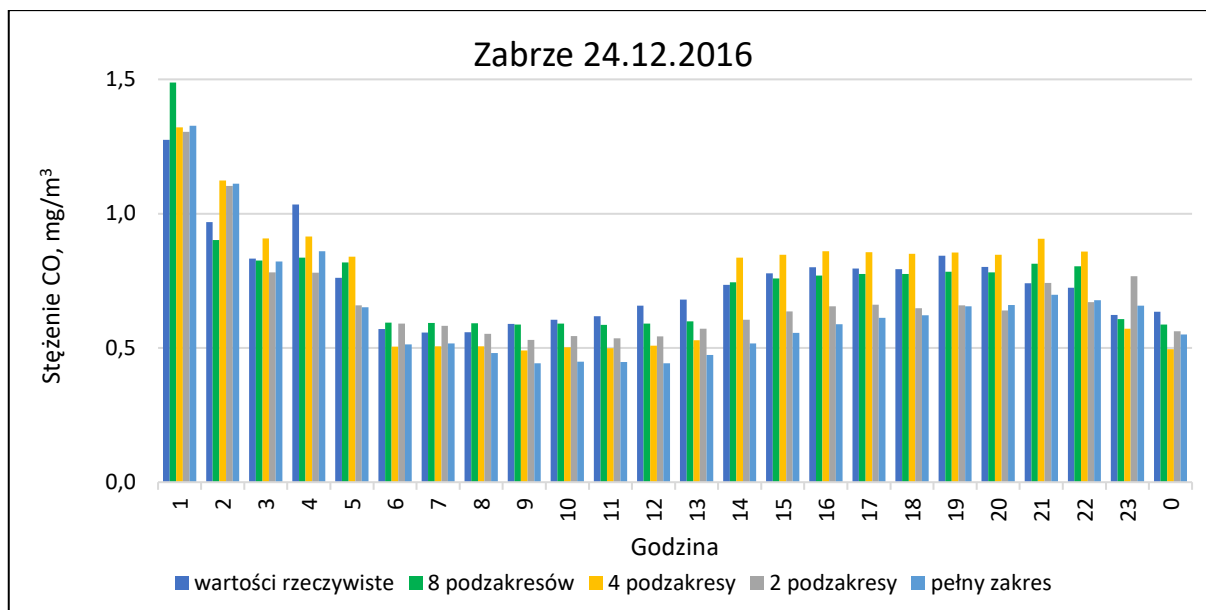


Rys. 37. Wartości błęd MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń CO, Zabrze 2011-2016

Na rys. 38 i 39 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele RVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W obu rozpatrywanych dniach najbardziej zbliżone do wartości rzeczywistych były stężenia aproksymowane modelami 8-podzakresowymi.



Rys. 38. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze w wybranym dniu w okresie wiosenno-letnim



Rys. 39. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze w wybranym dniu w okresie jesienno-zimowym

5.2. Modele PVS

W tym rozdziale przedstawiono wyniki modelowania w podzakresach uzyskanych po sortowaniu przypadków według rosnących stężeń predykcyjnych wybranego zanieczyszczenia. Modele PVS powstały na bazie pełnego zakresu stężeń predykcyjnych, który następnie został podzielony na węższe podzakresy stężeń, zgodnie ze schematami przedstawionymi na rys. 15 i 16. Dla wszystkich zanieczyszczeń powietrza modele PVS okazały się mniej dokładne niż analogiczne modele RVS.

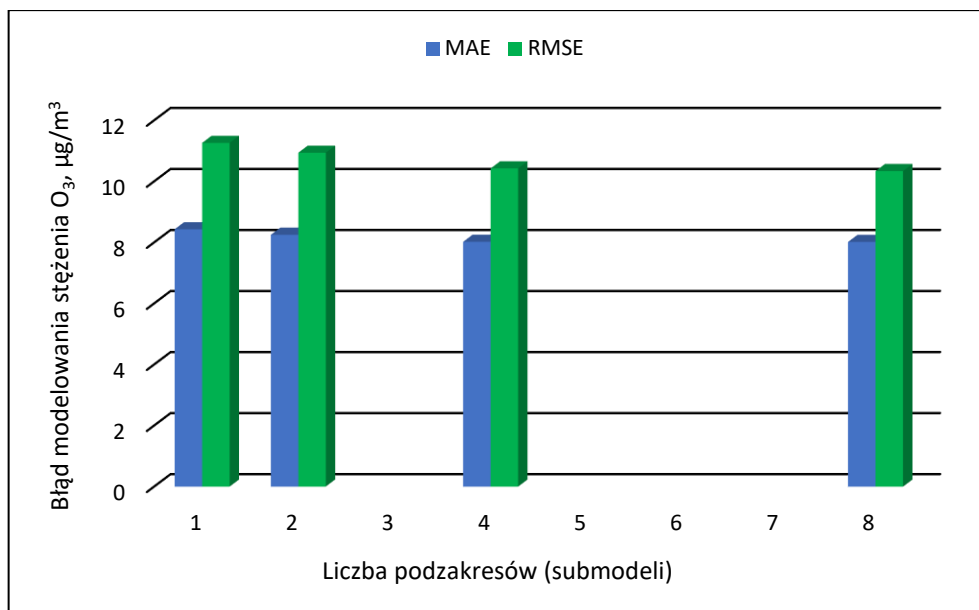
5.2.1. Wyniki modelowania stężeń O_3

W tabelach 19 i 20 przedstawiono obliczone błędy modelowania stężeń O_3 dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń predykcyjnych CO. Podział na coraz większą liczbę podzakresów (i tym samym submodeli) na ogół powodował zwiększenie dokładności predykcji, przejawiające się w obniżeniu średniej wartości błędów MAE i RMSE. Wyjątkiem są modele 8-podzakresowe uzyskane dla obu stacji monitoringu - w przypadku stacji Zabrze odnotowano taką samą średnią wartość błędów MAE jak w przypadku modeli uzyskanych dla 4 podzakresów ($8,02 \mu\text{g}/\text{m}^3$), a przypadku stacji Złoty Potok zaobserwowano wzrost średnich wartości błędów MAE i RMSE. Zastosowanie modeli 2-podzakresowych

skutkowało obniżeniem średniej wartości błędu MAE względem modelu pełnozakresowego z 8,40 do 8,25 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 2 %) na stacji Zabrze i z 8,41 do 7,91 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 6 %) na stacji Złoty Potok. Zastosowanie modeli 4-podzakresowych skutkowało obniżeniem średniej wartości błędu MAE dla danych zarejestrowanych na stacji Zabrze z 8,40 do 8,02 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 5 %), a błędu RMSE – z 11,30 do 10,43 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 7 %). W modelach 8-podzakresowych obniżono wartości błędu RMSE do 10,34 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 8 %). Na stacji Złoty Potok zastosowanie modeli 4-podzakresowych spowodowało obniżenie średniej wartości błędu MAE z 8,41 do 7,55 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 10 %), a średniej wartości błędu RMSE – z 10,70 do 9,68 $\mu\text{g}/\text{m}^3$ (zmiana o ok. 9 %). Na rysunkach 40 i 41 zobrazowano średnie wartości błędów MAE i RMSE w zależności od liczby podzakresów.

Tab. 19. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych O_3 , Zabrze 2011-2016

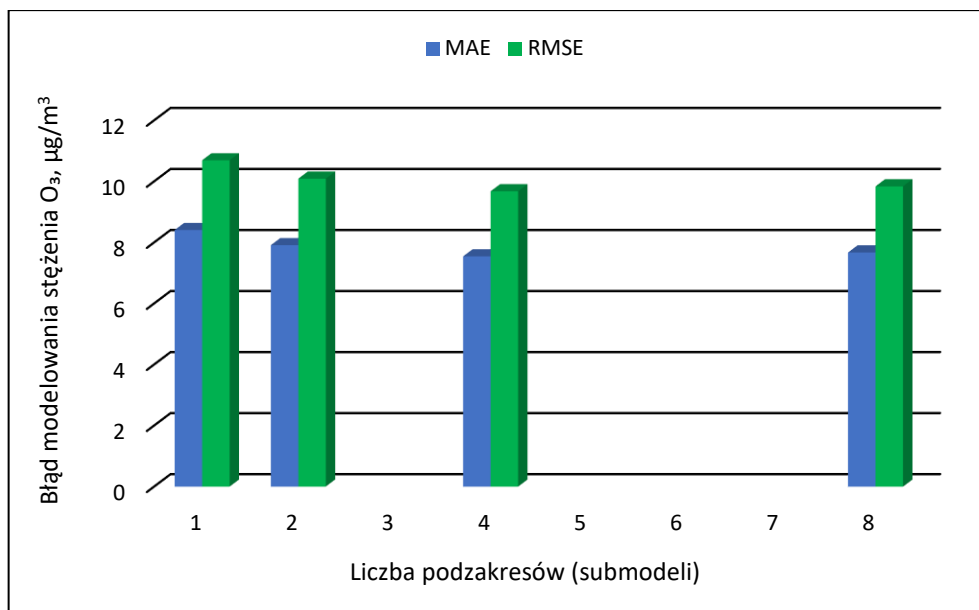
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych O_3 , $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	PVS-1/1-ZAB	1,0-165,0	36460	8,43	8,40	11,26	11,30
2 podzakresy	PVS-1/2-ZAB	1,0-37,4	18230	6,10	8,25	8,51	10,94
	PVS-2/2-ZAB	37,4-165,0	18230	10,40		13,37	
4 podzakresy	PVS-1/4-ZAB	1,0-17,7	9115	3,50	8,02	5,01	10,43
	PVS-2/4-ZAB	17,7-37,4	9115	8,45		10,79	
	PVS-3/4-ZAB	37,4-59,9	9115	9,84		12,74	
	PVS-4/4-ZAB	59,9-165,0	9115	10,3		13,17	
8 podzakresów	PVS-1/8-ZAB	1,0-8,6	4558	1,98	8,02	2,85	10,34
	PVS-2/8-ZAB	8,6-17,7	4557	4,81		6,33	
	PVS-3/8-ZAB	17,7-27,5	4558	7,47		9,56	
	PVS-4/8-ZAB	27,5-37,4	4557	9,4		11,91	
	PVS-5/8-ZAB	37,4-47,6	4558	9,91		12,73	
	PVS-6/8-ZAB	47,6-59,9	4557	9,98		12,87	
	PVS-7/8-ZAB	59,9-79,5	4558	9,78		12,73	
	PVS-8/8-ZAB	79,5-165,0	4557	10,81		13,74	



Rys. 40. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O₃, Zabrze 2011-2016

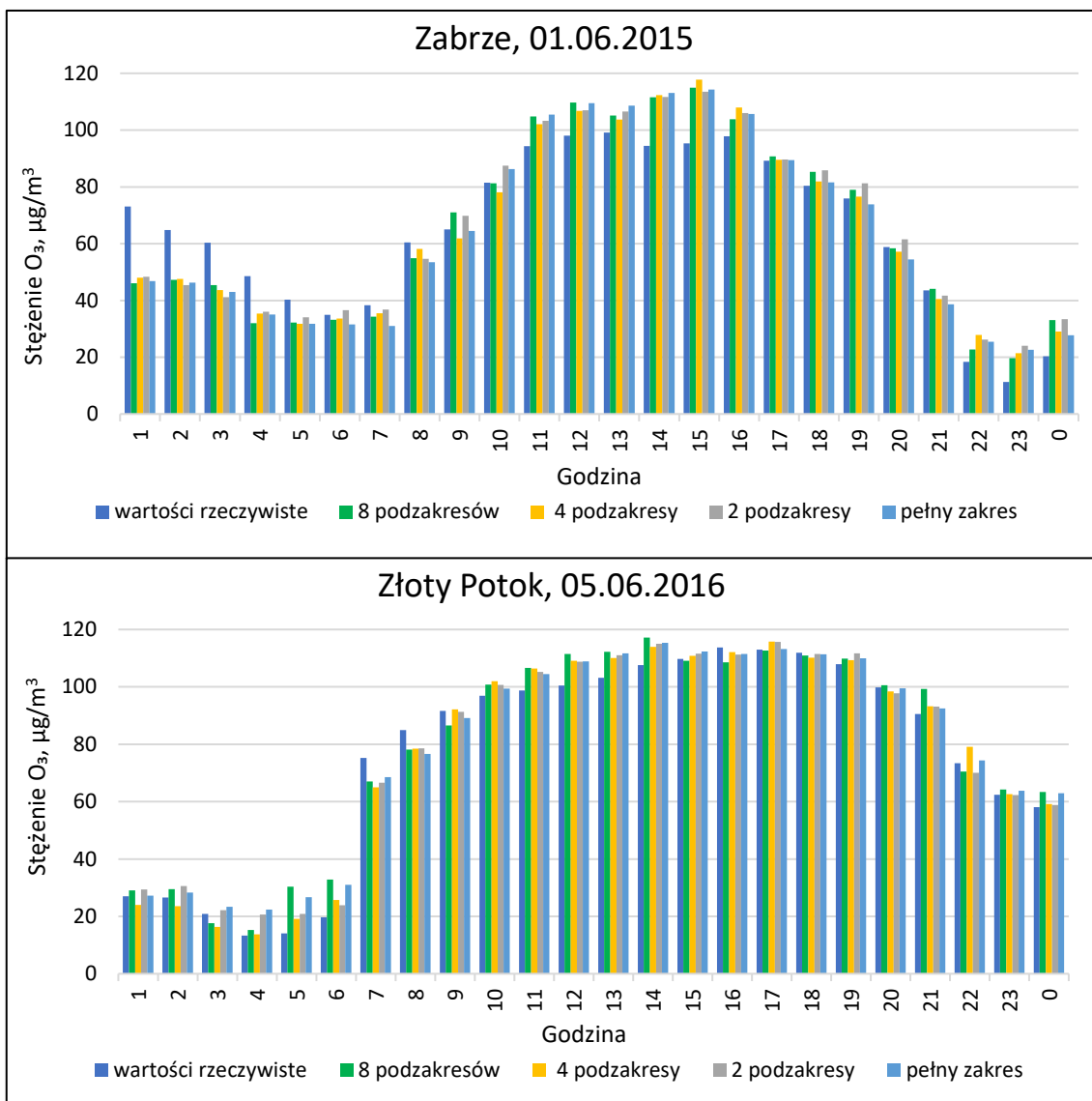
Tab. 20. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych O₃, Złoty Potok 2011-2016

Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych O ₃ , µg/m ³	Liczba przypadków	MAE, µg/m ³	Średnia wartość MAE, µg/m ³	RMSE, µg/m ³	Średnia wartość RMSE, µg/m ³
Pełny zakres	PVS-1/1-ZLP	1,6-147,5	15536	8,41	8,41	10,69	10,70
2 podzakresy	PVS-1/2-ZLP	1,6-50,3	7768	7,66	7,91	9,90	10,09
	PVS-2/2-ZLP	50,3-147,5	7768	8,16		10,28	
4 podzakresy	PVS-1/4-ZLP	1,6-34,8	3884	6,22	7,55	8,22	9,68
	PVS-2/4-ZLP	34,8-50,3	3884	8,52		10,92	
	PVS-3/4-ZLP	50,3-68,7	3884	7,81		9,97	
	PVS-4/4-ZLP	68,7-147,5	3884	7,64		9,61	
8 podzakresów	PVS-1/8-ZLP	1,6-25,3	1942	5,46	7,67	7,22	9,84
	PVS-2/8-ZLP	25,3-34,8	1942	8,26		10,55	
	PVS-3/8-ZLP	34,8-42,9	1942	8,57		10,88	
	PVS-4/8-ZLP	42,9-50,3	1942	8,69		11,26	
	PVS-5/8-ZLP	50,3-58,4	1942	7,86		10,13	
	PVS-6/8-ZLP	58,4-68,7	1942	7,57		9,61	
	PVS-7/8-ZLP	68,7-89,9	1942	7,60		9,74	
	PVS-8/8-ZLP	90,0-147,5	1942	7,36		9,31	



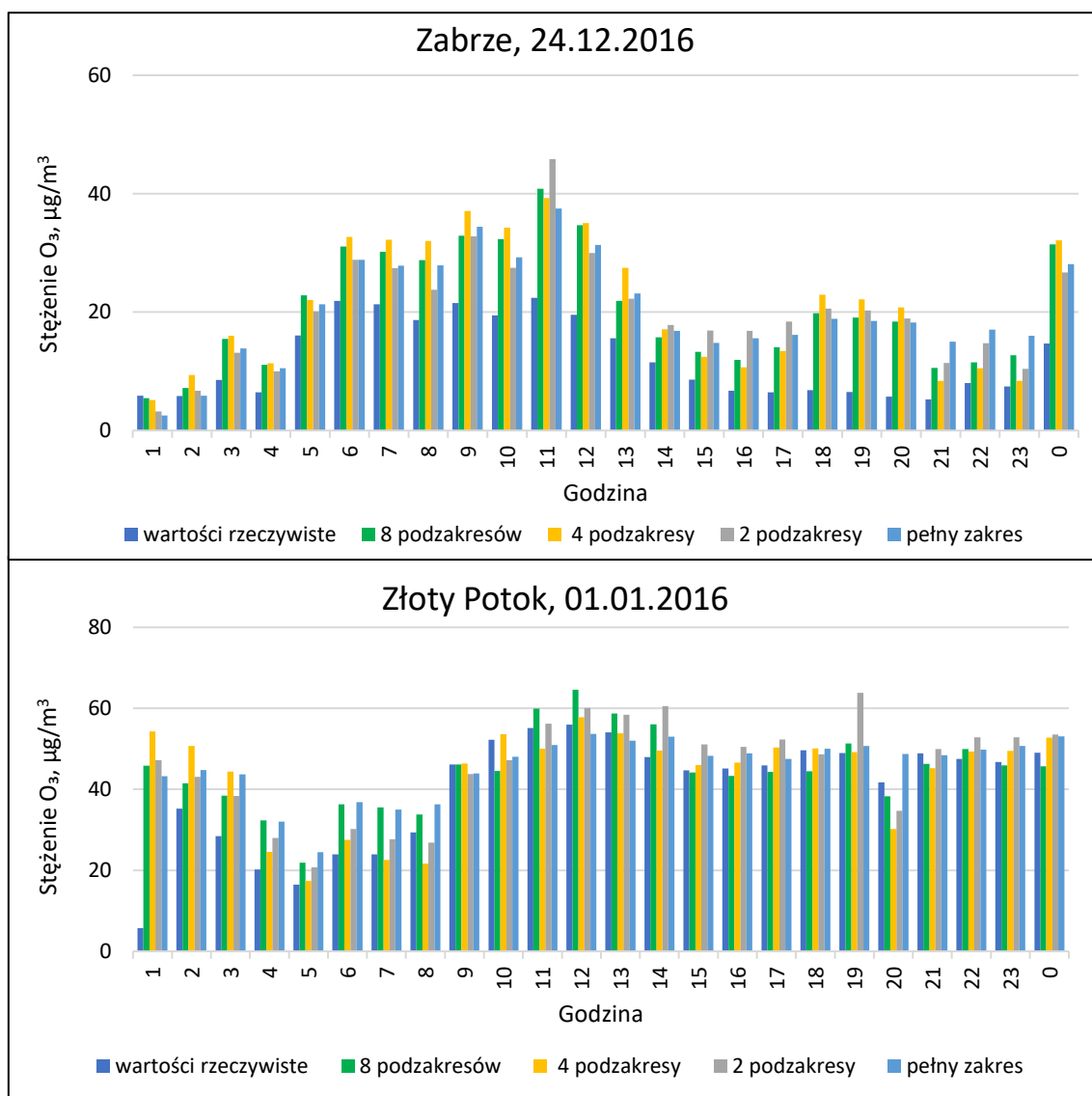
Rys. 41. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O₃, Złoty Potok 2011-2016

Na rys. 42 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele PVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W przypadku danych zarejestrowanych na stacji Zabrze w początkowych godzinach (1:00-5:00) można zaobserwować niezbyt dokładne dopasowanie modeli predykcyjnych do rzeczywistych stężeń. Z kolei w godzinach 14:00-15:00 można dostrzec niedopasowanie objawiające się zbyt wysokimi poziomami stężeń uzyskanymi przez modele predykcyjne. Na stacji Złoty Potok w czasie całego epizodu obserwowano dość dobrą jakość modelowania.



Rys. 42. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim

Na rys. 43 przedstawiono dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele PVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu jesienno-zimowego. Jakość predykcji była wyraźnie gorsza niż w epizodach wiosenno-letnich. Na stacji Zabrze stężenia aproksymowane modelami predykcyjnymi znacznie przewyższały rzeczywiste wartości. Szczególnie widoczne przewartościowanie stężenia rzeczywistego ozonu odnotowano o godzinie 1:00, na stacji Złoty Potok.



Rys. 43. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

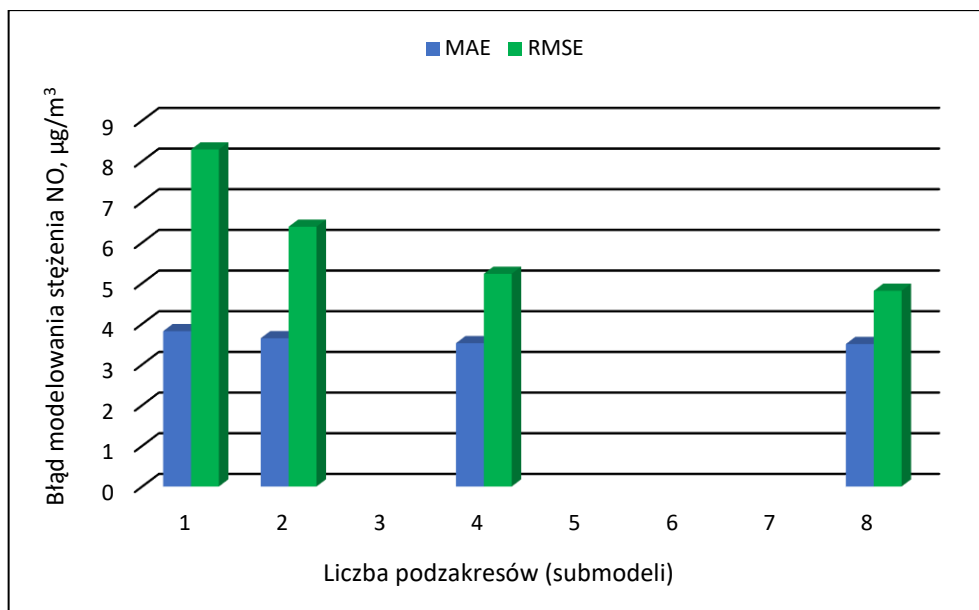
5.2.2. Wyniki modelowania stężeń NO

W tabelach 21 i 22 przedstawiono obliczone błędy modelowania stężeń NO dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń predykcyjnych NO. Wraz ze wzrostem liczby podzakresów zmniejszały się średnie miary błędów modelowania. Wyjątkiem były submodele stworzone dla podzakresów na stacji Złoty Potok, dla których odnotowano takie same średnie wartości błędu MAE ($0,43 \mu\text{g}/\text{m}^3$). W przypadku modeli 2-podzakresowych dla stacji Zabrze, odnotowano zmniejszenie średniej wartości błędu MAE względem modelu pełnozakresowego z $3,82$ do $3,65 \mu\text{g}/\text{m}^3$ (zmiana o ok. 5 %). W przypadku średniej wartości błędu RMSE zmiana

wynosiła dla stacji Zabrze ok. 23 % (spadek z 8,29 do 6,39 $\mu\text{g}/\text{m}^3$), a dla stacji Złoty Potok ok. 11 % (spadek z 0,79 do 0,70 $\mu\text{g}/\text{m}^3$). Zastosowanie modeli 4-podzakresowych skutkowało zmniejszeniem średniej wartości błędu MAE z 3,82 do 3,50 $\mu\text{g}/\text{m}^3$, a w przypadku średniej wartości błędu RMSE z 8,29 do 5,23 $\mu\text{g}/\text{m}^3$ (stacja Zabrze) i z 0,79 do 0,66 $\mu\text{g}/\text{m}^3$ (stacja Złoty Potok). W przypadku modeli 8-podzakresowych poprawa dokładności aproksymacji przejawiała się zmniejszeniem średniej wartości błędu MAE z 3,82 do 3,50 $\mu\text{g}/\text{m}^3$ i średniej wartości RMSE z 8,29 do 4,81 $\mu\text{g}/\text{m}^3$ dla stacji Zabrze, a dla stacji Złoty Potok zmniejszeniem średniej wartości błędu MAE z 0,46 do 0,43 $\mu\text{g}/\text{m}^3$, a średniej wartości RMSE z 0,79 do 0,63 $\mu\text{g}/\text{m}^3$. Na rysunkach 44 i 45 przedstawiono graficznie średnie wartości błędów modelowania dla poszczególnych podzakresów.

Tab. 21. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO, Zabrze 2011-2016

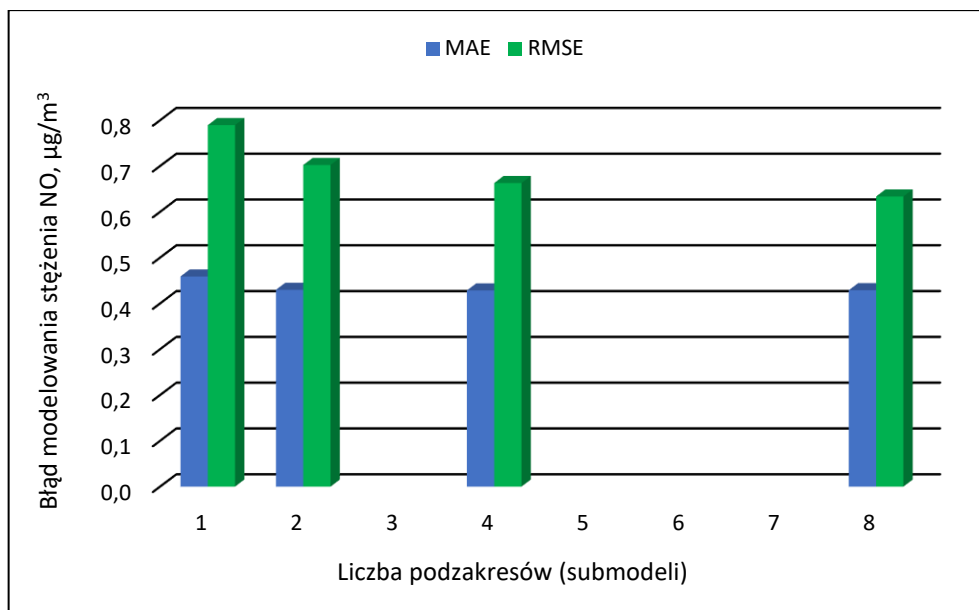
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych NO, $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	PVS-1/1-ZAB	0.0-661,2	36460	3,82	3,82	8,29	8,29
2 podzakresy	PVS-1/2-ZAB	0.5-2,8	18230	0,86	3,65	1,19	6,39
	PVS-2/2-ZAB	2,8-661,2	18230	6,44		11,59	
4 podzakresy	PVS-1/4-ZAB	0,5-2,0	9115	0,71	3,53	0,97	5,23
	PVS-2/4-ZAB	2,0-2,8	9115	1,01		1,37	
	PVS-3/4-ZAB	2,8-8,3	9115	1,89		2,63	
	PVS-4/4-ZAB	8,3-661,2	9115	10,5		15,95	
8 podzakresów	PVS-1/8-ZAB	0,5-1,7	4558	0,66	3,50	0,90	4,81
	PVS-2/8-ZAB	1,7-2,0	4557	0,77		1,05	
	PVS-3/8-ZAB	2,0-2,3	4558	0,90		1,22	
	PVS-4/8-ZAB	2,3-2,8	4557	1,18		1,58	
	PVS-5/8-ZAB	2,8-4,0	4558	1,48		2,01	
	PVS-6/8-ZAB	4,0-8,3	4557	2,30		3,14	
	PVS-7/8-ZAB	8,3-21,2	4558	5,15		7,22	
	PVS-8/8-ZAB	21,2-661,2	4557	15,59		21,38	



Rys. 44. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO, Zabrze 2011-2016

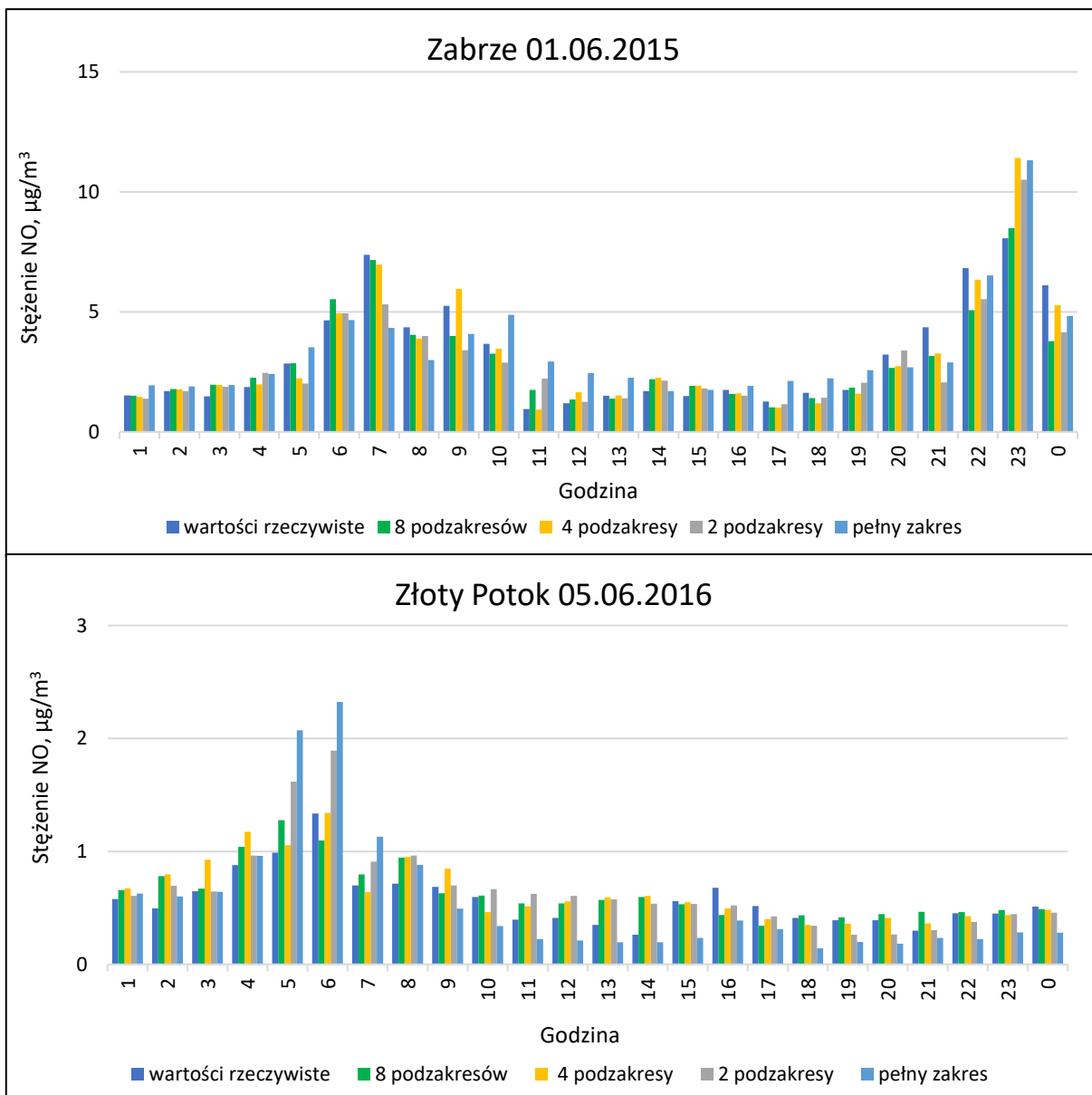
Tab. 22. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO, Złoty Potok 2011-2016

Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych NO, µg/m³	Liczba przypadków	MAE, µg/m³	Średnia wartość MAE, µg/m³	RMSE, µg/m³	Średnia wartość RMSE, µg/m³
Pełny zakres	PVS-1/1-ZLP	0,1-49,8	15536	0,46	0,46	0,79	0,79
2 podzakresy	PVS-1/2-ZLP	0,1-0,6	7768	0,32	0,43	0,43	0,70
	PVS-2/2-ZLP	0,6-49,8	7768	0,54		0,98	
4 podzakresy	PVS-1/4-ZLP	0,1-0,4	3884	0,30	0,43	0,37	0,66
	PVS-2/4-ZLP	0,4-0,6	3884	0,32		0,43	
	PVS-3/4-ZLP	0,6-1,1	3884	0,33		0,50	
	PVS-4/4-ZLP	1,1-49,8	3884	0,76		1,34	
8 podzakresów	PVS-1/8-ZLP	0,1-0,3	1942	0,26	0,43	0,33	0,63
	PVS-2/8-ZLP	0,3-0,4	1942	0,30		0,39	
	PVS-3/8-ZLP	0,4-0,5	1942	0,31		0,40	
	PVS-4/8-ZLP	0,5-0,6	1942	0,36		0,50	
	PVS-5/8-ZLP	0,6-0,8	1942	0,35		0,50	
	PVS-6/8-ZLP	0,8-1,1	1942	0,32		0,52	
	PVS-7/8-ZLP	1,1-1,8	1942	0,39		0,59	
	PVS-8/8-ZLP	1,8-49,8	1942	1,15		1,86	

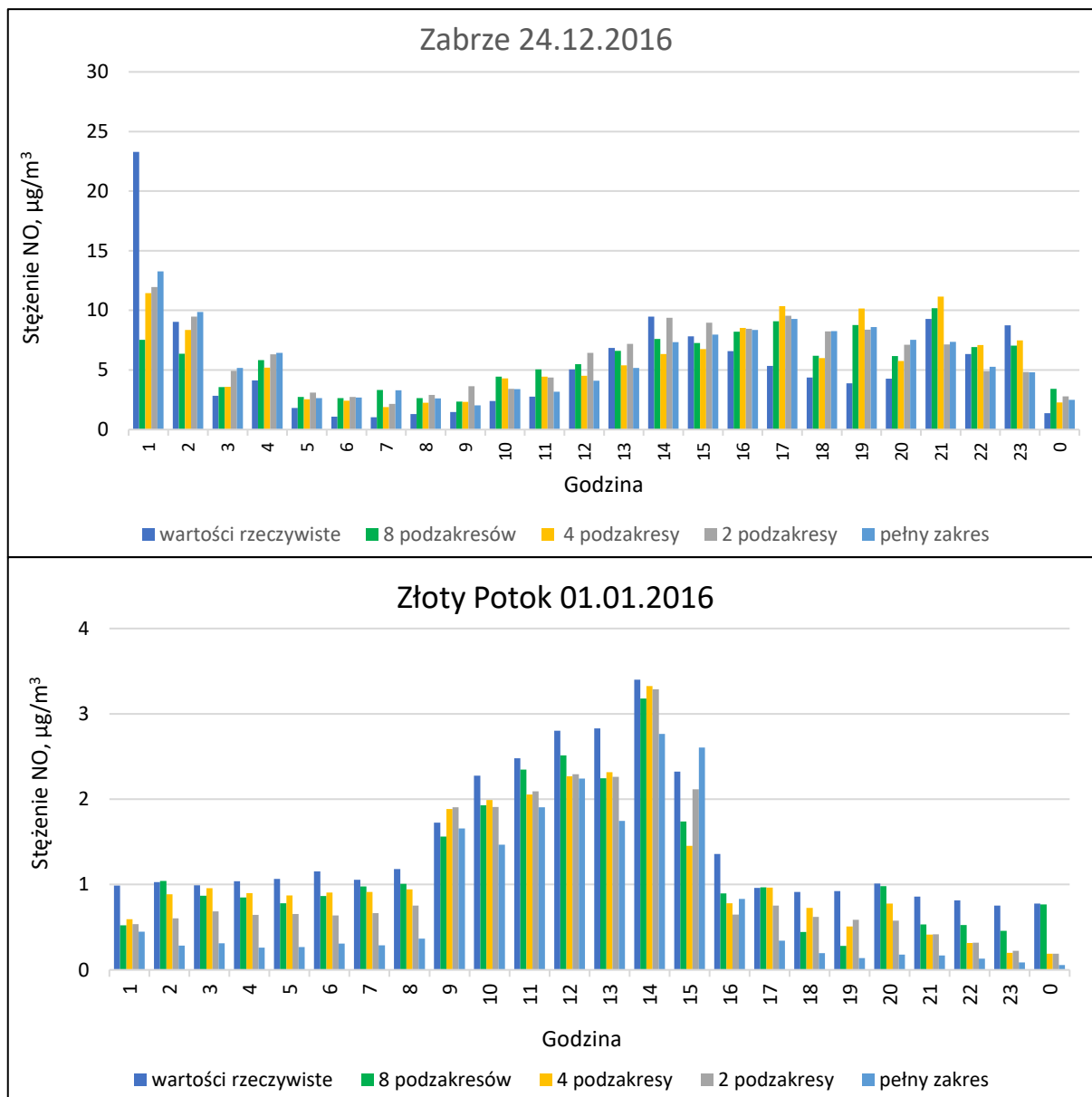


Rys. 45. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO, Złoty Potok 2011-2016

Na rys. 46 i 47 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele PVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. Dla wybranych epizodów występują niedopasowania do stężeń rzeczywistych, wyraźne zwłaszcza w przypadku modeli 2-podzakresowych i pełnozakresowego dla stacji Złoty Potok w godzinach 5:00 i 6:00 w dniu 5.06.2016. Szczególnie wyraźne niedopasowanie stężeń aproksymowanych do stężeń rzeczywistych można zaobserwować na stacji Zabrze o godzinie 1:00 w dniu 24.12.2016.



Rys. 46. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



Rys. 47. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

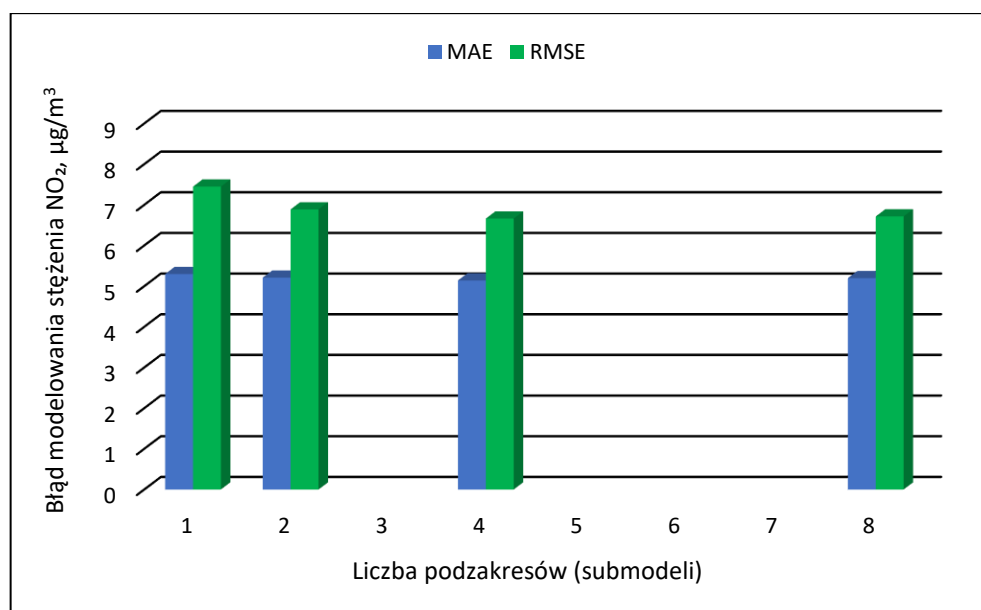
5.2.3. Wyniki modelowania stężeń NO₂

W tabelach 23 i 24 przedstawiono obliczone błędy modelowania stężeń NO₂ dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń predykcyjnych NO₂. Zastosowanie 8-podzakresowych modeli PVS na obu stacjach skutkowało uzyskaniem wyższych średnich wartości błędów MAE i RMSE niż zastosowanie modeli 4-podzakresowych (5,20 względem 5,14 $\mu\text{g}/\text{m}^3$ w przypadku stacji Zabrze i 1,81 względem 1,69 $\mu\text{g}/\text{m}^3$ w przypadku stacji Złoty Potok oraz 6,72 względem

6,66 $\mu\text{g}/\text{m}^3$ w przypadku stacji Zabrze i 2,34 względem 2,21 $\mu\text{g}/\text{m}^3$ w przypadku stacji Złoty Potok). Na rysunkach 48 i 49 przedstawiono graficznie średnie wartości błędów modelowania dla poszczególnych podzakresów.

Tab. 23. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO_2 , Zabrze 2011-2016

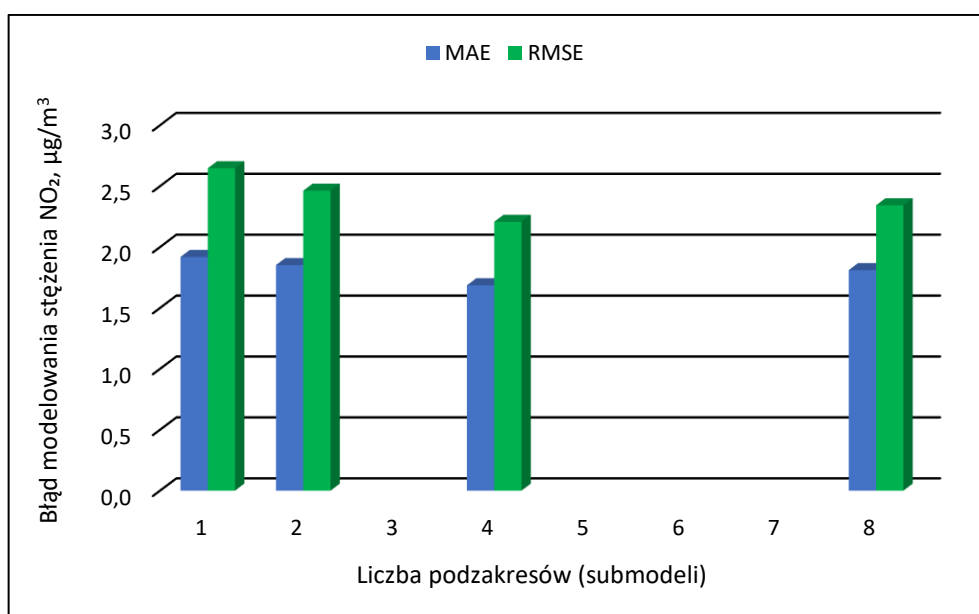
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych NO_2 , $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	PVS-1/1-ZAB	3,6-134,4	36460	5,30	5,30	7,45	7,4
2 podzakresy	PVS-1/2-ZAB	3,6-20,6	18230	3,13	5,21	4,18	6,89
	PVS-2/2-ZAB	20,6-134,4	18230	7,29		9,59	
4 podzakresy	PVS-1/4-ZAB	3,6-12,2	9115	2,29	5,14	3,03	6,66
	PVS-2/4-ZAB	12,3-20,6	9115	3,85		4,97	
	PVS-3/4-ZAB	20,6-32,8	9115	5,89		7,58	
	PVS-4/4-ZAB	32,8-134,4	9115	8,54		11,08	
8 podzakresów	PVS-1/8-ZAB	3,6-9,2	4558	1,80	5,20	2,35	6,72
	PVS-2/8-ZAB	9,2-12,2	4557	2,88		3,71	
	PVS-3/8-ZAB	12,3-16	4558	3,43		4,40	
	PVS-4/8-ZAB	16-20,6	4557	4,33		5,56	
	PVS-5/8-ZAB	20,6-26,2	4558	5,71		7,33	
	PVS-6/8-ZAB	26,2-32,8	4557	6,33		8,21	
	PVS-7/8-ZAB	32,8-41,7	4558	7,65		9,86	
	PVS-8/8-ZAB	41,7-134,4	4557	9,47		12,31	



Rys. 48. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO_2 , Zabrze 2011-2016

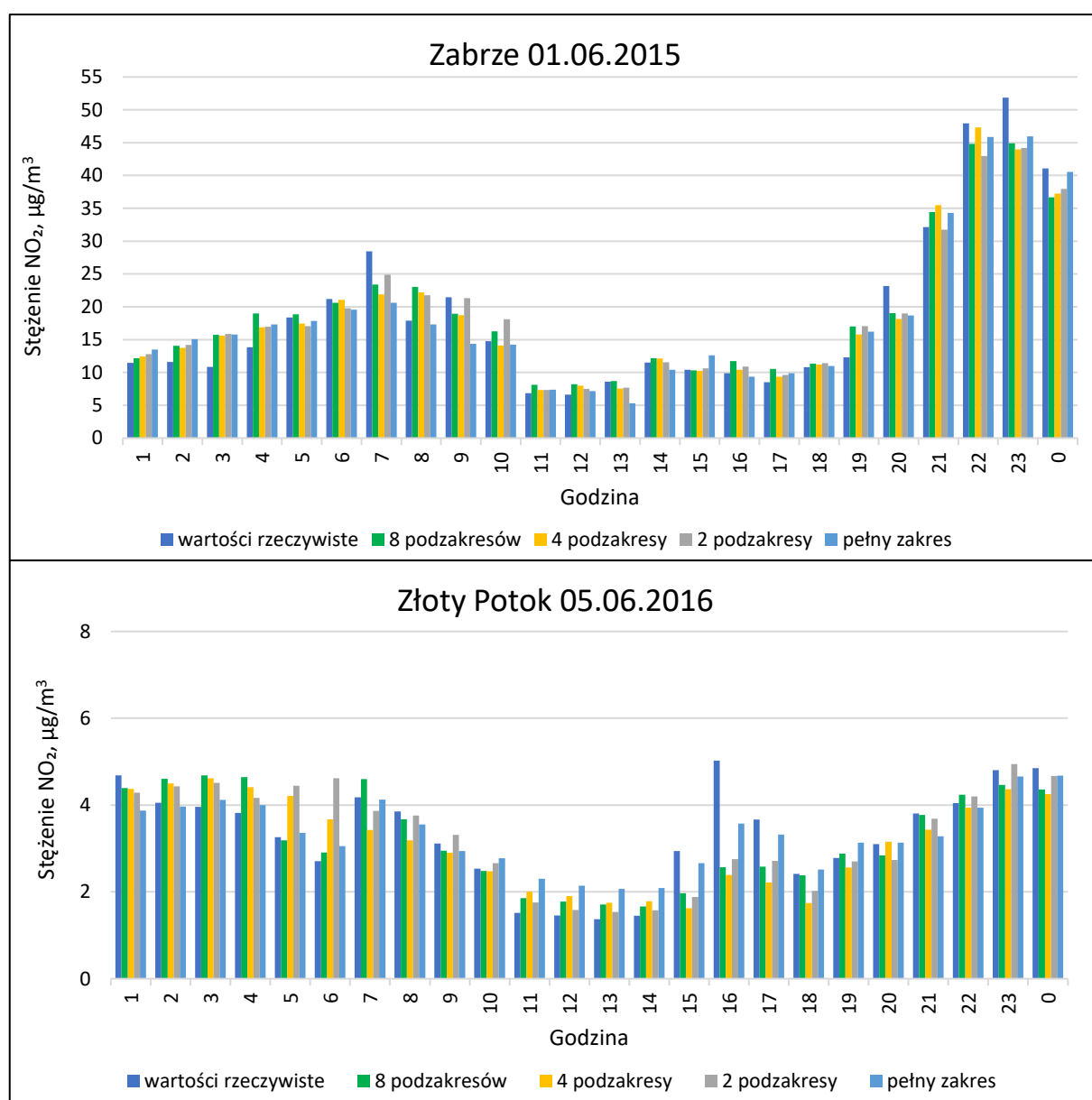
Tab. 24. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO₂, Złoty Potok 2011-2016

Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych NO ₂ , µg/m ³	Liczba przypadków	MAE, µg/m ³	Średnia wartość MAE, µg/m ³	RMSE, µg/m ³	Średnia wartość RMSE, µg/m ³
Pełny zakres	PVS-1/1-ZLP	1,6-52,9	15536	1,92	1,92	2,65	2,65
2 podzakresy	PVS-1/2-ZLP	1,6-6,7	7768	1,27	1,85	1,69	2,46
	PVS-2/2-ZLP	6,7-52,9	7768	2,44		3,24	
4 podzakresy	PVS-1/4-ZLP	1,6-4,7	3884	0,99	1,69	1,33	2,21
	PVS-2/4-ZLP	4,7-6,7	3884	0,99		1,33	
	PVS-3/4-ZLP	6,7-10,8	3884	1,85		2,39	
	PVS-4/4-ZLP	10,8-52,9	3884	2,91		3,78	
8 podzakresów	PVS-1/8-ZLP	1,6-3,7	1942	0,81	1,81	1,08	2,34
	PVS-2/8-ZLP	3,7-4,7	1942	1,27		1,66	
	PVS-3/8-ZLP	4,7-5,6	1942	1,37		1,73	
	PVS-4/8-ZLP	5,6-6,7	1942	1,67		2,17	
	PVS-5/8-ZLP	6,7-8,2	1942	1,79		2,26	
	PVS-6/8-ZLP	8,2-10,8	1942	1,97		2,55	
	PVS-7/8-ZLP	10,8-16,0	1942	2,36		3,06	
	PVS-8/8-ZLP	16,1-52,9	1942	3,25		4,22	

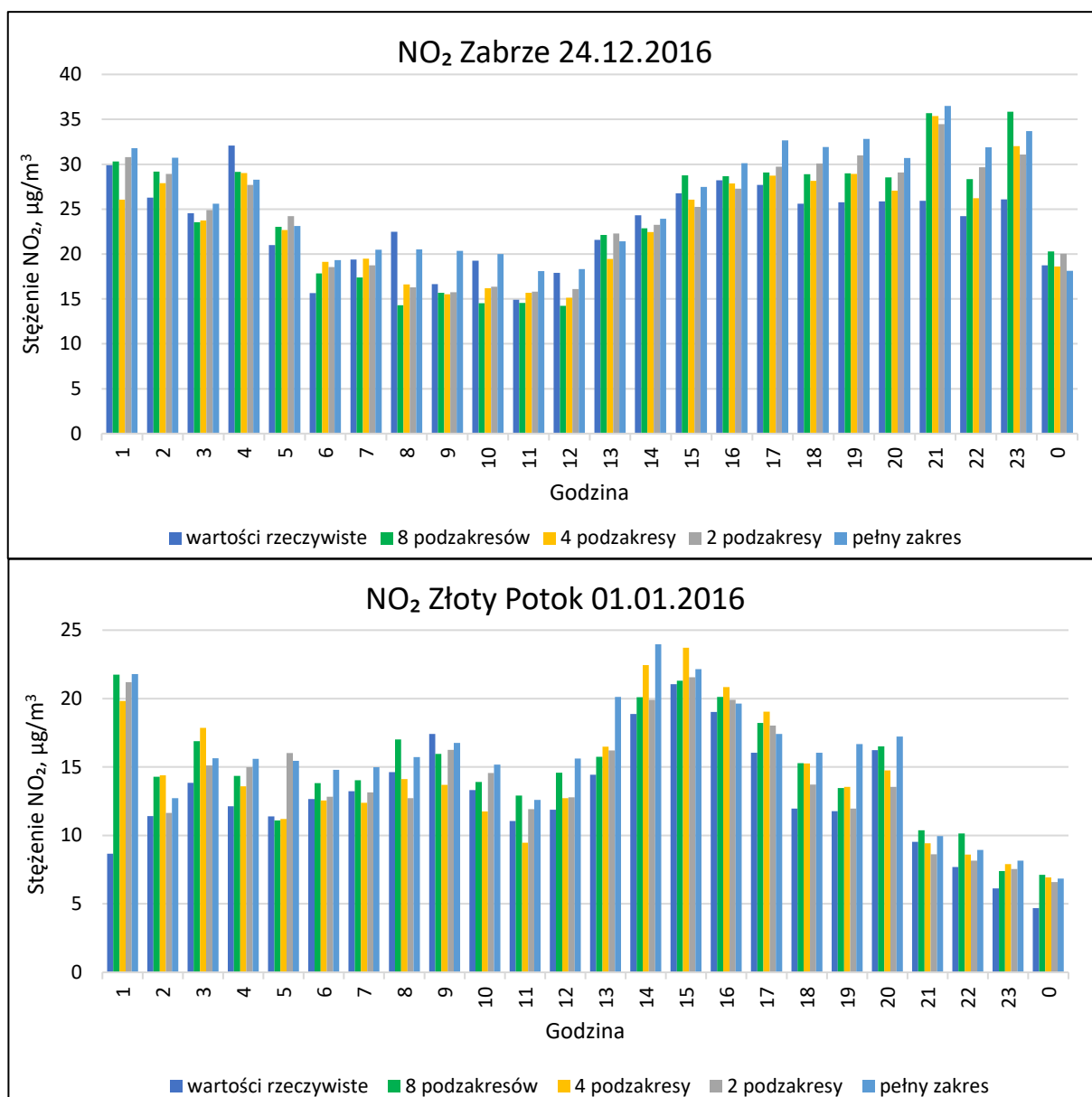


Rys. 49. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO₂, Złoty Potok 2011-2016

Na rys. 50 i 51 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele PVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. Na rys. 50 można zaobserwować niedopasowanie stężeń predykcyjnych w godzinach, w których pojawiają się gwałtowne wzrosty stężeń rzeczywistych. Na rys. 51 podobne niedopasowanie pojawia się na stacji Złoty Potok o godzinie 1:00 – stężenia predykcyjne osiągają tam znacznie wyższe poziomy w porównaniu do stężeń rzeczywistych.



Rys. 50. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



Rys. 51. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

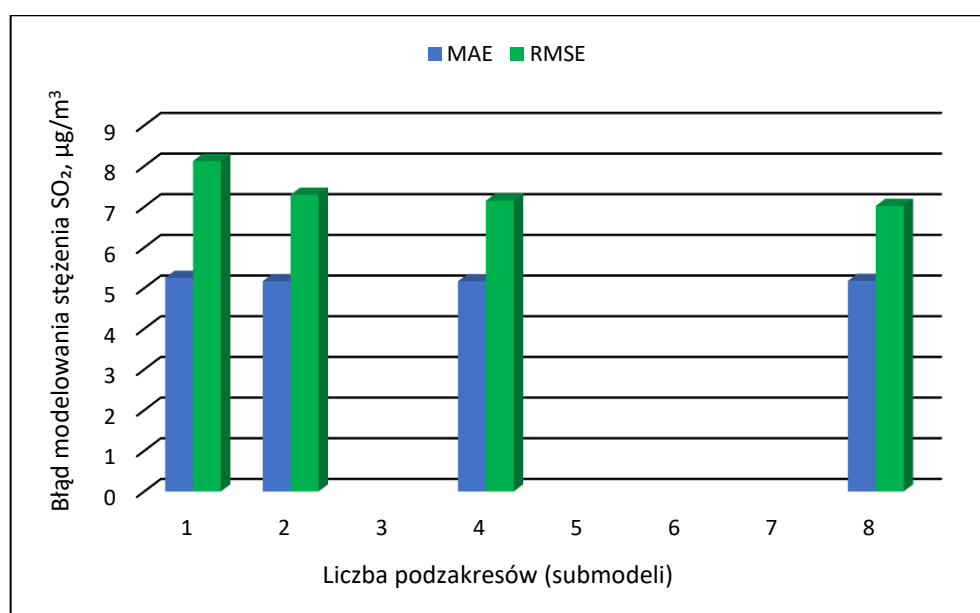
5.2.4. Wyniki modelowania stężeń SO₂

W tabelach 25 i 26 przedstawiono obliczone błędy modelowania stężeń NO₂ dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń predykcyjnych SO₂. Średnie wartości błędu MAE obliczone dla modeli PVS dla stacji Zabrze mały po podziale na 2 podzakresy, przy 4 podzakresach uzyskano taką samą wartość jak dla 2 podzakresów (5,17 µg/m³), przy 8 podzakresach zaobserwowano wzrost średniej wartości błędu MAE (do 5,18 µg/m³). Średnie wartości błędu RMSE mały

wraz z przyrostem liczby podzakresów. W przypadku modeli PVS dla stacji Złoty Potok wraz z przyrostem liczby podzakresów malały średnie wartości błędów MAE i RMSE. Na rysunkach 52 i 53 przedstawiono graficznie średnie wartości błędów modelowania dla poszczególnych podzakresów.

Tab. 25. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych SO₂, Zabrze 2011-2016

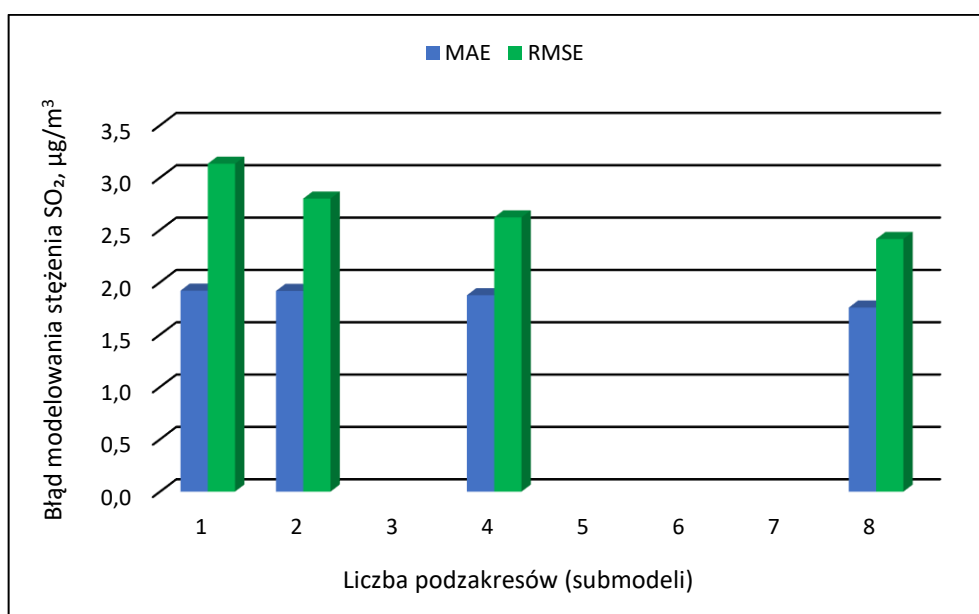
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych SO ₂ , µg/m ³	Liczba przypadków	MAE, µg/m ³	Średnia wartość MAE, µg/m ³	RMSE, µg/m ³	Średnia wartość RMSE, µg/m ³
Pełny zakres	PVS-1/1-ZAB	1,3-321,9	36460	5,25	5,25	8,13	8,1
2 podzakresy	PVS-1/2-ZAB	1,3-10,6	18230	2,57	5,17	3,69	7,30
	PVS-2/2-ZAB	10,6-321,9	18230	7,78		10,90	
4 podzakresy	PVS-1/4-ZAB	1,3-6,4	9115	1,81	5,17	2,58	7,15
	PVS-2/4-ZAB	6,4-10,6	9115	3,30		4,58	
	PVS-3/4-ZAB	10,6-21,3	9115	5,88		8,14	
	PVS-4/4-ZAB	21,3-321,9	9115	9,69		13,29	
8 podzakresów	PVS-1/8-ZAB	1,3-4,9	4558	1,49	5,18	2,17	7,02
	PVS-2/8-ZAB	4,9-6,4	4557	2,19		3,02	
	PVS-3/8-ZAB	6,4-8,2	4558	2,85		3,89	
	PVS-4/8-ZAB	8,2-10,6	4557	3,85		5,29	
	PVS-5/8-ZAB	10,6-14,4	4558	5,46		7,52	
	PVS-6/8-ZAB	14,4-21,3	4557	6,69		9,12	
	PVS-7/8-ZAB	21,3-34,0	4558	7,08		9,25	
	PVS-8/8-ZAB	34,0-321,9	4557	11,82		15,94	



Rys. 52. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO₂, Złoty Potok 2011-2016

Tab. 26. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych SO₂, Złoty Potok 2011-2016

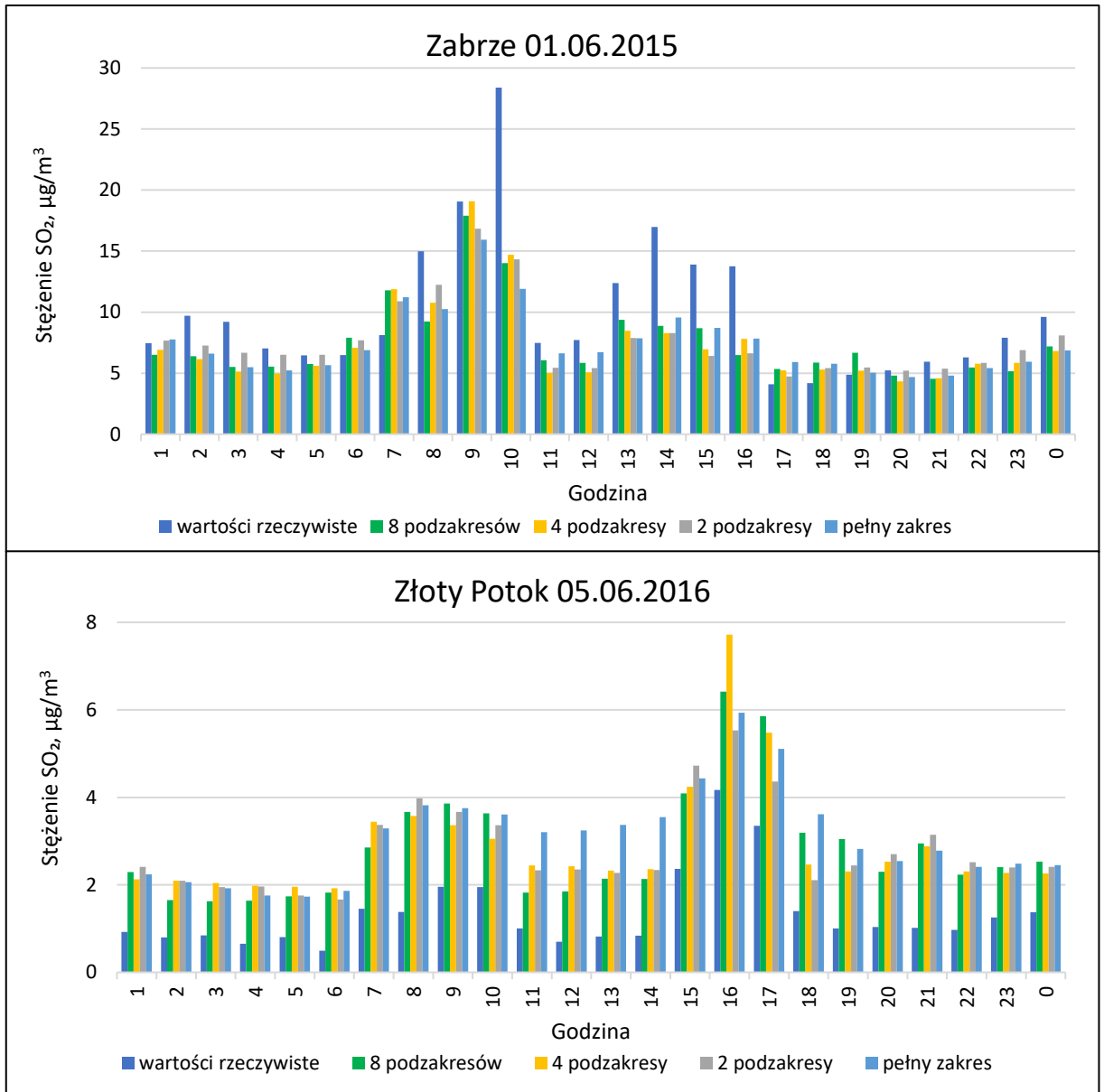
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych SO ₂ , µg/m ³	Liczba przypadków	MAE, µg/m ³	Średnia wartość MAE, µg/m ³	RMSE, µg/m ³	Średnia wartość RMSE, µg/m ³
Pełny zakres	PVS-1/1-ZLP	0,8-74,2	15536	1,92	1,92	3,13	3,13
2 podzakresy	PVS-1/2-ZLP	0,8-4,0	7768	0,94	1,92	1,30	2,80
	PVS-2/2-ZLP	4,0-74,2	7768	2,90		4,30	
4 podzakresy	PVS-1/4-ZLP	0,8-2,8	3884	0,82	1,88	1,10	2,62
	PVS-2/4-ZLP	2,8-4,0	3884	1,14		1,60	
	PVS-3/4-ZLP	4,0-7,2	3884	1,71		2,40	
	PVS-4/4-ZLP	7,2-74,2	3884	3,84		5,39	
8 podzakresów	PVS-1/8-ZLP	0,8-2,2	1942	0,70	1,76	0,90	2,41
	PVS-2/8-ZLP	2,2-2,8	1942	0,85		1,13	
	PVS-3/8-ZLP	2,8-3,3	1942	1,01		1,38	
	PVS-4/8-ZLP	3,3-4,0	1942	1,25		1,78	
	PVS-5/8-ZLP	4,0-5,1	1942	0,06		0,08	
	PVS-6/8-ZLP	5,1-7,2	1942	2,02		2,81	
	PVS-7/8-ZLP	7,2-11,8	1942	3,11		4,55	
	PVS-8/8-ZLP	11,8-74,2	1942	5,07		6,68	



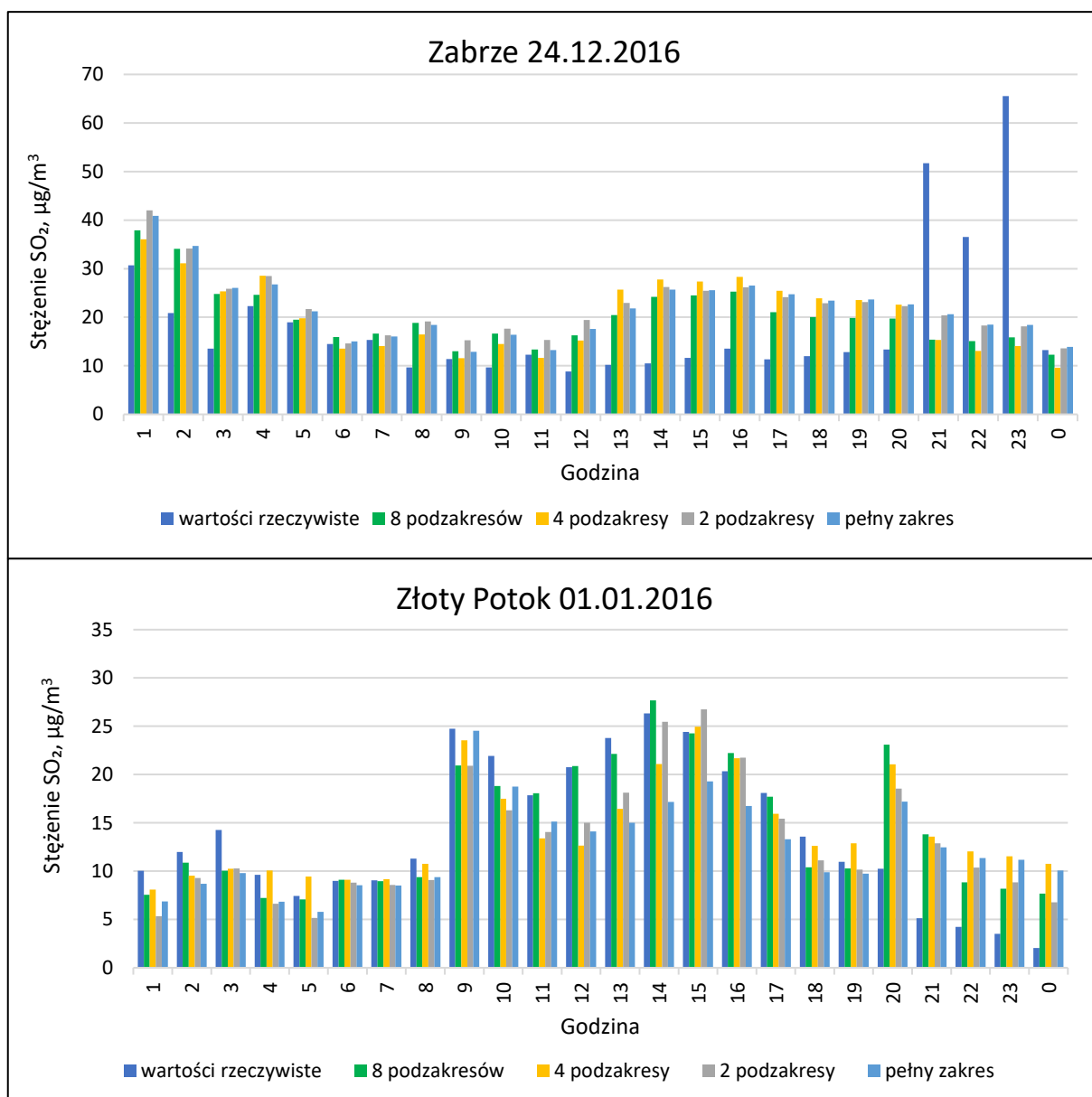
Rys. 53. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO₂, Złoty Potok 2011-2016

Na rys. 54 i 55 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele PVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W dniach reprezentujących oba okresy na stacji Zabrze zarejestrowano nadzwyczajnie wysokie stężenia rzeczywiste, do których modele predykcyjne nie zdołały dopasować podobnych wartości aproksymowanych. Na stacji Złoty Potok w dniu 5.06.2016 modele predykcyjne miały tendencję do zawyżania

poziomów stężeń, nieco lepiej było 1.01.2016, kiedy to w godzinach 11:00-17:00 modele 8-podzakresowe stosunkowo dobrze dopasowały wartości stężeń aproksymowanych do rzeczywistych.



Rys. 54. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



Rys. 55. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

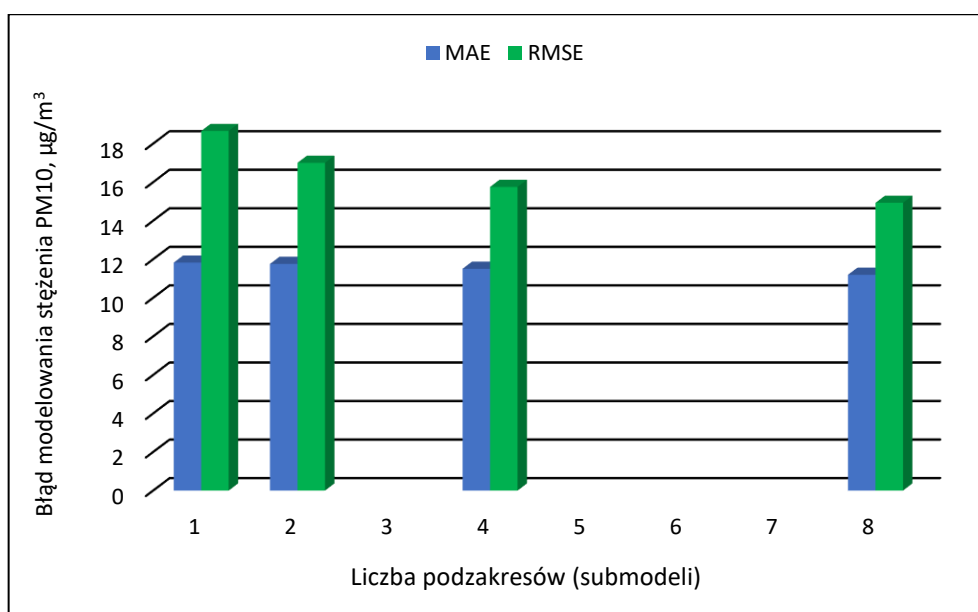
5.2.5. Wyniki modelowania stężeń PM10

W tabelach 27 i 28 przedstawiono obliczone błędy modelowania stężeń PM10 dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń predykcyjnych PM10. Na stacji Zabrze w modelach predykcyjnych uzyskiwano coraz niższe średnie wartości błędów MAE i RMSE wraz z podziałem na coraz węższe podzakresy. W przypadku średniej wartości błędu MAE poprawa względem modelu pełnozakresowego wyniosła 5 % (spadek z 11,83 do 11,20 $\mu\text{g}/\text{m}^3$), a w przypadku średniej wartości błędu RMSE 20 % (spadek z 18,66 do 14,93 $\mu\text{g}/\text{m}^3$). W przypadku stacji

Złoty Potok odnotowano wzrost średnich wartości błędów aproksymacji w modelach 8-podzakresowych względem modeli 4-podzakresowych (wzrost z 6,44 do 6,64 $\mu\text{g}/\text{m}^3$ w MAE i z 8,28 do 8,52 $\mu\text{g}/\text{m}^3$ w przypadku RMSE). Na rysunkach 56 i 57 przedstawiono graficznie średnie wartości błędów modelowania dla poszczególnych podzakresów.

Tab. 27. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych PM10, Zabrze 2011-2016

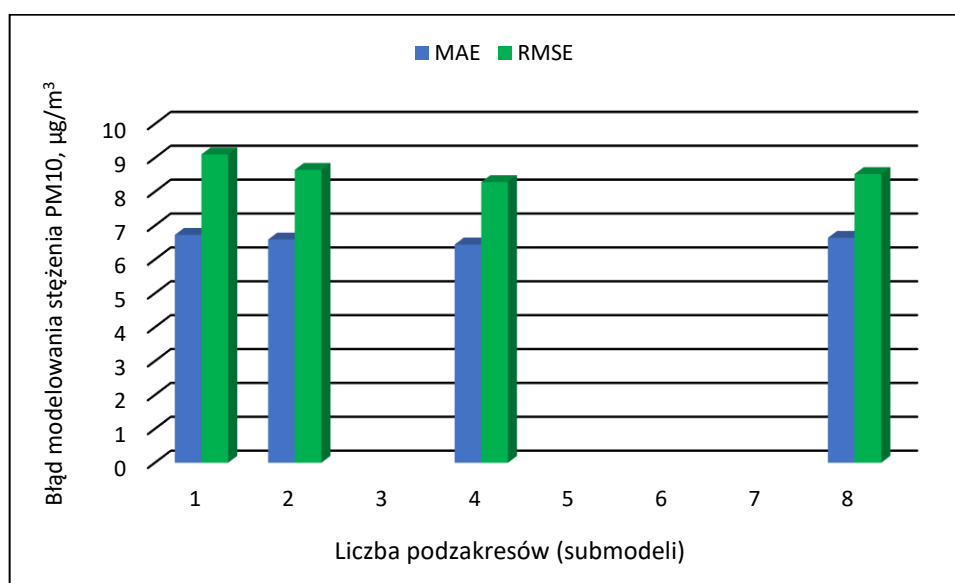
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych PM10, $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	PVS-1/1-ZAB	8,2-980,5	36460	11,83	11,83	18,66	18,66
2 podzakresy	PVS-1/2-ZAB	8,2-33,1	18230	6,53	11,76	8,79	17,00
	PVS-2/2-ZAB	33,1-980,5	18230	16,99		25,22	
4 podzakresy	PVS-1/4-ZAB	8,2-22,1	9115	5,52	11,51	7,56	15,76
	PVS-2/4-ZAB	22,1-33,1	9115	7,42		9,73	
	PVS-3/4-ZAB	33,1-55,6	9115	10,01		13,18	
	PVS-4/4-ZAB	55,6-980,5	9115	23,07		32,56	
8 podzakresów	PVS-1/8-ZAB	8,2-17,9	4558	4,47	11,20	5,88	14,93
	PVS-2/8-ZAB	17,9-22,1	4557	6,39		8,80	
	PVS-3/8-ZAB	22,1-26,9	4558	6,72		8,91	
	PVS-4/8-ZAB	26,9-33,1	4557	8,24		10,66	
	PVS-5/8-ZAB	33,1-41,2	4558	9,18		12,14	
	PVS-6/8-ZAB	41,2-55,6	4557	11,11		14,39	
	PVS-7/8-ZAB	55,6-90,1	4558	14,67		18,98	
	PVS-8/8-ZAB	90,2-980,5	4557	28,79		39,70	



Rys. 56. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń PM10, Zabrze 2011-2016

Tab. 28. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych PM10, Złoty Potok 2011-2016

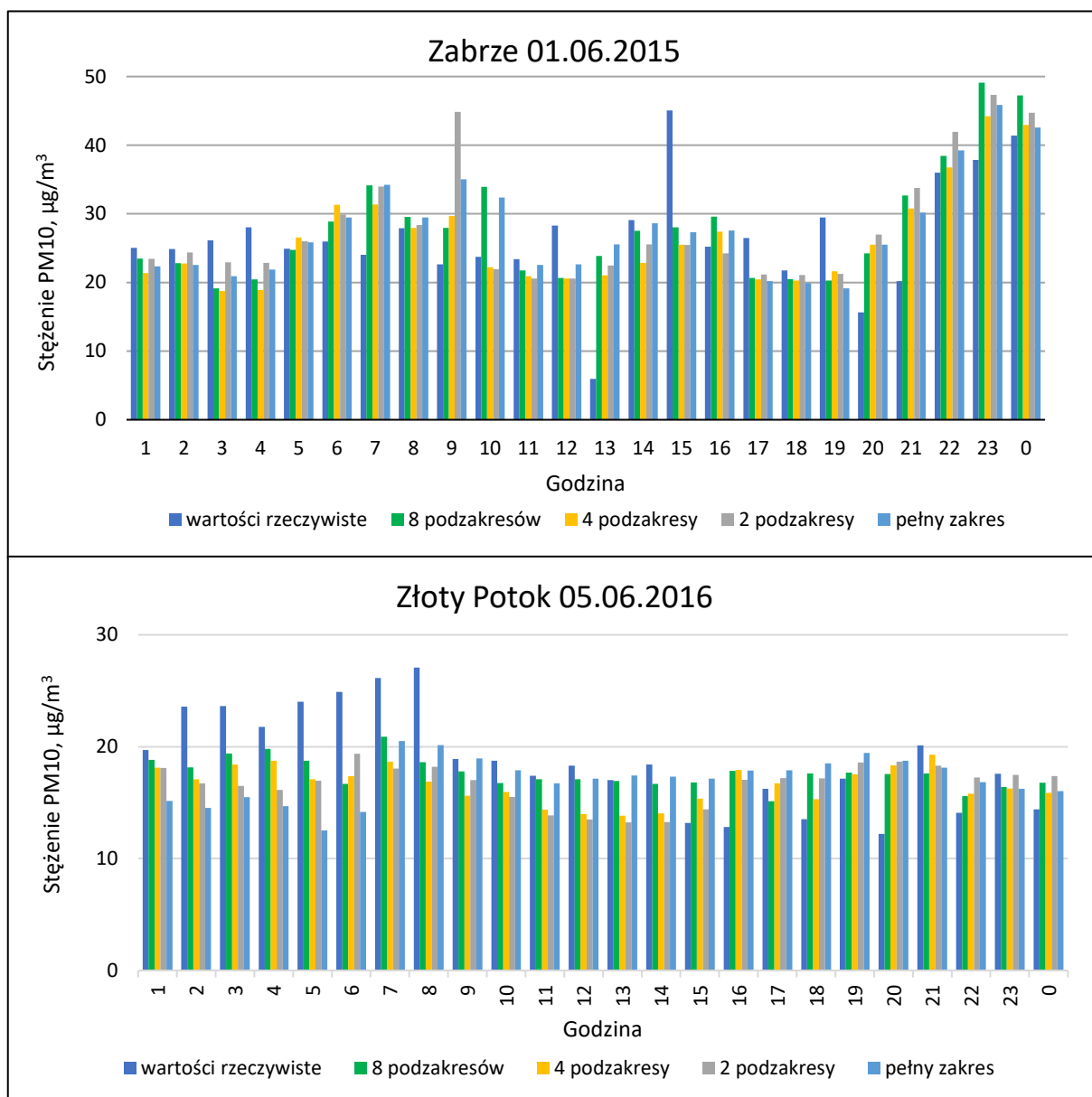
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych PM10, $\mu\text{g}/\text{m}^3$	Liczba przypadków	MAE, $\mu\text{g}/\text{m}^3$	Średnia wartość MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	Średnia wartość RMSE, $\mu\text{g}/\text{m}^3$
Pełny zakres	PVS-1/1-ZLP	5,1-120,4	15536	6,73	6,73	9,10	9,10
2 podzakresy	PVS-1/2-ZLP	5,1-22,7	7768	4,64	6,59	6,11	8,65
	PVS-2/2-ZLP	22,7-120,4	7768	8,54		11,18	
4 podzakresy	PVS-1/4-ZLP	5,1-18,0	3884	3,96	6,44	5,12	8,28
	PVS-2/4-ZLP	18,0-22,7	3884	5,23		6,76	
	PVS-3/4-ZLP	22,7-30,9	3884	6,60		8,57	
	PVS-4/4-ZLP	30,9-120,4	3884	9,97		12,70	
8 podzakresów	PVS-1/8-ZLP	5,1-15,7	1942	3,41	6,64	4,39	8,52
	PVS-2/8-ZLP	15,7-18,0	1942	4,63		6,01	
	PVS-3/8-ZLP	18,0-20,2	1942	4,96		6,43	
	PVS-4/8-ZLP	20,2-22,7	1942	5,61		7,20	
	PVS-5/8-ZLP	22,7-25,9	1942	6,75		8,74	
	PVS-6/8-ZLP	25,9-30,9	1942	7,20		9,28	
	PVS-7/8-ZLP	30,9-40,5	1942	8,03		10,23	
	PVS-8/8-ZLP	40,5-120,4	1942	12,52		15,88	



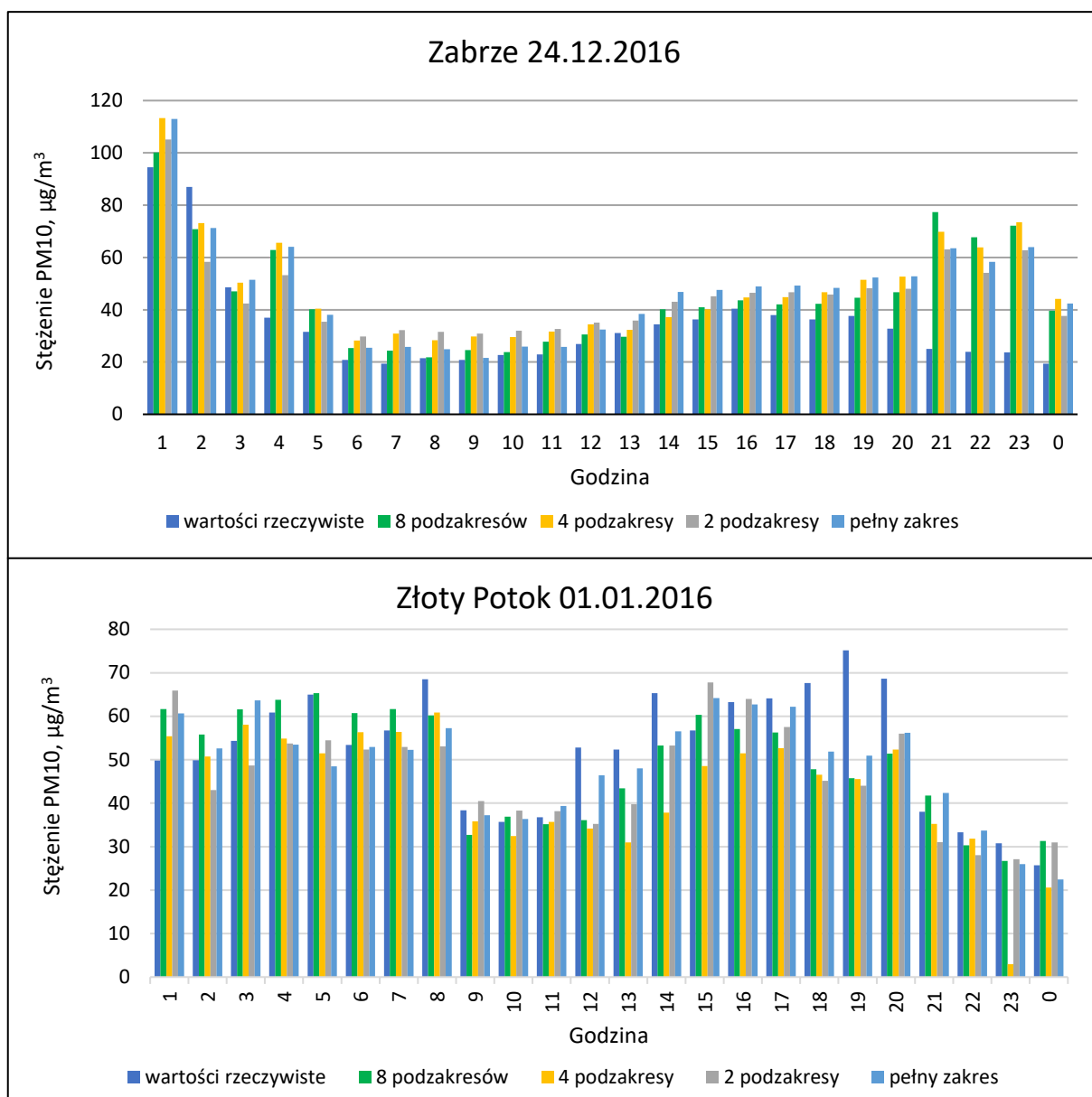
Rys. 57. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń PM10, Złoty Potok 2011-2016

Na rys. 58 i 59 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele PVS w zestawieniu ze stężeniami rzeczywistymi w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W czasie omawianych epizodów odnotowano duże zmiany stężeń rzeczywistych, a także niedokładne dopasowanie wartości aproksymowanych. Na stacji Zabrze w dniu 24.12.2016 przez większość dnia można zaobserwować dobre dopasowanie wartości aproksymowanych modelami 8-

podzakresowymi, jednak w godzinach wieczornych (21:00-0:00) modele predykcyjne miały tendencję do zawyżania poziomów stężeń PM10.



Rys. 58. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim



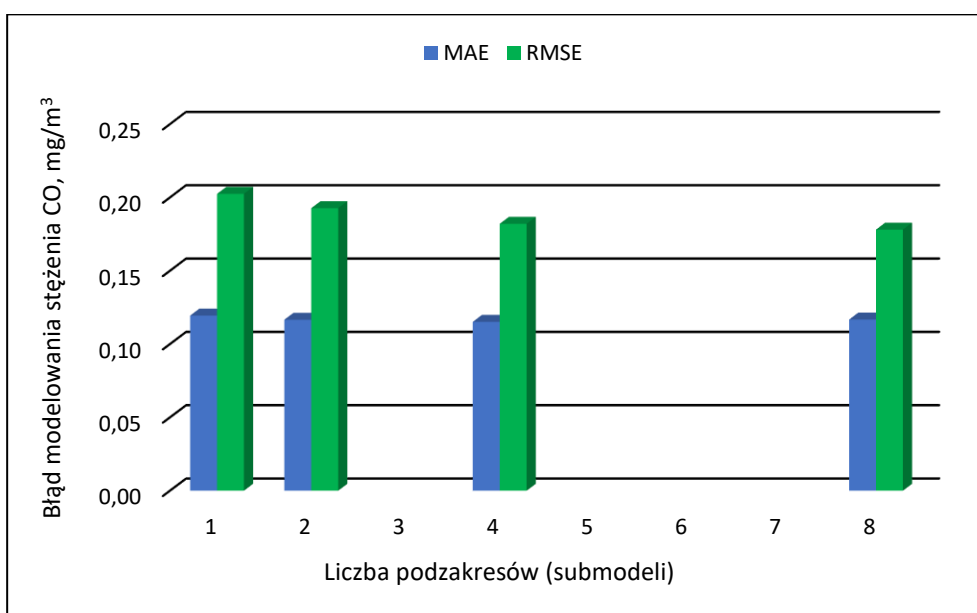
Rys. 59. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym

5.2.6. Wyniki modelowania stężeń CO

W tabeli 29 przedstawiono obliczone błędy modelowania stężeń CO dla podzakresów uzyskanych w wyniku podziału przypadków względem stężeń predykcyjnych CO. Przedstawione wyniki dotyczą wyłącznie stacji Zabrze, na stacji Złoty Potok stężenia CO nie były rejestrowane. Zastosowanie modeli 8-podzakresowych skutkowało wzrostem średniej wartości błędu MAE w stosunku do modeli 4-podzakresowych – z 0,115 do 0,117 mg/m³. Średnie wartości błędu RMSE malały wraz ze zwiększeniem liczby podzakresów. Zobrazowano to na rysunku 60.

Tab. 29. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych CO, Zabrze 2011-2016

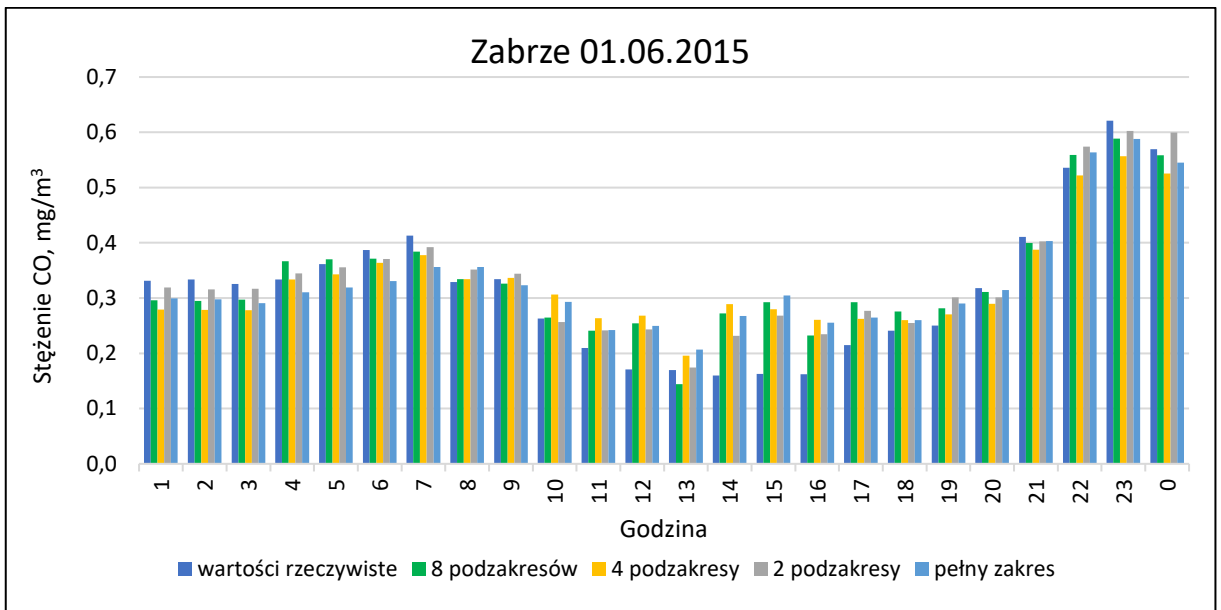
Liczba podzakresów	Nazwa modelu	Zakres stężeń predykcyjnych CO, mg/m ³	Liczba przypadków	MAE, mg/m ³	Średnia wartość MAE, mg/m ³	RMSE, mg/m ³	Średnia wartość RMSE, mg/m ³
Pełny zakres	PVS-1/1-ZAB	0,16-8,28	36460	0,119	0,119	0,202	0,202
2 podzakresy	PVS-1/2-ZAB	0,16-0,42	18230	0,071	0,117	0,155	0,193
	PVS-2/2-ZAB	0,42-8,28	18230	0,162		0,230	
4 podzakresy	PVS-1/4-ZAB	0,16-0,30	9115	0,066	0,115	0,196	0,182
	PVS-2/4-ZAB	0,30-0,42	9115	0,074		0,098	
	PVS-3/4-ZAB	0,42-0,69	9115	0,111		0,149	
	PVS-4/4-ZAB	0,69-8,28	9115	0,210		0,285	
8 podzakresów	PVS-1/8-ZAB	0,16-0,26	4558	0,072	0,117	0,198	0,178
	PVS-2/8-ZAB	0,26-0,30	4557	0,069		0,177	
	PVS-3/8-ZAB	0,30-0,35	4558	0,065		0,085	
	PVS-4/8-ZAB	0,35-0,42	4557	0,079		0,106	
	PVS-5/8-ZAB	0,42-0,52	4558	0,099		0,132	
	PVS-6/8-ZAB	0,52-0,69	4557	0,121		0,158	
	PVS-7/8-ZAB	0,69-1,03	4558	0,153		0,203	
	PVS-8/8-ZAB	1,03-8,28	4557	0,275		0,365	



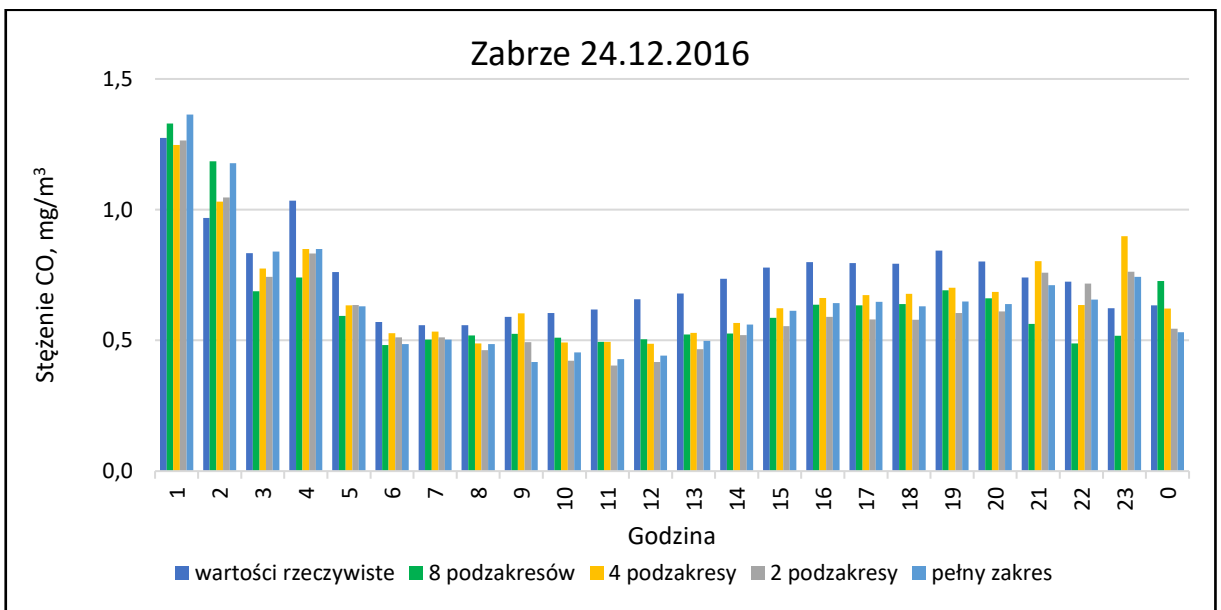
Rys. 60. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń CO, Zabrze 2011-2016

Na rys. 61 i 62 zilustrowano dobowe przebiegi stężeń predykcyjnych uzyskanych przez modele PVS w zestawieniu ze stężeniami rzeczywistymi zarejestrowanymi na stacji Zabrze w wybranych dniach okresu wiosenno-letniego i jesienno-zimowego. W dniu 01.06.2015 w godzinach 1.00-11.00 uzyskano stężenia predykcyjne stosunkowo dobrze dopasowane do wartości rzeczywistych. Niezbyt precyzyjne dopasowanie można

zaobserwować w dniu 24.12.2016, przez większość dnia modele przewidują niższe poziomy stężen w porównaniu do wartości rzeczywistych.



Rys. 61. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze w wybranym dniu w okresie wiosenno-letnim



Rys. 62. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze w wybranym dniu w okresie wiosenno-letnim

6. PODSUMOWANIE I DYSKUSJA

Archiwalne dane pomiarowe stanowią dobrą bazę do modelowania stężeń zanieczyszczeń powietrza. Przy pomocy nowoczesnego narzędzia, za jakie uznaje się sztuczne sieci neuronowe, można w stosunkowo łatwy sposób tworzyć regresyjne modele aproksymujące stężenia zanieczyszczeń powietrza na podstawie stężeń innych zanieczyszczeń i parametrów meteorologicznych zarejestrowanych na danej stacji monitoringu powietrza. W porównaniu do tradycyjnych metod statystycznych nie wymagają od użytkownika zagłębiania się w matematyczne i statystyczne ograniczenia i dzięki temu mogą być użyteczne również dla osób bez odpowiedniego przygotowania matematycznego. Utworzone modele stężeń mogą zostać wykorzystane, na przykład, do uzupełniania luk występujących w danych pomiarowych. Braki danych są często efektem nieuniknionym w procesie eksploatacji systemu monitoringu powietrza, stąd potrzeba stosowania modelowania, które może być użyteczne do zapewnienia kompletności wymaganej do oceny jakości powietrza.

Predykcja stężeń zanieczyszczeń powietrza w modelach regresyjnych jest tematem rozwijanym od kilkudziesięciu lat (Milionis, Davies, 1994; Prybutok i in., 2000; Chaloulakou i in., 2003). W ostatnim czasie można zaobserwować trend, który polega na dążeniu do poprawy dokładności predykcji modeli tworzonych za pomocą sztucznych sieci neuronowych (Hoffman, 2015; Hoffman, 2021). Jednym ze sposobów zmniejszenia błędów predykcji może być zastąpienie jednej sieci neuronowej – kilkoma innymi, stworzonymi specjalnie dla wąskiego zakresu stężeń danego zanieczyszczenia. Przeprowadzone badania dowodzą, że wprowadzenie takich rozwiązań przekłada się na zwiększenie jakości predykcji. Jedynie 8-podzakresowe modele PVS okazują się często mniej dokładnie niż modele 4-podzakresowe, co pozwala wnioskować że podział na aż tak wąskie zakresy stężeń nie zawsze jest zasadny.

W tab. 30. zamieszczono ocenę procentowej poprawy dokładności modelowania w podzakresach względem pojedynczego modelu regresyjnego stworzonego dla pełnego zakresu stężeń. Do obliczenia procentowej zmiany dokładności predykcji wykorzystano dane opisujące średnie wartości błędów MAE i RMSE zawarte w tabelach 8-29. Obliczono średnią wartość błędów MAE oraz RMSE dla wszystkich modeli z danego podzakresu i porównano go do wartości błędów uzyskanych w analogicznych modelach pełnozakresowych.

Tab. 30. Procentowa poprawa dokładności modelowania w podzakresach, względem wartości błędów uzyskanych w modelach pełnozakresowych

		Zabrze				Złoty Potok			
		Modele PVS		Modele RVS		Modele PVS		Modele RVS	
		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
O ₃	2 podzakresy	2,1%	2,9%	19,2%	21,3%	5,9%	5,6%	22,3%	23,0%
	4 podzakresy	4,9%	7,4%	39,5%	43,4%	10,2%	9,5%	44,3%	45,5%
	8 podzakresów	4,9%	8,2%	60,7%	64,1%	8,7%	8,0%	63,5%	65,1%
NO	2 podzakresy	4,6%	22,9%	9,9%	25,2%	6,4%	11,0%	27,9%	23,5%
	4 podzakresy	7,7%	36,9%	26,1%	47,7%	6,7%	16,0%	46,8%	45,3%
	8 podzakresów	8,3%	41,9%	40,6%	62,0%	6,5%	19,7%	-	-
NO ₂	2 podzakresy	1,7%	7,5%	16,8%	21,2%	3,5%	7,0%	20,3%	23,4%
	4 podzakresy	3,0%	10,5%	40,2%	45,3%	12,1%	16,6%	40,1%	44,4%
	8 podzakresów	1,9%	9,8%	61,6%	65,7%	5,6%	11,5%	59,7%	63,8%
SO ₂	2 podzakresy	1,5%	10,2%	17,7%	24,7%	0,1%	10,6%	15,4%	23,2%
	4 podzakresy	1,6%	12,0%	39,5%	47,3%	2,3%	16,4%	37,2%	46,2%
	8 podzakresów	1,4%	13,6%	57,2%	63,5%	8,4%	23,0%	57,2%	64,2%
PM10	2 podzakresy	0,5%	8,9%	16,1%	21,3%	2,0%	5,0%	26,6%	28,2%
	4 podzakresy	2,7%	15,5%	33,9%	41,7%	4,2%	9,0%	49,6%	52,0%
	8 podzakresów	5,3%	20,0%	51,3%	59,1%	1,3%	6,4%	69,3%	71,5%
CO	2 podzakresy	2,4%	4,8%	11,8%	18,9%	-	-	-	-
	4 podzakresy	3,5%	10,1%	33,4%	47,3%	-	-	-	-
	8 podzakresów	2,2%	12,1%	53,6%	64,5%	-	-	-	-

Przedstawione wyniki dowodzą zwiększenia dokładności predykcji, zwłaszcza w przypadku modeli RVS. W przypadku modeli PVS, modele 2-podzakresowe, stworzone na podstawie danych zarejestrowanych na stacji Zabrze, umożliwiły zmniejszenie wartości błędu MAE od 0,5 % do 4,6 %, a błędu RMSE od 2,9 % do 22,9 %. W modelach 4-podzakresowych uzyskano poprawę dokładności wynoszącą od 1,6 % do 7,7 % w przypadku błędu MAE oraz od 7,4 % do 36,9 % w przypadku błędu RMSE. Dla modeli 8-podzakresowych odnotowano zmniejszenie błędu MAE w zakresie 1,4 % - 8,3 %, a błędu RMSE – od 8,2 % do 41,9 %.

Znacznie większą poprawę dokładności zarejestrowano w przypadku modeli RVS. Modele 2-podzakresowe wykazały poprawę dokładności modelowania zmniejszając wartość błędu MAE w zakresie od 9,9 % do 19,2 % oraz błędu RMSE w zakresie od 18,9 % do 25,2 %. W modelach 4-podzakresowych błąd MAE zmniejszył się od 26,1 % do 40,2 %, a błąd RMSE od 41,7 % do 47,7 %. Modele 8-podzakresowe wykazały zmniejszenie błędu MAE o od 40,6 % do 61,6 %, a błędu RMSE od 59,1 % do 65,7 %.

Podobnie wyniki predykcji stężeń zanieczyszczeń powietrza uzyskano dla danych zarejestrowanych na stacji w Złotym Potoku. W modelach 2-podzakresowych PVS odnotowano zmniejszenie błędu MAE w przedziale 0,1 % - 6,4 %, a błędu RMSE w przedziale 5 % - 11 %. Dla modeli 4-podzakresowych PVS osiągnięto zmniejszenie błędu MAE w zakresie 2,3 % - 12,1 %, a błędu RMSE – w zakresie 9 % - 16,6 %. W przypadku modeli 8-podzakresowych błąd MAE został obniżony w zakresie 1,3 % - 8,7 %, a błąd RMSE w zakresie 6,4 % - 23 %.

Modele RVS dla Złotego Potoku, podobnie jak w przypadku modeli stworzonych dla stacji Zabrze, osiągnęły znacznie większą poprawę dokładności modelowania, w porównaniu z modelami PVS. Dla modeli 2-podzakresowych RVS odnotowano zmniejszenie błędu MAE o 15,4 % - 27,9 %, a błędu RMSE o 23 % - 28,2 %. Błąd MAE w przypadku modeli 4-podzakresowych zmniejszył się o 37,2 % - 49,6 %, a błąd RMSE o 44,4 % - 52 %. W przypadku modeli 8-podzakresowych odnotowano zmniejszenie błędu MAE zawierające się w przedziale 57,2 % - 69,3 %, natomiast w przypadku błędu RMSE – w przedziale 63,8 % - 71,5 %.

Podsumowując, należy stwierdzić, że zastosowanie kilku sieci neuronowych do węższych zakresów stężeń wpływa na obniżenie wartości błędów modelowania. Prawidłowość ta została potwierdzona dla wszystkich zanieczyszczeń powietrza, monitorowanych na stacjach w Zabrzu i Złotym Potoku. Modele podzakresowe wykazują zmienną poprawę dokładności predykcji, w porównaniu do dokładności modeli pełnozakresowych. Dokładność modelowania okazuje się różna w zależności od liczby submodeli (modeli podzakresowych) i poziomu stężeń charakteryzujących dany podzakres. Dla modeli RVS uzyskano znacznie lepsze wyniki predykcji w porównaniu z modelami PVS. Liczba podzakresów stężeń dla modeli PVS nie może być zbyt duża. Podział pełnozakresowego zbioru na 2 lub 4 podzakresy zapewniał poprawę jakości predykcji w każdym rozpatrywanym przypadku. Dla modeli PVS, podział na 8 podzakresów nie zawsze gwarantował spadek błędu modelowania, w porównaniu do modeli 4-podzakresowych.

7. WNIOSKI

7.1. Wnioski główne

Na podstawie uzyskanych wyników sformułować można kilka najważniejszych wniosków, korespondujących z tezami i celami rozprawy:

1. Archiwalne dane pomiarowe monitoringu powietrza stanowią dobrą bazę do aproksymowania stężeń zanieczyszczeń powietrza.
2. Sieci neuronowe stanowią praktyczne narzędzie do modelowania predykcyjnego poziomów stężeń zanieczyszczeń powietrza.
3. W porównaniu do jednej sieci neuronowej dla pełnego zakresu stężeń zanieczyszczeń powietrza, zastosowanie kilku sieci neuronowych do węższych zakresów stężeń wpływa na obniżenie wartości błędów modelowania. Prawdliwość ta została potwierdzona dla 6 podstawowych zanieczyszczeń powietrza, monitorowanych w sposób automatyczny.

7.2. Wnioski szczegółowe

Na podstawie otrzymanych wyników sformułować można następujące wnioski dodatkowe:

1. Modele podzakresowe wykazują zróżnicowaną poprawę dokładności predykcji, w porównaniu do dokładności modeli pełnozakresowych.
2. Dokładność modelowania zależy od liczby modeli podzakresowych i poziomu stężeń charakteryzujących dany podzakres. Na ogół największe wartości błędów występują przy modelach dedykowanych do podzakresów z najwyższymi poziomami stężeń, a najniższe – przy modelach dedykowanych dla najniższych podzakresów reprezentujących najniższe poziomy stężeń.
3. Dla modeli podzakresowych utworzonych po wstępnym sortowaniu przypadków według rosnących wartości stężeń rzeczywistych modelowanego zanieczyszczenia (modele RVS) uzyskano znacznie lepsze wyniki predykcji w porównaniu do modeli podzakresowych utworzonych po wstępnym sortowaniu przypadków według rosnących wartości stężeń predykcyjnych modelowanego zanieczyszczenia (modele PVS).

4. Liczba podzakresów stężeń dla modeli PVS nie może być zbyt duża. Podział pełnozakresowego zbioru przypadków na 8 podzakresów nie zawsze skutkuje poprawą dokładności aproksymacji w tak utworzonych modelach podzakresowych.

BIBLIOGRAFIA

1. Adamczak A., (2021): Metoda detekcji wad spawalniczych w stanowisku zrobotyzowanym z wykorzystaniem głębokiej sieci neuronowej, *Pomiary Automatyka Robotyka*, R. 25, Nr 1/2021, 67–72, DOI: 10.14313/PAR_239/67.
2. Atkinson R. W., Ross Anderson H., Sunyer J., Ayres J., Baccini M., Vonk J. M., I in., (2001): Acute Effects of Particulate Air Pollution on Respiratory Admissions. *American Journal of Respiratory and Critical Care Medicine*, 164(10), 1860–1866, DOI:10.1164/ajrccm.164.10.2010138.
3. Armstrong B., Hutchinson E., Unwin J., Fletcher T., (2004): Lung Cancer Risk after Exposure to Polycyclic Aromatic Hydrocarbons: A Review and Meta-Analysis, *Environmental Health Perspectives*, 112(9), 970–978, DOI:10.1289/ehp.6895.
4. Bank danych pomiarowych GIOŚ, (2021): <https://powietrze.gios.gov.pl/pjp/archives> (dostęp: 03.2021).
5. Bishop Ch. M., (1995): *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford.
6. Bell M. L., Davis D. L., (2001): Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution, *Environmental Health Perspectives*, 109 (suppl 3), 389–394, DOI:10.1289/ehp.01109s3389.
7. Błaszczak J., Janowski R., (2021): System automatycznej detekcji tablic rejestracyjnych i identyfikacji numerów pojazdów wykorzystujących sieci neuronowe, *Zeszyty Naukowe WWSI*, Vol. 15, No 25, 7-31, DOI: 10.26348/znwwsi.25.7.
8. Brunekreef B., (1997): Air pollution and life expectancy: is there a relation? *Occupational and Environmental Medicine*, 54(11), 781–784, DOI:10.1136/oem.54.11.781.
9. Bogacki M., Oleniacz R., (2004): Referencyjna metodyka modelowania poziomów substancji w powietrzu na tle innych modeli obliczeniowych, *Inżynieria Środowiska*, Akademia Górniczo-Hutnicza im. S. Staszica w Krakowie, 9(1), 35-45.
10. Chaloulakou A., Grivas G., Spyrellis N., (2003): Neural Network and Multiple Regression Models for PM10 Prediction in Athens: A Comparative Assessment.

- Journal of the Air & Waste Management Association, 53(10), 1183–1190, DOI:10.1080/10473289.2003.1046627.
11. Cembrzyńska J., Krakowiak E., Brewczyński P. Z., (2012): Zanieczyszczenie powietrza pyłem zawieszonym PM10 oraz PM2,5 w warunkach silnej antropopresji na przykładzie miasta Sosnowiec, *Environmental Medicine*, Vol. 15, No. 4, 31–38.
 12. Chlebowska-Styś A., Sówka I., Pachurka Ł., (2016): Analiza składu pyłu zawieszonego PM10 na stacji tła miejskiego w Pile, *Interdyscyplinarne Zagadnienia w Inżynierii i Ochronie Środowiska*, Tom 8, 36-53, Oficyna Wydawnicza Politechniki Wrocławskiej.
 13. Ciechulski T. Osowski S., (2014): Badanie jakości predykcji obciążeń elektroenergetycznych za pomocą sieci neuronowych SVM, RBF i MLP, *Przegląd Elektrotechniczny*, 90(8), 148-151, DOI:10.12915/pe.2014.08.3.
 14. Clancy L., Goodman P., Sinclair H., Dockery D. W., (2002): Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study, *The Lancet*, 360(9341), 1210–1214, DOI:10.1016/s0140-6736(02)11281-5.
 15. Collins J. F., Brown J. P., Dawson S. V., Marty M. A., (1991): Risk assessment for benzo[a]pyrene. *Regulatory Toxicology and Pharmacology*, 13(2), 170–184, DOI:10.1016/0273-2300(91)90020-v.
 16. Daly A., Zannetti P., (2007): Air pollution modeling – an overview, *Ambient air pollution*, Chapter 2, 15-28.
 17. De Vries W., (2021): Impacts of nitrogen emissions on ecosystems and human health: A mini review, *Current Opinion in Environmental Science & Health*, 21, 100249, DOI:10.1016/j.coesh.2021.100249.
 18. Drozdek A., (2000): Sieci neuronowe i metodologia nauki, *Filozofia Nauki*, 8(2), 89-99.
 19. Dyrektywa, (2008): Dyrektywa Parlamentu Europejskiego i Rady 2008/50/WE z dnia 21 maja 2008 r. w sprawie jakości powietrza i czystszej powietrza dla Europy.
 20. Dyrektywa, (2004): Dyrektywa 2004/107/WE Parlamentu Europejskiego i Rady z dnia 15 grudnia 2004 r. w sprawie arsenu, kadmu, rtęci, niklu i wielopierścieniowych węglowodorów aromatycznych w otaczającym powietrzu.
 21. Dyrektywa, (2015): Dyrektywa Komisji (UE) 2015/1480 z dnia 28 sierpnia 2015 r. zmieniająca niektóre załączniki do dyrektywy Parlamentu Europejskiego i Rady

2004/107/WE i 2008/50/WE ustanawiających przepisy dotyczące metod referencyjnych, zatwierdzania danych i lokalizacji punktów pomiarowych do oceny jakości powietrza.

22. Farzad K., Khorsandi B., Khorsandi M., Bouamra O., Maknoon R., (2021): Estimating short-term mortality benefits associated with a reduction in tropospheric ozone, *Atmospheric Environment*, 252, 118342, DOI:10.1016/j.atmosenv.2021.118342.
23. Fausett L., (1994): *Fundamentals of Neural Networks*, Prentice Hall, New York.
24. Firket J., (1936): Fog along the Meuse valley, *Transactions of the Faraday Society*, 32, 1192, DOI:10.1039/tf9363201192.
25. Flasiński M., (2011): *Wstęp do sztucznej inteligencji*, Wydawnictwo Naukowe PWN, s. 176, Warszawa.
26. Gardner M. W., Dorling S. R., (1998): Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences, *Atmos. Environ.*, 32, 14/15, 2627-2636.
27. Hadjiiski L., Geladi P., Hopke P., (1999): A comparison of modeling nonlinear systems with artificial neural networks and partial least squares, *Chemometrics and Intelligent Laboratory Systems*, 49, 91-103.
28. Hajto M. J., Godłowska J., Kaszowski W., Tomaszewska A. M., (2012): System prognozowania rozprzestrzeniania zanieczyszczeń powietrza FAPPS – założenia, możliwości, rozwój, *Ochrona powietrza w teorii i praktyce*, 2, Instytut Podstaw Inżynierii Środowiska PAN, 89-96, Zabrze.
29. Hauck H., Kromp-Kolb H., Petz E., (1999): Requirements for the completeness of ambient air quality data sets with respect to derived parameters, *Atmospheric Environment*, Volume 33, Issue 13, 1 June 1999, Pages 2059-2066, [https://doi.org/10.1016/S1352-2310\(98\)00401-4](https://doi.org/10.1016/S1352-2310(98)00401-4).
30. Haykin S., (1994): *Neural Networks: A Comprehensive Foundation*, Macmillan Publishing, New York.
31. Heaviside C., Witham C., Vardoulakis S., (2021): Potential health impacts from sulphur dioxide and sulphate exposure in the UK resulting from an Icelandic effusive volcanic eruption, *Science of The Total Environment*, 774, 145549, DOI:10.1016/j.scitotenv.2021.14554.

32. Hippert H. S., Pedreira C. E., Souza R. C., (2001): Neural networks for short-term load forecasting: a review and evaluation, *IEEE Transactions on Power Systems*, 16(1), 44–55, DOI:10.1109/59.910780.
33. Hoffman S., (2004): Zastosowanie sieci neuronowych w modelowaniu regresyjnym stężeń zanieczyszczeń powietrza, Wydawnictwo Politechniki Częstochowskiej, Częstochowa.
34. Hoffman S., Jasiński R., (2009): Uzupełnienie brakujących danych w systemach monitoringu powietrza, Wydawnictwo Politechniki Częstochowskiej, Częstochowa.
35. Hoffman S., (2013): Modelowanie poziomu imisji na stacjach monitoringu powietrza za pomocą perceptronowych sieci neuronowych, Wydawnictwo Politechniki Częstochowskiej, Częstochowa.
36. Hoffman S., (2015): Assessment of prediction accuracy in autonomous air quality models, *Desalination and Water Treatment*, 57(3), 1322–1326, DOI:10.1080/19443994.2014.1002283.
37. Hoffman S., (2021): Estimation of Prediction Error in Regression Air Quality Models, *Energies*, 14, 7387, <https://doi.org/10.3390/en1421738>.
38. Hornik K., Stinchcombe M., White H., (1989): Multilayer feedforward networks are universal approximators, *Neural Networks*, 2, 359-366.
39. GIOŚ, (2021): strona internetowa Głównego Inspektoratu Ochrony Środowiska: <https://www.gios.gov.pl/pl/> (dostęp: 20.11.2021 r.).
40. Hunt A., Abraham J. L., Judson B., Berry C. L. (2003): Toxicologic and epidemiologic clues from the characterization of the 1952 London smog fine particulate matter in archival autopsy lung tissues, *Environmental Health Perspectives*, 111(9), 1209–1214, DOI:10.1289/ehp.6114.
41. IMGW, (2022): <http://smog.imgw.pl/content/model> (dostęp: 12.02.2022 r.).
42. Jamhari A. A., Sahani M., Latif M. T., Chan K. M., Tan H. S., Khan M. F., Mohd Tahir N., (2014): Concentration and source identification of polycyclic aromatic hydrocarbons (PAHs) in PM10 of urban, industrial and semi-urban areas in Malaysia. *Atmos. Environ.*, 86, 16–27, DOI:10.1016/j.atmosenv.2013.12.019.
43. Jędrak J., Konduracka E., Badyda A. J., Dąbrowiecki P., (2017): Wpływ zanieczyszczeń powietrza na zdrowie, *Krakowski Alarm Smogowy*, ISBN: 978-83-943065-0-2, Kraków.

44. Juda-Rezler K., (2010): New Challenges in Air Quality and Climate Modeling, *Archives of Environmental Protection*, Vol. 36, no. 1, 3-28.
45. Karppinen A., Kukkonen J., Elolähde T., Konttinen M., Koskentalo T., (2000): A modelling system for predicting urban air pollution: *Atmospheric Environment*, 34(22), 3735–3743, DOI:10.1016/s1352-2310(00)00073-x.
46. Khomenko S., Cirach M., Pereira-Barboza E., Mueller N., Barrera-Gómez J., Rojas-Rueda D., De Hoogh K., Hoek G., Nieuwenhuijsen M., (2021): Premature mortality due to air pollution in European cities: a health impact assessment, *The Lancet Planetary Health*, 5(3), e121–e134, DOI:10.1016/s2542-5196(20)30272-2.
47. Kozielska B., Rogula-Kozłowska W., (2014): Polycyclic aromatic hydrocarbons in particulate matter in the cities of Upper Silesia, *Archiwum Gospodarki Odpadami I Ochrony Środowiska*, vol. 16, 2, 75-84.
48. Krewski D., (2009): Evaluating the Effects of Ambient Air Pollution on Life Expectancy, *New England Journal of Medicine*, 360(4), 413–415, DOI:10.1056/nejme0809178.
49. Krogh A., (2008): What are artificial neural networks? *Nature Biotechnology*, 26(2), 195–197, DOI:10.1038/nbt1386.
50. Kumar U., Jain V. K., (2009): ARIMA forecasting of ambient air pollutants (O₃, NO, NO₂ and CO), *Stochastic Environmental Research and Risk Assessment*, 24(5), 751–760, DOI:10.1007/s00477-009-0361-8.
51. Kwasigroch A., Grochowski M., (2018): Rozpoznawanie obiektów przez głębokie sieci neuronowe, *Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki PG*, ISSN 2353-1290, Nr 60/2018, DOI: 10.32016/1.60.12.
52. Latini G., Passerini G., (2004): *Handling Missing Data: Applications to Environmental Analysis*, WIT Press, Southampton, Boston.
53. Lee R. G. M., Coleman P., Jones J. L., Jones K. C., Lohmann R., (2005): Emission Factors and Importance of PCDD/Fs, PCBs, PCNs, PAHs and PM10 from the Domestic Burning of Coal and Wood in the U.K., *Environmental Science & Technology*, 39(6), 1436–1447, DOI:10.1021/es048745i.
54. Li M., Dong H., Wang B., Zhao W., Sakhvidi M. J. Z., Li L., Lin G., Yang, J., (2020): Association between ambient ozone pollution and mortality from a spectrum of causes in Guangzhou, China, *Science of The Total Environment*, 142110, DOI:10.1016/j.scitotenv.2020.142110.

55. Li D., Ren X., (2022): Prediction of Ozone Hourly Concentrations Based on Machine Learning Technology, *Sustainability*, 14, 5964, <https://doi.org/10.3390/su14105964>.
56. Logan W., (1953): Mortality in the London fog incident, 1952, *The Lancet*, 261(6755), 336-338, DOI:10.1016/s0140-6736(53)91012-5.
57. Mantis J., Chaloulakou A., Samara C., (2005): PM10-bound polycyclic aromatic hydrocarbons (PAHs) in the Greater Area of Athens, Greece. *Chemosphere*, 59(5), 593–604, DOI:10.1016/j.chemosphere.2004.10.
58. McCulloch W.S., Pitts W.H., (1943): A logical calculus of ideas immanent in nervous activity, *Bull. Math. Biophysics*, 5, 115-119.
59. Milionis A. E., Davies T. D., (1994): Regression and stochastic models for air pollution - I. Review, comments and suggestions, *Atmospheric Environment*, 28(17), 2801–2810, DOI:10.1016/1352-2310(94)90083-3.
60. Minsky M., Papert S., (1969): *Perceptrons: An introduction to computational geometry*, MIT Press, Cambridge.
61. Nabney I., (2001): *Algorithms for Pattern Recognition*, *Advances in Pattern Recognition*, Springer, ISBN: 1-85233-440-1.
62. Nadeem I., Ilyas A.M., Uduman P.S., (2020): Analyzing and forecasting ambient air quality of Chennai City in India, *Geography, Environment, Sustainability*, 2020;13(3):13-21, <https://doi.org/10.24057/2071-9388-2019-97>.
63. Nagendra S.M.S., Khare M., (2005): Modelling urban air quality using artificial neural network, *Clean Technologies and Environmental Policy* 7, 116–126, <https://doi.org/10.1007/s10098-004-0267-6>.
64. Norma PN-EN 14211:2013-02 „Powietrza atmosferyczne – Standardowa metoda pomiaru stężenia ditlenku azotu i tlenku azotu za pomocą chemiluminescencji”.
65. Norma PN-EN 14212:2013-02 „Powietrza atmosferyczne – Standardowa metoda pomiaru stężenia ditlenku siarki za pomocą fluorescencji w nadfiolecie”.
66. Norma PN-EN 14625:2013-02 „Powietrze atmosferyczne – Standardowa metoda pomiaru stężenia ozonu z wykorzystaniem fotometrii w nadfiolecie”.
67. Norma PN-EN 12341:2014-07 „Powietrze atmosferyczne – Standardowa grawimetryczna metoda pomiarowa do określenia stężeń masowych frakcji PM10 lub PM2,5 pyłu zawieszonego”.

68. Norma PN-EN 14626:2013-02 „Powietrze atmosferyczne – Standardowa metoda pomiaru stężenia tlenku węgla za pomocą niedyspersyjnej spektroskopii w podczerwieni”.
69. Nuvolone D., Petri D., Voller F., (2017): The effects of ozone on human health, *Environmental Science and Pollution Research*, 25(9), 8074–8088, DOI:10.1007/s11356-017-9239-3.
70. Orellano P., Reynoso J., Quaranta N., (2021): Short-term exposure to sulphur dioxide (SO₂) and all-cause and respiratory mortality: A systematic review and meta-analysis. *Environment International*, 150, 106434, DOI:10.1016/j.envint.2021.106434.
71. Osowski S., (1996): *Sieci neuronowe w ujęciu algorytmicznym*, WNT, Warszawa
72. Osowski S., Stodolski M., Bojarczak P., (1996): Fast second order learning algorithm for feedforward MLP, *Neural Networks*, vol. 9, 1583-1597.
73. Patterson D., (1996): *Artificial Neural Networks*, Prentice Hall, Singapore.
74. Pawlak I., Jarosławski J., (2019): Forecasting of Surface Ozone Concentration by Using Artificial Neural Networks in Rural and Urban Areas in Central Poland, *Atmosphere*, 10, 52, <https://doi.org/10.3390/atmos10020052>.
75. Pawul M., Śliwka M., (2016): Application of artificial neural networks for prediction of air pollution levels in environmental monitoring, *Journal of Ecological Engineering*, 17(4), 190-196.
76. Pawul M., (2019): Application of neural networks to the prediction of gas pollution of air, *New Trends in Production Engineering*, Vol. 2, Issue 1, 515-523, DOI:10.2478/ntpe-2019-0055.
77. Plaia A., Bondi A., (2006): Single imputation method of missing values in environmental pollution data sets, *Atmospheric Environment*, 40(38), 7316–7330, DOI:10.1016/j.atmosenv.2006.06.040.
78. Pluciński M., (2002): Application of the Probabilistic RBF Neural Network in Multidimensional Classification Problems, *Advanced Computer Systems*, 49–57, DOI:10.1007/978-1-4419-8530-9_4.
79. Pohoata A., Lungu E., (2017): A complex analysis employing ARIMA model and statistical methods on air pollutants recorded in Ploiesti, Romania, *Rev Chim*, 68(4), 818-823.

80. Polivka B. J., (2018): The Great London Smog of 1952. *AJN, American Journal of Nursing*, 118(4), 57–61, DOI:10.1097/01.naj.0000532078.72372.c3.
81. Poloczek Ł., Wilkosz M., Czech P., Saternus M., Kania H., (2021): Wykorzystanie sztucznych sieci neuronowych typu MLP do predykcji zanieczyszczenia powietrza na podstawie danych pogodowych ze stacji pomiarowej, *Projektowanie, badania i eksploatacja*, 223-240, DOI:10.53052/9788366249837.20.
82. Pope C. A., (1991): Respiratory Hospital Admissions Associated with PM10 Pollution in Utah, Salt Lake, and Cache Valleys, *Archives of Environmental Health: An International Journal*, 46(2), 90–97, DOI:10.1080/00039896.1991.9937434.
83. Pope C. A., Dockery D. W., (2006): Health Effects of Fine Particulate Air Pollution: Lines that Connect. *Journal of the Air & Waste Management Association*, 56(6), 709–742, DOI:10.1080/10473289.2006.10464485.
84. Program Państwowego Monitoringu Środowiska na lata 2007-2009, (2006), Główny Inspektor Ochrony Środowiska, Warszawa.
85. Program Państwowego Monitoringu Środowiska na lata 2010-2012, (2009), Główny Inspektor Ochrony Środowiska, Warszawa.
86. Program Państwowego Monitoringu Środowiska na lata 2013-2015, (2012), Główny Inspektor Ochrony Środowiska, Warszawa.
87. Program Państwowego Monitoringu Środowiska na lata 2016-2020, (2015), Główny Inspektor Ochrony Środowiska, Warszawa.
88. Prybutok V. R., Yi J., Mitchell D., (2000): Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations, *European Journal of Operational Research*, 122(1), 31–40, DOI:10.1016/s0377-2217(99)00069-7.
89. Punetha A., Saraswat S., Rai J. P. N.: An insight on microbial degradation of benzo[a]pyrene: current status and advances in research, *World Journal Microbiology and Biotechnology* 38, article numer: 61, <https://doi.org/10.1007/s11274-022-03250-3>.
90. Raport EEA, (2021): Air quality in Europe 2021, European Environment Agency, Report no. 15/2021, DOI: 10,2800/549289.
91. Raport EEA, (2020): Air quality in Europe – 2020 report, European Environment Agency, Report no. 09/2020.

92. Rekhi J. K., Nagrath P., Jain R., (2020): Forecasting Air Quality of Delhi Using ARIMA Model, in: *Advances in Data Sciences, Security and Applications*, 315-325, Springer, Singapore.
93. Rogula-Kozłowska W., (2015): Size-segregated urban particulate matter: mass closure, chemical composition, and primary and secondary matter content, *Air Quality, Atmosphere & Health*, 9(5), 533-550, DOI:10.1007/s11869-015-0359-y.
94. Rojek I., Studziński J., (2012): Sieci neuronowe w lokalizacji awarii w sieci wodociągowej, *Studia i Materiały Informatyki Stosowanej*, Tom 4, Nr 9, 29-34.
95. Rosenblatt F., (1958): The perceptron: a probabilistic model for information storage and organization in the brain, *Psychological Review*, 65(6), 386-408, <https://doi.org/10.1037/h0042519>
96. Rosienkiewicz M., (2019): Prognozowanie w produkcji – przykłady zastosowania, prezentacja na XXIII Konferencji, Zastosowania statystyki i data mining w badaniach naukowych, Warszawa, 24.10.2019.
97. Rozporządzenie, (2012): Rozporządzenie Ministra Środowiska z dnia 13 września 2012 r. w sprawie dokonywania oceny poziomów substancji w powietrzu, Dz. U. 2012 poz. 1032.
98. Rozporządzenie, (2018): Rozporządzenie Ministra Środowiska z dnia 8 czerwca 2018 r. w sprawie dokonywania oceny poziomów substancji w powietrzu, Dz. U. 2018 poz. 1119.
99. Rozporządzenie, (2020): Rozporządzenie Ministra Klimatu i Środowiska z dnia 11 grudnia 2020 r. w sprawie dokonywania oceny poziomów substancji w powietrzu, Dz. U. 2020 poz. 2279
100. Rumelhart D. E., Hinton G. E., Williams R. J., (1986): Learning representations by back-propagating errors, *Nature*, 323(6088), 533-536.
101. Schrenk H. H., Heimann H., Clayton G. D., Gafafer W. M., Wexler H., (1949): Air pollution in Donora, Pa: epidemiology of unusual smog episode of October 1948: preliminary report, Federal Security Agency, Public Health Service, Bureau of State Services, Division of Industrial Hygiene, Washington.
102. Shams S. R., Jahani A., Moeinaddini M., Khorasani N., (2020): Air carbon monoxide forecasting using an artificial neural network in comparison with multiple regression, *Modeling Earth Systems and Environment*, DOI:10.1007/s40808-020-00762-5.

103. Siderska J., (2021): Pomiar wartości kapitału społecznego z wykorzystaniem sztucznych sieci neuronowych, Oficyna Wydawnicza Politechniki Białostockiej, ISBN: 978-83-67185-00-4, Białystok.
104. Siljama, O., Koskinen, T., Jessen-Juhler, O. I in., (2021): Automated Flaw Detection in Multi-channel Phased Array Ultrasonic Data Using Machine Learning, *Journal of Nondestructive Evaluation*, vol. 40, article number 67, DOI.org/10.1007/s10921-021-00796-4.
105. Sims P., Grover P. L., Swaisland A., Pal K., Hewer A., (1974): Metabolic activation of benzo(a)pyrene proceeds by a diol-epoxide, *Nature*, 252(5481), 326–328, DOI:10.1038/252326a0.
106. Siwek K., Osowski S., (2001): Regularization of neural networks for improved load forecasting in power system, ICECS 2001, 8th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.01EX483), DOI:10.1109/icecs.2001.957443.
107. Siwek K., Osowski S., (2012): Improving the accuracy of prediction of PM₁₀ pollution by the wavelet transformation and an ensemble of neural predictors, *Engineering Applications of Artificial Intelligence*, Vol. 25, 6, 1246-1258, DOI.org/10.1016/j.engappai.2011.10.013.
108. Słota-Valim M., (2018): Określanie mechanicznych właściwości skał na podstawie właściwości fizycznych przy użyciu sztucznych sieci neuronowych, *Nafta-Gaz* 2018, nr 5, s. 343–355, DOI:10.18668/NG.2018.05.01.
109. Sówka I., Chlebowska-Styś A., Pachurka Ł., Rogula-Kozłowska W., Mathews B., (2019): Analysis of Particulate Matter Concentration Variability and Origin in Selected Urban Areas in Poland. *Sustainability*, 11, 5735, DOI.org/10.3390/su11205735
110. StatSoft, (2022), Internetowy Podręcznik Statystyki, <https://www.statsoft.pl/textbook/stathome.html> (dostęp: 19.01.2022 r.).
111. Strategiczny program Państwowego Monitoringu Środowiska na lata 2020-2025, (2020), Główny Inspektor Ochrony Środowiska, Warszawa.
112. Tadeusiewicz R., (1993): Sieci neuronowe, Akademicka Oficyna Wydawnicza, Warszawa.
113. Tadeusiewicz R., (2009): O celowości zastosowania sieci neuronowych w problemach związanych z elektrotechniką, *Przegląd Elektrotechniczny*, 85, 200-211.

114. Tadeusiewicz R., (2015): Neural Networks In Mining Sciences – General Overview And Some Representative Examples. Archives of Mining Sciences, 60(4), 971–984, DOI:10.1515/amsc-2015-0064.
115. Tadeusiewicz R., Szaleniec M., (2015): Leksykon sieci neuronowych, Wydawnictwo Fundacji „Projekt Nauka”, Wrocław.
116. Tétréault L.-F., Doucet M., Gamache P., Fournier M., Brand A., Kosatsky T. Smargiassi A., (2016): Childhood Exposure to Ambient Air Pollutants and the Onset of Asthma: An Administrative Cohort Study in Québec, Environmental Health Perspectives, 124(8), 1276–1282, DOI:10.1289/ehp.1509838.
117. Tumidajski T., Foszcz D., Niedoba T., Siewior J., (2009): Modele stochastyczne zanieczyszczeń powietrza w aglomeracjach przemysłowych, Rocznik Ochrona Środowiska, 11, 543-554.
118. Ustawa, (1991): Ustawa z dnia 20 lipca 1991 r. o Państwowej Inspekcji Ochrony Środowiska, Dz. U. 1991, nr 77, poz. 335.
119. Ustawa, (2001): Ustawa z dnia 27 kwietnia 2001 r. Prawo Ochrony Środowiska, Dz. U. 2001, nr 62, poz. 627.
120. Ustawa, (2017): Ustawa z dnia 14 grudnia 2017 r. o zmianie ustawy o Inspekcji Ochrony Środowiska oraz ustawy – Prawo ochrony środowiska, Dz.U. 2018, poz. 88.
121. Vallero D. A., (2019): Air Pollution Calculations: Quantifying Pollutant Formation, Transport, Transformation, Fate and Risks, Elsevier, DOI:10.1016/c2017-0-02742-8.
122. Vallero D., (2014): Fundamentals of Air Pollution, Fifth Edition, Elsevier,.
123. Wang W., Guo Y., (2009): Air pollution PM2.5 data analysis in Los Angeles long beach with seasonal ARIMA model, IEEE, Vol. 3, 7-10.
124. Wang M., Li H., Huang S., Qian Y., Steenland K., Xie Y., Papatheodorou S., Shi, L., (2021): Short-term exposure to nitrogen dioxide and mortality: A systematic review and meta-analysis. Environmental Research, 202, 111766, DOI:10.1016/j.envres.2021.111766.
125. World Health Organization, (2006): WHO Air Quality Guidelines, Global Update 2005, Particulate matter, ozone, nitrogen dioxide and sulfur dioxide.
126. World Health Organization, (2021): WHO Global Air Quality Guidelines, Particulate matter (PM2,5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.

127. Wieczorek J., Wieczorek Z., Mozolewski W., Pomianowski J., (2011): Wielopierścieniowe węglowodory aromatyczne w pyłe PM10, *Inż. Ap. Chem.* 2011, 50, 2, 26-27.
128. Wielgosiński G., Czerwińska J., Namiecińska O., Cichowicz R., (2018): Smog episodes in the Lodz agglomeration in the years 2014-17, *E3S Web of Conferences*, 28, 01039, DOI:10.1051/e3sconf/20182801039.
129. Wielgosiński G., Czerwińska J., (2020): Smog Episodes in Poland, *Atmosphere*, 11, 277, DOI.org/10.3390/atmos11030277.
130. Wilkosz M., Poloczek Ł., Czech P., Saternus M., Kania H., (2021): Prognozowaniu zanieczyszczenia powietrza atmosferycznego przy użyciu szeregów czasowych i różnych typów sztucznych sieci neuronowych, *Projektowanie, badania i eksploatacja*, 295-308, DOI:10.53052/9788366249837.28.
131. Willmott C. J., (1981): On the validation of models, *Physical Geography*, 2, 184-194.
132. Willmott C. J., Ackleson S. G., Davis R. E., Feddema J. J., Klink K. M., Legates D. R., O'Donnell J., Rowe C. M., (1985): Statistics for the evaluation and comparison of models, *J. Geophys. Res.*, 90, C5, 8995-9005.
133. Yamazaki S., Shima M., Yoda Y., Oka K., Kurosaka F., Shimizu S. i in., (2015): Exposure to air pollution and meteorological factors associated with children's primary care visits at night due to asthma attack: case-crossover design for 3-year pooled patients, *BMJ Open*, 5(4), e005736, DOI:10.1136/bmjopen-2014-005736.
134. Zhang H., Zhang S., Wang P., Qin Y., Wang H., (2017): Forecasting of particulate matter time series using wavelet analysis and wavelet-ARMA/ARIMA model in Taiyuan, China, *Journal of the Air & Waste Management Association*, 67(7), 776-788, DOI:10.1080/10962247.2017.1292968.

Spis zamieszczonych tabel

Tab. 1. Charakterystyka wybranych stacji monitoringu powietrza.....	strona 15
Tab. 2. Zestawienie danych zarejestrowanych na poszczególnych stacjach monitoringu powietrza (+ - mierzone, - - niemierzone).....	strona 16
Tab. 3. Podstawowe parametry statystyczne danych opisujących poziomy stężenia zarejestrowanych zanieczyszczeń powietrza.....	strona 16
Tab. 4. Procentowa średnia kompletność danych zarejestrowanych w okresie 2011-2016 w poszczególnych miesiącach na stacji Zabrze oraz Złoty Potok.....	strona 17
Tab. 5. Minimalne i maksymalne średnie wartości godzinowe zanieczyszczeń powietrza zarejestrowane w okresie letnim i zimowym.....	strona 21
Tab. 6. Wartości błędów predykcji obliczone dla każdego typu uzyskanych sieci.....	strona 23
Tab. 7. Opis zmiennych w modelach stężeń poszczególnych zanieczyszczeń dla obu stacji monitoringu powietrza.....	strona 29
Tab. 8. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych O ₃ , Zabrze 2011-2016.....	strona 36
Tab. 9. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych O ₃ , Złoty Potok 2011-2016.....	strona 37
Tab. 10. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO, Zabrze 2011-2016.....	strona 42
Tab. 11. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO, Złoty Potok 2011-2016.....	strona 43
Tab. 12. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO ₂ , Zabrze 2011-2016.....	strona 46
Tab. 13. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych NO ₂ , Złoty Potok 2011-2016.....	strona 47
Tab. 14. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych SO ₂ , Zabrze 2011-2016.....	strona 51
Tab. 15. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych SO ₂ , Złoty Potok 2011-2016.....	strona 52
Tab. 16. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych PM ₁₀ , Zabrze 2011-2016.....	strona 55

Tab. 17. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych PM10, Złoty Potok 2011-2016.....	strona 56
Tab. 18. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń rzeczywistych CO, Zabrze 2011-2016.....	strona 60
Tab. 19. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych O ₃ , Zabrze 2011-2016.....	strona 63
Tab. 20. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych O ₃ , Złoty Potok 2011-2016.....	strona 64
Tab. 21. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO, Zabrze 2011-2016.....	strona 68
Tab. 22. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO, Złoty Potok 2011-2016.....	strona 69
Tab. 23. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO ₂ , Zabrze 2011-2016.....	strona 73
Tab. 24. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych NO ₂ , Złoty Potok 2011-2016.....	strona 74
Tab. 25. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych SO ₂ , Zabrze 2011-2016.....	strona 77
Tab. 26. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych SO ₂ , Złoty Potok 2011-2016.....	strona 78
Tab. 27. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych PM10, Zabrze 2011-2016.....	strona 81
Tab. 28. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych PM10, Złoty Potok 2011-2016.....	strona 82
Tab. 29. Wartości błędów aproksymacji obliczonych dla podzakresów utworzonych w wyniku podziału przypadków względem stężeń predykcyjnych CO, Zabrze 2011-2016.....	strona 85
Tab. 30. Procentowa poprawa dokładności modelowania względem wartości błędów uzyskanych przez model pełnozakresowy.....	strona 88

Spis zamieszczonych rysunków

Rys. 1. Lokalizacja wybranych stacji monitoringu powietrza z zaznaczonymi największymi miastami na mapie Polski.....	strona 15
Rys. 2. Procentowa średnia kompletność danych zarejestrowanych na stacji Zabrze z okresu 2011-2016 dla poszczególnych miesięcy.....	strona 18
Rys. 3. Procentowa średnia kompletność danych zarejestrowanych na stacji Złoty Potok z okresu 2011-2016 dla poszczególnych miesięcy.....	strona 18
Rys. 4. Przebiegi średniomiesięcznych stężeń zanieczyszczeń zarejestrowanych na stacji Zabrze w miesiącach styczeń i luty.....	strona 19
Rys. 5. Przebiegi średniomiesięcznych stężeń zanieczyszczeń zarejestrowanych na stacji Złoty Potok w miesiącach styczeń i luty.....	strona 19
Rys. 6. Przebiegi średniomiesięcznych stężeń zanieczyszczeń zarejestrowanych na stacji Zabrze w miesiącach czerwiec i lipiec.....	strona 20
Rys. 7. Przebiegi średniomiesięcznych stężeń zanieczyszczeń zarejestrowanych na stacji Złoty Potok w miesiącach czerwiec i lipiec.....	strona 20
Rys. 8. Porównanie wartości rzeczywistych i predykcyjnych uzyskanych przez sieć liniową, MLP oraz RBF, stężenie O ₃ zarejestrowane na stacji Zabrze, 06.04.2015 r.	strona 24
Rys. 9. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia O ₃ , Zabrze 2011-2016.....	strona 25
Rys. 10. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia NO, Zabrze 2011-2016.....	strona 25
Rys. 11. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia NO ₂ , Zabrze 2011-2016.....	strona 26
Rys. 12. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia SO ₂ , Zabrze 2011-2016.....	strona 26
Rys. 13. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia PM ₁₀ , Zabrze 2011-2016.....	strona 27
Rys. 14. Wartości błędu MAE i współczynnika korelacji Pearsona w zależności od liczby neuronów w warstwie ukrytej sieci MLP, model stężenia CO, Zabrze 2011-2016.....	strona 27
Rys. 15. Schemat podziału na podzakresy dla stacji Zabrze ze wskazaniem liczby przypadków w poszczególnych podzakresach modelowania.....	strona 32
Rys. 16. Schemat podziału na podzakresy dla stacji Złoty Potok ze wskazaniem liczby przypadków w poszczególnych podzakresach modelowania.....	strona 32

Rys. 17. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O_3 , Zabrze 2011-2016.....	strona 37
Rys. 18. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O_3 , Złoty Potok 2011-2016.....	strona 38
Rys. 19. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 39
Rys. 20. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 40
Rys. 21. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO , Zabrze 2011-2016.....	strona 42
Rys. 22. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO , Złoty Potok 2011-2016.....	strona 43
Rys. 23. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 44
Rys. 24. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 45
Rys. 25. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO_2 , Zabrze 2011-2016.....	strona 47
Rys. 26. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO_2 , Złoty Potok 2011-2016.....	strona 48
Rys. 27. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym w okresie wiosenno-letnim.....	strona 49
Rys. 28. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 50
Rys. 29. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO_2 , Zabrze 2011-2016.....	strona 51
Rys. 30. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO_2 , Złoty Potok 2011-2016.....	strona 52

Rys. 31. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 53
Rys. 32. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 54
Rys. 33. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń PM10, Zabrze 2011-2016.....	strona 56
Rys. 34. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń PM10, Zabrze 2011-2016.....	strona 57
Rys. 35. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 58
Rys. 36. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 59
Rys. 37. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń CO, Zabrze 2011-2016.....	strona 61
Rys. 38. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze w wybranym dniu w okresie wiosenno-letnim.....	strona 61
Rys. 39. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami RVS dla stacji Zabrze w wybranym dniu w okresie jesienno-zimowym.....	strona 62
Rys. 40. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O ₃ , Zabrze 2011-2016.....	strona 64
Rys. 41. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń O ₃ , Złoty Potok 2011-2016.....	strona 65
Rys. 42. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 66
Rys. 43. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 67
Rys. 44. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO, Zabrze 2011-2016.....	strona 69

Rys. 45. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO ₂ , Złoty Potok 2011-2016.....	strona 70
Rys. 46. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 71
Rys. 47. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 72
Rys. 48. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO ₂ , Zabrze 2011-2016.....	strona 73
Rys. 49. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń NO ₂ , Złoty Potok 2011-2016.....	strona 74
Rys. 50. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 75
Rys. 51. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 76
Rys. 52. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO ₂ , Zabrze 2011-2016.....	strona 77
Rys. 53. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń SO ₂ , Złoty Potok 2011-2016.....	strona 78
Rys. 54. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 79
Rys. 55. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 80
Rys. 56. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń PM ₁₀ , Zabrze 2011-2016.....	strona 81
Rys. 57. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń PM ₁₀ , Złoty Potok 2011-2016.....	strona 82
Rys. 58. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Złoty Potok w wybranych dniach w okresie wiosenno-letnim.....	strona 83

Rys. 59. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze i Żłoty Potok w wybranych dniach w okresie jesienno-zimowym.....	strona 84
Rys. 60. Wartości błędu MAE oraz RMSE obliczone dla wszystkich przypadków w zależności od liczby podzakresów (submodeli) wykorzystanych w procesie modelowania stężeń CO, Zabrze 2011-2016.....	strona 85
Rys. 61. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze w wybranym dniu w okresie wiosenno-letnim.....	strona 86
Rys. 62. Porównanie przebiegów dobowych stężeń rzeczywistych oraz aproksymowanych modelami PVS dla stacji Zabrze w wybranym dniu w okresie wiosenno-letnim.....	strona 86