

## WPŁYW RUCHU SAMOCHODOWEGO ORAZ WARUNKÓW METEOROLOGICZNYCH NA STĘŻENIA TLENKÓW AZOTU

*W pracy określono wpływ transportu drogowego mierzonego natężeniem ruchu samochodowego oraz warunków meteorologicznych (temperatura, kierunek i prędkość wiatru, ciśnienie powietrza, wilgotność względna) na stężenia  $\text{NO}_2$  oraz  $\text{NO}_x$  w powietrzu w kanionie komunikacyjnym. Badania przeprowadzono na podstawie godzinowych wartości zmiennych z okresu dwóch lat 2015 i 2016 w jednym punkcie Wrocławia. Określono 5 różnych podzbiorów danych wyznaczonych przez warunki klimatyczne Wrocławia (pory roku). Modelowano zależności regresyjne metodą regresji krokowej postępującej oraz metodą wzmacnianych drzew decyzyjnych. Analizowano dopasowanie modeli tworzonych dla tych podzbiorów oraz udział poszczególnych zmiennych w wyjaśnianiu zmienności stężeń zanieczyszczeń. Wykazano wpływ pory roku zarówno na dopasowanie modelu jak i na udziały poszczególnych predyktorów w wyjaśnianiu wariancji stężeń zanieczyszczeń. Najlepsze dopasowanie wykazały modele tworzone dla okresu zimowego (XII-II). Najważniejszą zmienną objaśniającą w modelowaniu stężeń tlenków azotu była prędkość wiatru oraz natężenie ruchu. Oba rozważane modele z podobną dokładnością opisują badane zależności.*

### WSTĘP

Zanieczyszczenia powietrza atmosferycznego pochodzenia antropogenicznego są od dawna faktem niezaprzecznym. Występujące w powietrzu nienaturalnie wysokie stężenia związków chemicznych i pyłów wpływa negatywnie na całe środowisko. Szczególnie w dużych aglomeracjach miejskich, gdzie liczne źródła zanieczyszczeń są różnego rodzaju, dąży się do redukcji ich emisji na wielu płaszczyznach. Jednym z bardzo istotnych źródeł zanieczyszczeń powietrza w aglomeracji miejskiej jest ruch samochodowy. Według Wojewódzkiego Inspektoratu Ochrony Środowiska we Wrocławiu 56% emisji  $\text{NO}_2$ , 44% emisji CO oraz 16% emisji  $\text{PM}_{2.5}$  pochodzi z transportu drogowego [7]. Mniejszy niż w innych krajach udział spalin samochodowych w emisji szkodliwych substancji, w szczególności  $\text{PM}_{2.5}$ , jest efektem struktury mieszkaniowej miasta. Historyczne losy miasta sprawiły, że na jego terenie znajduje się wiele stuletnich domów i kamienic do dziś opalanych paliwami stałymi (węgiel, drewno). Szacuje się, że 81% emisji  $\text{PM}_{2.5}$ , 54% CO i 9%  $\text{NO}_2$  we Wrocławiu pochodzi z sektora komunalno-bytowego. We Wrocławiu jest aktualnie zarejestrowanych 270 tys. pojazdów z czego 200 tys. stanowią samochody osobowe. Struktura drogowa w połączeniu ze znaczną liczbą pojazdów poruszających się po mieście skutkuje powstawaniem kongestii co zwiększa emisję spalin z poruszających się pojazdów. Wciąż prowadzone są działania mające na celu ograniczenie emisji powierzchniowej zarówno z sektora komunalno-bytowego jak i transportu.

W niniejszej pracy jako dane wejściowe wykorzystano informacje dotyczące warunków meteorologicznych oraz natężenia ruchu w latach 2015-2016. Zagadnienie doboru właściwego modelu do opisu zależności między stężeniem zanieczyszczeń w powietrzu i zmiennymi objaśniającymi staje się coraz trudniejsze wraz z rozwojem technik obliczeniowych i uczenia maszynowego. Popularne, najpierw liniowe, współcześnie bardziej złożone, wielowymiarowe modele regresyjne, wciąż rozwijane, skutecznie opisują te zależności. González-Aparicio i in. [6] przedstawili trzy różne modele regresyjne do opisu zależności między stężeniem  $\text{PM}_{10}$  a zmiennymi meteorologicznymi, natężeniem ruchu oraz zmiennymi czasowymi. Betraccini i in. [2] oraz Aldrin i Haff [1] zaproponowali zastosowanie uogólnionego

modelu addytywnego (GAM - generalized additive model). Wciąż rozwijane uczenie maszynowe również znalazło zastosowanie w modelowaniu stężenia zanieczyszczeń. Wzmacniane drzewa regresyjne będące jedną z metod klasyfikacyjno-regresyjnych bazujących na drzewach decyzyjnych zostały zastosowane między innymi przez Sayegh i in. [12]. Jeszcze bardziej złożoną obliczeniowo metodą jest zastosowane w pracy [9] losowy las (RF - random forest) zawierający w swej procedurze kompilację informacji z wielu drzew decyzyjnych jednocześnie. W niniejszej pracy zastosowano dwa modele regresyjne: krokowej regresji wielowymiarowej oraz wzmacnianych drzew decyzyjnych. Porównano dopasowanie każdego z modeli z podziałem na sezony klimatologiczne oraz zinterpretowano wpływ każdej ze zmiennych objaśniających na stężenia zanieczyszczeń.

Odrębnym od doboru sposobu modelowania zależności zagadnieniem jest wybór okresu czasu, dla którego wyznaczone będą współczynniki modelu, bądź budowane drzewa decyzyjne. W pierwszej chwili wydaje się, że najlepiej wykorzystać najdłuższy możliwy szereg czasowy. Jednak pewne relacje mogą zachodzić jedynie w wybranych podokresach, a zanikać w błędach losowych przy danych obejmujących pełne lata. Podział roku kalendarzowego na okresy pozwala wyeliminować z modelowania w okresie ciepłym zanieczyszczenia pochodzące z sektora komunalno-bytowego (głównie emisja niska z indywidualnych systemów grzewczych), które w okresie zimnym stanowią w modelu znaczną część błędu losowego. W pracy [2] dla modelowania stężenia  $\text{NO}_2$  oraz  $\text{PM}_{10}$  zastosowano modelowanie dla całego roku oraz z podziałem na 4 pory roku. Szczególnie dla temperatury, prędkości wiatru i ciśnienia powietrza atmosferyczne goefekty estymacji dla różnych podokresów znacząco się różniły.

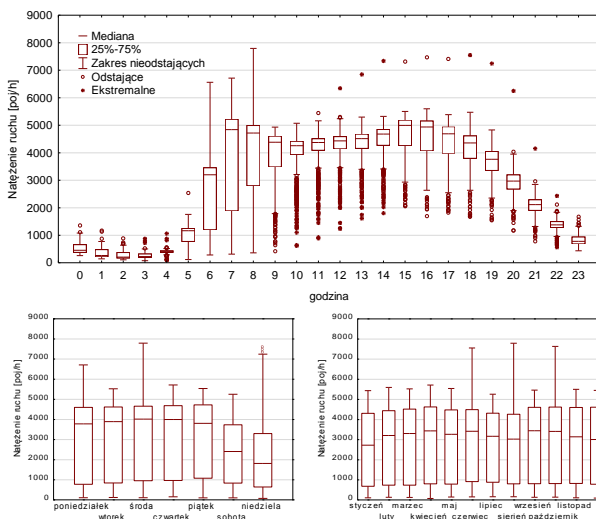
W niniejszej pracy porównano skuteczność modelowania wpływu natężenia ruchu oraz czynników meteorologicznych na stężenie  $\text{NO}_2$  oraz  $\text{NO}_x$  w kanionie komunikacyjnym w zastosowaniu dwóch metod: krokowej regresji wielorakiej oraz regresyjnych wzmacnianych drzew losowych. Modele budowano niezależnie dla każdego rodzaju zanieczyszczenia dla całego okresu badawczego 2015-2016 oraz z podziałem na 4 pory roku (łącznie 20 modeli).

## 1. CHARAKTERYSTYKA DANYCH

### 1.1. Ruch samochodowy

Dane dotyczące natężenia ruchu samochodowego zostały udostępnione przez Dział ds. Centrum Zarządzania Ruchem i Transportem Publicznym Zarządu Dróg i Utrzymania Miasta we Wrocławiu, który administruje 921 kamerami z wideodekacją firmy Autoscope. Kamery wraz z oprogramowaniem stosowane są do monitoringu ruchu miejskiego w ramach systemu ITS (Intelligent Transport System) we Wrocławiu. Jedną z informacji pozyskiwanych metodą wideodekacji w ramach ITS jest liczba pojazdów przejeżdżających przez przekrój pomiarowy na określonym pasie lub pasach ruchu. Zliczane są wszystkie pojazdy przecinające przekrój pasa ruchu (osobowe, ciężarowe, komunikacji miejskiej). Monitorowaniem objętych jest 68 skrzyżowań. Liczebności sumowane są dla 15-minutowych kwantów czasu. Na potrzeby modelowania wartości sumowano dla pełnych godzin aby odpowiadały częstotliwości pozyskiwania informacji o wysokości stężeń zanieczyszczeń. Braki danych zastępowano wartością średnią z danej godziny dla dnia tygodnia, w którym brak wystąpił. Modelowanie przeprowadzono dla obszaru skrzyżowania ulic Hallera-Powstańców Śląskich, na którym znajduje się stacja pomiaru jakości powietrza i jest ono objęte monitoringiem systemu ITS.

Poniżej (rys.1.) przedstawiono zmienność dobową, tygodniową i miesięczną natężenia ruchu na skrzyżowaniu ulic Hallera-Powstańców Śląskich we Wrocławiu.



**Rys.1.** Wykresy pudełkowe natężenia ruchu na skrzyżowaniu Hallera-Powstańców Śl. we Wrocławiu w różnej skali czasowej

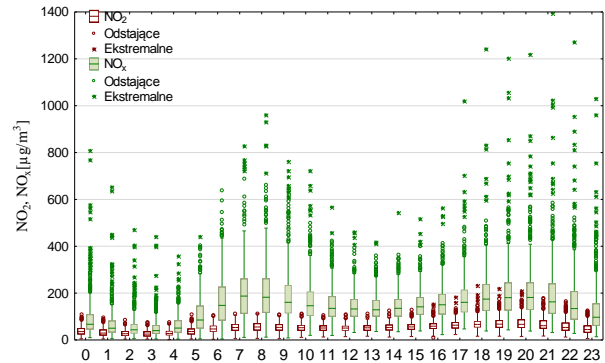
Dobowa zmienność natężenia ruchu wykazuje klasyczny charakter dwumodalny ze szczytami porannym oraz popołudniowym. Wartości ekstremalne pochodzą z jednego dnia 21 czerwca 2015. Zaznacza się mniejsze natężenie ruchu w weekendy. Nie występuje wyraźna sezonowość miesięczna.

### 1.2. Zanieczyszczenia powietrza: NO<sub>2</sub>, NO<sub>x</sub>

Dane dotyczące jakości powietrza pozyskano z Wojewódzkiego Inspektoratu Ochrony Środowiska zarządzającego pięcioma stacjami pomiarowymi z różnym zakresem mierzonych stężeń zanieczyszczeń. Wybrano stację położoną w centrum miasta w kanionie komunikacyjnym. Pozostałe stacje odpowiadają za pomiar tła.

W Polsce dopuszczalna wartość stężenia dwutlenku azotu w powietrzu, która nie może być przekroczona więcej niż 18 razy w roku, wynosi 200 µg/m<sup>3</sup>. Poziom alarmowy stężenia NO<sub>2</sub> w powietrzu wynosi 400 µg/m<sup>3</sup> utrzymujące się co najmniej 3 kolejne godziny [11]. W rozpatrywanym okresie poziom dopuszczalny został przekroczony

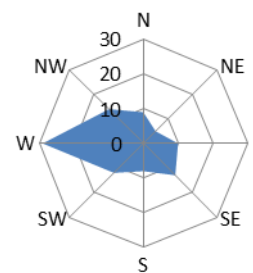
dwukrotnie, a alarmowy nie został przekroczony. Stężenie NO<sub>2</sub> wykazuje wyraźną zmienność dobową (Rys.2.). Wyższe wartości pokrywają się z okresami szczytów komunikacyjnych. Szczyt poranny stężenia tlenków azotu przypada na godziny 7-9, a popołudniowy na godziny 18-21. Zauważalne są niższe stężenia w weekendy również w związku ze zredukowanym ruchem samochodowym w tym czasie. Podobną zmienność dobową oraz tygodniową wykazują stężenia NO<sub>x</sub> przy czym jego wartości są nawet siedmiokrotnie wyższe.



**Rys.2.** Wykresy pudełkowe stężenia tlenków azotu na skrzyżowaniu Hallera-Powstańców Śl. we Wrocławiu w kolejnych godzinach doby

### 1.3. Dane meteorologiczne

Dane meteorologiczne zostały pozyskane z Instytutu Meteorologii i Gospodarki Wodnej i obejmują godzinowe wartości temperatury powietrza, prędkości wiatru, ciśnienia atmosferycznego oraz wilgotności względnej. Ze względu na położenie meteorologicznej stacji pomiarowej na obrzeżach miasta nie uwzględniono w badaniach opadu. Warunki klimatyczne Wrocławia wykazują cechy typowe dla dużych aglomeracji miejsko-przemysłowych, w których wpływ gospodarczej działalności przejawia się zmianą cech fizycznych podłoża, zanieczyszczeniem atmosfery, a także emisją ciepła sztucznego, wytworzonego w procesach komunalnych i przemysłowych. Obserwowane podwyższenie temperatury powietrza (nawet o 5°C podczas pogodnych nocy letnich), określane jest jako „miejska wyspa ciepła”. Średnia temperatura roczna we Wrocławiu wynosi 9,1°C i jest to najcieplejsze miasto w Polsce. We Wrocławiu nie występują silne wiatry. Maksymalnie zaobserwowano wiatr wiejący z prędkością 15 m/s. Średnia prędkość wiatru wynosi 3,1 m/s. Procentowy udział kierunków wiatru przedstawiono na róży wiatrów (Rys. 3.).



**Rys.3.** Róża wiatrów dla Wrocławia, okres 2015-2016

## 2. METODYKA

### 2.1. Krokowa regresja wielowymiarowa

Do modelowania zależności stężenia zanieczyszczenia w powietrzu od natężenia ruchu oraz warunków meteorologicznych zastosowano wielowymiarową regresję krokową z estymacją parametrów równania (1) metodą najmniejszych kwadratów. Umożliwia ona bowiem ocenę udziału każdej zmiennej objaśniającej w wyjaśnianiu zmienności zmiennej objaśnianej.

$$Y = a_0 + a_1X_1 + \dots + a_nX_n \quad (1)$$

$a_i$  – współczynniki modelu

$X_i$  – zmienne objaśniające  $i=1, \dots, n$

$n$  – liczba zmiennych predykcyjnych (objaśniających)

$Y$  – zmienna objaśniana – tu stężenie zanieczyszczenia

Wyznaczono modele wielowymiarowe dla każdego rodzaju zanieczyszczeń dla 5 zbiorów danych: pełne dwa lata 2016-2017 (17113 przypadków), połączone pory roku: lato (VI-VIII), jesień (IX, XI), zima (XII, I), wiosna (III, V). Dla każdego modelu wyznaczono zmienne statystycznie istotne w wyjaśnianiu zmienności stężenia zanieczyszczenia, oceniono możliwość nadmiarowości zmiennych, wyznaczono współczynniki równań i dokonano ich interpretacji.

## 2.2. Wzmacnianie drzewa regresyjnego

Główną ideą wzmacnianych drzew regresyjnych jest tworzenie ciągu prostych drzew binarnych składających się z korzenia i dwóch potomków (jeden podział), z których każde kolejne drzewo jest zbudowane do predykcji reszt generowanych przez poprzednie [4,5]. Metoda buduje drzewa binarne, czyli dzieli dane na dwie próby w każdym węźle podziału. W kolejnych krokach wzmacniania algorytmu określany jest pojedynczy (najlepszy) podział danych i obliczane są odchyłki wartości obserwowanych od średnich (reszty w każdym podziale). Zastosowana w algorytmie metoda stochastycznego wzmacniania gradientowego oznacza, że każde kolejne drzewo budowane jest na podstawie losowej próby zawierającej 50% z całego zbioru danych. Zatem kolejne drzewa tworzone są do predykcji reszt w próbach wylosowanych niezależnie. Wprowadzenie pewnego stopnia losowości do analizy ma na celu zabezpieczenie przeciwko przeuczeniu i zwraca modele, które mają własność generalizacji i mają dobrą trafność predykcyjną. Opisany algorytm daje w efekcie dobre dopasowanie wartości przewidywanych do wartości obserwowanych, nawet jeśli sama natura relacji pomiędzy predyktorami a zmienną zależną jest bardzo złożona (np. nieliniowa). Zastosowanie drzew decyzyjnych z metodą podziału C&RT wyczerpującego poszukiwania podziałów jednowymiarowych umożliwia ocenę ilościową ważności zmiennych jako sumę - po wszystkich węzłach drzewa - wzrostów oceny resubstytucji, oraz wyrażenia tej wartości jako ułamek maksymalnej sumy (po wszystkich zmiennych predykcyjnych).

Analogicznie jak w przypadku krokowej regresji wielowymiarowej dla każdego rodzaju zanieczyszczenia wyznaczono 5 modeli dla następujących zbiorów danych: pełne dwa lata 2016-2017, połączone pory roku: lato (VI-VIII), jesień (IX, XI), zima (XII, I), wiosna (III, V). Wyznaczono podstawowe miary dopasowania modelu do danych testowych stanowiących 30% próby oraz określono ważność zmiennych zgodnie z procedurą opisaną w [3]. Bardzo ważną zaletą drzew regresyjnych jest możliwość wprowadzania do zbioru zmiennych objaśniających zmiennych jakościowych. W pracy, korzystając z tej możliwości, wprowadzono dodatkowe zmienne w porównaniu z regresją krokową: kierunek wiatru (N, NE, E, itp.), miesiąc, dzień tygodnia oraz pracujący/wolny dzień.

## 2.3. Ocena modeli

Do oceny dopasowania modeli posłużono się popularnym współczynnikiem determinacji wskazującym część wariacji zmiennej objaśnianej wyjaśnioną przez model (2).

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (2)$$

gdzie:  $N$  – liczebność próby

Zastosowanie regresji krokowej pozwoliło ocenić przyrost współczynnika determinacji po dodaniu do modelu kolejnej zmiennej, a zatem określić część wariacji wyjaśnianej przez określoną zmienną

objaśniającą. Do porównania dwóch zastosowanych metod regresyjnych zastosowano dodatkowo Bayesian Information Criterion (BIC) według wzoru (3).

$$BIC = N \cdot \ln\left(\frac{RSS}{N}\right) + k \cdot \ln(N) \quad (3)$$

gdzie:  $k$  – liczba parametrów modelu,

$RSS$  – suma kwadratów reszt.

Kryterium BIC różnicuje ocenę dopasowania modelu w zależności od liczby jego parametrów. Wzrost liczby parametrów poprawia jego dopasowanie, ale zwiększa stopień skomplikowania modelu i ryzyko jego przeuczenia. Kryterium BIC „karze” za dodanie każdej kolejnej zmiennej zwiększeniem wartości o logarytm naturalny z liczebności próby. Wartości wskaźników dopasowania modeli przedstawiono w tabelach 1-4.

## 3. WYNIKI I DISKUSJA

### 3.1. Krokowa regresja wielowymiarowa

Wartości współczynników modeli wraz z oceną dopasowania dla każdego okresu przedstawiono w tabeli 1.

Tab.1. Współczynniki regresji modeli dla stężeń  $NO_2$

	2015-2016	lato	jesień	zima	wiosna
Wyraz wolny	167,7	223,8	208,0	58,34	318,02
Natężenie ruchu	0,01 (23)	0,005 (15)	0,01 (24)	0,01 (36)	0,01 (19)
Prędkość wiatru	-4,73 (13)	-4,94 (8)	-5,1 (12)	-3,89 (14)	-5,97 (18)
Temperatura powietrza	-	0,99 (4)	0,12 (<1)	-0,18 (<1)	-0,19 (<1)
Ciśnienie atmosferyczne	-0,11 (<1)	-0,19 (<1)	-0,14 (<1)	-	-0,27 (1)
Wilgotność względna	-0,11 (1)	-	-0,26 (3)	-0,21 (1)	-
$R^2$	0,36	0,26	0,39	0,5	0,38
BIC	5429	6063	6141	5432	6050
N	17113	4293	4344	4305	4171

W nawiasach umieszczono procentowy udział zmiennej w wyjaśnianiu wariacji  $NO_2$

Największą część zmienności stężenia  $NO_2$  w powietrzu (do 36%) wyjaśnia natężenie ruchu, czyli jest głównym czynnikiem modelującym stężenie tlenu azotu w powietrzu w aglomeracji miejskiej. Drugą, co do ważności, zmienną jest prędkość wiatru, jednak jej wpływ jest o połowę mniejszy. Zdecydowanie najmniejsze znaczenie ma temperatura powietrza, ciśnienie atmosferyczne oraz wilgotność powietrza, których udział w wyjaśnianiu wariacji jest znikomy, jednak statystycznie istotny. Żadna z rozważanych zmiennych nie wykazała cech nadmiarowości, zatem modele są poprawnie zbudowane. Wzrost natężenia ruchu powoduje wzrost stężenia  $NO_2$  o około  $10 \mu g/m^3$  na każde 1000 pojazdów. Zatem ruch samochodowy odpowiada za ładunek nawet za  $80 \mu g/m^3$   $NO_2$  w powietrzu w godzinach szczytu komunikacyjnego w bezpośrednim sąsiedztwie skrzyżowania. Ujemny współczynnik kierunkowy przy prędkości wiatru w każdym modelu oznacza, że wraz ze wzrostem prędkości wiatru stężenie  $NO_2$  maleje. Zjawisko to jest nazywane efektem przewietrzania miasta. 25 lutego 2014 Sejmik Województwa Dolnośląskiego przyjął uchwałę „Program ochrony powietrza dla województwa dolnośląskiego” [10], w związku z przekroczeniem poziomów dopuszczalnych pyłu zawieszonego  $PM_{10}$ , pyłu zawieszonego  $PM_{2,5}$ , dwutlenku azotu oraz poziomów docelowych benzo(a)pirenu i ozonu w powietrzu w 2011 r. Jedno z działań ujętych w tym programie brzmi: „stosowanie odpowiednich zapisów w planach zagospodarowania przestrzennego, umożliwiających ograniczenie stężeń dwutlenku azotu oraz

pyłu zawieszonego PM<sub>10</sub>, PM<sub>2,5</sub> i B(a)P, w miejscowych planach zagospodarowania przestrzennego dotyczących np. układu zabudowy zapewniającego przewietrzanie miasta (...). Od wprowadzenia tego planu upłynęły 3 lata. Jakość powietrza we Wrocławiu w tym czasie uległa poprawie [8] jednak głównie dzięki redukcji liczby indywidualnie wykorzystywanych pieców węglowych oraz wzrostowi świadomości obywateli. Właściwe planowanie przestrzenne zabudowy miast mogłoby wzmocnić ten efekt, niestety decyzje planistyczne w miastach są najczęściej dyktowane maksymalizacją zysku głównie finansowego. Najlepsze dopasowanie do danych empirycznych wykazał model zbudowany dla okresu zimowego, a najslabsze dla okresu letniego.

**Tab.2. Współczynniki regresji modeli dla stężeń NO<sub>x</sub>**

	2015-2016	lato	jesień	zima	wiosna
Wyraz wolny	-33,1*	24,21	131,6	26,24	881,35
Natężenie ruchu	0,04 (18)	0,02 (17)	0,04 (17)	0,04 (29)	0,03 (17)
Prędkość wiatru	-16,31 (10)	-11,02 (8)	-21,69 (12)	-17,09 (14)	-17,64 (13)
Temperatura powietrza	-2,42 (7)	0,68 (<)	-3,64 (4)	-0,87 (<1)	-2,59 (3)
Ciśnienie atmosferyczne	0,09* (<1)	-	-	-	-0,79 (<1)
Wilgotność względna	0,93 (1)	0,70 (2)	0,40 (<1)	1,20 (1)	0,55 (1)
R <sup>2</sup>	0,372	0,270	0,325	0,440	0,361
BIC	8476	8049	8978	8500	8952

W nawiasach umieszczono procentowy udział zmiennej w wyjaśnieniu wariancji NO<sub>x</sub>, \* - nieistotne statystycznie na poziomie  $\alpha=0,05$

W modelowaniu stężenia NO<sub>x</sub> występują podobne zależności (tab.2.) jak dla samych NO<sub>2</sub>. Największą część zmienności wyjaśnia natężenie ruchu – do 29% zatem nieznacznie mniej niż dla dwutlenków azotu. Następną w kolejności ważności zmienną jest prędkość wiatru wyjaśniająca do 14% zmienności stężenia NO<sub>x</sub>. Pozostałe zmienne meteorologiczne choć w większości statystycznie istotne, nie wnoszą znaczącego wkładu w wyjaśnianie zmienności NO<sub>x</sub>. Dopasowanie modelu jest nieco słabsze niż dla dwutlenków azotu. Podobnie najlepsze dopasowanie otrzymano dla okresu zimowego.

### 3.2. Wzmacniane drzewa regresyjne

Stosując metodę przedstawioną w punkcie 2.2. wyznaczono 10 modeli określających zależności między stężeniami NO<sub>2</sub> oraz NO<sub>x</sub> w powietrzu dla różnych okresów. Każdorazowo określono ważność zmiennych. Wyniki zestawiono w tabelach 3 i 4.

**Tab.3. Ważność zmiennych w modelach dla stężeń NO<sub>2</sub>**

	2015-2016	lato	jesień	zima	wiosna
Prędkość wiatru	100	100	100	91	100
Natężenie ruchu	88	75	92	100	85
Dzień wolny/pracujący	87	81	73	65	84
Kierunek wiatru	82	94	66	68	80
Dzień tygodnia	70	80	67	58	76
Miesiąc	68	48	66	40	23
Wilgotność względna	57	68	83	43	45
Temperatura powietrza	56	80	90	59	28
Ciśnienie atmosferyczne	17	46	40	53	34
R <sup>2</sup>	0,514	0,474	0,484	0,600	0,572
BIC	5639	5852	6019	5337	5583

Podobnie jak w przypadku modelowania za pomocą regresji wielorakiej najlepsze dopasowanie modelu otrzymano dla okresu zimowego, następnie dla wiosny i pełnych dwóch lat. Najważniejszą zmienną okazała się prędkość wiatru, z wyjątkiem okresu zimowego, kiedy rolę tę przejęła zmienna „natężenie ruchu”. Bardzo ważnymi zmiennymi są również „kierunek wiatru” oraz „dzień wolny/pracujący”. Ważność pozostałych zmiennych jest różna, dla różnych pór roku. Czynniki meteorologiczne: wilgotność względna oraz temperatura powietrza są ważniejsze latem i jesienią niż zimą i wiosną. Najmniej istotną zmienną jest ciśnienie powietrza atmosferycznego.

**Tab.4. Ważność zmiennych w modelach dla stężeń NO<sub>x</sub>**

	2015-2016	lato	jesień	zima	wiosna
Prędkość wiatru	100	100	100	100	100
Natężenie ruchu	81	75	85	97	83
Dzień wolny/pracujący	92	96	81	66	89
Kierunek wiatru	83	92	62	76	69
Dzień tygodnia	76	88	77	64	70
Miesiąc	94	26	40	28	67
Wilgotność względna	91	91	72	73	76
Temperatura powietrza	87	34	87	63	64
Ciśnienie atmosferyczne	56	40	57	66	34
R <sup>2</sup>	0,482	0,468	0,388	0,551	0,540
BIC	8698	7915	9536	8837	8320

Dla modelowania stężeń NO<sub>x</sub> w każdym okresie najważniejszą zmienną jest prędkość wiatru następnie „natężenie ruchu” oraz „dzień wolny/pracujący”. Podobnie jak w modelowaniu stężeń NO<sub>2</sub> najmniejsze wartości ważności otrzymano dla ciśnienia atmosferycznego. Współczynniki determinacji są nieco niższe niż dla modeli wykonanych dla dwutlenku azotu. Fakt ten jest związany ze znacznie większą zmiennością stężeń NO<sub>x</sub> w porównaniu ze stężeniami NO<sub>2</sub> (współczynnik zmienności 73 i 46% odpowiednio).

## PODSUMOWANIE

W pracy przedstawiono wyniki badania zależności stężenia NO<sub>x</sub> oraz NO<sub>2</sub> od natężenia ruchu, czynników meteorologicznych oraz jakościowych zmiennych czasowych. Wykorzystano dwa zupełnie różne w swej idei sposoby modelowania zależności: regresję liniową wieloraką krokową oraz wzmacniane drzewa decyzyjne. Największy wpływ na stężenia obu rodzajów zanieczyszczeń w powietrzu ma natężenie ruchu oraz prędkość wiatru. Regresja krokowa wskazuje natężenie ruchu jako zmienną wyjaśniającą największą część zmienności, natomiast metoda drzew wzmacnianych największą ważność przypisuje prędkości wiatru. Jako drugie co do ważności zmienne występują te same naprzemiennie. Do porównania modeli między sobą zastosowano kryterium Bayesa BIC. Wartości tego kryterium są dla obu rodzajów modeli mniejsze przy modelowaniu stężeń NO<sub>2</sub>, a większe dla modelowania stężeń NO<sub>x</sub>. Kryterium BIC zależy nie tylko od wartości reszt (różnica między wartością empiryczną, a modelową) ale również od liczby zmiennych w modelu. W modelach z zastosowaniem drzew wzmacnianych użyto 9 zmiennych podczas gdy w modelach regresyjnych liniowych 5 lub 6. W okresach letnich oraz wiosennych lepsze dopasowanie mierzone wskaźnikiem BIC otrzymano

w modelach drzew wzmacnianych. Dla całego okresu 2015-2016 lepsze dopasowanie wykazał model regresji liniowej wielowymiarowej. Dla okresów jesiennego oraz zimowego preferencje wskazywane przez BIC są rozbieżne dla różnych zanieczyszczeń. Trudno zatem jednoznacznie rozstrzygnąć o preferowaniu jednego z opisywanych modeli. Wyższe wskaźniki współczynnika determinacji dla modeli drzew wzmacnianych wynikają z większej ilości parametrów tych modeli. Koszt zwiększenia liczby parametrów jest jednak porównywalny ze wzrostem dobroci dopasowania dzięki nim otrzymanej.

## BIBLIOGRAFIA

1. Aldrin M., Haff I.H., 2005. *Generalized additive modelling of air pollution, traffic volume and meteorology*. Atmos. Environ. Vol. 39 is.11 p.2145-2155.
2. Bertaccini P., Dukic V., Ignaccolo R., *Modeling the Short-Term Effect of Traffic and Meteorology on Air Pollution in Turin with Generalized Additive Models*. Advances in Meteorology 2012 p.1-16.
3. Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J., *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software 1984.
4. Friedman, J. H., *Greedy function approximation: A gradient boosting machine*. IMS 1999 Reitz Lecture. 1999a.
5. Friedman, J. H., *Stochastic gradient boosting*. Stanford University 1999b.
6. González-Aparicio I., Hidalgo J., Baklanov A., Padró A., Santa-Coloma O., *An hourly PM10 diagnosis model for the Bilbao metropolitan area using a linear regression methodology*. Environ. Sci. Pollut. Res., 2013. 20(7), p.4469-4483.
7. *Informacja o jakości powietrza na terenie miasta Wrocławia*, Wojewódzki Inspektorat Ochrony Środowiska, Wrocław 2016.
8. Kamińska J.A., *Zjawisko smogu na tle jakości powietrza we Wrocławiu w latach 2012-2016*. Inżynieria Ekologiczna, Ecological Engineering. 2017, vol.18 Iss.5 p.66-76
9. Laña I., Del Ser J., Pedró A., Vélez M., Casanova-Mateo C., *The role of local urban traffic and meteorological conditions in air pollution: A data-based study in Madrid, Spain*. Atmos. Environ. 2016, 145, p.424-438.
10. *Program ochrony powietrza dla województwa dolnośląskiego*, uchwała nr XLVI/1544/14 Sejmiku Województwa Dolnośląskiego z dnia 25 lutego 2014 r.
11. *Rozporządzenie Ministra Środowiska z dnia 24 sierpnia 2012 r. w sprawie poziomów niektórych substancji w powietrzu* (Dz.U.2012.1031)
12. Sayegh, A., Tate, J.A., Ropkins, K., *Understanding how roadside concentrations of NO<sub>x</sub> are influenced by the background levels, traffic density, and meteorological conditions using Boosted Regression Trees*. Atmos. Environ. 2016, 127, p.163-175.

## The impact of traffic flow and meteorological conditions on nitrogen oxides concentration

*The influence of road transport, measured by traffic intensity, and meteorological conditions (temperature, direction and velocity, air pressure, relative humidity) on NO<sub>2</sub> and NO<sub>x</sub> concentrations in air in the communication canal were determined. The research was conducted on the basis of hourly values of variables from the period of two years 2015, 2016 at one place in Wrocław. There were 5 different time subdivisions designated by Wrocław's climatic conditions (seasons). Regression was modeled using the stepwise regression method and the method of boosted decision trees. The fit of models created for these subsets and the share of individual variables in explaining variations in pollutant concentrations were analyzed. The influence of the seasons on the fit of the model and on the shares of the individual predictors in explaining the variance of pollutant concentrations was shown. The best fit was the models created for the winter (XII-II). The most important explanatory variable in the modeling of nitrogen oxide concentrations was wind speed and traffic intensity. Both models with similar accuracy describe the studied relationships.*

Autorzy:

dr inż. **Joanna A. Kamińska** – Uniwersytet Przyrodniczy we Wrocławiu, Katedra Matematyki, joanna.kaminska@upwr.edu.pl  
dr **Mieczysław Chalfen** – Uniwersytet Przyrodniczy we Wrocławiu, Katedra Matematyki, mieczyslaw.chalfen@upwr.edu.pl  
dr inż. **Bożena Szczucka-Lasota** – Wyższa Szkoła Zarządzania Ochroną Pracy w Katowicach, bszczucka-lasota@wszop.edu.pl