

**Artur GRAMACKI, Jarosław GRAMACKI**  
 UNIwersytet Zielonogórski, Instytut Informatyki i Elektroniki

## Szacowanie emisji tlenków azotu (NO<sub>x</sub>) na podstawie danych eksploatacyjnych rzeczywistego obiektu przemysłowego

Dr inż. Artur GRAMACKI

Pracuje w Instytucie Informatyki i Elektroniki Uniwersytetu Zielonogórskiego na stanowisku adiunkta. Jego zainteresowania koncentrują się wokół zagadnień związanych z bazami danych oraz eksploatacją danych.



e-mail: a.gramacki@iie.uz.zgora.pl

Dr inż. Jarosław GRAMACKI

Pracuje w Instytucie Informatyki i Elektroniki Uniwersytetu Zielonogórskiego na stanowisku adiunkta. Jego zainteresowania koncentrują się wokół zagadnień związanych z bazami danych oraz eksploatacją danych.



e-mail: j.gramacki@iie.uz.zgora.pl

### Streszczenie

W pracy podjęto próbę szacowania emisji tlenków azotu (NO<sub>x</sub>) na podstawie danych eksploatacyjnych zbieranych podczas normalnej pracy Elektrociepłowni w Zielonej Górze. Pomimo że zanieczyszczenia te monitorowane są z użyciem specjalistycznych czujników, to wydaje się, że dodatkowa ich kontrola może być bardzo pożądana. Zwłaszcza, że ilość emitowanych zanieczyszczeń (w tym NO<sub>x</sub>) jest ściśle limitowana i kontrolowana. Wspomniane szacowanie wykonano wykorzystując technikę statystyczną zwaną regresją liniową.

**Słowa kluczowe:** NO<sub>x</sub>, elektrociepłownia Zielona Góra, regresja liniowa, szacowanie.

### Estimation of nitrogen oxides NO<sub>x</sub> emission based on data from real industry plant

#### Abstract

In the paper there is analysed the possibility of estimating nitrogen oxides NO<sub>x</sub> emissions based on the data taken from an industry plant (gas-based combined heat and power (CHP) plant in Zielona Góra, Poland). Even though the emission of NO<sub>x</sub> is monitored on-line by dedicated sensors installed on the plant, we believe that some additional monitoring would be desirable. This is especially important because emission of different kind of air-pollutants (including NO<sub>x</sub>) is strictly limited and monitored. The above mentioned estimation is performed by the statistical technique known as linear regression. In Section 2 the gas-based combined heat and power plant in Zielona Góra is briefly described. A simplified chart of the plant is shown in Fig. 1. The list of measurement points taken from the DCS system (distributed control system) of the greatest influence on the NO<sub>x</sub> emission is also given. In Section 3 there is shown the way how the raw data should be prepared, especially in the context of removing unwanted disturbances as well as the points which were selected for the final regression analysis. The results obtained are presented in Section 4. The best regression solution (row 3 in Table 3) is shown in Fig. 4. The accuracy is quite satisfactory and we believe that it is sufficient for the purpose of additional monitoring the main NO<sub>x</sub> sensors installed on the plant.

**Keywords:** NO<sub>x</sub> emission, combined heat and power plant, CHP, linear regression.

### 1. Wstęp

Problem emisji zanieczyszczeń oraz ich właściwe monitorowanie jest obecnie bardzo ważnym praktycznym zagadnieniem. Ma to szczególne znaczenie w elektrowniach i elektrociepłowniach, które w wyniku spalania paliw kopalnych, gazu, biopaliw oraz innych źródeł energii, emitują tych zanieczyszczeń często ogromne ilości. Każdy tego typu obiekt przemysłowy ma obowiązek rejestrować emisję generowanych zanieczyszczeń, gdyż odeszły (lub odchodzą) już czasy, gdy można było dowolnie i praktycznie poza wszelką kontrolą je produkować. Obecnie coraz intensywniej rozwija się rynek zanieczyszczeń, którymi można (lub będzie można w przyszłości) handlować na podobieństwo handlu akcjami giełdowymi.

Rejestrowanie wspomnianych zanieczyszczeń wykonuje się z użyciem specjalizowanych czujników i rejestratorów. Żadne urządzenie nie jest jednak w 100% bezawaryjne i każda dodatkowa kontrola wydaje się być jak najbardziej pożądana.

W pracy zbadano, czy możliwe jest na akceptowalnym poziomie dokładności szacowanie zanieczyszczeń tlenkami azotu (zwanych potocznie NO<sub>x</sub>) na podstawie wybranych danych eksploatacyjnych zbieranych podczas normalnej pracy Elektrociepłowni w Zielonej Górze (ECZG) [6]. ECZG jest bardzo nowoczesnym obiektem, z mocno rozbudowaną infrastrukturą kontrolno-pomiarową w postaci tzw. systemu DCS (ang. *distributed control system*). Na bieżąco rejestrowane są wielkości w kilku tysiącach zmiennych eksploatacyjnych, a zebrane dane zapisywane są w wydajnych systemach bazodanowych.

Oszacowane wielkości NO<sub>x</sub> mogą być porównywane z tymi, które są rejestrowanymi za pomocą czujników i tym samym możemy otrzymać prosty i tani sposób ich kontroli i nadzoru (gdy dane z czujników będą bardzo różniły się od danych oszacowanych może to być znakiem jakiejś awarii).

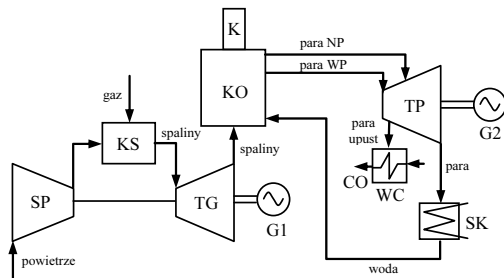
### 2. Elektrociepłownia Zielona Góra

Elektrociepłownia w Zielonej Górze powstała w roku 1974. Aż do roku 2004 jedynym wykorzystywanym rodzajem paliwa był węgiel kamienny. W roku 2004 wybudowano nowoczesny blok gazowo-parowy, który opalany jest wyłącznie krajowym gazem ziemnym (z pobliskiej kopalni w Kościanie). Od tego momentu blok węglowy stał się blokiem pomocniczym i rezerwowym zarazem, gdyż jest uruchamiany tylko w krótkich okresach remontowych bloku gazowo-parowego oraz czasami w okresach zimowych (przy niskich temperaturach), aby wspomagać blok gazowo-parowy w produkcji głównie energii cieplnej.

ECZG ma dość typową konstrukcję i wytwarza energię w tzw. skojarzeniu [3, 4]. Oznacza to, że produkowana jest zarówno energia cieplna jak i elektryczna. Przy czym im więcej produkowanej jest tej pierwszej, tym mniej tej drugiej i odwrotnie. W CZG zamontowano dwie turbiny: gazową oraz parową. Schemat obiektu, w dużym uproszczeniu, pokazano na rysunku 1.

Turbina gazowa TG opalana jest gazem ziemnym. Gaz spalany jest w komorze spalania KS, w obecności powietrza sprężonego w sprężarce SP. Powstałe w KS spaliny kierowane są na łopatki turbiny, która sprężona jest na jednym wale z generatorem (G1) produkującą energię elektryczną. Gazy spalinowe po opuszczeniu TG mają jednak ciągle ogromną energię i temperaturę, która może (i powinna) być wykorzystana. Dlatego też spaliny kierowane są do tzw. kotła odzysknicowego KO (kocioł dwuciśnieniowy), w którym, wykorzystując energię spalin, wytwarzana jest para wodna o wysokich parametrach (temperatura i ciśnienie). Dalej spaliny kierowane są do komina K. Tak wytworzona para wodna jest kierowana do turbiny parowej TP na jej łopatki. Para, rozprężając się w TP, uprawia ją w ruch i napędza sprężony z nią na jednym wale generator G2 produkujący energię elektryczną. Para

po wykonaniu pracy w TP kierowana jest do skraplacza SK, gdzie jest skraplana i w postaci tzw. kondensatu trafia z powrotem do kotła odzysknicowego KO. Część pary wodnej pobierana jest z upustu do wymienników ciepła WC i w postaci gorącej wody kierowana do odbiorców CO wykorzystujących ją do celów grzewczych.



Rys. 1. Uproszczony schemat elektrociepłowni w Zielonej Górze  
Fig. 1. A simplified layout of the CHP plant in Zielona Góra

Jedynym źródłem emisji tlenków azotu NOx jest proces spalania gazu w TG. Od strony chemicznej jest to proces bardzo złożony i nie do końca jeszcze poznany. Na wielkość emisji ma wpływ wiele parametrów, takich jak ilość spalanej paliwa, jego jakość, temperatura spalania, warunki otoczenia i inne. Wiele z nich jest rejestrowanych w systemie DCS [7]. Po konsultacjach z pracownikami ECZG dokonano wyboru tych zmiennych eksploatacyjnych, które mają największy wpływ na wielkość emisji NOx i w eksperymentach numerycznych posługiwano się tylko nimi.

Tab. 1. Wybrane zmienne eksploatacyjne z systemu DCS  
Tab. 1. Selected operational variables from the DCS system

Symbol	Nazwa
wilg	Wilgotność powietrza (°C)
T-zewn	Temperatura zewnętrzna (°C)
T-pks	Temperatura sprężonego powietrza podawanego do komory spalania turbiny gazowej (°C)
F-spaliny	Przepływ (ilość) spalin (Nm <sup>3</sup> /h, tzw. normalne m <sup>3</sup> ) opuszczających turbinę gazową
T-spp	Temperatura spalin przed tzw. przegrzewaczem wysokiego ciśnienia w kotle odzysknicowym KO (°C). De facto jest to też temperatura spalin opuszczających turbinę gazową.
T-sk	Temperatura spalin w kominie (°C)
Q-gazu	Przepływ (ilość) gazu (m <sup>3</sup> /h)
P-TG	Moc elektryczna turbiny gazowej (MW)
nox	Wielkość emisji tlenków azotu (mg/Nm <sup>3</sup> , , tzw. normalne m <sup>3</sup> )

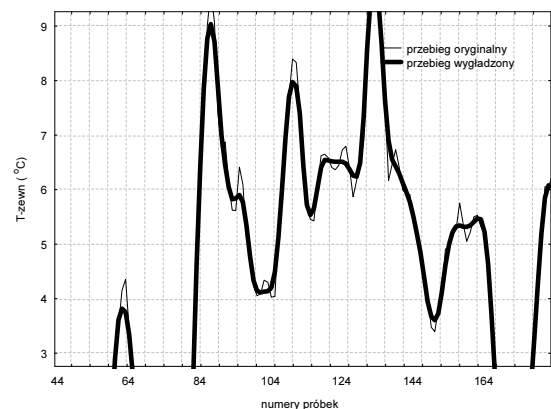
### 3. Wstępne przygotowanie danych

Szacowania wielkości emisji NOx dokonano wykorzystując jedną z podstawowych technik statystycznych a mianowicie regresję liniową. Jest ona bardzo obszernie i dokładnie opracowana. Jej opis można znaleźć w ogromnej ilości pozycji literaturowych. Jako przykład podajmy [1, 2]. Jest to metoda estymowania wartości oczekiwanej zmiennej  $y$  przy znanych wartościach innej zmiennej  $x$  lub zmiennych  $x_k$ . Szukana zmienna  $y$  jest tradycyjnie nazywana zmienną objaśnianą, lub zależną. Zmienne  $x_k$  nazywa się zmiennymi objaśniającymi lub niezależnymi.

Do badań wybrano zmienne zestawione w tabeli 1. Dane zebrano dla okresu od 01-02-2008 do 28-02-2008. Okres próbkowania ustalono na 1 godzinę (ze względu na bardzo dużą inercję obiektu ECZG wydaje się, że częstsze próbkowanie jest niecelowe). Daje to w sumie 648 próbek pomiarowych.

Zanim rozpoczniemy faktyczną analizę danych należy poddać je oczyszczeniu. Powinny zostać usunięte tzw. punkty oddalone, które są najczęściej wynikiem błędów pomiarowych, uszkodzeń czujników itp. Analiza przebiegów czasowych zmiennych eksploatacyjnych z tabeli 1 pokazała jednak, że w wybranym okresie nie występują znaczące punkty oddalone. W dalszej kolejności należy postarać się o usunięcie przypadkowych błędów pomiarowych,

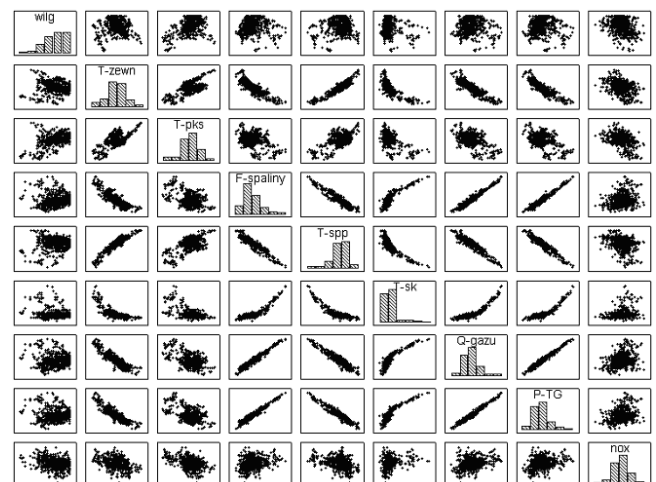
które zwykle objawiają się jako „poszarpania” na wykresach. Błędy te mają zwykle charakter losowy i można je zaniedbać. Na rysunku 2 pokazano fragment przebiegu zmiennej eksploatacyjnej  $T_{zewn}$ . Linia cieniową pokazano oryginalny przebieg, natomiast linią pogrubioną przebieg wygładzony, który w efekcie pozbawiony został błędów przypadkowych. Do wygładzenia zastosowano dość mało chyba znany filtr 4253H [5]. Zachowuje on bardzo dobrze charakter oryginalnego przebiegu. Niwelowane są przypadkowe oscylacje, co bardzo dobrze widać na rysunku. Jego działanie opiera się na pomysłowym kilkukrotnym zastosowaniu wygładzania z użyciem średniej ruchomej oraz mediany ruchomej z różnymi wielkościami okna oraz użycia tzw. filtra Hanninga (4, 2, 5, 3 oraz H; stad nazwa metody).



Rys. 2. Efekt działania filtra wygładzającego 4253H  
Fig. 2. Effect of using 4253H smoothing filter

Aby wykonać poprawnie analizę regresji muszą być spełnione pewne warunki. Należy w tym miejscu zauważyć, że nawet dla zupełnie przypadkowych danych równanie regresji zawsze można wyznaczyć, choć oczywiście nie będzie miało ono żadnej wartości. W literaturze chętnie podawany jest przykład szacowania ilości urodzeń w danym roku (zmienna objaśniana  $y$ ) od wielkości populacji bocianów na danym terenie (zmienna objaśniająca  $x$ ). Przy sprzyjającym układzie danych można liczyć na otrzymanie wyników świadczących o silnej zależności, czego oczywiście nikt nie uzna za wynik racjonalny.

Pierwszy podstawowy warunek, jaki musi być spełniony to wybór takich zmiennych objaśniających  $x_k$ , które są możliwie jak najbardziej skorelowane ze zmienną objaśnianą  $y$ , a z drugiej strony maksymalnie słabo skorelowane ze sobą nawzajem. Bardzo pomocne może być tutaj wykonanie tzw. wykresu rozrzutu, gdzie wyraźnie widać zależności między danymi. Na rysunku 3 pokazano taki wykres dla danych z tabeli 1.



Rys. 3. Wykres rozrzutu wybranych zmiennych eksploatacyjnych  
Fig. 3. Scatter plot of selected operational variables

Widać na nim, że pewne zmienne są bardzo silnie ze sobą skorelowane ( $T\text{-zewn}$ ,  $F\text{-spaliny}$ ,  $T\text{-spp}$ ,  $T\text{-sk}$ ,  $Q\text{-gazu}$ ,  $P\text{-TG}$ ). Nie ma przy tym znaczenia, że w niektórych przypadkach korelacja jest dodatnia a w niektórych ujemna. Wydaje się więc, że można wybrać tylko jedną z nich. Ale którą? Tu dość oczywistym kryterium będzie ocena współczynnika korelacji i wybór tej zmiennej, która najlepiej koreluje ze zmienną objaśnianą ( $nox$ , tabela 1). Ponadto, aby poprawić jakość wyników regresji, należałoby wybrać dodatkowe zmienne spośród jeszcze niewybranych. Ponieważ pozostały jeszcze tylko 2 zmienne ( $wilg$  oraz  $T\text{-pks}$ ), więc je wybieramy.

Ponadto analizując wielkości współczynnika korelacji między zmienną objaśnianą  $nox$  a pozostałymi zmiennymi widać, że omawiane powyżej wygładzanie przebiegów w każdym przypadku dość znacznie go poprawiło. Pokazano to w tabeli 2.

Tab. 2. Porównanie współczynników korelacji między zmiennymi objaśniającymi a zmienną objaśnianą ( $NO_x$ ) dla przebiegów bez oraz z wygładzaniem za pomocą filtra 4253H

Tab. 2. Comparison of correlation coefficients between independent variables and dependent variable ( $NO_x$ ) for both not smoothed and smoothed curves after using 4253H filter

	wilg	T-zewn	T-pks	F-spaliny	T-spp	T-sk	Q-gazu	P-TG
Bez wygl.	-0,31	-0,38	-0,40	0,27	-0,39	0,28	0,32	0,34
z wygl.	-0,35	-0,46	-0,48	0,34	-0,46	0,35	0,39	0,41

Po analizie danych zdecydowano się więc wybrać 3 zmienne, które będą zmiennymi objaśniającymi w równaniu regresji:  $T\text{-pks}$ ,  $T\text{-spp}$ ,  $wilg$ .

## 4. Wyniki eksperymentów

Wykonano szereg eksperymentów numerycznych. Wybrane wyniki zamieszczono w tabeli 3. Potwierdzono, że zastosowanie wstępnego wygładzania danych wejściowych poprawia dokładność uzyskanego modelu (mierzonego współczynnikiem determinacji  $r^2$  oraz błędem standardowym estymacji  $s$ ).

Regresja liniowa w czystej postaci zakłada liniową zależność między danymi. Analiza wykresu rozrzutu z rysunku 3 raczej nie pozwala mówić w omawianym przypadku o liniowej zależności między danymi. Należałoby więc dokonać pewnych transformacji danych w celu osiągnięcia tej liniowości [1, 2]. Najczęściej dokonuje się tego w ten sposób, że zmienne objaśnianą i/lub objaśniającą(e) poddaje się transformacji za pomocą takich funkcji jak: logarytm, pierwiastek, odwrotność, potęgowanie. Wówczas zależność nieliniowa ma szansę „uliniowić się”.

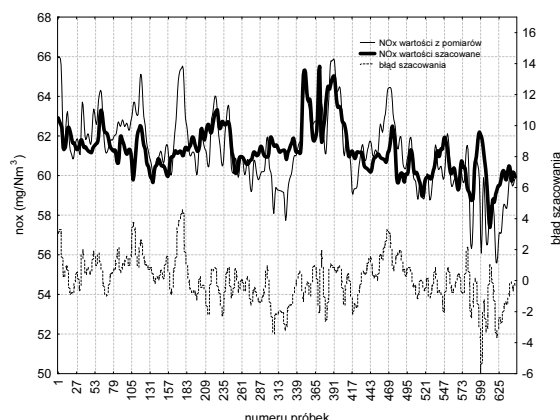
Tab. 3. Wyniki analizy regresji  
Tab. 3. Results of regression analysis

Równanie regresji	$r^2$	$s$	Uwagi
$nox = 177,92 - 0,036 * wilg - 0,055 * T\text{-pks} - 0,17 * T\text{-spp}$	0,229	1,919	bez wygładzania
$nox = 183,012 - 0,035 * wilg - 0,055 * T\text{-pks} - 0,18 * T\text{-spp}$	0,308	1,583	z wygładzaniem
$nox = 5400170 - 932,80 * \ln(T\text{-pks}) - 1,89 * (T\text{-spp})^2 + 0,0071 * (wilg)^2 + 4124,54 * T\text{-spp} - 1123870 * \ln(T\text{-spp}) + 0,0043 * (T\text{-pks})^2 + 68,3658 * \ln(wilg) - 2,06552 * wilg$	0,424	1,450	z wygładzaniem, dobór zmiennych metodą regresji krokowej

Stosując transformacje, przy obecności nawet tylko 3 zmiennych objaśniających, mamy wiele różnych wariantów doboru zmiennych (każda zmienna może być lub też nie poddana różnym transformacjom). W takich sytuacjach dobrze jest zastosować

jedną z dostępnych metod automatycznego wyboru najwłaściwszej transformacji i najwłaściwszych zmiennych objaśniających. Mamy dostępne takie metody jak: dołączania (ang. *forward selection procedure*), eliminacji (ang. *backward elimination procedure*) oraz krokowa (ang. *stepwise procedure*) [1, 2].

Na rysunku 4 pokazano przebieg oryginalnej zmiennej  $nox$  (linia cienioka) oraz przebieg wynikający z równania regresji z tabeli 3, wiersz 3 (linia gruba). Widać dość dobre dopasowanie obu krzywych, choć należy podkreślić, że nie jest ono idealne. Wynika to oczywiście z tego, że prosta regresja liniowa ma jednak swoje ograniczenia i nie jest możliwe dowolne zwiększanie dokładności.



Rys. 4. Przebieg oryginalnej zmiennej  $nox$  oraz jej estymacja  
Fig. 4. Original values of  $NO_x$  and their estimates

## 5. Wnioski

W artykule pokazano, że jest możliwe zastosowanie prostej regresji liniowej do szacowania emisji  $NO_x$  w wybranym rzeczywistym zakładzie przemysłowym. Oczywiście od regresji liniowej nie możemy spodziewać się rewelacyjnych wyników, gdy dane wejściowe mają ewidentnie silnie nieliniowy charakter. Otrzymane szacowanie może jednak być użyte jako kolejny element kontrolny i weryfikujący działanie rzeczywistych czujników  $NO_x$  zamontowanych na obiekcie.

Ograniczona wielkość artykułu nie pozwoliła pokazać wyników pełnej analizy sprawdzającej założenia regresji (liniowość, niezależność, normalność oraz stałość wariancji). Dopiero pozytywny wynik tej analizy daje prawo uważać otrzymany model (modele) za wiarygodny. Ponadto nie zamieszczono precyzyjnych wyników dotyczących dokładności regresji (błędy szacowania, przedziały ufności itd).

Należy również wspomnieć o innym podejściu do wyboru zmiennych objaśniających. Można mianowicie, przed właściwą analizą regresji, wykonać na oryginalnych danych wejściowych analizę składowych głównych (ang. *Principal Components Analysis*, PCA). Otrzymamy wówczas nowy zestaw zmiennych, na których będziemy pracować. Informacje na temat PCA można znaleźć w bardzo bogatej literaturze, przykładowo patrz [1, 2].

## 6. Literatura

- [1] Koronacki J., Mielniczuk J.: Statystyka dla studentów kierunków technicznych i przyrodniczych, WNT, Warszawa, 2006.
- [2] Larose D. T.: Metody i modele eksploracji danych, PWN, Warszawa, 2008.
- [3] Laudyn D., Pawlik M., Strzelczyński F.: Elektrownie, WNT Warszawa, 2006.
- [4] Szargut J.: Termodynamika Techniczna, Wydawnictwo Politechniki Śląskiej, wydanie 4, 2005.
- [5] Velleman P. F.: Definition and comparison of robust nonlinear data smoothing algorithms. *Journal of the American Statistical Association*, 75, 609-615, 1980.
- [6] <http://www.ec.zgora.pl>
- [7] <http://pl.wikipedia.org/wiki/DCS>