

Sopot, 15 września 2017 r.

Prof. dr hab. Jacek Piskozub
Instytut Oceanologii PAN
ul. Powstańców Warszawy 55
81-712 Sopot
email: piskozub@iopan.gda.pl

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr Izabeli Pawlak

pt.: „Analiza zmienności zawartości ozonu w przyziemnej warstwie atmosfery w województwie mazowieckim”

Ozon jest bardzo ważnym składnikiem atmosfery planety, na której żyjemy. Powszechnie wiadomo, jaką rolę pełni on w blokowaniu, już w stratosferze, strumienia promieniowania ultrafioletowego przychodzącego od Słońca. Pewnie jednak nie wszyscy zdają sobie sprawę z tego, że samo istnienie stratosfery (obszaru atmosfery gdzie temperatura rośnie z wysokością) zawdzięczamy właśnie ozonowi, a dokładniej intensywnej absorpcji ultrafioletu w tej części atmosfery, chociaż wiemy to już od 50 lat, czyli odkąd stworzono pierwsze modele radiacyjno-konwekcyjne. Ten związek chemiczny ważny jest jednak także w troposferze, którą oddychamy (choć znajdziemy tu tylko 10% atmosferycznego ozonu). Ważny jest przez negatywne efekty dla naszego zdrowia, ale także dlatego że wzrost stężenia troposferycznego ozonu odpowiedzialny jest za ok. 20% antropogenicznego wymuszenia radiacyjnego netto, liczonego w stosunku do czasów przedprzemysłowych, czyli mówiąc prościej za podwyższenie temperatury Ziemi prawie o 0,2 °C.

Tematem recenzowanej pracy doktorskiej mgr Izabeli Pawlak są badania zmienności stężenia ozonu przy powierzchni ziemi, oraz próby ulepszenia jego prognozowania w skali doby, w oparciu o materiał badawczy pochodzący z sześciu stacji pomiarowych w województwie mazowieckim. Praca, licząca (po odjęciu bibliografii) 135 stron podzielona jest na osiem rozdziałów, które omówię kolejno. Już na wstępie zauważę jednak, że układ pracy jest absolutnie „nieklasyczny” czyli metodologia przeplata się z wynikami. Zaznaczam to jednak, aby podkreślić, że moim zdaniem w pracy, zawierającej różnorodne metody, tego typu układ sprawia, że „czyta się ją” znacznie lepiej i moim zdaniem nie można traktować tego jako wadę. Dodam, że tekst jest napisany bardzo dobrze i

praktycznie nie znalazłem w nim żadnych błędów językowych (dodałbym jedynie gdzieś kilka przecinków, ale ponieważ już w szkole średniej dowiedziałem się, że nie jestem ekspertem od polskiej interpunkcji, więc mogę się w tej sprawie mylić). Redakcja pracy jest w ogóle na bardzo wysokim poziomie, tzn. ciężko znaleźć w niej także merytoryczne pomyłki. Natomiast, jak będzie widać z dalszej części opinii, pewnym problemem bywa w pracy bardziej to czego w niej nie ma, niż to co w niej jest.

Mam jednak dwie uwagi terminologiczne:

- w pracy stężenie i koncentracja stosowane są wymiennie jako synonimy. Osobiście żadna z nich mi nie przeszkadza, chociaż jako „aerolog” preferuję koncentrację gdy mówimy o policzalnej ilości cząstek w danej objętości. Jednak uważam że w przypadku masy lub ilości moli na jednostkę objętości lepiej używać typowo chemicznego terminu stężenie. Na pewno jednak trzeba wybrać jeden z nich, bo w tekście naukowym jeden parametr powinien mieć tylko jedną nazwę.
- W jednym z rozdziałów pracy używany jest termin odchyłka w znaczeniu anomalii (odstępstwa od średniej). Chwaliłem w zeszłym roku użycie tego słowa jako odpowiednika angielskiego *bias*, dla którego nie mamy innego odpowiednika. Jednak tym bardziej nie należy go stosować w sensie anomalii. Jeśli jednak ktoś nie lubi anomalii to już lepsze byłoby „odchylenie”.

Rozdział pierwszy „Wprowadzenie” łączy w sobie wprowadzenie do zagadnienia z przedstawieniem celu i zakresu pracy. To pierwsze kontynuowane jest w rozdziale następnym (o procesach), jednak umieszczenie celu pracy w tym miejscu jest uzasadnione „narracją”. Nie mam szczególnych uwag do tego rozdziału (*nie licząc zwrócenia uwagi na drobne powtórzenia np. w kwestii wczesnych pomiarów Bojkova*) oprócz jednej. Na str. 10 opisane są możliwe powody zmian stężenia ozonu obserwowanego na przełomie XIX i XX wieku. Jednym z nich jest wzrost stężenia metanu. Niestety pominięto opis mechanizmu w jaki wzrost stężenia metanu wpływa na wzrost stężenia ozonu. Może dlatego, że wynika to częściowo z opisu chemii ozonu w rozdziale następnym. Jednak nie do końca, gdyż ozon jest nie tylko produktem utleniania metanu w obecności tlenków azotu (co opisano w pracy), ale także jego destrukcja w atmosferze spowalniana jest w obecności metanu, który redukuje chlor, będący katalizatorem tego procesu do mniej aktywnego chemicznie kwasu solnego (patrz np. artykuł przeglądowy o roli metanu w chemii atmosfery Wuebbles i Hayhoe, 2002, [https://doi.org/10.1016/S0012-8252\(01\)00062-9](https://doi.org/10.1016/S0012-8252(01)00062-9)). Ma to znaczenie nawet w troposferze, w związku z opisywanym w pracy transportem powietrza ze

stratosfery, w stronę powierzchni ziemi.

Rozdział drugi, jak już wspomniano, poświęcony jest procesom „determinującym” zawartość ozonu w dolnej troposferze. Jest on najtrudniejszym intelektualnie w całej pracy (dla wszystkich niechemików) ale napisany jest bardzo dobrze. Te skomplikowane procesy, opisane po raz pierwszy przez noblistę Crutzena w pracach z 1973 i 1974 roku (odpowiednio <https://doi.org/10.1007/BF00881092> i <https://doi.org/10.1111/j.2153-3490.1974.tb01951.x> – nawiasem mówiąc mogłyby one zostać zacytowane w recenzowanej pracy!) zostały przedstawione w formie uproszczonej bez utraty istotnych informacji – co nie jest łatwą sztuką! W tej części występuje dość nietypowe w pracach naukowych powtarzanie tych samych równań (z tym samym ich numerem) ale i to sprzyja moim zdaniem płynności narracji. Paradoksalnie, procesy fizyczne opisane są w nieco większym uproszczeniu *i tu mam jedyne drobne uwagi:*

- „cut off lows” ze strony 29 chyba najlepiej się tłumaczy jako „nizę odcięte”, <http://lowcyburz.pl/slownik-skywarn-polska/niz-odciety/>
- Na rysunku 2.1 nierówność $NO_x > LZO$ jest nieprawdziwa ze względu na różne skale x i y . Można by temu zaradzić przesuwać ten napis w lewo lub zmienić go na coś innego¹.

W tym miejscu (między rozdziałami 2 i 3) wspomnę natomiast o dość istotnym braku recenzowanej pracy. Nie ma mianowicie żadnej dyskusji jakości omawianych serii pomiarów ozonu. Czy wszystkie stacje używały mierników tego samego typu? Jakiego? Czy nie zmieniano niczego w trakcie trwania serii pomiarowych? Jakie są błędy używanych przyrządów i od czego zależą? Omówienie tego powinno znaleźć się w pracy nawet jeśli nie wszystkie te informacje są dostępne (to też istotna wiadomość).

Rozdział trzeci stanowi analizę wyżej wspomnianych serii pomiarowych. Analiza ta jest zasadniczo poprawna, chociaż mam szereg uwag (niektóre z nich natury polemicznej):

- Rozdział jest pełen analiz statystycznych, lecz nie przeprowadzono najprostszej: czy średnie wartości stężenia ozonu dla poszczególnych stacji są różne od siebie w sposób istotny statystycznie, co jest warunkiem koniecznym do stwierdzenia, że stacje są różnorodne. Przeprowadziłem sam taką analizę przy pomocy dwustronnego testu-t Studenta dla danych z Tabeli 3.1. Wynika z niej, że jedynie dwie stacje nie mają istotnie statystycznie ($p < 0.05$) różnych średnich: Legionowo i Warszawa Ursynów.

¹ Uwag pisanych kursywą nie planuję czytać dlatego nie muszą być one uwzględniane w wygłoszonej na obronie odpowiedzi na recenzję.

- Nie przeprowadzono też nigdzie, ani na danych rocznych z Tabeli 3.1, ani dla zależności NO_x i O₃ z Rys. 3.6, analizy istotności statystycznej trendu czasowego czy uzyskanej zależności liniowej (matematycznie jest to to samo). W przypadku trendu czasowego dysponowałem danymi z Tabeli 3.1 i sam sprawdziłem, że jedyną stacją dla której trend spadkowy w okresie 2005 -2012 był istotny statystycznie jest Belsk, a i to ledwo ($p < 0.041$). Na szczęście analizy zmian czasowych w podrozdziale 3.5 dotyczą dłuższego okresu ale i tu nie pokazano istotności statystycznej dla żadnego z trzech wyróżnionych okresów o różnych trendach.
- Wiele z rozważań z tego rozdziału (str. 36, 40 oraz rys 3.5) staje się łatwiejsze do zrozumienia w kontekście omawianej później antykorelacji NO_x i O₃, świadczącej o tym, że przynajmniej w warunkach miejskich dominującymi procesami są reakcje powodujące zamianę w siebie tych związków. Czy nie lepiej byłoby zamienić kolejność, wprowadzając wcześniej antykorelację z Rys. 3.6, szczególnie, że jest ona znana z literatury, np. z cytowanej pracy Chou i inni, 2006 dla danych z Taipei na Tajwanie.
- Ta antykorelacja (a w zasadzie nawet kwazi-zachowawczość sumy NO_x i O₃) byłaby jeszcze wyraźniej widoczna na rys 3.3 gdyby stężenia związków przedstawić molowo, a nie wagowo – zamiana NO₂ na O₃ (i odwrotnie) występuje w relacji „mol za mol”, a nie „gram za gram”.
- Argument że przyczyną dodatniej anomalii ozonu w poniedziałek na niektórych stacjach (podobnie jak w weekend) jest brak tam intensywnego ruchu ulicznego nie jest przekonujący. Brak ten zapewne dotyczy także wtorku i następnych dni (z odwrotną anomalią). Jest to raczej „pamięć” wartości ozonu z poprzedniego dnia (lokalnej lub podlegającej adwekcji). Potwierdzają to wyniki analiz przy pomocy sieci neuronowych (Rozdział 6), w których wartość stężenia ozonu z poprzedniego dnia jest przydatnym elementem predykcji dla dnia następnego.
- Jedną z przyczyn dużych amplitud dobowych na leśnej stacji Granica może być również znaczna produkcja lotnych związków organicznych przez lasy iglaste i skomplikowane procesy jakim ulegają one w obecności światła słonecznego (patrz np. Andreae 2013, <https://doi.org/10.1126/science.1233798>).
- Czy w związku z dyskusją (str. 61) dotyczącą wpływu fali upałów na maksimum serii wypadające w 2003 roku nie warto było zastosować korelacji wartości ozonu z jakąś miarą nasłonecznienia lub zachmurzenia?
- *Wspomniane na str. 55 „procesy pionowego mieszania powietrza i w efekcie transport w dół powietrza” występujące w godzinach dziennych to oczywiście konwekcja. Natomiast*

podczas warunków stabilnych (niżej na tej samej stronie) z definicji nie należy spodziewać się pionowego mieszania.

- *Fala upałów z roku 2003 dotyczyła raczej Europy Zachodniej niż Środkowej.*
- *LOWESS to metoda wygładzania, a nie regresja.*
- *NO w punkcie 4 na stronie 40 to zapewne Nox.*
- *Prawidłowa pisownia to test-t Studenta (a nie test t-Studenta).*

Kolejny, czwarty rozdział, stanowi analizę wpływu warunków meteorologicznych (ale nie synoptycznych jak stwierdzono na pierwszej jego stronie) na zmienność stężenia ozonu w warstwie przyziemnej. Omówiono w nim teoretyczne przyczyny wpływu na tę zmienność następujących parametrów: temperatury, nasłonecznienia, wilgotności względnej oraz prędkości wiatru. Przedstawiono również wartości korelacji tych parametrów ze stężeniem ozonu dla każdego miesiąca i każdej stacji pomiarowej. Wyniki wskazują na duży wpływ temperatury (z wyjątkiem miesięcy zimowych), promieniowania słonecznego (z wyjątkiem grudnia i stycznia gdy w naszych szerokościach jest jego niewiele), ujemną korelację z wilgotnością względną, oraz niewielki i różne (nawet co do znaku) dla różnych stacji i miesięcy korelacje z siłą wiatru. Są to ciekawe wyniki chociaż można było pokazać na ile te parametry korelują się z sobą nawzajem (a oczywiście korelują się dość silnie). Do zasadniczo poprawnego omówienia wyników mam jednak parę uwag:

- Co to jest globalne promieniowanie słoneczne? I nie mówię tu nawet, że powinno ono być raczej „całkowite” (angielskie „total”, nie „global”). Chodzi o niepodanie granic spektralnych (widzialne, krótkofalowe, czy jeszcze inne)?
- W pracy rozpatrywane jest oświetlenie „globalne”, ale w produkcji ozonu bierze udział jedynie światło o fali światła krótszej niż 424 nm (fioletowe i ultrafioletowe). Jak dobrze koreluje się ono z „globalnym”?
- Ilość molekuł pary wodnej dostępnych w atmosferze dla reakcji z wzbudzonymi atomami tlenu (str. 69) to zasadniczo wilgotność bezwzględna. Stosowanie wilgotności względnej może być zatem nie najlepszym wyborem. Szkoda, że tego nie sprawdzono.
- Ujemne korelacje temperatury i ozonu w miesiącach zimowych mogą być efektem występującego wówczas efektu inwersji termicznej. Nie da się ustalić na ile to inwersja, a na ile adwekcja powietrza arktycznego (mechanizm proponowany na str. 67) bez rozpatrzenia danych w skali dziennej, zamiast miesięcznej, czego niestety nie zrobiono.
- *W pracy stosowane są zamiennie symbole p i α jako parametr istotności statystycznej.*

Powyższa analiza aż prosi się o zwięźczenie przez przeprowadzenie analizy regresji wielu

zmiennych. Mgr Izabela Pawlak idzie dalej i przeprowadza, w dwóch następnych rozdziałach, analizę przy pomocy sieci neuronowych. Jest to metoda lepsza bo umożliwia także zbadanie relacji nieliniowych. Stwierdzenie to ma oparcie w literaturze przedmiotu, także dla badań ozonu w warstwie przyziemnej². Jednak regresja wielu zmiennych pokazuje w bardziej bezpośredniej formie, które zmienne są najważniejsze i dlatego trochę szkoda, że jej nie przeprowadzono.

Rozdział 5 streszcza istotę działania sieci neuronowych oraz metody ich uczenia. W rozdziale 6 wykorzystuje je w praktyce. Niestety brakuje omówienia literatury na temat historii używania tej metody w badaniach ozonu i osiągniętych wcześniej wyników. Autor (to konwencja przyjęta w samej pracy) wymienia tu jedynie (na stronie 79) cztery wcześniejsze prace, bez wchodzenia w jakiegokolwiek szczegóły. A szkoda, bo sprawdzenie w Web of Science³ pokazuje, że prac takich było wcześniej już 114, prawie wszystkie o ozonie w warstwie przy powierzchniowej, szczególnie w miastach (w Polskiej literaturze jedynie raz, dla pionowych profili aerozolu, przez promotora recenzowanej pracy: Jarosławski, 2013, <https://10.1080/01431161.2013.793463>).

Tu pozwolę sobie na szerszą uwagę o bibliografii recenzowanej pracy. O ile w kwestii procesów chemicznych i fizycznych, wpływających na ozon jest ona bogata, to w innych jest o wiele gorzej. Teorię sieci neuronowych autor opisuje w oparciu o dwa podręczniki (Tadeusiewicz, 1993 oraz Tadeusiewicz i inni 2007), nie korzystając z bardzo bogatej literatury „pierwotnej”. Jeszcze gorzej jest w przypadku meteorologii, gdzie jedynym źródłem jest podręcznik dość popularnej natury, a co gorsza bardzo stary (Kaczorowska 1986, w istocie przedruk z roku 1977).

Wracając jednak do sieci neuronowych, mgr Izabela Pawlak skutecznie zastosowała moduł programu Statistica 10 o nazwie „Automatyczne Sieci Neuronowe”, dzieląc posiadany zestaw danych losowo na zbiory uczący, testowy i walidacyjny. Sieci neuronowe miały jedną ukrytą warstwę (najprostszy model sieci co jest uzasadnione ilością posiadanych danych). Dla wszystkich stacji uzyskano wysokie korelacje wyliczonych wartości stężeń ozonu z mierzonymi, także dla zbioru testowego, a co najważniejsze walidacyjnego (wszystkie były istotne statystycznie z wyjątkiem stacji Ursynów, gdzie korelacje były najmniejsze – rozumiem że brak informacji o tym w pracy oznacza brak istotności!). Wartości błędów przedstawionych w Tabeli 6.4 są jednak mało czytelne, gdyż przedstawiono w nich wartości „SOS” (sumy kwadratów odchyłeń), których

2 Na przykład Bandyopadhyay and Chattopadhyay 2007, <https://doi.org/10.1007/BF03325972> albo Sousa i inni, 2007, <https://doi.org/10.1016/j.envsoft.2005.12.002>

3 Szukałem w jednym polu „neural network” OR neural networks”, a w drugim połączonym z nim logicznym AND słowa „ozone”

wartość zależny od ilości danych n (nie ma w jego wzorze dzielenia przez n). Teoretycznie podzielenie tych wartości przez n , a następnie wyliczenie pierwiastka kwadratowego powinno dać błąd RMSE, którego wartości podano dla każdej stacji na kolejnych tabelach, jednak w praktyce tak nie jest (ciekawe dlaczego?).

W kolejnych podrozdziałach przedstawiono wyniki analizy wrażliwości globalnej i lokalnej sieci neuronowej. Z pierwszej (właśnie jej wyniki można by porównać z współczynnikami regresji liniowej wielu zmiennych) wynika, że najważniejszym parametrem wpływającym na stężenie ozonu jest temperatura, mniej ważnymi kolejno miesiąc, promieniowanie słoneczne i wartość z dnia poprzedniego, a najmniej wilgotność i prędkość wiatru. Z drugiej wynika jakie części zakresu zmienności każdego parametru wpływają najsilniej na wartość stężenia ozonu. Przy czym nie zgadzam się, „że wartości temperatury powietrza do ok. 15 °C nie mają istotnego znaczenia dla wyniku predykcji” (str. 92), gdyż są one i tak wyższe niż dla wilgotności, gdziekolwiek w jej zakresie zmienności. Natomiast wartości czułości dla wiatru, wyglądają na losowe i nie dziwię się, że przy pominięciu tego parametru uzyskiwano nie gorsze dopasowanie niż z nim.

Kolejny, szósty rozdział przedstawia wartości szeregu różnie zdefiniowanych miar błędu dla wyników analizy z wykorzystaniem sieci neuronowych. Analizy te są w pełni poprawne. Ostatni, siódmy rozdział stosuje uzyskane modele sieci neuronowych do danych z innego roku, z dobrym skutkiem, chociaż błędy są zauważalnie większe niż dla danych z roku, na którym trenowano sieć. Może to świadczyć o tym, że trenowanie sieci na danych z tylko jednego roku nie „pokazało” im całego możliwego zakresu zmienności danych. Wyniki te porównano z prognozami z modelu cyrkulacji z modułem chemii atmosfery GEM-AQ, pokazując iż sieci neuronowe są w stanie dać lepszą prognozę na następny dzień niż tego typu model.

Podsumowując, mgr Izabela Pawlak opracowała kompleksowo dostępne dane o stężeniach ozonu w warstwie przy powierzchniowej w Województwie Mazowieckim, stosując szereg metod, z których żadna nie jest wprawdzie nowatorska, ale łącznie pozwalają one na uzyskanie wyników interesujących, które moim zdaniem powinny zostać opublikowane również w czasopiśmie recenzowanym. Przedstawione analizy nie budzą wątpliwości, ich dyskusja (z pewnymi wyjątkami) również. Wadami pracy są braki opisu metod pomiaru i związanych z tym niepewności oraz brak dyskusji wyników wcześniejszych podobnych prac używających metodę sieci neuronowych (i nie tylko) do badania ozonu powierzchniowego w innych rejonach świata.

Uważam zatem, że recenzowana praca doktorska mgr Izabeli Pawlak stanowi oryginalne opracowanie, pierwsze tego typu dla rejonu Mazowsza, a nawet Polski, przy użyciu szeregu metod statystycznych, w tym ciągle jeszcze zbyt rzadko u nas stosowanej współczesnej metodologii sieci neuronowych. Metody analizy danych nie budzą wątpliwości, a wyniki badań są interesujące i cytowalne. Recenzowana praca doktorska, pomimo pewnych niedociągnięć, spełnia zatem wymóg ustawy stanowiąc „oryginalne rozwiązanie problemu naukowego”.

To stwierdziwszy, mogę przejść do oficjalnej konkluzji:

Rozprawa przedstawiona do recenzji spełnia wymagania stawiane rozprawom doktorskim w świetle obowiązujących przepisów. Stawiam zatem wniosek o dopuszczenie jej do publicznej obrony.

Na zakończenie chciałbym życzyć mgr Izabeli Pawlak dalszych sukcesów w prowadzonych badaniach i karierze zawodowej.

Z poważaniem

