

Krzysztof SIWEK*, Mateusz BARANOWSKI*, Tomasz GRZYWACZ*

PROGNOZOWANIE ZAWARTOŚCI PM_{2,5} W POWIETRZU ATMOSFERYCZNYM PRZY UŻYCIU SIECI ELM

Artykuł przedstawia nowe podejście do predykcji zawartości pyłów zawieszonych w powietrzu atmosferycznym wykorzystujące sieć neuronową typu ELM (Extreme Learning Machine). Predykcja ta dotyczy obliczenia średniego poziomu zanieczyszczenia powietrza pyłem PM_{2,5} na warszawskim Ursynowie z jednodniowym wyprzedzeniem. Do zrozumienia wagi problemu zawarto zwięzły opis zagrożeń, jakie niesie za sobą zanieczyszczenie powietrza drobnym pyłem zawieszonym PM_{2,5}.

W pracy przedstawiono krótki opis sieci ELM oraz zaprezentowano uzyskane wyniki prognozy. Przeprowadzono analizę uzyskanych wyników. Omówiono zasadność prognozowania oraz wskazano możliwe środki zapobiegawcze i ochronne.

SŁOWA KLUCZOWE: zanieczyszczenie powietrza atmosferycznego, PM_{2,5}, sieci neuronowe, Extreme Learning Machine, ELM.

1. WSTĘP

1.1. Postawienie problemu

W ostatnich latach obserwuje się stałe i rosnące zanieczyszczenia środowiska naturalnego, w którym żyjemy. Zanieczyszczeniu ulegają między innymi woda, gleba oraz powietrze atmosferyczne. Obserwując informacje dostarczane nam przez media jak i dostępne bezpośrednio ze stacji pomiarowych widzimy, że problem nie dotyczy tylko Polski, jest problemem całego świata.

W wyniku postępu cywilizacyjnego i rozwoju przemysłu wpływ działalności człowieka na środowisko naturalne jest coraz większy i niestety w większości przypadków jest dla niego niekorzystny. Bezpośrednim skutkiem, który odczuwamy, jest niekorzystny wpływ bezpośrednio na człowieka.

Czynnikiem, o którym jest głośno w ostatnich latach jest jakość powietrza atmosferycznego. Głównym problemem jest zanieczyszczenie powietrza pyłem zawieszonym, które stanowi zagrożenie dla życia i zdrowia człowieka, i które może wpłynąć niekorzystnie na długość i jakość życia człowieka [1, 2].

* Politechnika Warszawska

1.2. Zaproponowane rozwiązanie

Znajomość przyszłych wartości zanieczyszczenia powietrza atmosferycznego pyłami zawieszonymi może być pomocna do przeciwdziałania zjawiskom powodującym te zanieczyszczenia lub wczesnego ostrzegania i zastosowania środków ochronnych zarówno dla pojedynczych ludzi (np. maseczki, oczyszczacze powietrza w mieszkaniach) lub na poziomie miast i krajów (np. darmowe przejazdy komunikacją miejską).

Autorzy mają wieloletnie doświadczenie w pracy naukowej związanej z predykcją ciągów czasowych, w szczególności przy wykorzystaniu sztucznych sieci neuronowych. W pracy tej zastosowano sieci neuronowe typu ELM (Extreme Learning Machine) do przewidywania wartości średniodobowej zanieczyszczenia PM_{2,5} na dzień następny.

2. PYŁY ZAWIESZONE W POWIETRZU ATMOSFERYCZNYM

2.1. Klasyfikacja pyłów zawieszonych

Pył zawieszony w powietrzu atmosferycznym klasyfikujemy ze względu na jego średnicę. Wyróżnia się:

- PM₁₀ – pyły o średnicy mniejszej niż 10 μ m, które mogą docierać do górnych dróg oddechowych i płuc,
- PM_{2,5} – pyły o średnicy mniejszej niż 2,5 μ m, które wnikają głęboko do płuc i przedostają się do krwioobiegu,
- PM₁ – pyły o średnicy mniejszej niż 1 μ m, od tej średnicy pyły usuwane są bardzo powoli z pęcherzyków płucnych lub wcale,
- PM_{0,1} – pyły ultradrobne, nanopyły, o średnicy mniejszej niż 0,1 μ m, mogą przedostawać się z nosa wzdłuż nerwu węchowego do mózgu, wpływając na pojemność pamięci i inteligencję [2].

2.1. Wpływ pyłów zawieszonych na środowisko i człowieka

Pył zawieszony składa się z nieokreślonej mieszaniny cząstek substancji nieorganicznych oraz organicznych. Pył zawieszony może zawierać różnorodne substancje toksyczne na przykład wielopierścieniowe węglowodory aromatyczne – benzo(a)pireny i benzo(e)pireny, dioksyne, furany, metale ciężkie oraz wiele innych substancji. Obecnie nie bada się składu pyłów zawieszonych.

Zanieczyszczenia powietrza są wchłaniane przez ludzi głównie w trakcie oddychania, przyczyniają się do powstawania schorzeń układu oddechowego i alergii, chorób nowotworowych, a także chorób serca i naczyń krwionośnych.

Działają niekorzystnie również na świat roślinny, zaburzając procesy fotosyntezy, transpiracji i oddychania (poprzez osadzający się pył).

Zanieczyszczenia powietrza powodują ponadto korozję metali i niszczenie materiałów budowlanych (kwaśne deszcze). Wtórnie skażają wody i gleby oraz zwiększają kwasowość wody pitnej. Zmniejszają przejrzystość powietrza, częstym efektem zanieczyszczenia powietrza pyłem zawieszonym jest smog.

W skali globalnej mają wpływ na zmiany klimatyczne (dziura ozonowa, efekt cieplarniany).

Pył PM₁₀ zawierający cząstki o średnicy mniejszej niż 10 mikrometrów, w procesie oddychania dociera do górnych dróg oddechowych i płuc. Może zmniejszać powierzchnię pęcherzyków płucnych. Ustawodawstwo krajowe i Europejskie definiuje poziom dopuszczalny dla stężenia średniodobowego oraz średniorocznego. Poziom średniodobowy wynosi 50 $\mu\text{g}/\text{m}^3$ i nie może być przekraczany nie więcej niż przez 35 dni w ciągu roku. Poziom dopuszczalny dla stężenia średniorocznego wynosi 40 $\mu\text{g}/\text{m}^3$. Poziom alarmowy w Europie wynosi, w zależności od kraju od 50 do 100 $\mu\text{g}/\text{m}^3$, w Polsce 300 $\mu\text{g}/\text{m}^3$. Jest to wartość, przy przekroczeniu której należy informować społeczeństwo o zagrożeniu.

Pył PM_{2,5} należy do kategorii pyłów drobnych, zawiera cząstki o średnicy mniejszej niż 2,5 mikrometra. Ustawodawstwo Polskie i Europejskie dla pyłu PM_{2,5} definiuje wartość średnioroczną, która jest jednocześnie poziomem dopuszczalnym w wysokości 25 $\mu\text{g}/\text{m}^3$. Nie definiuje się pojęcia wartości średniodobowej.

Małe drobiny pyłów mogą z górnych dróg oddechowych i płuc pokonywać barierę tkanka-krew, przenikać do krwi, a z krwią rozprzestrzeniać się po całym organizmie. Najmniejsze cząsteczki pokonują barierę krew-komórka i dostają się do wnętrza komórek ciała. Mogą tam zalegać i być źródłem niekorzystnych zmian w organizmie (np. nowotworów). Dłuższe narażenie na wysokie stężenia pyłu może mieć istotny wpływ na występowanie chorób serca i układu naczyniowego (nadciśnienie, zawał). Szacuje się, że choroby spowodowane przez zanieczyszczenia powietrza atmosferycznego w Polsce przyczyniają się do śmierci 45 tysięcy osób w ciągu roku, to jest 15 razy więcej niż ginie w wypadkach (ok. 3000 osób rocznie) [3].

2.2. Źródła zanieczyszczeń powietrza pyłami zawieszonymi

Źródła zanieczyszczeń powietrza atmosferycznego są dwojakiego rodzaju: naturalne oraz antropogeniczne. Do źródeł naturalnych należą między innymi:

- wybuchy wulkanów,
- erozja wietrzna skał,
- pył nawiewany z pustyń,
- pożary lasów i stepów,

- pył kosmiczny,
- niektóre procesy biologiczne.

Na emisję zanieczyszczeń z takich źródeł człowiek nie ma większego wpływu.

Do drugiej kategorii należą zanieczyszczenia powietrza będące skutkiem działalności człowieka:

- chemiczna konwersja paliw (spalanie),
- wydobywanie i transport surowców,
- motoryzacja,
- przemysł chemiczny,
- przemysł rafineryjny,
- przemysł metalurgiczny,
- cementownie,
- składowiska surowców i odpadów,
- wypalanie lasów i stepów.

Rozróżnia się następujące rodzaje źródeł zanieczyszczeń powietrza:

- punktowe – są to głównie duże zakłady przemysłowe emitujące pyły, dwutlenku siarki, tlenku azotu, tlenku węgla, metale ciężkie itp.
- powierzchniowe (rozproszone) – są to paleniska domowe, lokalne kotłownie, niewielkie zakłady przemysłowe, emitujące głównie pyły oraz dwutlenek siarki,
- liniowe – są to głównie zanieczyszczenia komunikacyjne odpowiedzialne za emisję tlenków azotu, tlenków węgla, metali ciężkich (głównie ołów) oraz pyłów mechanicznych.

Przyczyny przekroczenia norm pyłów zawieszonych w powietrzu są właściwie tożsame z ich źródłami:

- indywidualne ogrzewanie budynków (niska emisja: małe miasta, osiedla),
- motoryzacja (centra dużych miast),
- napływ zanieczyszczeń spoza terenu (miasta, województwa, kraju),
- produkcja przemysłowa, prace remontowe, wydobywanie surowców,
- naturalne źródła emisji,
- warunki meteorologiczne (temperatura, wilgotność, ciśnienie).

W Polsce zanieczyszczenia powietrza pyłami pochodzą najczęściej z dwóch źródeł: spalania słabej jakości węgla w starych, nieefektywnych kotłach i piecach domowych oraz ze źródeł komunikacyjnych – ze spalin silnikowych i pyłów unoszonych spod kół samochodów.

Pojęcie „niskiej emisji” pochodzi od spalania węgla, drewna i odpadów w domach jednorodzinnych, w każdym w małej ilości, posiadających wyloty kominów na małej wysokości. Jednakże domy takie usytuowane na dużych osiedlach mogą stanowić poważne źródło zanieczyszczeń, które stanowią największe zagrożenie właśnie dla ich mieszkańców.

Zanieczyszczenia powstałe w wyniku innej działalności ludzkiej nie stanowią tak dużego zagrożenia, ponieważ przemysł, szczególnie energetyczny, ograni-

czony jest ścisłymi przepisami, a stosowane tam filtry oraz używane wysokie kominy znacznie ograniczają emisję i mają zwykle dużo mniejszy wpływ na jakość powietrza.

3. DANE POMIAROWE PM_{2,5}

3.1. Metody pomiaru pyłu zawieszonego

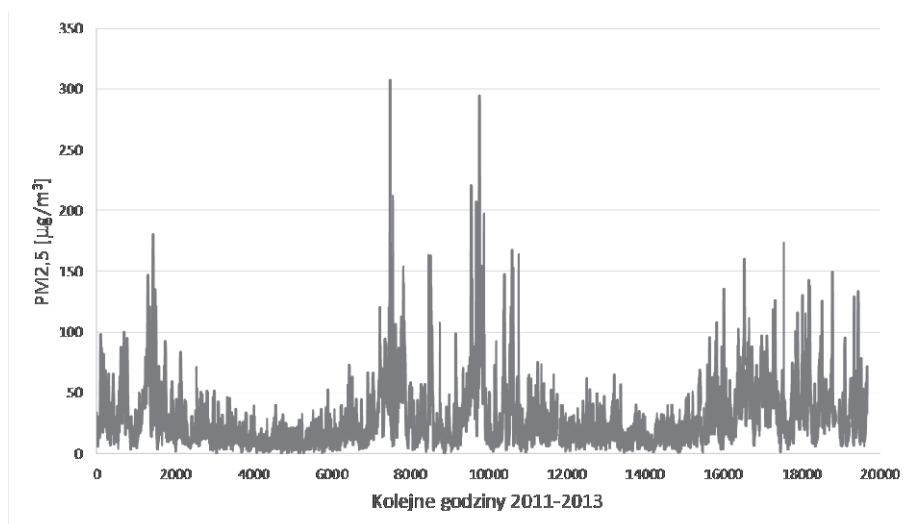
Na jakość informacji o stężeniu PM_{2,5} i PM₁₀, a tym samym na niepewność pomiarową, decydujący wpływ ma metodyka pomiarów, wskazana przez dyrektywę Parlamentu Europejskiego i Rady Europejskiej 2008/50/WE oraz rozporządzenie Ministra Środowiska z dnia 13 września 2012 r., która została określona w normie PN-EN 12341:2014 r.

Metoda grawimetryczna jest uznawana na świecie za najbardziej miarodajną i precyzyjną. W tej metodzie używa się filtrów z oczkami o średnicy mierzonej frakcji pyłów. Co 2 tygodnie w stacjach pomiarowych instaluje się 14 filtrów, a co 24 godziny na ekspozycje wystawiany jest nowy filtr. Przed pomiarem czyste filtry są ważone w laboratorium. Po 14 dniach od instalacji filtry są wyjmowane i ponownie ważone w laboratorium. Z obliczonej różnicy masy odniesionej do objętości powietrza przepuszczonego przez filtr, obliczany jest poziom stężenia pyłu.

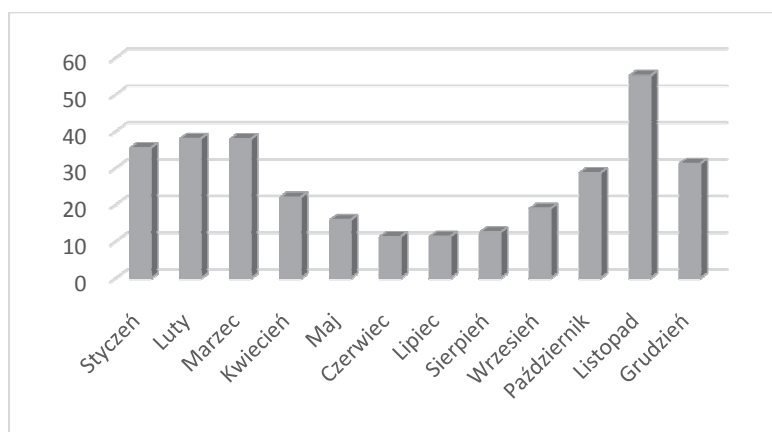
Metoda ta gwarantuje bardzo wysoką dokładność wyników, wadą natomiast jest czas oczekiwania na rezultat, który wynosi do 3 tygodni. Obecnie w Polsce GIOŚ – Główny Inspektorat Ochrony Środowiska – posiada ok. 180 tego typu stanowisk do pomiaru PM₁₀ oraz ok. 70 do pomiaru PM_{2,5}.

3.2. Baza danych pyłu zawieszonego PM_{2,5}

Danymi pomiarowymi użytymi do badań eksperymentalnych były dane pochodzące ze stacji pomiarowej WIOŚ (Warszawski Inspektorat Ochrony Środowiska) zlokalizowanej w mieszkaniowej dzielnicy Ursynów w Warszawie z lat 2011-2013.

Rys. 1. Dane pomiarowe PM_{2,5} z lat 2011÷2013

Jak widać na rysunkach 1 i 2 wartości dopuszczalne PM_{2,5} ($25 \mu\text{g}/\text{m}^3$) w stacji Warszawa-Ursynów są stale przekraczane. Szczególnie widoczne jest to w miesiącach chłodnych (jesień-zima), kiedy budynki indywidualne są ogrzewane piecykami węglowymi.

Rys. 2. Średni poziom PM_{2,5} w poszczególnych miesiącach lat 2011-2013

Przy głębszej analizie można również dostrzec przekroczenia norm w dni długich weekendów gdy istnieje zwiększony ruch samochodami. To samo zjawisko dostrzegalne jest w poniedziałki rano i piątki po południu, a także codziennie w szczytach komunikacyjnych.

4. SIEĆ NEURONOWA TYPU EXTREME LEARNING MACHINE

Sieć neuronowa typu Extreme Learning Machine jest rozwinięciem sieci jednokierunkowej wielowarstwowej MLP. Zaprojektowana została do rozwiązywania zadań optymalizacyjnych dla bardzo dużej liczby danych. Założenie to jednocześnie wymusiło zastosowanie bardzo szybkiego algorytmu uczącego.

Czas uczenia sieci jednokierunkowych jest zazwyczaj dosyć długi z uwagi na stosowanie wolnozbieżnych metod gradientowych. W przeciwieństwie do tradycyjnych implementacji, w pracy [5] został zastosowany nowy algorytm uczenia nazwany Extreme Learning Machine (ELM) do SLFN (Single Hidden Layer Feedforward Neural Network, tj. sieć MLP z jedną warstwą ukrytą).

4.1. Algorytm ELM

Uczenie sieci ELM polega na losowym doborze parametrów początkowych sieci (wag) oraz obliczeniu w sposób analityczny jedynie wag wyjściowych. Teoretycznie algorytm ELM oferuje bardzo wysoki stopień generalizacji i ekstremalnie szybki czas uczenia [4].

Algorytm ELM dotyczy sieci SLFN z \tilde{N} neuronów w warstwie ukrytej, $\tilde{N} \leq N$, dla N różnych par

$$(\mathbf{x}_i, \mathbf{t}_i), \mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n \text{ i } \mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \mathbf{R}^m.$$

Zwykły SLFN z \tilde{N} neuronów w warstwie ukrytej i funkcją aktywacyjną $\varphi(x)$ może być przedstawiony matematycznie jako:

$$\mathbf{o}_j = \sum_{i=1}^{\tilde{N}} \beta_i \varphi(\mathbf{w}_i \cdot \mathbf{x}_j + b_i), \quad j = 1, \dots, N \quad (1)$$

gdzie $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ – wektor wag pomiędzy i -tym neuronem ukrytym a neuronami wejściowymi, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ – wektor wag pomiędzy i -tym neuronem ukrytym a neuronami wyjściowymi, b_i – bias i -tego neuronu ukrytego.

Neurony wyjściowe z uwagi na rozważany problem posiadają liniowe funkcje aktywacji.

SLFN z \tilde{N} neuronów w warstwie ukrytej i funkcją aktywacyjną $\varphi(x)$ może aproksymować N par $(\mathbf{x}_i, \mathbf{t}_i)$, z zerowym błędem – $\sum_{i=1}^{\tilde{N}} \|\mathbf{o}_j - \mathbf{t}_j\| = 0$, więc istnieje takie \mathbf{w}_i, β_i i b_i , że:

$$\sum_{i=1}^{\tilde{N}} \beta_i \varphi(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, \quad j = 1, \dots, N \quad (2)$$

Powyzsze równanie może zostać zapisane w formie:

$$\mathbf{H}\beta = \mathbf{T} \quad (3)$$

$$\begin{aligned}
\mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_{\tilde{N}}, b_1, \dots, b_{\tilde{N}}, \mathbf{x}_1, \dots, \mathbf{x}_N) &= \\
&= \begin{bmatrix} \varphi(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \dots & \varphi(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \dots & \vdots \\ \varphi(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \dots & \varphi(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \quad (4) \\
\boldsymbol{\beta} &= \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \quad \text{i} \quad \mathbf{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m}
\end{aligned}$$

Macierz \mathbf{H} to macierz wyjściowa warstwy ukrytej, i -ta kolumna macierzy \mathbf{H} reprezentuje wartość wyjściową i -tego neuronu w warstwie ukrytej zależną od wartości wejściowych $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$.

Algorytm uczenia SLFN zwany ELM można zapisać w 3 krokach:

Na danym zbiorze uczącym $\mathbf{X} = \{(\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m, i = 1, \dots, N\}$, z funkcją aktywacyjną $\varphi(x)$ i liczbą neuronów ukrytych \tilde{N} należy:

- 1: Ustalić wagi \mathbf{w}_i , wartości biasów b_i , $i = 1, \dots, \tilde{N}$,
- 2: Obliczyć macierz wyjściową warstwy ukrytej \mathbf{H} ,
- 3: Obliczyć wektor wag wyjściowych $\boldsymbol{\beta}$:

$$\boldsymbol{\beta} = \mathbf{H}^T \mathbf{T} \quad (5)$$

4.2. Algorytm OS-ELM

Podstawowy algorytm ELM może zostać użyty tylko dla uczenia off-line, tzn. wszystkie dane uczące muszą być dostępne przy uczeniu. Jeżeli w danej aplikacji wymagane jest uczenie on-line, tzn. potrzebna jest aktualizacja wag do danych nadchodzących w czasie, można użyć algorytmu OS-ELM [6].

W algorytmie OS-ELM (Online Sequential Extreme Learning Machine) tylko specyficzny rodzaj macierzy \mathbf{H} brany jest pod uwagę. Taki, że $\mathbf{H} \in \mathbf{R}^{N \times \tilde{N}}$, $N \geq \tilde{N}$ i $\text{rank}(\mathbf{H}) = \tilde{N}$. W tym przypadku można przedstawić macierz pseudo-odwrotną \mathbf{H} w następujący sposób:

$$\mathbf{H}^+ = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \quad (6)$$

Podstawiając równanie (6) do równania (5) otrzymujemy:

$$\boldsymbol{\beta} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{T} \quad (7)$$

Powyższe równanie (7) jest rozwiązaniem metodą minimalizacji najmniejszych kwadratów równania $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$. Sekwencyjna implementacja rozwiązania najmniejszych kwadratów równania znana jest pod nazwą RLS (Recursive Least Squares).

Bazując na algorytmie RLS algorytm OS-ELM można przedstawić w następujący sposób:

Etap 1: Faza wzmocnienia: znając mały początkowy zbiór danych uczących $X = \{(\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m, i = 1, \dots, \tilde{N}\}$ należy najpierw:

- Ustalić wagi \mathbf{w}_i , wartości biasów b_i i szerokości funkcji RBF σ_i dla $i = 1, \dots, \tilde{N}$.
- Obliczyć początkową macierz wyjścia warstwy ukrytej $\mathbf{H}_0 = [\mathbf{h}_1, \dots, \mathbf{h}_{\tilde{N}}]^T$, gdzie $\mathbf{h}_i = [\varphi(\mathbf{w}_1 \cdot \mathbf{x}_i + b_1), \dots, \varphi(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_i + b_{\tilde{N}})]^T$, $i = 1, \dots, \tilde{N}$.
- Wylczyć początkowy wektor wag wyjściowych $\beta^{(0)} = \mathbf{M}_0 \mathbf{H}_0^T \mathbf{T}_0$, gdzie $\mathbf{M}_0 = (\mathbf{H}_0^T \mathbf{H}_0)^{-1}$ i $\mathbf{T}_0 = [t_1, \dots, t_{\tilde{N}}]^T$,
- Ustawić $k = 0$.

Etap 2: Faza uczenie sekwencyjnego: dla każdej następnej obserwacji $(\mathbf{x}_i, \mathbf{t}_i)$, gdzie $\mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m$ dla $i = \tilde{N} + 1, \tilde{N} + 2, \tilde{N} + 3, \dots$, należy:

- Obliczyć $(k+1)$ -szy wektor macierzy wyjścia warstwy ukrytej: $\mathbf{h}_{k+1} = [\varphi(\mathbf{w}_1 \cdot \mathbf{x}_i + b_1), \dots, \varphi(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_i + b_{\tilde{N}})]^T$.
- Obliczyć nowy wektor wag wyjściowych

$$\mathbf{M}_{k+1} = \mathbf{M}_k - \frac{\mathbf{M}_k \mathbf{h}_{k+1} \mathbf{h}_{k+1}^T \mathbf{M}_k}{1 + \mathbf{h}_{k+1}^T \mathbf{M}_k \mathbf{h}_{k+1}} \quad (8)$$

$$\beta^{(k+1)} = \beta^{(k)} + \mathbf{M}_{k+1} \mathbf{h}_{k+1} (\mathbf{t}_i^T - \mathbf{h}_{k+1}^T \beta^{(k)}) \quad (9)$$

- ustawić k na $k=k+1$.

W pierwszym etapie (faza wzmocnienia) sieć SLFN uczona jest przy użyciu podstawowego algorytmu ELM na pewnej, niewielkiej, części zbioru par danych uczących. W drugim etapie OS-ELM sieć uczona jest sekwencyjnie na danych dostępnych pojedynczo lub w pewnych częściach zbioru uczącego.

Algorytm ELM zapewnia mały błąd generalizacji przy bardzo krótkim czasie uczenia. W pracy [6] pokazano porównanie szybkości uczenia sieci MLP zawierającej 20 neuronów w warstwie ukrytej, algorytmami gradientowymi opartymi na back propagation (BP) oraz ELM, algorytm ELM zadziałał 170 razy szybciej niż algorytm BP, przy dużo niższym błędzie. Rozwiązująca to samo zadanie maszyna wektorów podtrzymujących SVR była ponad 10 tysięcy razy wolniejsza od ELM. to

5. WYNIKI EKSPERYMENTÓW NUMERYCZNYCH

5.1. Cechy modelu matematycznego

Do uczenia sieci ELM użyto danych pomiarowych PM2,5 z lat 2011-2013 podzielonych losowo na części uczącą i testującą w stosunku 70:30. Dodatkowo użyto danych prognoz godzinowych pogody z ICM.

Pojedyncza para ucząca ma postać $[\mathbf{x}_i, y_i]$, gdzie y_i to średnia wartość PM2,5 na dzień następny $d+1$, natomiast \mathbf{x}_i jest to wektor cech i -tego dnia opisujący modelowany problem, przedstawiony jako:

$$\mathbf{x}_i = [PM2,5_d(h_1, \dots, h_{24}), temp_{d+1}(h_1, \dots, h_{24}), wilg_{d+1}(h_1, \dots, h_{24}), \\ ci\acute{s}n_{d+1}(h_1, \dots, h_{24}), zachm_{d+1}(h_1, \dots, h_{24}), opady_{d+1}(h_1, \dots, h_{24}), \\ si\acute{l}aw_{d+1}(h_1, \dots, h_{24}), typdnia_{d+1}, poraroku_{d+1}] \quad (10)$$

gdzie: dla dnia $(d+1)$: $temp_{d+1}(h)$ – jest temperaturą o godzinie h , $wilg_{d+1}(h)$ – jest wilgotnością o godzinie h , $ci\acute{s}n_{d+1}(h)$ – jest ciśnieniem o godzinie h , $zachm_{d+1}(h)$ – jest procentowym zachmurzeniem o godzinie h , $opady_{d+1}(h)$ – jest sumą opadów o godzinie h , $si\acute{l}aw_{d+1}(h)$ – jest siłą wiatru o godzinie h .

Dodatkowe znaczniki to: $typdnia_{d+1}$: 1 – dzień roboczy, 0 – weekend, $poraroku_{d+1}$: 00 – lato, 01 – wiosna, 10 – jesień, 11 – zima.

Przykładowo $PM2,5_d(h_1, \dots, h_{24})$ oznacza zbiór 24 danych godzinowych dla dnia o numerze d . Pełny wektor wejściowy zawiera zatem 171 pozycji co oznacza, że $\mathbf{x} \in \mathbf{R}^{171 \times p}$ a wektor $\mathbf{y} \in \mathbf{R}^{1 \times p}$ gdzie p to ilość par uczących.

5.2. Wyniki predykcji PM2,5

Eksperymenty numeryczne przeprowadzane były 10-krotnie przy każdorazowym losowym podziale danych na część uczącą i testującą. Po przeprowadzeniu uczenia i testowania wyliczane były średnie wartości błędów MAPE i RMSE [1] oraz ich odchylenia standardowe.

Tabela 1. Błędy predykcji dla różnej liczby neuronów n dla sigmoidalnej funkcji aktywacji.

n	$r = 0$		$r = 0,3$		$r = 0,4$		$r = 0,5$		$r = 0,6$	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
5	80,86 ± 5,95	22,29 ± 1,15	67,81 ± 4,40	21,05 ± 1,43	41,54 ± 2,92	13,87 ± 0,81	44,47 ± 3,45	14,27 ± 1,04	30,59 ± 1,44	11,47 ± 0,48
10	64,10 ± 2,91	18,74 ± 1,18	39,57 ± 2,79	12,68 ± 0,84	34,19 ± 1,82	11,49 ± 0,98	38,31 ± 1,92	13,22 ± 1,09	27,44 ± 1,25	10,33 ± 0,42
15	63,43 ± 4,77	17,91 ± 1,54	51,46 ± 2,64	16,92 ± 0,83	52,7 ± 4,09	15,02 ± 0,84	44,15 ± 4,15	14,37 ± 1,31	32,80 ± 1,57	11,53 ± 0,68
20	59,42 ± 3,13	18,70 ± 1,11	45,31 ± 2,13	14,68 ± 1,03	43,07 ± 3,45	14,68 ± 0,55	26,80 ± 1,18	11,01 ± 0,72	28,83 ± 1,02	11,18 ± 1,0
25	59,42 ± 5,22	17,78 ± 1,13	41,15 ± 3,18	14,79 ± 1,33	35,92 ± 1,76	13,04 ± 1,23	31,34 ± 1,99	11,66 ± 0,74	–	–
30	54,59 ± 3,01	15,60 ± 1,17	35,82 ± 2,18	11,93 ± 0,90	30,59 ± 1,23	12,01 ± 0,83	25,63 ± 1,33	10,31 ± 0,98	–	–
35	51,94 ± 3,49	15,07 ± 1,20	36,63 ± 1,98	12,81 ± 0,67	28,39 ± 1,45	10,69 ± 1,07	27,02 ± 1,37	10,55 ± 1,07	–	–
40	54,49 ± 4,08	5,61 ± 1,06	28,81 ± 1,81	10,32 ± 0,98	27,57 ± 1,66	10,65 ± 0,75	27,89 ± 1,66	10,83 ± 0,7	–	–
45	45,68 ± 2,02	14,34 ± 0,94	31,57 ± 1,41	11,07 ± 1,22	28,31 ± 1,45	11,32 ± 0,87	26,68 ± 1,14	10,68 ± 1,77	–	–
50	47,45 ± 3,41	14,30 ± 0,73	32,86 ± 1,57	11,08 ± 0,86	29,11 ± 2,00	1,02 ± 0,56	–	–	–	–

Przeprowadzono szereg badań przy różnych strukturach wektorów wejściowych, czyli przy różnych zestawach cech procesu. Przeprowadzono badania korelacji poszczególnych cech z wartością wyjściową – prognozowaną wartością średnią na następną dobę. Do wejścia sieci neuronowej włączano elementy wektora x (wzór 11), dla których współczynniki korelacji z wyjściem wynosiły odpowiednio: $r = 0.3; 0.4; 0.5; 0.6$ oraz dla $r = 0$ (wszystkie cechy). Tabela 1 przedstawia uzyskane wyniki prognozy w sensie błędów MAPE i RMSE dla sigmoidalnej funkcji aktywacji sieci ELM dla różnych liczb n neuronów w warstwie ukrytej sieci.

Brak wyników predykcji dla niektórych kombinacji parametrów wynikają z ilości danych wejściowych, ELM wykazuje działanie suboptymalne dla sieci z większą ilością neuronów w warstwie ukrytej niż danych wejściowych.

5. WNIOSKI KOŃCOWE

W pracy przedstawiono metodę predykcji średniodobowego poziomu PM_{2,5} na następny dzień wykorzystującą algorytm ELM. Uzyskano wyniki predykcji ze średnim bezwzględnym błędem procentowym poniżej 26% oraz RMSE na poziomie 10.3. Dla porównania podobne badania przeprowadzono przy użyciu sieci SVR (Support Vector machine for Regression) i uzyskano zbliżone lub gorsze rezultaty [1].

Złożoność problemu predykcji PM_{2,5} wynika między innymi z zależności lokalnego poziomu zanieczyszczeń powietrza atmosferycznego z lokalnymi warunkami pogodowymi, które zmieniają się w sposób bardzo szybki i przypadkowy. Kolejnym krokiem wykonania dokładniejszej predykcji będzie agregacja wyników wielu prognoz obliczonych różnymi metodami – neuronowymi jak i klasycznymi. Ciekawym podejściem jest również zastosowanie selekcji cech opartej na algorytmach genetycznych a nie współczynnika korelacji liniowej Pearsona [1].

Zastosowanie zaawansowanych metod neuronowych predykcji daje dobre rezultaty, a prognozy mogą mieć zastosowanie praktyczne.

LITERATURA

- [1] Siwek K., Osowski S., Data mining methods for prediction of air pollution, *Int. J. Appl. Math. Comput. Sci.*, Vol. 26, No. 2, 467–478, 2016.
- [2] Rückerl R., Schneider A., Breitner S., Cyrys J., Peters A., Health effects of particulate air pollution: A review of epidemiological evidence, *Inhalation Toxicology*, 23(10): 555–592, 2011.
- [3] Raport Europejskiej Agencji Środowiska (EEA), Air quality in Europe, 2016.
- [4] Huang G.B., Zhu Q.Y., Siew C.K., Extreme Learning Machine: Theory and applications, 1-10, 2016.

- [5] Huang G.B., Zhu Q.Y., Siew C.K., Extreme learning machine: a new learning scheme of feedforward neural networks, IEEE International Joint Conference on Neural Networks, Budapest, Hungary, 2004.
- [6] Huang G.B., Liang N.Y., Rong H.J., Saratchandran P., Sundararajan N., On-Line Sequential Extreme Learning Machine, The IASTED International Conference on Computational Intelligence (CI 2005), Calgary, Canada, July 4-6, 2005.

PM_{2.5} PARTICULATE POLLUTION FORECASTING USING ELM

The article presents a new approach to atmospheric dust prediction using an ELM (Extreme Learning Machine) neural network. This prediction concerns the calculation of the average level of PM_{2.5} air pollution in Warsaw's Ursynów one day ahead. To understand the significance of the problem, a brief description of the hazards posed by PM_{2.5} air pollution is included.

The work presents a short description of the ELM network and presents the obtained forecast results. The analysis of the obtained results was carried out. The validity of forecasting was discussed and possible preventive and protective measures were indicated.

(Received: 18.02.2018, revised: 6.03.2018)