

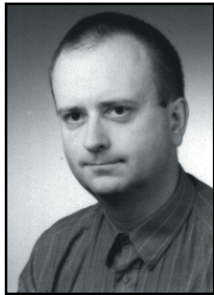
Tomasz GUSZKOWSKI, Janusz MROCZKA

KATEDRA METROLOGII ELEKTRONICZNEJ I FOTONICZNEJ, POLITECHNIKA WROCŁAWSKA

Rozwiązanie problemu odwrotnego w turbidymetrii z wykorzystaniem sieci neuronowej

Mgr inż. Tomasz GUSZKOWSKI

Absolwent Wydziału Elektroniki Politechniki Wrocławskiej, obecnie pracownik Katedry Metrologii Elektronicznej i Fotonicznej Politechniki Wrocławskiej w zespole prof. Mroczy. Jego zainteresowania naukowe obejmują metody sztucznej inteligencji, sposoby rozwiązania problemu odwrotnego oraz konstrukcję inteligentnej aparatury pomiarowej.



e-mail: tomasz.guszkowski@pwr.wroc.pl

Prof. dr hab. inż. Janusz MROCZKA

Kierownik Katedry Metrologii Elektronicznej i Fotonicznej Politechniki Wrocławskiej. Zajmuje się metodologią obserwacji i eksperymentu, algorytmizacją problemu odwrotnego, modelowaniem matematycznym pól fizycznych, analizą spektralną i polaryzacyjną promieniowania rozproszonego, reprezentacjami czasowo-częstotliwościowymi w przetwarzaniu danych.



e-mail: janusz.mrocza@pwr.wroc.pl

Streszczenie

Praca prezentuje metodę rozwiązania problemu odwrotnego w turbidymetrii z wykorzystaniem sieci neuronowej. Problem odwrotny jest często problemem źle postawionym i/lub źle uwarunkowanym numerycznie, a jego rozwiązanie przeważnie możliwe jest po nałożeniu na możliwy wynik dodatkowych warunków. Jako przykład zaprezentowano rozwiązanie problemu pomiaru wielkości cząstek na podstawie pomiarów turbidymetrycznych wykonanych dla kilku długości fali świetlnej.

Słowa kluczowe: problem odwrotny, sieć neuronowa, turbidymetria, rozkład wielkości cząstek

A Solution to Inverse Problem in Turbidimetry with Use of Artificial Neural Network

Abstract

The paper presents a proposed solution to an inverse problem in turbidimetry with Artificial Neural Network (ANN). The inverse problem is mostly ill-posed and/or ill-conditioned, and its solution is usually possible after introduction of additional constrain to the proposed result. As an example inversion a determination of water particle size distribution based on multispectral turbidity measurements was shown.

Keywords: inverse problem, Artificial Neural Network, turbidimetry, particle size distribution

1. Wstęp

W badaniu turbidymetrycznym wiązka światła o określonej długości fali jest osłabiana podczas przejścia przez badaną próbkę. Stopień osłabienia zależy od długości padającej fali, rozmiaru cząstki oraz względnego współczynnika załamania cząstki względem otaczającego ją ośrodka. Jednakże tradycyjne metody określania rozkładu wielkości cząstek bazujące na wartościach mętności, zarówno bezpośrednie, jak i iteracyjne, np. [1-5], nie zawsze dobrze działają, np.: wprowadzają silne oscylacje, zezwalają na wyniki, które są nierealne fizycznie, lub wymagają kłopotliwego doboru parametrów. Cechą wspólną tych metod jest również duży koszt obliczeniowy – związane jest to ze złym uwarunkowaniem numerycznym problemu. W pracy podjęto próbę wykorzystania sztucznej sieci neuronowej (SSN) jako bloku odwracającego przy wykorzystaniu znajomości typu rozkładu jako ograniczenia zabezpieczającego przed oscylacjami wyniku odwracania. Dzięki temu rozwiązanie problemu odwrotnego sprowadza się do identyfikacji parametrów rozkładu.

1.1. Założenia metody

W niniejszej pracy założono, że rozkład wielkości cząstek może być aproksymowany rozkładem normalnym lub logarytmiczno-normalnym. SSN została wykorzystana do estymacji parametrów tych rozkładów. Na wejście sieci neuronowej podawano względne wartości mętności dla kilku długości fali w odniesieniu do mętności dla długości fali odniesienia, na wyjściu uzyskując średnią i odchylenie standardowe dla rozkładu normalnego oraz geometryczną średnią i geometryczne odchylenie standardowe dla rozkładu logarytmiczno-normalnego.

Operowanie mętnością względną, czyli stosunkiem mętności dla danej długości fali do mętności dla długości fali odniesienia, pozwoliło zminimalizować wpływ koncentracji cząstek na wyniki odwracania.

2. Mętność

Wiązka światła napotykać cząstkę o współczynniku refrakcji różnym od współczynnika refrakcji otaczającego ośrodka podlega rozproszeniu i absorpcji, co objawia się tłumieniem tej wiązki. Dla próbki składającej się z cząstek o niewielkim stężeniu (< 5%), dla której efekty wielokrotnego rozproszenia oraz interakcje pomiędzy cząstkami (np. tworzenie agregatów) mogą być zaniebane, mętność τ dla długości fali λ jest dana wyrażeniem:

$$\tau(\lambda) = \frac{c\pi}{4} \int_0^{\infty} D^2 Q_{ext}(D, \lambda, m) f(D) dD \quad (1)$$

gdzie c jest stężeniem (koncentracją), D średnicą cząstek sferycznych, m względnym zespolonym współczynnikiem załamania, Q_{ext} współczynnikiem ekstynkcji, a $f(D)$ rozkładem wielkości cząstek. Współczynnik ekstynkcji Q_{ext} zależy od względnego współczynnika załamania i parametru rozmiaru α , $\alpha = \pi D / \lambda$ i może zostać wyznaczony z teorii rozproszenia światła Mie. Wykorzystując α równanie (1) można przekształcić do postaci:

$$\tau(\lambda) = \frac{c\lambda^3}{4\pi^2} \int_0^{\infty} \alpha^2 Q_{ext}(\alpha, \lambda) f(\alpha) d\alpha \quad (2)$$

co jest postacią wygodniejszą do obliczeń od (1) ze względu na dostępność bibliotek zawierających procedury obliczania Q_{ext} w zależności od α .

3. Określanie rozkładu wielkości cząstek

Założono, że rozkład wielkości cząstek jest reprezentowany przez rozkład Gaussa:

$$f(D) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(D-D_m)^2}{2\sigma^2}\right) \quad (3)$$

gdzie D_m jest średnią, a σ odchyleniem standardowym rozkładu, lub rozkład logarytmiczno-normalny:

$$f(D) = \frac{1}{\ln(\sigma)\sqrt{2\pi}D} \exp\left(-\frac{(\ln D - \ln D_m)^2}{2\sigma^2}\right) \quad (4)$$

gdzie D_m jest średnią geometryczną, a σ geometrycznym odchyleniem standardowym.

3.1. Mętność względna

Mętność dla określonej długości fali zależy nie tylko od rozkładu wielkości cząstek, ale również od ich koncentracji (2). Aby wyeliminować wpływ stężenia, jako dane wejściowe dla sieci neuronowej zostały wybrane stosunki mętności dla określonej długości fali do mętności dla długości fali odniesienia λ_0 .

$$R(\lambda, \lambda_0) = \frac{\int_0^{\infty} D^2 Q_{ext}(D, \lambda, m) f(D) dD}{\int_0^{\infty} D^2 Q_{ext}(D, \lambda_0, m_0) f(D) dD} \quad (5)$$

Dla danej długości fali odniesienia λ_0 i długości fali pomiarowej λ stosunek $R(\lambda, \lambda_0)$ jest zdeterminowany tylko przez rozkład wielkości cząstek $f(D)$. Zależność od koncentracji została wyeliminowana. Stwierdzenie to jest prawdziwe dla małych stężeń, gdzie nie występuje wielokrotne rozproszenie.

W pracy założono, że medium rozpraszającym są cząstki wodnego aerozolu. W przypadku szerokich zmian długości fali światła λ istotne staje się uwzględnienie zmian wartości współczynnika refrakcji m . Wartości funkcji $m(\lambda)$ dla wody zostały wyznaczone na podstawie [6-7]. Założono temperaturę $T=300$ K. Do obliczeń współczynnika Q_{ext} wykorzystano przetłumaczoną na język C procedurę opublikowaną w [8]. Wykorzystano procedurę całkowania metodą Simpsona.

3.2. Architektura SSN

Wykorzystano jednokierunkową sieć neuronową z nieliniową funkcją aktywacji i jedną warstwą ukrytą. Liczba wejść sieci neuronowej (równa liczbie zmierzonych mętności względnych) jest istotnym parametrem, gdyż zbyt mała ich liczba spowoduje utratę informacji i mało dokładne dopasowanie sieci neuronowej. Zbyt duża liczba wejść sprawi, że sieć będzie trudna do nauczenia. Na podstawie analizy symulacji mętności wykonanych dla długości fali λ od 100 nm do 1 μ m wybrano pięć długości fali pomiarowej λ (300, 400, 450, 550, 650 nm) przy długości fali odniesienia $\lambda_0 = 500$ nm, zakładając, że pięć pomiarów $R(\lambda, \lambda_0)$ wystarczy do odtworzenia pierwotnych parametrów rozkładów. Ważna jest również prawidłowo dobrana liczba neuronów w warstwie ukrytej – zbyt mała powoduje mało dokładne dopasowanie, zbyt duża – problemy z przeuczeniem.

3.3 Zbiory uczące, walidujące i testujące

Wygenerowano dwa zestawy zbiorów uczących – po jednym dla każdego z rozważanych rozkładów. Zasympulowano 910 wektorów uczących dla rozkładu normalnego obejmującego zmiany D_m od 100 nm do 1 μ m oraz zmiany σ od 1 do 19 nm rozmieszczonych liniowo. Dla rozkładu logarytmiczno-normalnego wygenerowano 500 wektorów uczących w zakresie zmian D_m od 100nm do 1 μ m, rozmieszczonych logarytmicznie, oraz zmian σ od 1,02 do 1,52 nm rozmieszczonych liniowo. Ze względu na całkowanie numeryczne wymagane do obliczenia zależności (5) oraz złożoność procedur obliczających Q_{ext} było to zadanie o dużym koszcie obliczeniowym. Obliczenia wykonano z wykorzystaniem kompilatorów: dla języka C – gcc, oraz dla języka FORTRAN – g77.

W celu bieżącej kontroli stopnia nauczenia sieci i ustalenia liczby neuronów w warstwie ukrytej wygenerowano po 100 zestawów walidujących działające sieci dla każdego z rozkładów. Rozmieszczenie parametrów rozkładów w ciągu walidującym w podanych wyżej przedziałach było zrandomizowane.

Ostateczny zbiór testowy obejmował 20 wybranych losowo zestawów danych i służył do końcowej oceny działania sieci neuronowej. Zbiory uczące, walidujące i testujące były rozłączne.

Łącznie przygotowanie wszystkich zbiorów wymaganych do uczenia i oceny jakości nauczenia sieci trwało około 3 godzin na standardowym komputerze PC z zegarem 1,3 GHz.

3.4. Uczenie sieci neuronowej

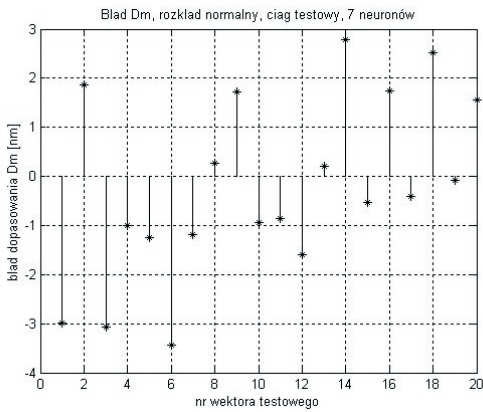
Jako kryterium dopasowania przyjęto błąd średniokwadratowy MSE. Ostatecznie przeprowadzono uczenie sieci algorytmem Levenberga-Marquardta dla liczby neuronów ukrytych od 4 do 10, po 20 serii uczących dla każdej liczby neuronów ukrytych, startując za każdym razem z innego, losowo wybranego zestawu wag sieci (łącznie 560 uczeń sieci – 20 razy 7 wariantów architektury sieci razy 2 parametry razy 2 rozkłady) i zapamiętując sieć o najmniejszym błędzie MSE dla ciągu walidującego. Wykorzystano dwuwarstwową sieć neuronową o 5 wejściach i jednym wyjściu – każdy parametr rozkładu estymowany był z użyciem oddzielnej sieci. Utworzono odrębne sieci do estymacji D_m i σ , ze względu na fakt, iż sieć o jednym wyjściu jest łatwiej nauczyć, niż sieć o dwóch wyjściach [9]. Warstwa ukryta każdej z sieci miała funkcję aktywacji tangens hiperboliczny, warstwa wyjściowa charakteryzowała się liniową funkcją aktywacji.

Najmniejszy błąd MSE ciągu walidującego dla rozkładu normalnego i parametru D_m uzyskano dla sieci o 7 neuronach w warstwie ukrytej, dla parametru σ dla sieci o 10 neuronach w warstwie ukrytej. Dla rozkładu logarytmiczno-normalnego wartości te wynosiły odpowiednio 6 dla parametru D_m i 10 dla σ .

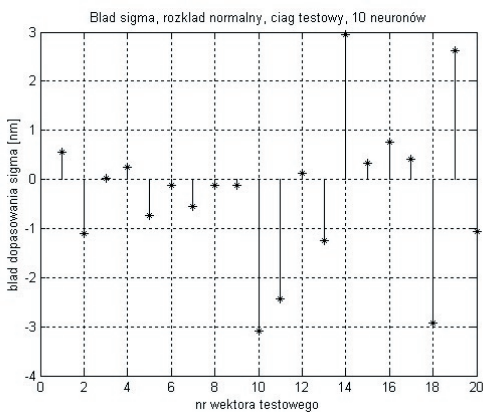
Ze względu na powtarzanie cyklu uczenia dla dobowanych losowo wag sieci oraz ze względu na automatyczny dobór najlepszej liczby neuronów w warstwie ukrytej pełne uczenie sieci neuronowych trwało kilka godzin z wykorzystaniem programu MATLAB.

4. Wyniki

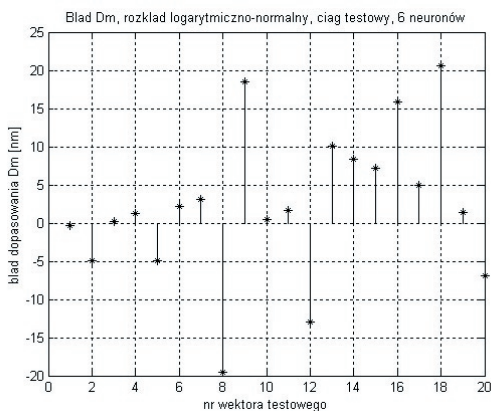
Po etapie uczenia działanie sieci ostatecznie przetestowano z użyciem 20 wektorów testujących, odrębnie dla każdego parametru. Błąd bezwzględny uzyskany dla ostatecznego ciągu testowego zaprezentowano na rys. 1-2 dla rozkładu normalnego i 3-4 dla rozkładu logarytmiczno-normalnego. Testowanie przeprowadzono dla najlepszej znalezionej liczby neuronów ukrytych i najlepszych zestawów wag znalezionych podczas uczenia (kryterium oceny – zachowanie sieci dla ciągu walidującego).



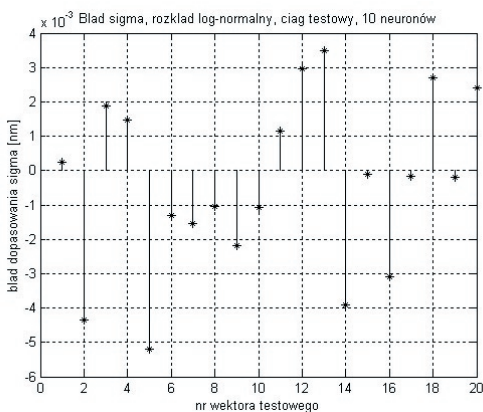
Rys. 1. Błąd odtworzenia D_m , rozkład normalny
Fig. 1. Error of D_m estimation, normal distribution



Rys. 2. Błąd odtworzenia σ , rozkład normalny
Fig. 2. Error of σ estimation, normal distribution



Rys. 3. Błąd odtworzenia D_m , rozkład logarytmiczno-normalny
Fig. 3. Error of D_m estimation, logarithmic-normal distribution



Rys. 4. Błąd odtworzenia σ , rozkład logarytmiczno-normalny
Fig. 4. Error of σ estimation, logarithmic-normal distribution

5. Dyskusja i wnioski

Zaproponowano nową metodę określania parametrów rozkładu wielkości cząstek na podstawie kilku pomiarów mętności względem mętności dla wybranej długości fali odniesienia λ_0 . Symulacje przeprowadzone dla wodnego aerozolu w powietrzu wykazały przydatność opracowanej techniki i celowość wykorzystania sztucznej sieci neuronowej do określania parametrów rozkładu. Zarówno parametry rozkładu normalnego, jak i logarytmiczno-normalnego były określane przez sieć neuronową. Sieci łatwiej uczyły się określania parametru D_m , niż σ . Mogło to wynikać z większej wrażliwości mętności względnej $R(\lambda, \lambda_0)$ na zmiany tego parametru. Zmiany σ objawiały się niewielkim wpływem na wartości $R(\lambda, \lambda_0)$. Metoda charakteryzuje się prostotą użycia, niezależnością od koncentracji cząstek w zakresie obowiązywania prawa Lamberta-Beera, oraz możliwością wykorzystania w aparatach mierzących rozkład wielkości cząstek ze względu na bardzo niski koszt obliczeniowy, związany z symulowaniem nauczonej sieci, co może być istotne w zastosowaniach przemysłowych. Niestety, duży koszt obliczeniowy związany jest zarówno z symulacyjną generacją danych niezbędnych do uczenia sieci, jak i z samym uczeniem. Jednak zarówno przygotowanie danych do uczenia, jak i uczenie wykonuje się dla danych warunków pomiarowych tylko raz, a odpowiednio wytrenowanej sieci można używać wiele razy.

Prawdopodobnie możliwa jest dalsza poprawa dokładności estymacji parametrów w przypadku użycia bardziej licznego zbioru uczącego (szczególnie dla rozkładu logarytmiczno-normalnego) i przeprowadzenia większej liczby prób uczenia sieci. Wiąże się to jednak z dużo większym kosztem obliczeniowym poniesionym w trakcie uczenia.

Istnieje możliwość zaadaptowania proponowanej metody tak, aby ciąg uczący składał się z danych pomiarowych, a nie symulowanych. Pozwoliłoby to, między innymi, wyeliminować problemy związane np. z koniecznością znajomości zespolonego współczynnika refrakcji m w funkcji długości fali λ , co jest niezbędne do obliczeniowej generacji ciągu uczącego.

6. Literatura

- [1] Phillips D.L.: A Technique for the Numerical Solution of Certain Integral Equations of the First Kind; J. Assoc. Comp. Mech., 1962, 9, 84-97.
- [2] Twomey S.: On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature; J. Assoc. Comp. Mech., 1963, 10, 97.
- [3] Hansen P.C.: Truncated SVD Solutions to Discrete Ill-posed Problems with Ill-Determined Numerical Rank; SIAM J. Sci. Stat. Comput., 1990, 11, 503.
- [4] Markowski G.R.: Improving Twomey's Algorithm for Inversion of Aerosol Measurement Data; Aerosol Sci. and Tech. 1987, 7:127.
- [5] Chahine M.T.: Determination of the Temperature profile in an Atmosphere from its Outgoing Radiance; J. Opt. Soc. Am, 1968, 58, 1634.
- [6] Hale G.M., Query M.R.: Optical Constant of Water in the 200-nm to 2000nm Wavelength Region; Applied Optics, 1973, 2, 555.
- [7] Pope R.M., Fry E.S.: Absorption spectrum (380-700 nm) of Pure Water. II. Integrating Cavity Measurements; Applied Optics, 1997, 36(33): 8710-23.
- [8] Bohren C.F., Huffman D.R.: Absorption and Scattering of Light by Small Particles; Wiley, New York, 1983.
- [9] Blum E.K., Li L.K.: Approximation Theory and Feedforward Networks; Neural Networks, 1991, 4(4), 511.