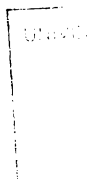

UNIVERSITATEA "POLITEHNICA" TIMIȘOARA
FACULTATEA DE CHIMIE INDUSTRIALĂ ȘI INGINERIA MEDIULUI

**CONTRIBUȚII LA UTILIZAREA
REȚELEOR NEURALE
ÎN
CHIMIE ȘI INGINERIE CHIMICĂ**

Ing. Mircea Geantă

TEZĂ DE DOCTORAT

BIBLIOTECA CENTRALĂ
UNIVERSITATEA "POLITEHNICA"
TIMIȘOARA



618 615
366 A

CONDUCĂTOR ȘTIINȚIFIC:
Prof. dr. ing. Delia Perju

TIMIȘOARA
1997

INTRODUCERE

Crearea și utilizarea Rețelelor Neurale Artificiale a fost inițial motivată de încercarea de a simula prin calcul modalitatea de percepție și recunoaștere a formelor și de memorare a informației de către cortexul cerebral. La momentul elaborării primului model matematic (McCulloch&Pitts - 1943), trecuseră 42 de ani de la accepția *neuronului* ca și constituent structural al creierului fără a se fi clarificat însă exact funcționalitatea sa.

Ceea ce este curios însă, după relevarea modului de funcționare a cortexului cerebral (1963), până la apariția unor schimbări majore în principiul de bază al algoritmilor ce descriu Rețelele Neurale Artificiale a fost necesară o perioadă îndelungată de timp (Minsky -1969- prin perceptronul multistrat și Rumelhart - 1986- prin algoritmul de propagare înapoi).

Pentru întârzierea reacției există două posibile explicații:

- ❖ disiparea atenției cercetătorilor în domenii diferite de motivația inițială, algoritmilor de calcul fiindu-le găsite câmpuri de aplicare de o remarcabilă diversitate.
- ❖ încercarea de a modela procesarea informației de către cortex a dus la sesizarea imensei diferențe care există între randamentul și capacitatea sistemului neural natural și cel artificial. În acest mod, propunerea unui scop mult prea îndepărtat nu constituie un gradient stimulatoriu pentru abordarea acestei direcții de cercetare.

Progresul conceptual și matematic al algoritmilor care descriu Rețelele Neurale Artificiale s-a datorat încercării de adaptare a acestora la aplicațiile pentru care au fost creați. Astfel, având în vedere că majoritatea aplicațiilor rețelelor neurale se îndepărtează de scopul inițial, s-a propus chiar înlocuirea denumirii generice de Rețele Neurale Artificiale cu altele considerate a fi mai adecvate, cum ar fi *modele conexiuniste* sau *modele de procesare paralel-distribuită*. Denumirea inițială, este însă atât de uzitată, încât renunțarea la ea ar presupune un “mic haos” în literatura de profil.

În ce privește motivația inițială, s-au efectuat progrese majore, domeniul recunoașterii formelor găsiindu-și aplicații la roboții industriali în identificarea și prelucrarea adecvată a subsansamblelor într-un proces complet automatizat, comanda verbală a roboților și a computerelor, sisteme expert în aplicații civile și militare, sisteme de operare pentru calculatoare dedicate aplicațiilor speciale. Din acest punct de vedere, alături de alte tehnici de vârf, Rețelele Neurale Artificiale fac parte din domeniul intitulat *Inteligență Artificială*.

Inclusiv în ceea ce privește denumirea actuală există două alternative - Rețele Neurale și Rețele Neuronale - opțiunea pentru una din ele fiind o chestiune de interpretare personală. În aceasta lucrare s-a utilizat denumirea de Rețele Neurale mizându-se pe diferența de nuanță dintre termeni. Astfel, chiar dacă ambele denumiri provin de la *neuron*, termenul *neuronale* se referă cu predilecție la o unitate nervoasă, pe când termenul *neural* se referă la sistemul în ansamblu, deci la o rețea

de neuroni, fapt ce presupune o analogie mai apropiată cu modalitatea de utilizare a acestora.

Diversitatea mare de câmpuri aplicative pentru Rețelele Neurale Artificiale a dus la o dezvoltare concurențială dar de multe ori paralelă a algoritmilor, fapt ce a determinat o adevărată explozie a publicisticii în domeniu.

Având în vedere publicațiile specializate pe domenii restrânse, acestea abundă în lucrări cu caracter pretins inovator, deși se referă nu numai la un domeniu restrâns de aplicabilitate ci este specific acestuia doar pe subdomenii limitate. Pretinsele inovații sun în general observații de bun simț coroborate cu seturi de date de o specificitate restrânsă, fără a se putea efectua o generalizare a concluziilor.

În acest context, firmele producătoare de soft comercial au excelat în realizarea unor programe cu aplicabilitate extinsă, oferite unei palete cât mai largi de utilizatori. Caracteristica de baza a acestor produse soft nu este însă profesionalismul abordării, ci accesibilitatea utilizatorului la funcțiile de bază ale programului. Chiar dacă accesul la informație este rapid, promovarea celor mai performanți algoritmi sau a tehnicilor de operare proprii unor clase de aplicații se regăsesc în programele comerciale cu o întârziere substanțială.

În lucrarea de față, sunt vizate acele aplicații ale Rețelelor Neurale Artificiale în care se urmărește atât capacitatea de învățare a acestora cât și mecanismul de învățare și capacitatea de predicție pentru sisteme în care dependența variabilelor de ieșire de cele de intrare este complexă.

Dacă pentru dependențe ieșire-intrare simple, corelarea prin regresie multiliniară sau modelele matematice rezolvate numeric sunt preferabile, pentru dependențe în care majoritatea factorilor au o variație de tipul sugerat în figura 1, prin interferență generând o interdependență complexă, modelele corelaționale clasice sunt nesatisfăcătoare.

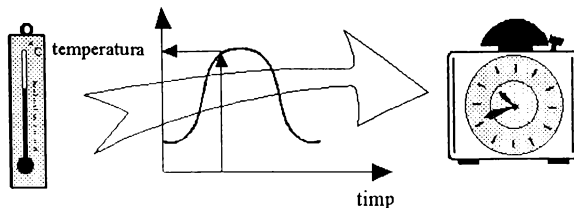


Fig. 1 Variație neliniară a temperaturii într-un proces ipotetic

Modelele neurale se adresează exact acestor tipuri de procese în care dependența intrare-ieșire este deosebit de complexă și în care selecția și cuprinderea într-un model analitic a variabilelor este dacă nu imposibilă, foarte dificilă.

Bazată pe programe de concepție proprie, deși studiază atât algoritmi de învățare cât și procesul de învățare în sine, **lucrarea de față nu își atribuie un caracter pregnant inovator în domeniul Rețelelor Neurale Artificiale** ci este axată pe o **utilizare rațională și atent motivată de particularitățile aplicației**, subliniindu-se superioritatea acestei forme de modelare în detrimentul modalităților clasice de abordare a fenomenelor specifice chimiei și ingineriei chimice.

I.1 Inspirația din neurologie

I.1.1 Celula nervoasă

Celula nervoasă sau **neuronul** - este elementul fundamental al sistemului nervos central; Schematic, în figura 1 este redat un lanț simplist de neuroni, având menționate componentele semnificative din punctul de vedere al utilității în modelele matematice ulterioare¹:

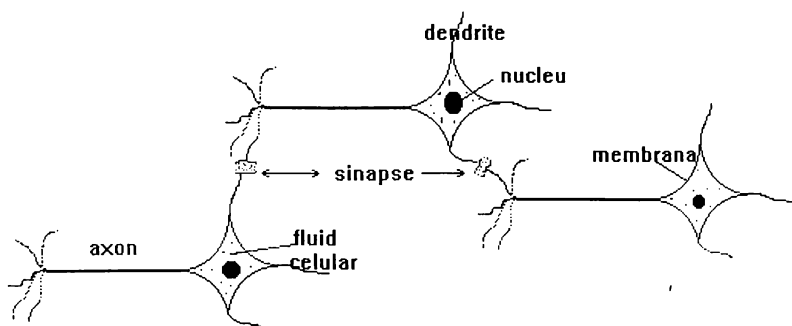


Fig. 1 Conexiune neurală simplistă

Celula nervoasă are diametrul de ordinul micronilor și este alcătuită din:

- - **corpul celular** - conține nucleul neuronului și guvernează transformările biochimice de sinteză a enzimelor necesare funcționalității celulei nervoase
- - **dendritele** - sunt extensii filiforme cu ramificații arborescente ale corpului celular, având diametrul de ordinul zecimilor de microni și lungimi de până la 10 microni. Dendritele sunt principalii receptori de semnal ai neuronului.
- - **axonul** - este o prelungire fibroasă - de lungime considerabilă în comparație cu restul celulei nervoase - variind între 0.2 -1 centimetri pentru neuronii din cortex și 10 centimetri pentru neuronii periferici. Axonul reprezintă conexiunea de ieșire a semnalului procesat de neuron.

Modalitatea de interconectare spațială a celulelor nervoase este realizată prin intermediul unor zone cu o conductanță electrică îmbunătățită denumite **sinapse** și care constituie practic o legătură între axonul unui neuron și o dendrită a unui alt neuron cu rol de comunicare între cele două celule.

Eficiența energetică² a creierului uman este de aproximativ 10^{-16} J/(operație·s), în timp ce, cele mai performante calculatoare au o eficiență de 10^{-6} J/(operație·s).

I.1.2. Sinapsa

În funcție de conexiuni, există două categorii de sinapse:

1. neuro-neuronală
2. neuro-musculară

Ca mărime a zonei de conexiune, ambele tipuri se încadrează între limitele: 0.5 - 3 μ . Numărul de butoni sinaptici de pe suprafața unui neuron acoperă aprox. 38% din suprafața acestuia și a dendritelor sale.

Cercetări recente au arătat că sinapsa convertește un semnal presinaptic de tip electric în semnal chimic și apoi din nou în semnal electric⁴.

Se admite că sinapsele pot fi excitatoare sau inhibitoare în funcție de mediatorul chimic care asigură trecerea fluxului nervos de la un neuron la altul.

Mediatori excitatori - Acetilcolina (ester al colinei) este depozitată intracelular în veziculele sinaptice sub forma unui complex acetilcolină-proteină. Mediatorul nu difuzează extracelular deoarece membrana celulară nu este permeabilă pentru bazele cuaternare de amoniu între care se află și colina și implicit derivatul ei acetilat.

În momentul în care influxul nervos presinaptic ajunge la nivelul sinapsei, acetil-colina este eliberată, difuzează în spațiul lacunar al sinapsei și se combină cu o proteină din membrana neuronului postsinaptic (receptorul de acetilcolină). Astfel, configurația moleculară a membranei postsinaptice se schimbă brusc și devine permeabilă pentru ioni. Ca urmare, are loc o pătrundere bruscă de Na⁺ în interiorul celulei, fapt ce produce o depolarizare a membranei (fenomenul de potențial postsinaptic excitator). Acest proces durează o milisecundă deoarece intervine o altă enzimă - colinesteraza (prezentă în toate celulele) - care descompune acetilcolina din cuplul acetilcolin-proteină în acid acetic și colină.

În urma acestei descompuneri, membrana celulară își reia configurația sa moleculară anterioară.

Impulsul neuronal se transmite de-a lungul membranei celulare din aproape în aproape prin depolarizări (+) și repolarizări (-) succesive cu o viteză care poate ajunge până la 120m/s.

Mediatori inhibitori - Sunt reprezentați de acidul gamma-amino-butaric provenit din decarboxilarea acidului glutamic și mai ales acidul gamma-amino-beta-hidroxi-butaric care are o acțiune de 50 de ori mai puternică decât precedentul.

I.1.3. Neuronul - unitate funcțională³

Neuronul apare la viermii platelminți, evoluând pe scara animală, numărul lor fiind de $\approx 4 \cdot 10^{11}$ în organismul uman dintre care aproximativ $1.6 \cdot 10^{11}$ în cortexul cerebral.

Intrarea în acțiune a neuronului se face printr-un fenomen electric de depolarizare, adică o negativare la exterior a membranei celulare. Membrana celulară are o constituție lipoproteică și este perforată de pori cu un diametru de 3Å

prin care poate trece apa și anumiți ioni. Având în vedere că diametrul hidratat al K^+ este de $2,2\text{\AA}$ iar cel al Na^+ de $3,4\text{\AA}$, permeabilitatea membranei față de Na^+ este de ≈ 100 ori mai mare decât față de K^+ .

Lichidul intracelular și cel extracelular sunt oarecum similare din punctul de vedere al numărului particulelor dizolvate în unitatea de volum, dar sunt diferite din punctul de vedere al concentrației ionice și al potențialului electric.

Concentrația Na^+ și Cl^- este de 12 respectiv 30 ori mai mare în lichidul extracelular decât în cel intracelular în timp ce concentrația K^+ este de 39 ori mai mare în lichidul intracelular decât în cel extracelular.

Astfel, există în repaus un potențial de membrană de $\approx 60-90$ mV. În repaus, membrana este mult mai permeabilă pentru K^+ decât pentru Na^+ . În momentul activării neuronului se produce un flux ionic: sub acțiunea stimulului nervos existent la acest nivel, prin mărirea diametrului porilor crește permeabilitatea pentru Na^+ . Se produce astfel un aflux de Na^+ în interiorul pericarionului (corpului celular) pe care îl pozitivează iar suprafața membranei se negativează. Acest aflux de Na^+ durează ≈ 1 milisecundă. Ca expresie a depolarizării (negativării) membranei se produce un potențial local de acțiune de $\approx 90-120$ mV, deci o diferență de 30 mV față de curentul de repaus.

Ruperea bruscă a echilibrului ionic este urmată de o fază de repolarizare (repozitivare) a suprafeței membranei celulare prin migrarea K^+ din interiorul pericarionului către exterior (eflux de K^+) care contrabalansează migrarea inițială a Na^+ . După trecerea influxului nervos are loc o redistribuire ionică, astfel Na^+ va trece în spațiul extracelular iar K^+ în cel intracelular printr-o pompă activă denumită **pompa de Na^+ și K^+** . S-a calculat că $\approx 20\%$ din metabolismul celular este utilizat de această pompă ionică. Potențialul de acțiune odată generat, este transmis mai departe de-a lungul fibrei nervoase.

Mecanismul propagării fluxului este interpretat conform teoriei membranei ca un mecanism de suprafață. La fel ca membrana celulară, și fibra nervoasă prezintă aceeași diferență de potențial între interiorul și exteriorul ei. Când stimulul ajunge într-un anumit punct al fibrei se produce - ca și la pericarion - o redistribuție ionică și deci, o depolarizare (negativare) a membranei. Punctul stimulat devine electronegativ față de porțiunea inactivă (polarizată) din vecinătate; apare un diferență de potențial locală care face să se depolarizeze regiunea apropiată, fenomenul continuând până la punctul de propagare final de-a lungul nervului.

Deci, conform teoriei membranei, impulsul nervos ar fi o propagare a unei unde de depolarizare din aproape în aproape de-a lungul fibrei.

1.1.4. Concluzii

Neuronul uman are caracteristici comune cu cele ale celulelor vii în general, însă ca organizare și sistem biochimic prezintă individualități ce permit specializarea lor în realizarea unor funcțiuni specifice ca:

- ☞ - primesc semnale de la neuronii vecini
- ☞ - conduc semnalul nervos
- ☞ - integrează semnalele primite
- ☞ - amplifică sau mediază semnalul nervos rezultat
- ☞ - transmit semnalul nervos pe canale preferențiale spre neuronii receptori

I.2 Modele neurale artificiale

În forma cea mai generală și în *accepția inițială*, o rețea neurală este un algoritm proiectat să modeleze modul în care creierul realizează o anumită funcțiune estimându-se empiric cu aproximații acceptabile modalitatea de procesare a informației de către cortex. Acest scop este valabil doar pentru cazul primelor modele, ulterior modelele neurale au părăsit ideea inițială de simulare a propagării și procesării semnalului nervos, păstrând însă specificul nomenclatorului biologic pentru desemnarea unor etape. Astfel, aplicațiile tratate în teză se vor referi la rețele neurale care își realizează scopul propus printr-o proces inițial de **învățare**. Pentru o bună performanță aplicativă, rețeaua își crează ponderi de interconectare între unitățile de intrare și cele de ieșire, unități care vor fi denumite *neuroni* sau *unități de procesare*. O definiție simplistă a noțiunii de rețea neurală este dată de Alecsander & Morton⁵:

O rețea neurală este un procesor cu distribuție paralelă care are o mare capacitate de a stoca cunoștințe dobândite din experimente în scopul utilizării active ale acestora; Simularea acțiunii cerebrale este elocventă în cele două acțiuni de bază:

- cunoștințele sunt acumulate în timpul unui **proces de învățare**
- stocarea cunoștințelor se realizează prin valoarea conexiunilor interneuronale denumite și **ponderi sinaptice**.

In continuare, pornind de la primul model de rețea neurală, se va trata evoluția acestuia pas cu pas în ordine cronologică, subliniind caracteristicile, avantajele și dezavantajele majore. Având în vedere diversitatea modelelor neurale provenită din aplicațiile cărora le sunt dedicate, se va urmări doar filonul ce duce la modelele conexioniste cu propagare înapoi utilizate în teză.

I.2.1. Modelul McCulloch & Pitts

Înainte de descoperirea modalității de funcționare a sistemului nervos celular⁶ pentru care Hodgkin, Huxley & Eccles au luat premiul Nobel în 1963, în 1943, McCulloch & Pitts au elaborat un model simplist al funcționării unui neuron⁷ plecând de la funcționalitatea acestuia. Lucrarea lor a fost utilizată de von Newman la realizarea elementelor de comutare cu întârziere necesare primelor prototipuri de computer ENIAC (Electronic Numerical Integrator And Computer) și EDVAC

(Electronic Discrete Variable Automatic Computer) construite la Universitatea din Pennsylvania.

McCulloch & Pitts au presupus că neuronul se comportă în modul următor:

- din sinapse, prin dendrite, acesta acumulează o sarcină electrică ce provine de la axonii neuronilor vecini; în momentul în care valoarea sarcinii depășește o valoare anume - numită **prag** - neuronul devine activ transmițând prin axonul său un curent constant. Această funcționare poate fi descrisă matematic astfel:

$$o = \begin{cases} U & \text{dacă } \sum_{i=1}^n x_i \geq T \\ 0 & \text{dacă } \sum_{i=1}^n x_i < T \end{cases} \quad (1)$$

unde x_i = intrarea i

o = ieșire

Schematic, modelul este redat în figura 2:

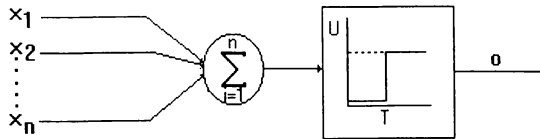


Fig. 2. Schema modelului matematic McCulloch & Pitts

O astfel de unitate este denumită **unitate de procesare**. Considerând valoarea pragului $T=1$ și valoarea tensiunii de ieșire $U=1$ se pot construi cu aceste unități de procesare unitare circuite logice de tipul **AND**, **NAND**, **OR**, și **celule de memorie** în modul sugestiv sugerat în figura 3 a,b,c, și respectiv d.

Cei doi matematicieni au sugerat posibilitatea ca creierul uman să fie o aglomerare de asemenea circuite logice care ar procesa informația într-un mod asemănător cu calculatoarele digitale al căror model era definitivat în aceeași perioadă.

Această prezumție s-a dovedit a fi greșită, însă are valoare intuitivă deoarece s-a dovedit ulterior că neuronii funcționează oarecum după modelul activării la atingerea unui prag.

Acest model neural este caracterizat prin formalism și printr-o definire matematică elegantă și precisă, având un potențial de calcul considerabil.

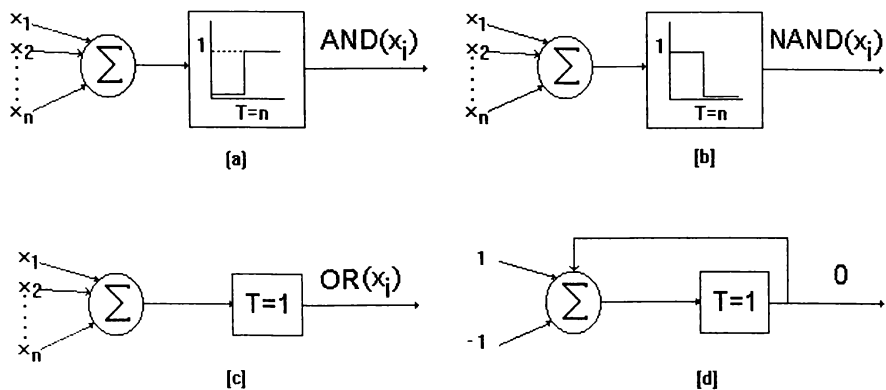


Fig. 3. Circuite logice realizate cu unități de procesare McCulloch & Pitts

Simplicitatea sa presupune:

- stări binare de tip 0 și 1
- asumarea operării în timp discret
- sincronicitatea operării tuturor neuronilor

1.2.2. Modelul Hebb

În 1949, Daniel Hebb, un biolog care se ocupa cu studiul comportamentului primatelor presupune că prin modificarea sinapselor, conectivitatea celulelor nervoase este într-o continuă schimbare pe măsură ce organismul își exercită diferite aptitudini. Având la bază prezumțiile date de Ramon y Cajal⁸, în cartea *The Organization of Behavior. A Neuropsychological Theory* a enunțat o **regulă a învățării**⁹ conform căreia **valoarea efectivă a variabilei de conexiune între doi neuroni crește prin activarea repetată a conexiunii**. Cartea a avut un răsunset imens printre psihologi dar a fost neglijată de ingineri, fapt ce a dus la ignorarea regulii de învățare în domeniul rețelelor neuronale. Abia în 1956 Rochester & All¹⁰ au utilizat această regulă de învățare evidențiind fenomenul de **inhibiție** a conexiunii interneuronale.

Conform regulii de învățare a lui Hebb:

- sinapsele care conectează neuronii nu transmit axonului neuronului receptor decât o parte a sarcinii electrice primite de la axonul neuronului transmițător. Redarea matematică a acestei **leme** se obține utilizând un coeficient subunitar de transmisie numit **pondere** și notat cu w .

Ponderile sunt asociate neuronului receptor și își pot modifica valorile în timp conform regulii:

- dacă neuronul transmițător și neuronul receptor sunt ambii activi în același interval de timp, valoarea ponderii asociate conexiunii crește

- dacă fenomenul de activare concomitentă se repetă de mai multe ori sau dacă intervalul de timp al dublei activări este mare, valoarea ponderii asociate va crește considerabil

- dacă neuronul transmițător și neuronul receptor nu sunt simultan activi într-o perioadă de timp mare, valoarea ponderii asociate conexiunii va scăde

Ulterior, s-a dovedit că într-adevăr această regulă este valabilă și stă la baza mecanismului de învățare și uitare în creierile naturale.

Plecând de la această regulă, se poate descrie o unitate de procesare cu ponderi asociate care este modelul matematic utilizat și în prezent cu variațiuni nesemnificative.

Astfel, considerând o unitate cu indicele j cu n intrări, fiecare având asociată o pondere w_j , funcționarea acesteia se poate descrie prin:

$$o_i = \begin{cases} 1 & \text{dacă } \sum_{j=1}^n x_j \cdot w_j \geq T \\ 0 & \text{dacă } \sum_{j=1}^n x_j \cdot w_j < T \end{cases} \quad (2)$$

sau prin schema din figura 4:

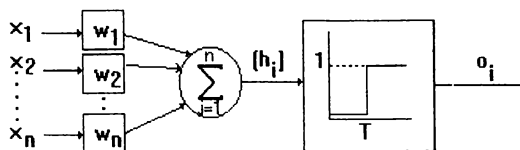


Fig. 4 Schema modelului matematic Hebb

Pentru simplificare, se va nota suma ponderată care reprezintă intrarea totală în unitatea de procesare cu h_i :

$$h_i = \sum_{j=1}^n w_j \cdot x_j \quad (3)$$

O asemenea unitate poate fi privită ca un procesor de informație foarte primitiv. Dacă legăm un număr de astfel de procesoare între ele, obținem o *rețea neurală artificială* sau mai corect spus o **structură de calcul conexiionistă**.

Modul de legare al acestor procesoare între ele reprezintă topologia rețelei și determină modul de calcul al ponderilor și implicit funcționalitatea rețelei.

Calculul ponderilor este denumit **învățare** iar metodele iterative de calcul sunt cunoscute sub denumirea de **algoritmi de învățare** sau antrenament.

Expresia **structură de calcul conexiunistă** reprezintă de fapt noțiunea descriptivă la care s-a ajuns prin încercarea de a exprima matematic teoriile de moment ce sugerează posibilități de modelare a propagării semnalului prin sistemul neural uman. Este practic o expresie mult mai adecvată atribuită algoritmilor astfel obținuți, care nu își mențin scopul inițial ci îl adoptă pe cel al aplicațiilor cărora s-au dedicat.

În 1952, cartea lui Ashby¹¹ "Design for a Brain: The Origin of Adaptive Behavior" este destinată comportamentului adaptiv ce presupune și argumentează că un sistem natural nu se naște învățat ci ajunge la anumite performanțe prin corelarea variabilelor ce descriu evenimentele percepute sistemocentric. Cartea tratează aspectele dinamice ale organismului viu ca un sistem abstract, mecanic și evaluează conceptul de stabilitate.

În 1954, teza de doctorat a lui Minsky¹² intitulată "Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model Problem" se constituie ca prolog al cărții "Steps Toward Artificial Intelligence" apărută în 1961¹³ ce constituie și azi unul din capitolele fundamentale în domeniul **rețelelor neurale**.

În 1961, Steinbuch¹⁴ continuă ideea lansată de Taylor¹⁵ în 1956 despre o **memorie asociativă** realizabilă prin așa numitele **matrici de învățare** care ar fi interpusă între câmpuri de senzori receptori și efectori motori.

Contribuții semnificative la teoria memoriei asociative au adus ulterior și Anderson¹⁶, Kohonen¹⁷, și Nakano¹⁸ în același an (1972); independent au introdus ideea unei matrici corelative de memorie (correlation matrix memory) care utilizează o regulă de învățare bazată pe produsul de ieșire al algoritmului.

O figură particulară în contextul evoluției conceptului rețelelor neurale a fost von Neumann care a utilizat ideea de redundanță pentru proiectarea unei rețele neurale privită ca un ansamblu unitar cu un comportament cert alcătuit din neuroni care pot fi priviți separat ca și componente cu comportament incert. Această idee a fost dezvoltată de Winograd & Cowan care au utilizat¹⁹ o reprezentare redundant distribuită a rețelelor neurale.

1.2.3. Modelul Rosenblatt - Perceptronul

După 15 ani de la modelul McCulloch&Pitts devenit clasic, un prim pas spre rezolvarea problemei recunoașterii tiparelor îl face Rosenblatt²⁰ prin lucrarea despre

perceptron în care enunță teorema convergenței perceptronului și prin lucrările adiacente²¹⁻²³ în care dezvoltă această teorie.

Topologia cea mai simplă este asociată unei structuri conexiunite denumită **perceptron** a cărei structură schematică este arătată în figura 5.

În forma cea mai simplă, constă dintr-un vector de intrare n -dimensional, o unitate de asociere și o singură ieșire. Menirea unei astfel de unități este de a învăța o transformare de tipul $T: \{-1,1\}^N \rightarrow \{-1,1\}$ utilizând seturi de tipare cu intrări de tip x și ieșirile corespunzătoare $o = T(x)$.

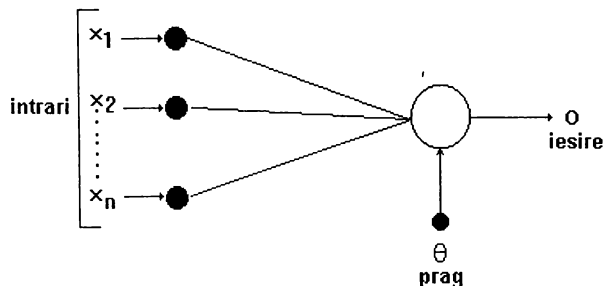


Fig. 5 Structura schematică a perceptronului

Unitatea de ieșire a unui perceptron este un element liniar cu prag care calculează ieșirea o conform relației (4):

$$o = F\left(\sum_{i=1}^n w_i \phi_i + \theta\right) \quad (4)$$

în care funcția F este funcția lui Heaviside sau sigum²⁴ iar θ poate fi orice funcție sau valoare care ajustează conexiunile astfel încât să se obțină ieșirea dorită.

În acest caz, funcția F este:

$$F(i) = \begin{cases} \text{dacă } i > 0 & = 1 \\ \text{altfel} & = -1 \end{cases} \quad (5)$$

Un **tipar** de învățare binar este o pereche de vectori cu valori binare, unul de lungime n (numărul de unități de pe nivelul 1) iar celălalt de lungime m (numărul de unități de pe nivelul 2). Numărul total de tipare alcătuiește o *bază de învățare*.

Algoritmul de învățare este următorul:

- Se prezintă la intrarea rețelei vectorul 1 al unui tipar, activând unitățile corespunzătoare valorii 1,
- Se prezintă la ieșirea rețelei vectorul 2 al aceluiași tipar, activând unitățile de pe nivelul doi corespunzător valorilor 1
- Se caută perechi de unități care sunt conectate și activate simultan și se mărește valoarea ponderii conexiunii dintre acestea cu o valoare mică η . Valoarea lui η se alege în funcție de numărul de tipare de învățare.
- După prezentarea tuturor tiparelor din baza de învățare, ponderile se vor modifica astfel încât prezentând la intrare un vector 1 dintr-un tipar, activând corespunzător unitățile de pe primul nivel și calculând ieșirile din unitățile de pe nivelul 2, se obține vectorul 2 din tiparul de învățare respectiv.

În anii 60, **perceptronul** a constituit punctul de interes al grupului celor care studiau acest domeniu.

I.2.3.1 Regula Delta

În 1960, Widrow&Hoff au introdus **algoritmul celor mai mici patrate** (Least Mean-Square Algorithm - LMS) cunoscut și ca **Regula Delta** (delta rule) ca algoritm de corecție a ponderilor în rețea²⁵. Astfel au luat naștere elementul liniar adaptiv denumit și ADALINE (ADaptive LINear Element).

Mecanismul de funcționare care utilizează această regulă de corecție a ponderilor este sugestiv redat în figura 6.

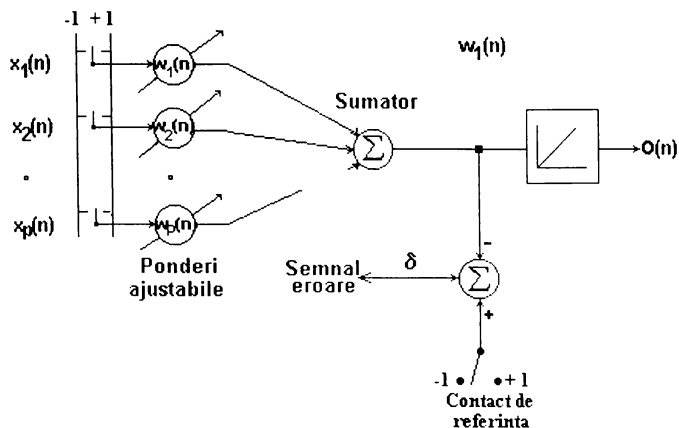


Fig. 6 Schema algoritmului ce utilizează LMS pentru corecția ponderilor

Widrow și Hoff realizează practic o generalizare importantă a algoritmului de antrenare a perceptronului extinzând tehnica intrărilor și ieșirilor continue. **Regula delta** a fost aplicată mai des unităților cu ieșiri pur liniare.

Pentru un astfel de perceptron cu o singură unitate de ieșire, ieșirea este dată de relația:

$$o = \sum_j \omega_j x_j + \theta \quad (6)$$

Funcția erorii așa cum este indicată de numele “cel mai mic patrat” este suma patratelor erorilor. Deci eroarea totală E este definită ca:

$$E = \sum_p E^p = \frac{1}{2} \sum_p (d^p - a^p)^2 \quad (7)$$

unde indicele p evoluează pe setul datelor de intrare și E^p reprezintă eroarea setului p . Variabila d^p este ieșirea dorită (ținta) la aplicarea setului p , iar a^p este ieșirea actuală pentru acest set. Procedura LMS găsește valorile tuturor ponderilor care minimizează eroarea funcției prin metoda numită a “**gradientului descendent**”. Ideea este de a face o schimbare în pondere, direct proporțională cu negativul derivatei erorii determinată pe setul curent în concordanță cu fiecare pondere:

$$\Delta^p \omega_j^p = -\eta \frac{\partial E^p}{\partial E^j} \quad (8)$$

unde η este o constantă de proporționalitate. Derivata este:

$$\frac{\partial E^p}{\partial E^j} = \frac{\partial E^p}{\partial a^p} \frac{\partial a^p}{\partial \omega_j} \quad (9)$$

Datorită unităților liniare,

$$\frac{\partial a^p}{\partial \omega_j} = x_j \quad \text{și} \quad \frac{\partial E^p}{\partial a^p} = -(d^p - a^p) \quad (10)$$

astfel că

$$\Delta^p \omega_j^p = \eta \delta^p x_j \quad (11)$$

unde

$$\mathcal{J} = \mathcal{J}' - \mathcal{J}'' \quad (12)$$

este diferența dintre ieșirea (țintă) dorită și actuala ieșire pentru setul \mathbf{p} . Regula delta modifică ponderea corespunzător țintei și actualei ieșiri pentru orice polaritate și pentru unități cu intrări-ieșiri continue sau binare. Aceste caracteristici au deschis un bogat câmp de noi aplicații.

Spre deosebire de algoritmul inițial ce utilizează incrementarea progresivă, acest algoritm ia în considerare și greșelile pe care la face rețeaua și încearcă să le corecteze prin schimbarea valorilor ponderilor.

Acest algoritm funcționează mai bine decât regula lui Hebb simplă, mai ales în cazul prezentării repetate a bazei de învățare, luând de fiecare dată tiparele într-o altă ordine - de preferință aleatoare.

Etapă de traversare o dată a rețelei de către întreaga baza de învățare poartă numele de *epocă*, numărul epocilor în care rețeaua învață cu precizia dorită întreaga bază de date fiind de fapt un factor de performanță al rețelei.

O slăbiciune a modelului conexiunist este faptul că precizia de învățare a tiparelor variază de la un tipar la altul în funcție de modalitatea de încadrare a tiparului în logica ponderată a bazei.

În cazul în care se dorește ca rețeaua să obțină la ieșire vectorul caracteristic unui vector de intrare care nu face parte din baza de învățare, dar este asemănător ca distribuție cu vectorii de intrare ai bazei de învățare, este nevoie de definirea unui parametru ce caracterizează performanța de învățare: *eroarea totală*-relativă la toate tiparele:

$$E = \sum_{\mu=1}^p \sum_{i=1}^m \delta_i^{\mu} \quad (13)$$

Eroarea maximă a rețelei:

$$E_{max} = \mathbf{m} \cdot \mathbf{p} \text{ dacă vectorii au valori binare de tipul } (0,1)$$

și

$$E_{max} = 2 \cdot \mathbf{m} \cdot \mathbf{p} \text{ dacă valorile sunt de tipul } (-1,1).$$

$$\text{Raportul } \frac{E}{E_{max}} \cdot 100 = E_{\%} \text{ este eroarea relativă procentuală a rețelei și este}$$

utilizat ca și criteriu de stopare a etapei de învățare a rețelei.

Având în vedere avantajul acestui model de a recunoaște vectorul de ieșire în cazul prezentării la intrare a unui vector care nu face parte din baza de învățare, perceptronul a fost utilizat pe scară largă.

Concluziv, diferențele principale dintre ADALINE și Perceptron sunt:

- la ADALINE, are loc o ponderare liniară ce se formează adaptiv (printr-o metodă de antrenare supervizată) care produce semnalul de răspuns dorit pentru vectori de intrare diferiți; rezultatul este optim în sensul celor mai mici pătrate;

- la Perceptron, există o unitate de procesare suplimentară pentru extracția combinațiilor de semnal din tiparele de intrare. Circuitele de ieșire au ca singur scop sesizarea diferenței dintre vectorii de intrare, fiecare dintre ei corespunzând unei clase diferite de tipare. Ulterior ADALINE, Widrow propune²⁶ și realizează pentru prima dată o rețea neurală stratificată și antrenabilă cu elemente adaptive multiple. În perioada anilor 60 se părea că s-a ajuns la rețele neurale care în vederea realizării obiectivului propus, puteau face orice tip de conexiuni.

1.2.3.2 Problema XOR

Euforiei generate de perceptron i-au pus capăt în 1969 Minsky&Papert²⁷ care utilizând argumente matematice simple dar elegante, au arătat că un perceptron uni-strat are limitele sale. În acest context, una dintre problemele care au rămas drept istorice este problema imposibilității unui perceptron uni-strat de a reprezenta o **funcție exclusiv-sau (XOR)**.

În perceptronul din figura 7, valoarea totală a intrării este dată conform relației 14 de:

$$i = w_1 x_1 + w_2 x_2 + \theta \quad (14)$$

În figura 7 este redată reprezentarea geometrică a spațiului de intrare.

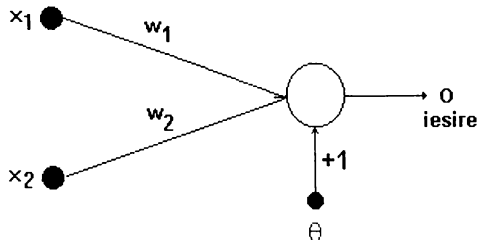


Fig. 7 Perceptron uni-strat cu 2 intrări

Pentru o constantă θ , ieșirea perceptronului este egală cu una din cele două părți ale liniei de diviziune definită de ecuația:

$$w_1 x_1 + w_2 x_2 = -\theta \quad (15)$$

și egală cu zero de cealaltă parte a liniei.

Conform figurii 8, spațiul de intrare constă din cele 4 puncte (cercuri). cercurile pline localizate la (1,-1) și (-1,1) neputând fi separate printr-o linie dreaptă de cercurile goale localizate în (-1,-1) și (1,1).

I.2.3.3 Perceptronul multistrat

Tot Minsky&Papert au dat și soluția - introducând un nou strat conectat la toate intrările, deci extinzând rețeaua la *perceptron multi-strat* se poate rezolva problema XOR.

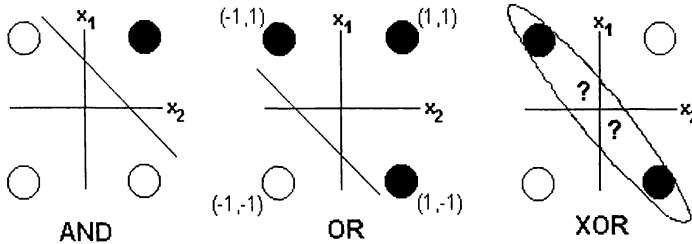


Fig. 8 Reprezentarea geometrică a spațiului de intrare pentru perceptronul unistrat din fig. 6

În figura 9a cu valorile ponderilor indicate pe liniile de conexiune și ale pragurilor indicate în cercuri se poate rezolva în mod elegant problema XOR. În figura 9b este redată reprezentarea spațială a separării celor două grupuri de puncte printr-un plan de separare.

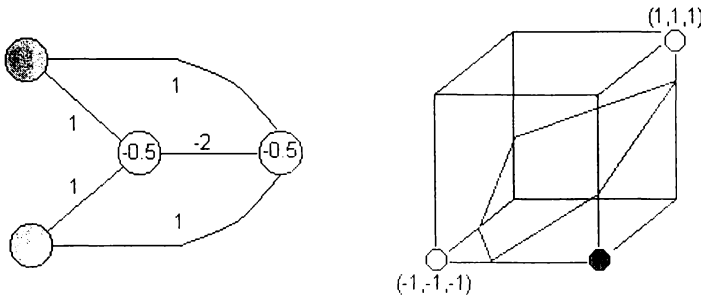


Fig. 9

- a) Perceptronul cu două intrări și o unitate ascunsă ce rezolvă problema XOR b) Reprezentarea spațială a separării celor două grupuri de puncte

Acest exemplu simplu demonstrează că adăugarea unor *unități ascunse* crește clasa aplicațiilor rezolvabile de către perceptronul multistrat.

În figura 10 este reprezentată schema unui perceptron având la intrare un vector n-dimensional iar la ieșire unul m-dimensional.

Fiecare unitate i este multiplicată pentru toate unitățile din rețea care vor avea n intrări.

O problemă importantă întâlnită în proiectarea rețelelor tip perceptron multistrat a fost problema **acordării de credit neuronilor din stratul ascuns** al rețelei (credit assignment theory). Cel care a introdus pentru prima dată terminologia de credit assignment a fost Minsky²⁸, ulterior căruia în anii '60 strădaniile cercetătorilor din domeniu au fost dedicate problemelor legate de această teorie care aveau rezultate contradictorii fără a o rezolva definitiv. A trebuit să vină perioada anilor '80 pentru a găsi soluțiile de ieșire din impas.

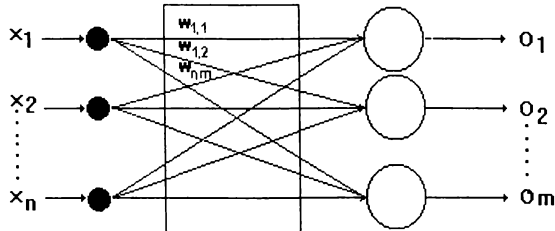


Fig. 10. Structura schematică a unui perceptron multistrat

Motivele unei asemenea întârzieri au fost de natură diferite dintre care:

- tehnologic - nu existau calculatoare personale sau stații de lucru pentru experimentare
- psihologic - monografia Minsky&Papert nu a fost deloc încurajatoare pentru continuarea studiilor în domeniu
- financiar - este derivat din cel psihologic care nu încuraja nici o agenție în lansarea unor programe de cercetare în domeniu

Din nefericire, aceste motive sau altele au dus la sistarea interesului pentru studiul rețelelor neurale. Singurele cercetări ale anilor '70 (totuși colaterale domeniului) au fost cele în psihologie și neurologie, de aceea decada acestor ani a fost caracterizată ca *dormancy decade* (decada de somnolență în domeniu).

1.2.4. Modelul Hopfield

În 1982, Hopfield²⁹ a utilizat ideea unei *funcții de energie* pentru a reformula un nou mod de înțelegere a calculelor dintr-o rețea neurală folosind conexiuni sinaptice simetrice și stabilind o analogie între o rețea neurală recurentă și *modelul Ising* utilizat în fizica statistică. Aceste tipuri de rețele sunt cunoscute sub denumirea de rețele Hopfield.

1.2.4.1 Modelul fizic Ising

Conform modelului Ising, un material magnetic este compus dintr-o rețea regulată de atomi cu proprietăți magnetice, fiecare având asociată o mărime numită **spin magnetic** orientat vectorial în sensul dat de compunerea forțelor magnetice la care este supus atomul. Pentru cazul ipotetic în care este permisă orientarea spinului pe o singură direcție în două sensuri, valoarea acestuia poate fi memorată într-o variabilă binară s_i . Alegând pentru această variabilă valorile $-1, 1$, se obține un model simplificat, **modelul Ising** (fig. 11).

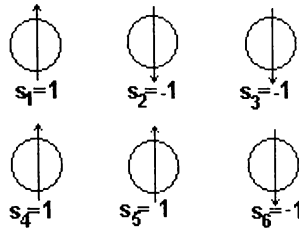


Fig. 11 Modelul fizic - Ising

Fiecare spin s_i este influențat de câmpul magnetic H_i compus din câmpul magnetic exterior H_{ext} și câmpul produs de atomii învecinați.

Ising a propus pentru calculul H_i relația :

$$H_i = \sum_j p_{i,j} \cdot s_j + H_{ext} \quad (16)$$

unde $p_{i,j}$ este un coeficient dat de poziția relativă dintre atomii i și j .

Materialul magnetic poate avea înmagazinată o energie potențială a cărei valoare poate fi calculată cu relația 17:

$$H = -\frac{1}{2} \sum_i \sum_j p_{i,j} \cdot s_i \cdot s_j - H_{ext} \sum_i s_i \quad (17)$$

Pentru simplificarea discuției în ce privește utilizarea modelului Ising la construcția unei rețele neurale, în continuare, se va considera $H_{ext}=0$.

1.2.4.2 Modelul Hopfield propriu-zis

Trecerea de la modelul Ising la o structură conexiionistă a fost definitivată de Hopfield în 1984³¹. Rețeaua conexiionistă bazată pe acest model este topologic diferită de perceptron, *având fiecare unitate legată de celelalte unități, toate*

legăturile fiind ponderate. Nu există diferențiere între unitățile de intrare și cele de ieșire, ieșirea fiecărei unități fiind multiplicată ca intrare către toate unitățile inclusiv ea însăși. Ponderile se memorează într-o matrice de dimensiune $N \times N$ (unde N este numărul de unități).

Funcționarea rețelei este deasemenea diferită de cea a perceptronului. Astfel, dacă la o rețea structurată pe straturi de intrare și respectiv ieșire aplicăm un vector de intrare la primul strat și calculăm activările din cel de-al doilea strat pentru a determina ieșirea, pentru noua rețea vom considera ca intrare o stare generală a întregii rețele, reprezentabilă printr-un vector de stare $S = [S_1, S_2, \dots, S_N]$.

Datorită câmpului magnetic exercitat asupra lor, atomii dintr-un material magnetic își pot schimba spinul. Acceptând prin convenție că o unitate este activă dacă are spinul egal cu 1 și este inactivă dacă are spinul egal cu -1, în modelul Ising, valoarea spinului se poate calcula cu funcția **signum**. Astfel, dacă semnul valorii câmpului din locația atomului este pozitiv, atunci și spinul este pozitiv, iar dacă este negativ, spinul este și el negativ. Considerând că $h_i = H_i$, (intrarea într-o unitate este egală cu câmpul) și $w_{ij} = p_{ij}$ (ponderea unei intrări este egală cu coeficientul de interacțiune dintre atomi), se poate afirma că unitățile acestei rețele conexiuniste au următoarea funcție de transfer:

$$s_i = \text{signum}(h_i) = \begin{cases} 1 & \text{dacă } h_i \geq 0 \\ -1 & \text{dacă } h_i < 0 \end{cases} \quad (18)$$

numită **lege dinamică de evoluție**, unde:

$$h_i = \sum_{j=1}^N w_{ij} \cdot s_j \quad (19)$$

Un astfel de sistem caută să ajungă întotdeauna într-o stare de energie potențială minimă. Unitățile își vor schimba valoarea de activare până la un moment în care schimbarea activării oricărei unități nu va mai duce la scăderea energiei potențiale a întregului sistem.

Se poate demonstra că un astfel de sistem ajunge întotdeauna într-o stare de minim energetic - stabilă. Considerând energia potențială dată de relația:

$$H = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{i,j} \cdot s_i \cdot s_j \quad (20)$$

și $s_i' = \text{sign}(h_i')$ ca fiind starea următoare a lui s_i , vor putea exista două situații:

1. $-s_i' = s_i \Rightarrow$ unitatea **i** nu își schimbă activarea $\Rightarrow \Delta H=0$
 2. $-s_i' = -s_i \Rightarrow$ unitatea își schimbă activarea
- (21)

În primul caz, $H'=H$, deci energia rămâne aceeași. În cel de-al doilea caz,

$$\Delta H = H' - H = -\frac{1}{2} \sum_{i,j} w_{i,j} \cdot s_i' \cdot s_j + \frac{1}{2} \sum_{i,j} w_{i,j} \cdot s_i \cdot s_j \quad (22)$$

Înlocuind s_i' cu $-s_i$, se obține:

$$\Delta H = \sum_i s_i \sum_j w_{i,j} \cdot s_j = \sum_i s_i h_i' \quad (23)$$

Având în vedere ca semnul lui h_i' este și semnul lui s' și ca $s_i = s_i'$, rezultă că ΔH este întotdeauna negativ atunci când o unitate își schimbă activarea, deci sistemul va evolua întotdeauna către o stare de minim energetic.

Pentru un singur tipar de învățare ζ constând dintr-un vector binar cu componente $(-1,1)$ de lungime $N \times N$, relația de calcul a ponderilor va fi:

$$w_{i,j} = 1/N \times \zeta_i \cdot \zeta_j \quad (i=1,N ; j=1,N) \quad (24)$$

deci matricea de ponderi va fi o matrice simetrică.

Pentru mai multe tipare de învățare, ζ^μ ($\mu=1,p$) relația de calcul a ponderilor va fi:

$$w_{i,j} = 1/N \times \sum_{\mu=1}^p \zeta_i^\mu \cdot \zeta_j^\mu \quad (i=1,N ; j=1,N) \quad (25)$$

Având în vedere capacitatea limitată de memorare și posibilitatea apariției stărilor mixate ca o combinație liniară a stărilor învățate, acest tip de rețele nu au aplicații practice.

Publicațiile lui Hopfield au generat numeroase controverse. În 1983 Cohen&Grossberg³² au stabilit un principiu general pentru proiectarea unei memorii cu conținut direct adresabil care a inclus și versiunea rețelei Hopfield ca și caz particular. O componentă distinctivă a convergenței unei astfel de rețele neurale este modul natural în care timpul ca și dimensiune esențială a învățării se manifestă în dinamica neliniară a rețelei.

În 1983, Kirkpatrick, Gellat & Vecchi³³ au descris o nouă procedură numită **algoritm de simulare a răcirii treptate** (simulated annealing algorithm), ideea fiind preluată de la Metropolis & all³⁴, și își are rădăcina în statistica termodinamică pentru rezolvarea unor probleme de optimizare combinatorială. Această idee a fost

exploatătă în 1985 de Ackley, Hinton&Sejnowski³⁵ în dezvoltarea algoritmului stohastic de învățare ce utilizează proprietăți ale distribuției Boltzmann denumit **învățarea Boltzmann**.

Metoda presupune pentru S_i o dependența de tipul:

$$S_i = \begin{cases} +1 & \text{cu probabilitatea } g(h_i) \\ -1 & \text{cu probabilitatea } 1-g(h_i) \end{cases} \quad (26)$$

unde:

$$h_i = \sum_j w_{i,j} \cdot S_j \quad (27)$$

Funcția de probabilitate $g(h)$ trebuie să îndeplinească condițiile:

- 1- să ia valori în intervalul $[0,1]$
- 2- să fie dependentă de temperatură

Cea mai utilizată funcție de probabilitate este cea sigmoidă:

$$F(x) = \frac{1}{1 + e^{-x}} \quad F(x): \mathbb{R} \rightarrow [0,1] \quad (28)$$

Pentru a simula dependența de temperatură a funcției, se utilizează factorul de temperatură β astfel:

$$F(x) = \frac{1}{1 + e^{-\beta x}} \quad \beta = \frac{1}{K_B \cdot T_K} \quad (29)$$

unde K_B este constanta lui Boltzmann iar T_K este temperatura absolută în K. Pentru diferite valori ale lui β se obține o variație a funcției de genul celui ilustrat în figura 12. Inițial se pornește de la o temperatură ridicată fapt ce induce o valoare mică a valorii lui β având ca și consecință o pantă mică a funcției de probabilitate, urmând ca la creșterea valorii temperaturii panta funcției să crească în mod considerabil.

Lăsând sistemul să evolueze “la o temperatură ridicată”, acesta va ajunge într-o stare stabilă care nu este una din stările învățate, dar este mai joasă din punct de vedere energetic. În acest moment, la scăderea temperaturii în mod progresiv, sistemul va ajunge la una din stările învățate sau la negativele acestora. Capacitatea unei astfel de rețele Hopfield stohastice este similară cu cea Hopfield discretă, evoluția fiind doar în ce privește lansarea în domeniu a *funcției de activare sigmoide* care a stat la baza trecerii la *rețele cu propagare înapoi*.

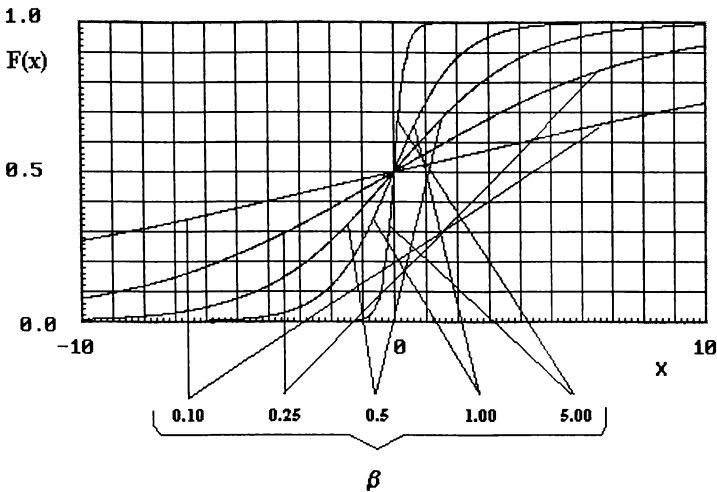


Fig. 12 Variația funcției sigmoide în varianta Boltzmann pentru diferite valori ale variabilei β

Tot în 1984, Breitenberg³⁶ publică o carte “*Experimente în psihologia sintetică*”, în care descrie diverse mecanisme cu capacitate de autoorganizare inspirate din comportamentul animal.

1.2.5 Rețele cu propagare înapoi (*Backpropagation Networks*)

Deși fusese descris de către Bryson&Ho în 1969 și de către Werbos³⁷ în 1974, abia în 1986 Rumelhart, Hinton&Williams³⁸ publică studiul algoritmului de propagare înapoi (the back-propagation algorithm) spre care se îndreaptă atenția cercetătorilor în domeniu.

În același an, Rumelhart&McClelland³⁹ publică în două volume cartea “*Procesare paralelă distributivă: Explorări în microstructurile cunoașterii*” ce se va constitui în perioada următoare ca referință de bază în studiul antrenării perceptronului multistrat. Aceeași idee dar cu o rezonanță mai mică o utilizează și Parker^{40,41} în 1985 - 1987 și Le Cun⁴² în 1985.

Ideea centrală a acestui model este că erorile pentru unitățile de pe stratul ascuns sunt determinate prin propagare înapoi (by back-propagation) a erorilor de pe stratul de ieșire. Propagarea înapoi este o generalizare a **regulei delta** pentru funcții de activare neliniare aplicate rețelelor multistrat.

O rețea neurală cu propagare înainte (figura 13) are o structură multistrat, fiecare strat primind intrările de pe stratul anterior și trimițând ieșirile stratului imediat următor. În cadrul unui strat nu există conexiuni între unitățile proprii. Cele N_i intrări alimentează primul strat cu $N_{h,1}$ unități ascunse. Unitățile de intrare sunt doar unități generatoare, în ele neavând loc nici un proces de prelucrare a datelor. Activarea unităților ascunse se face prin aplicarea unei funcții F asupra ponderilor care le preced și adăugarea unei constante denumită **bias** care nu este obligatorie. Ieșirea unități ascunse este distribuită stratului imediat următor de unități ascunse

$N_{h,2}$ care dacă este ultimul strat de unități ascunse, alimentează cele N_o unități de pe stratul de ieșire.

Algoritmul de propagare înapoi poate fi aplicat oricărui tip de rețele cu oricât de multe straturi. Pentru cazul rețelelor cu unități binare s-a demonstrat de către Hornik, K.; Stinchcombe, M.&White. H.^{43,44} în 1989 și de către Hartman⁴⁵ în 1991 că un singur strat de unități ascunse este suficient pentru a aproxima orice funcție cu un număr mare dar finit de discontinuități pentru funcții de activare neliniare aplicate pentru straturile ascunse.

Pentru a utiliza funcții neliniare de activare, este necesară generalizarea regulii *delta* care este specifică funcțiilor liniare.

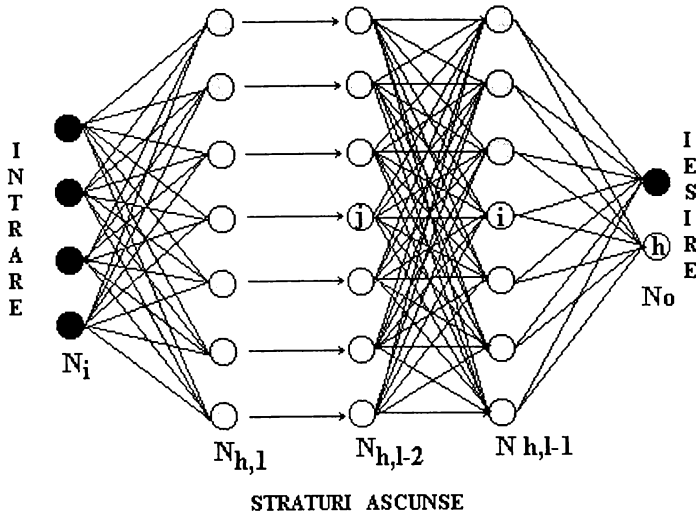


Fig 13. Rețea neurală cu l straturi

1.2.5.1 Generalizarea regulii delta

Funcția de activare este o funcție diferențiabilă a intrării dată de relația 30:

$$a_i^p = F(i_i^p) \quad (30)$$

în care a_i^p este ieșirea unității i obținută pentru intrarea i_i^p în cazul tiparului p . Valoarea de intrare este obținută prin relația :

$$i_i^p = \sum_j w_{i,j} a_j^p + \theta_i \quad (31)$$

în care θ_i este o valoare tip **bias**.

Pentru obținerea unei generalizări corecte a regulii *delta*, se admite:

$$\Delta^p w_{i,j} = -\eta \frac{\partial E^p}{\partial w_{i,j}} \quad (32)$$

Mărimea E^p este definită ca eroarea pătratică totală pentru tiparul p la ieșire:

$$E^p = \frac{1}{2} \cdot \sum_p^{N_{tip}} (d_i^p - a_i^p)^2 \quad (33)$$

expresie ce denotă termenul care trebuie minimizat pe parcursul rulării rețelei. Coeficientul 1/2 din fața expresiei este un coeficient de conveniență pentru simplificarea expresiei gradientului obținut prin derivare, acesta neafectând localizarea minimului de eroare și nici procesul de minimizare în sine.

Termenul erorii totale a rețelei la un anumit moment, este dat de suma erorilor tuturor tiparelor:

$$E_T = \sum_p^{N_{tip}} E^p \quad (34)$$

Se poate scrie:

$$\frac{\partial E^p}{\partial w_{i,j}} = \frac{\partial E^p}{\partial \tilde{a}_i^p} \frac{\partial \tilde{a}_i^p}{\partial w_{i,j}} \quad (35)$$

Din relația (35) se observă că cel de-al doilea factor este:

$$\frac{\partial \tilde{a}_i^p}{\partial w_{i,j}} = a_j^p \quad (36)$$

Definind:

$$\delta_i^p = -\frac{\partial E^p}{\partial \tilde{a}_i^p} \quad (37)$$

se obține o regulă nouă care este echivalentă cu regula *delta* constând într-un gradient descendent al suprafeței de eroare dacă schimbările ponderilor se vor face în concordanță cu relația:

$$\Delta^p w_{i,j} = \eta \cdot \delta_i^p \cdot a_j^p \quad (38)$$

Ceea ce rămâne de demonstrat este ce reprezintă δ_i^p pentru fiecare unitate i din rețea. Pentru a calcula δ_i^p se va aplica regula derivării înlănțuite pentru a scrie

derivata parțială ca un produs de doi factori, unul reflectând variația erorii ca funcție de ieșire iar celălalt reflectând schimbarea ieșirii ca funcție de schimbarea intrării. Astfel, rezultă:

$$\delta_i^p = -\frac{\partial E^p}{\partial \alpha_i^p} = -\frac{\partial E^p}{\partial \alpha_i^p} \cdot \frac{\partial \alpha_i^p}{\partial i_i^p} \quad (39)$$

Din relația (39) se observă că:

$$\frac{\partial \alpha_i^p}{\partial i_i^p} = F'(i_i^p) \quad (40)$$

expresie care este derivata funcției pentru unitatea **i**, evaluată la intrarea i_i^p a acestei unități. Pentru a calcula primul factor din relația (39) vor fi considerate două cazuri:

1. **Unitatea i este o unitate de ieșire a rețelei.** In acest caz, din definiția erorii tiparului E^p , se obține:

$$\delta_i^p = (d_i^p - a_i^p)F'_i(i_i^p) \quad (41)$$

pentru o unitate de ieșire **i**.

2. **Unitatea i nu este o unitate de ieșire.** In acest caz, nu se cunoaște contribuția unității la eroarea de ieșire a rețelei. Deci, valoarea erorii, poate fi scrisă ca funcție de intrările din stratul ascuns spre stratul de ieșire;

$$E^p = E^p(i_1^p, i_2^p, \dots, i_i^p, \dots) \quad (42)$$

Utilizând aceeași regulă de derivare înlănțuită, rezulta:

$$\frac{\partial E^p}{\partial \alpha_i^p} = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial \alpha_h^p} \frac{\partial \alpha_h^p}{\partial \alpha_i^p} = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial \alpha_h^p} \frac{\partial}{\partial \alpha_i^p} \sum_{k=1}^{N_h} w_{hk} a_k^p = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial \alpha_h^p} w_{h,i} = -\sum_{h=1}^{N_o} \delta_h^p w_{h,i} \quad (43)$$

Substituind rezultatul în ecuația (39), rezultă:

$$\delta_i^p = F'(i_i^p) \sum_{h=1}^{N_o} \delta_h^p w_{h,i} \quad (44)$$

Ecuatiile (41) și (43) redau o procedură de calcul a valorii lui δ pentru toate unitățile din rețea care la rândul lor sunt utilizate pentru a calcula schimbările în valoare ale ponderilor în concordanță cu ecuația (38). Această procedură constituie **generalizarea regulii delta** pentru o rețea cu unități neliniare.

1.2.5.2 Modul de lucru cu *back-propagation*

Aplicarea *regulii delta generalizată* implică două faze:

1. În decursul primei faze, se aplică intrarea și se propagă de-a lungul rețelei pentru calculul valorilor a_i^p pentru fiecare termen al unității de ieșire; această valoare este comparată cu valoarea dorită rezultând un semnal de eroare δ_i^p pentru fiecare termen al unității de ieșire;

2. Cea de-a doua fază implică o trecere înapoi prin rețea în decursul căreia semnalul de eroare este transmis fiecărui element din rețea, calculându-se schimbările de pondere adecvate.

Ponderea conexiunii este ajustată cu o cantitate proporțională cu produsul semnalului de eroare δ pe unitatea i primind intrarea iar ieșirea unității j trimițând acest semnal de-a lungul conexiunii:

$$\Delta^p \omega_{ij} = \eta_{ij} \delta_i^p a_j^p \quad (45)$$

- unde:
- ω_{ij} - ponderea conexiunii dintre unitatea i și unitatea j
 - $\Delta^p \omega_{ij}$ - termenul de ajustare a ponderii conexiunii
 - η_{ij} - rata de învățare asociată cu ponderea ω_{ij}
 - δ_i^p - semnalul de eroare asociat elementului i pentru vectorul \mathbf{p} aplicat la intrarea rețelei
 - a_j^p - valoarea de activare a elementului j al rețelei pentru vectorul \mathbf{p} aplicat la intrarea rețelei

Dacă unitatea este situată la ieșirea din rețea, semnalul de eroare este dat de:

$$\delta_k^p = (d_k^p - a_k^p) F'(z_k^p) \quad (46)$$

- unde:
- z_k^p - intrarea netă a neuronului k pentru un vector de intrare \mathbf{p}
 - d_k^p - al k -lea element al vectorului de ieșire dorit al rețelei pentru vectorul \mathbf{p} aplicat la intrarea rețelei

Definită ca *funcția de activare sigmoidă*, funcția F este:

$$a_j^p = F(y_j^p) = \frac{1}{1 + e^{-y_j^p}} \quad (47)$$

În cazul utilizării funcției de activare *temperată* a_j^p devine:

$$a_j^p = F(y_j^p) = \frac{1}{1 + e^{-\beta y_j^p}} \quad (48)$$

unde: $\beta = \frac{1}{K_B \cdot T_K}$, K_B este constanta lui Boltzmann iar T_K - temperatura absolută în K. Derivata acestei funcții este:

$$F'(y_j^p) = a_j^p(1 - a_j^p) \quad (49)$$

$$F'(i_i^p) = \frac{\partial}{\partial a_i^p} \frac{1}{1 + e^{-i_i^p}} = \frac{1}{(1 + e^{-i_i^p})^2} (-e^{-i_i^p}) \quad (50)$$

$$F'(i_i^p) = \frac{1}{(1 + e^{-i_i^p})} \frac{-e^{-i_i^p}}{(1 + e^{-i_i^p})} = a_i^p(1 - a_i^p)$$

astfel că semnalul de eroare pentru o unitate de ieșire poate fi scris ca:

$$\delta_j^p = (d_j^p - a_j^p) a_j^p(1 - a_j^p) \quad (51)$$

Semnalul de eroare pentru o unitate ascunsă este determinat recursiv în termenii semnalelor de eroare ale unității cu care este direct conectată și ponderile acestor conexiuni. Pentru funcția de activare sigmoidă:

$$\delta_j^p = F'(y_j^p) \sum_{k=1}^N \delta_k^p \omega_{j,k} = a_j^p(1 - a_j^p) \sum_{k=1}^N \delta_k^p \omega_{j,k} \quad (52)$$

I.2.5.3 Alte funcții de activare neliniare

Practic, numărul funcțiilor care pot fi utilizate ca funcții de activare este relativ nelimitat. Acestea se pot clasifica în două clase:

1. Funcții nesimetrice față de origine similare ca alură cu funcția sigmoidă

Din această categorie, alături de funcția sigmoidă, cea mai utilizată este funcția tangentă hiperbolică. Această funcție este definită de relația:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (53)$$

Reprezentarea grafică a funcției tangentă hiperbolică este redată în figura 14.

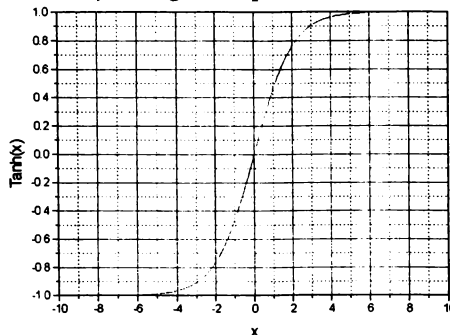


Fig. 14. Reprezentarea grafică a funcției tangentă hiperbolică

2. Funcții simetrice față de origine

Din această categorie fac parte:

a. Funcția Gaussiană

Aceasta este o funcție radială fiind definită de relația:

$$F_{Gauss}(x) = e^{-\frac{x^2}{v}} \quad (54)$$

în care v este o constantă predefinită numită **varianță**, cu ajutorul căreia este ajustată amplitudinea funcției.

Reprezentarea grafică a funcției Gaussiene este redată în figura 15.

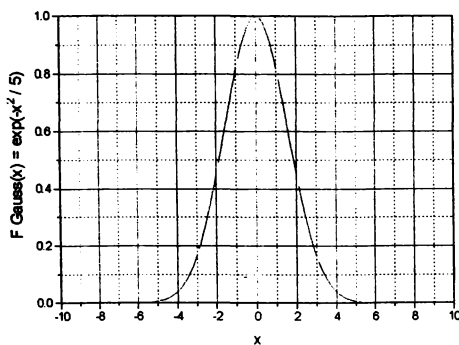


Fig. 15 Reprezentarea grafică a funcției Gaussiene

b. Funcția secantă hiperbolică

Și această funcție este o funcție radială fiind definită de relația:

$$\operatorname{sech}(x) = \frac{2}{e^x + e^{-x}} \quad (55)$$

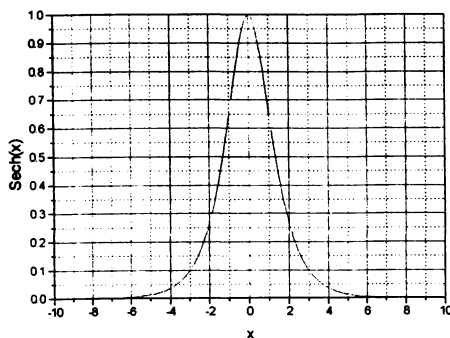


Fig. 16 Reprezentarea grafică a funcției secantă hiperbolică

Utilizarea unei anumite funcții de activare este o opțiune personală a programatorului, neexistând argumente tehnice privind indici de performanță care să justifice o selecție preferențială.

I.2.5.4 Rata de învățare și momentul

Procedura de învățare necesită ca schimbarea ponderii să fie proporțională cu $\frac{\partial E^p}{\partial \omega}$. Scăderea reală a gradientului necesită utilizarea unor pași infimezimali.

Constanta de proporționalitate este rata de învățare η . Pentru aplicațiile practice se alege o rată de învățare cât mai mare posibilă fără a se ajunge însă la **oscilație**. O modalitate de evitare a oscilației la valori mari ale η este efectuarea unei schimbări în pondere dependentă de schimbarea anterioară a ponderii prin adăugarea unui termen **moment**:

$$\Delta \omega_{i,j}(t+1) = \eta \delta_i^p a_j^p + \gamma \Delta \omega_{i,j}(t) \quad \gamma \in (0,1) \quad (56)$$

unde t indexează numărul epocii, iar γ este o constantă care determină efectul schimbării anterioare a ponderii. Valoarea lui γ poate fi de ex. 0.1, scăzând la creșterea erorii sau crescând la scăderea ei.

Rolul termenului **moment** este arătat în figura 17. Când nu este utilizat termenul **moment**, găsirea minimumului cu o rată mică de învățare necesită o perioadă lungă de timp (1), iar pentru o rată de învățare mare, minimumul nu poate fi găsit din cauza fenomenului de oscilație (2). Adăugarea unui termen **moment** duce la o găsire mai rapidă a valorii de **minim** (3).

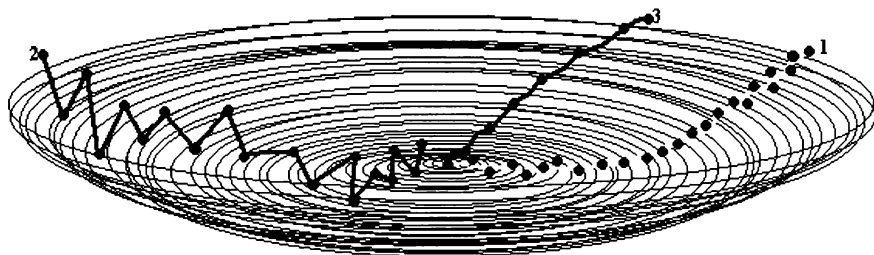


Fig. 17 Variația erorii rețelei funcție de valoarea momentului

I.2.5.5 Invățarea unui tipar

Teoretic, algoritmul *back-propagation* operează scăderea gradientului erorii totale a rețelei doar dacă ponderile sunt ajustate pe întregul set de tipare. Se va acorda o atenție mărită ordinii în care setul de tipare este prezentat rețelei deoarece, în cazul în care este utilizată aceeași secvență, rețeaua se poate focaliza pe ultimele tipare prezentate. Problema se poate depăși dacă se utilizează un algoritm aleator sau o tehnică de permutare la prezentarea tiparelor în etapa de antrenare.

1.2.5.6 Deficiențe ale algoritmului Back-propagation

Contrar succesului obținut de algoritmul back-propagation, există câteva aspecte care fac ca utilizarea algoritmului să nu se suprapună peste garanția succesului universal.

Se vor prezenta mai jos, trei dintre problemele cel mai des întâlnite oferindu-se și câteva din motivele posibile ale manifestării lor.

1. O antrenare de lungă durată cu un gradient al erorii totale scăzut - se poate datora unei selecții neoptimale a ratei de învățare și a termenului moment. Există algoritmi care pornesc cu o rată a învățării mică, în funcție de variația erorii aceasta modificându-se corespunzător- la o variație decrementală a erorii, rata de învățare crescând iar la o variație incrementală - scăzând.

2. Paralizia rețelei - se poate datora valorilor foarte mari în valoare absolută pe care le pot lua ponderile pe parcursul antrenării rețelei, valori care induc valori de activare apropiate de 0 sau 1. Conform relațiilor:

$$\delta_j^p = (d_j^p - a_j^p) a_j^p (1 - a_j^p) \quad (57)$$

$$\delta_j^p = F'(y_j^p) \sum_{k=1}^N \delta_k^p \omega_{j,k} = a_j^p (1 - a_j^p) \sum_{k=1}^N \delta_k^p \omega_{j,k} \quad (58)$$

termenul de ajustare a valorii ponderilor care este direct proporțională cu $a_j^p(1 - a_j^p)$ va avea valori nesemnificative fapt ce duce la un palier staționar de evoluție pentru gradientul de eroare.

3. Convergența spre un minim local - având în vedere că hypersuprafața de eroare pentru o rețea complexă este accidentată, există posibilitatea ca rețeaua să se afle într-un minim local de valoare nesatisfăcătoare în comparație cu minimul absolut al rețelei. Există metode probabilistice⁴⁶ de tipul "algoritmilor genetici" prin a căror utilizare se pot evita capcanele locale, dar acestea au dezavantajul de a fi lente. O altă posibilitate este de a mări numărul de unități de pe stratul ascuns, metodă care are același dezavantaj - este lentă.

Această problemă este cea mai spinoasă, existând în literatură o serie de articole care pretind rezolvarea parțială a ei, însă **rețelele** oferite nu au o eficiență generală, ci doar evită problema în contextul particular în care este studiată. Spre exemplificare, în 1994, Klawun&Wilkins⁴⁷ publică un articol denumit "Un nou algoritm pentru ieșirea dintr-un minim local" care constă în prezentarea în cazul sesizării unui minim local (sesizarea se face prin constatarea unei erori maxime pentru unele tipare) pentru câteva epoci doar a tiparelor cu eroare maximă (flashcard-algorithm). Pentru moment, eroarea totală a rețelei crește dar își urmează cursul descendent depășind impasul și convergând spre valoarea de minim global.

Această metodă este doar o observație de bun simț foarte des folosită de către utilizatorii de programe personale dedicate, programe care permit monitorizarea în timpul antrenării rețelei a parametrilor de rețea și intervenția în momentele de antrenare ineficace. Este de remarcat însă efortul logistic și financiar alocat pentru elaborarea metodei.

antrenare ineficace. Este de remarcat însă efortul logistic și financiar alocat pentru elaborarea metodei.

Același rezultat se obține și prin bruiajul rețelei prin impunerea unei valori anormale unei conexiuni din rețea, bruiajul urmând a se propaga prin întreaga rețea, fapt ce conduce pentru moment la o valoare mare a erorii totale concomitent cu scoaterea rețelei din minimul local.

I.2.5.7 Îmbunătățiri ale algoritmului backpropagation

Cercetători în domeniu au lansat variațiuni pe tema algoritmului **back-propagation** care s-au dorit a fi algoritmi îmbunătățiți în ce privește etapa de antrenare a rețelei. Evaluarea contribuției fiecăruia este prematură, valoarea noilor algoritmi trebuind să treacă testul de duranță al aplicării și verificării în condiții în care metodele actuale nu dau rezultate satisfăcătoare. În această secțiune se vor discuta câțiva dintre acești algoritmi.

O îmbunătățire semnificativă în domeniu este înlocuirea metodei de minimizare simplistă pe baza gradientului la fiecare pas, cu o metodă de minimizare direcționată cuplată cu minimizarea gradientului conjugat. Minimizarea unei funcții f de-a lungul unei direcții u aduce funcția f într-un loc în care gradientul ei este perpendicular pe direcția u . În caz contrar, procesul de minimizare de-a lungul direcției u este incomplet. În locul urmaririi gradientului pas cu pas, se construiește un set de n direcții conjugate între ele, astfel că minimizarea de-a lungul uneia dintre ele u_j nu se repercutează în mod distructiv asupra minimizării de-a lungul unei direcții anterioare, practic minimizările de-a lungul direcțiilor nu interferă. Sistemul de gradul n va evolua deci, spre un minim global. Metoda este diferită de metoda gradientului descendent ce presupune minimizarea în direcția cu gradient descendent maxim.

Acceptând aproximarea prin serie Taylor a funcției ce urmează a fi minimizată, se obține:

$$f(x) = f(p) + \sum_i \frac{\partial f}{\partial x_i} \Big|_p x_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_p x_i x_j + \dots \cong \frac{1}{2} x^T A x - b^T x + c \quad (59)$$

unde:

- T denotă operațiunea de transpunere

- $c = f(p)$

- $b = -\nabla f \Big|_p$

- $[A]_{i,j} \equiv \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_p$ este o matrice simetrică de dimensiuni $n \times n$, **pozitiv**

definită ($\forall y \neq 0, y^T A y > 0$), de fapt, Hessiana funcției f în punctul p :

$$A_p = \begin{bmatrix} \frac{\partial^2 f}{\partial x_i^2}(p) & \frac{\partial^2 f}{\partial x_i \partial x_j}(p) \\ \frac{\partial^2 f}{\partial x_i \partial x_j}(p) & \frac{\partial^2 f}{\partial x_j^2}(p) \end{bmatrix} \quad (60)$$

Gradientul funcției f este:

$$\nabla f = Ax - b \quad (61)$$

astfel că o schimbare în x induce o schimbare de gradient conform relației:

$$\delta(\nabla f) = A(\delta x) \quad (62)$$

Presupunând că funcția f a fost minimizată de-a lungul direcției u_i într-un punct în care gradientul $-g_{i+1}$ al funcției este perpendicular pe direcția u_i ,

$$u_i^T g_{i+1} = 0 \quad (63)$$

se trece la minimizarea pe o nouă direcție u_{i+1} . Pentru a ne asigura că parcurgerea direcției u_{i+1} nu afectează minimizarea de-a lungul direcției u_i este necesar ca gradientul funcției f să rămână perpendicular pe direcția u_i , astfel încât:

$$u_i^T g_{i+2} = 0 \quad (64)$$

Altfel, va mai trebui parcursă o dată etapa de minimizare în direcția u_i . Combinând relațiile (63) și (64) se obține:

$$0 = u_i^T (g_{i+1} - g_{i+2}) = u_i^T \delta(\nabla f) = u_i^T A u_{i+1} \quad (65)$$

Când relația (62) se aplică pentru doi vectori u_i și u_{i+1} , se spune despre aceștia că sunt **conjugați**.

Plecând de la același punct p_0 , prima direcție de minimizare u_0 se va lua egală cu: $g_0 = -\nabla f(p_0)$. Pentru $i \geq 0$, direcțiile se calculează cu relația:

$$u_{i+1} = u_i + \lambda_i^u u_i \quad (66)$$

unde valoarea coeficientului incremental λ_i^u este aleasă astfel încât $u_i^T A u_{i+1} = 0$ și gradientii succesivi să fie perpendiculari:

$$\gamma_i = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i} \quad \text{cu} \quad g(k) = -\nabla f|_{p^k} \text{ pentru } k \geq 0 \quad (64)$$

Următorul pas este calculul:

$$p_{i+2} = p_{i+2} + \lambda^p_{i+1} u_{i+1} \quad (68)$$

unde coeficientul λ^p_{i+1} se alege astfel încât să minimizeze $f(p_{i+2})^2$.

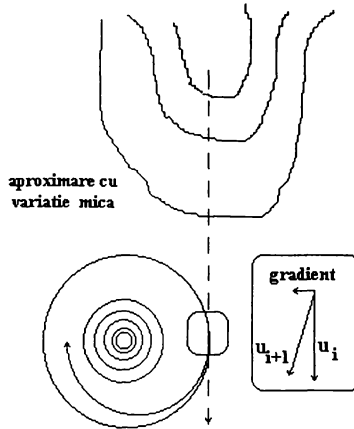


Fig.18. Coborârea cu pantă mică cu gradient conjugat pentru sistemele nepatratice

Se poate demonstra că direcțiile u astfel construite sunt toate interconjugate^{48,49}. Din figura 18 se observă că la o nouă selecție de direcție, aceasta se contruiește din prima direcție și gradient, rezultând o minimizare în spirală.

Astfel, pentru un sistem patratice cu n grade de libertate sunt necesare doar n iterații. Având în vedere că nu se urmărește minimizarea unui sistem patratice, este necesară parcurgerea celor n direcții de mai multe ori.

Powell⁵⁰ a introdus câteva îmbunătățiri semnificative pentru a corecta comportarea metodei în sistemele nepătratic.

Van den Boomgaard&Smeulders⁵¹ - 1989 au arătat că pentru o rețea neurală cu propagare înainte fără unități ascunse, procedura incrementală de aflare a matricii de ponderi optime necesită o ajustare a ponderilor cu:

$$\Delta W(t+1) = \eta(t+1) \cdot (d(t+1) - W(t)x(t+1)) x(t+1) \quad (69)$$

în care η nu este o constantă ci o variabilă tip matrice $(N_i+1) \cdot (N_i+1)$ care depinde de vectorul de intrare. Utilizând cunoștințe apriori despre semnalul de intrare, necesitățile de stocare pentru valorile termenului moment γ pot fi reduse substanțial.

Silva și Almeida⁵¹ au arătat deasemenea avantajele utilizării unui pas de valoare specifică pentru fiecare pondere din rețea. În algoritmul lor, rata de învățare este adaptată pentru fiecare tipar:

$$\eta_{i,j}(t+1) = \begin{cases} u\eta_{ij}(t) & \text{dacă } \frac{\partial E_{i+1}}{\partial w_{ij}} \text{ și } \frac{\partial E_t}{\partial w_{ij}} \text{ au același semn} \\ d\eta_{ij}(t) & \text{dacă } \frac{\partial E_{i+1}}{\partial w_{ij}} \text{ și } \frac{\partial E_t}{\partial w_{ij}} \text{ au semne diferite} \end{cases} \quad (70)$$

unde u și d sunt constante pozitive cu valori puțin peste și respectiv puțin sub unitate. Ideea de bază a acestei metode este de a descrește rata de învățare în cazul oscilațiilor valorilor ponderilor.

1.2.5.8 Parametrii de antrenare a rețelei

În etapa de antrenare a rețelei, conform algoritmului de învățare, există câțiva parametri care definesc această etapă. Enumerarea doar a unora nu conduce la o descriere satisfăcătoare nici a rețelei și nici a capacității de învățare a acesteia.

În figura 19 este prezentată variația funcției sigmoide unipolare și a derivatei pentru valori ale factorului de temperatură β cuprinse între 0.5 - 2 utilizând un pas de 0.1 unități.

Se observă că pentru două valori vecine ale variabilei, creșterea ambelor funcții este mai substanțială pentru valori ale factorului de temperatură k mai mari. De aici rezultă două consecințe:

- 1 - la normalizarea valorilor cu funcția sigmoidă, pentru două valori apropiate ale variabilei, valorile funcției sunt progresiv diferite odată cu creșterea factorului de temperatură
- 2 - având în vedere relația (56) de corecție a ponderilor, pentru două valori apropiate ale erorii, corecția este mai drastică la valori ale factorului de temperatură crescute.

Informația despre valoarea absolută a acestui factor nu este relevantă decât dacă este completată cu valorile ratei de învățare η și ale termenului moment γ .

Aceste variabile de rețea intervin doar în corecția ponderilor în vederea reducerii erorii rețelei de la o epocă la alta.

Rata de învățare η este de fapt o constantă de proporționalitate cu valori între 0 și 1 cu care se ajustează corecția ponderii. O valoare a ratei de învățare de 0.1 înseamnă practic că se ține seama doar de 10% din valoarea derivatei funcției sigmoide din termenul de corecție al ponderii.

Odată caracterizat termenul de corecție al ponderilor prin valoarea de factorului de temperatură β și prin valoarea ratei de învățare, în cazul utilizării termenului moment, se intervine din nou printr-un proces de ajustare a acestuia conform relației (56) în funcție de viteza de convergență a rețelei spre valoarea de minim de eroare.

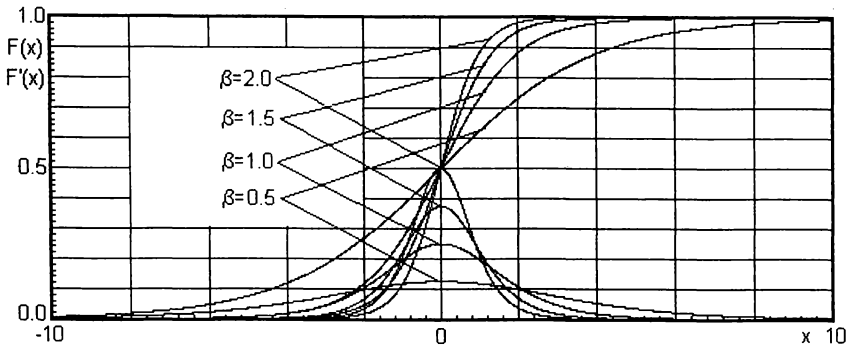


Fig. 19. Variația funcției sigmoide unipolare și a derivatei sale pentru valori ale factorului de temperatură β cuprinse între 0.5 - 2 utilizând un pas de 0.5 unități

Importanța termenului **moment** este discutabilă deoarece acesta aplică o ajustare a algoritmului de corecție, deci o scădere a vitezei de convergență spre minim chiar și atunci când rețeaua se află pe drumul descendent către minimul global. Oricum valoarea sugerată în literatură²⁴ $\gamma=0.9$ este mult prea mare.

Această valoare impune de fapt o corecție actuală nesemnificativă punând accentul pe corecția exercitată într-o epocă anterioară. Aplicarea termenului moment conduce de fapt la o mediere ponderată și consecutivă a factorului de corecție a ponderilor.

Deci, o comparare a funcționării unor rețele în etapa de antrenare este incompletă dacă nu sunt prezentați toți factorii discutați anterior⁵²⁻⁵⁴.

I.2.5.9 Accelerarea convergenței prin utilizarea unei rate de învățare adaptive

În această secțiune se vor descrie și analiza câteva metode euristice descrise de Jacobs⁵⁵ ca modalități de accelerare a convergenței învățării prin adaptarea ratei de învățare.

Metoda euristică 1. - Fiecare parametru ajustabil al rețelei care descrie și funcția de cost a învățării trebuie să fie evaluat în particular la o valoare optimă proprie valorii de moment a funcției de cost.

Această metodă propune o valoare diferită pentru fiecare pondere argumentând prin faptul că la un anumit moment, o anumită ajustare poate fi benefică pentru o anumită pondere și total necorespunzătoare pentru o altă pondere.

Metoda euristică 2. - Valoarea ratei de învățare trebuie să fie diferită de la o epocă la alta.

Această metodă are în vedere faptul că suprafața de eroare diferă de-a lungul regiunii corespunzătoare unei anumite ponderi.

Metoda euristică 3. - Când derivata funcției de cost în raport cu o anumită pondere are același semn pentru câteva iterații consecutive ale algoritmului, rata de învățare pentru acea pondere trebuie crescută.

Metoda are în vedere posibilitatea ca punctul de operare curent în spațiul ponderii vizate poate ajunge pe o porțiune plană a suprafeței de eroare, fapt rezultat din consecvența semnului pe parcursul mai multor epoci. Printr-o creștere a valorii ratei de învățare, numărul de iterații pentru traversarea acestei suprafețe plane va fi mai mic, fapt ce duce la o îmbunătățire a valorii funcției de cost.

Metoda euristică 3. - *Când semnul algebric al derivatei funcției de cost în raport cu o anumită pondere alternează consecutiv pe parcursul mai multor iterații ale algoritmului, rata de învățare pentru acea pondere trebuie micșorată.*

Metoda are în vedere posibilitatea ca punctul de operare curent în spațiul ponderii vizate să ajungă pe o porțiune foarte curbată a suprafeței de eroare, fapt rezultat din inconsecvența semnului pe parcursul mai multor epoci. Astfel, valoarea ponderii oscilează ca semn, rămânând relativ constantă în valoare absolută ajungându-se la o creștere a funcției de cost fără vreun real progres.

Este cert că utilizarea parametrilor de rețea în raport cu fiecare din ponderi conform acestor metode euristice duce inevitabil la o modificare fundamentală a algoritmului backpropagation care nu mai corespunde unei căutări în direcția descendentă spre minim.

Deci ajustările aplicate ponderilor se bazează pe:

- derivatele parțiale ale suprafeței de eroare în raport cu ponderile;
- estimarea cuadraturii suprafeței de eroare pe care se situează punctul de operare în spațiul ponderilor.

Astfel, toate patru metodele euristice satisfac doar necesități locale ale rețelei, fiind benefice pentru anumite puncte și total necorespunzătoare pentru altele.

I.2.6. Aplicații ale rețelelor neurale cu propagare înapoi în chimie și inginerie chimică

Algoritmul back-propagation a fost utilizat încă de la lansare într-o varietate relativ largă în domeniile chimiei și ingineriei chimice. O clasificare a subtipurilor de rețele din clasa back-propagation utilizate este redată schematic mai jos.

INTRARE	IESIRE	ALGORITM	
R sau I	I,R	uzual	I
R sau I	I	WTA	II
R sau I	R	uzual	III

unde notațiile corespund:

I=intreg ;

R=real ;

WTA -ieșirea cu valoare maximă =1, restul ieșirilor =0

(algoritm Winner Takes All)

Prima clasă de rețele neurale în care ieșirile sunt de tip întreg își găsește aplicații de extensie restrânsă în estimarea unor probabilități (ieșirile sunt de tip real) sau certitudini (ieșirile sunt de tip întreg) de apartenență la anumite clase.

Cea de-a doua clasă de rețele neurale are aplicații în gruparea (clustering) unor tipare în anumite categorii de apartenență care le caracterizează. Algoritmul "winner takes all" (învingătorul ia totul) presupune ca apartenența la o clasă să fie determinată de ieșirea de valoare maximă căreia i se atribuie valoarea 1, celorlalte ieșiri atribuindu-li-se valoarea 0. Astfel, valoarea maximă a unui element din vectorul de ieșire presupune apartenența vectorului de intrare la clasa determinată de acest element.

Cea de-a treia clasă este cea mai utilizată și vizează aplicații în care valorile ponderilor rețelei devin esențiale pentru a estima dependența vectorului de ieșire de vectorul de intrare. Toate aplicațiile din lucrarea de față sunt bazate pe acest ultim tip de rețele neurale, diversitatea lucrării derivând din abordarea unor domenii diferite.

În ce privește domeniile chimiei și ale ingineriei chimice abordate prin intermediul rețelilor neurale, acestea sunt de o diversitate mare, o încercare de grupare a acestora fiind redată în cele ce urmează⁵⁶.

1. Aplicații ale rețelilor neurale în chimie

- corelarea structurii - cu activitatea chimică, biologică sau cu proprietăți ale compușilor⁵⁷⁻⁸¹
- identificarea și clasificarea compușilor pe baza:
 - spectrelor de rezonanță magnetică nucleară⁸²⁻⁸⁶
 - spectrelor de mobilitate ionică^{87,88}
 - spectrelor IR⁸⁹⁻⁹³
 - spectrelor de masă^{94,95}
 - analizei cromatografice⁹⁶

2. Aplicații ale rețelilor neurale în inginerie chimică

- modelarea în regim dinamic și controlul proceselor din industria chimică^{97-106,112}

Domeniile industriei chimice abordate sunt:

- industria cimentului
- industria petrochimică
- tratarea apelor reziduale
- industriei electrochimice

și operații unitare întâlnite în majoritatea proceselor din industria chimică:

- reacții în reactoare de tipologii diferite,
- coloane de distilare
- filtrare
- clasare și sortare.

În ceea ce privește evoluția controlului proceselor din industria chimică, în ultimul deceniu s-a pus un accent deosebit pe conducerea de tip fuzzy și cea bazată pe rețelele neurale. Aceste două metode presupun într-o accepție inițială utilizarea acestor două tehnici pe baza algoritmilor clasici de tip PID, până la dezvoltarea tehnicilor într-o manieră de aplicare independentă.

II. CORELĂRI STRUCTURĂ ELECTRONICĂ - PROPRIETĂȚI FIZICE

II.1 Noțiuni generale despre proprietățile fizice ale elementelor

II.1.1 Căldura specifică

Este definită prin cantitatea de căldură necesară pentru ridicarea (fără schimbare de fază și la presiune constantă) a temperaturii unei unități de cantitate (1 mol sau 1 gram) de substanță cu 1°C. Notația consacrată pentru căldura specifică la presiune constantă este C_p .

Variația acestei proprietăți în tabelul periodic al elementelor a fost observată încă din anul 1818 de către Dulong și Petit¹⁰⁸ care au sesizat că la substanțele elementare solide cu greutate atomică peste 35, produsul dintre căldura specifică a unui gram și greutatea atomică (căldura specifică molară) este aproximativ constant, având valoarea de $\cong 26$ J/grad-mol.

Această regulă a fost utilizată înainte de anul 1900 pentru a obține greutatele atomice aproximative ale câtorva elemente, împărțind valoarea de 26 cu căldura specifică măsurată pentru 1 gram de substanță elementară solidă. De exemplu, pentru bismut, care are căldura specifică 0,123 kJ/kg·K, se obține pentru greutatea atomică valoarea de 211 kg/kmol ceea ce s-a dovedit a fi o valoare foarte apropiată de cea reală 209 kg/kmol.

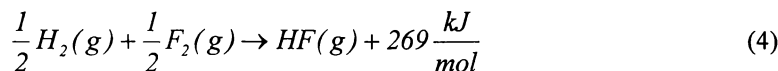
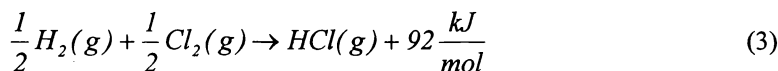
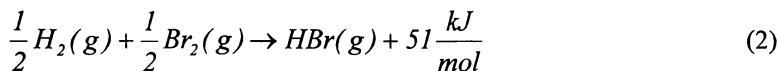
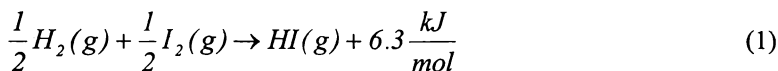
Importanța cunoașterii acestei proprietăți elementare rezidă din posibilitatea valorificării elementelor în stare pură sau aliaje la procese speciale de transfer termic. De aceea, predicția valorii căldurii specifice pentru elementele artificiale constituie o aplicație de interes major din punct de vedere teoretic și practic

II.1.2 Electronegativitatea elementelor

Elementelor le pot fi atribuite numere care în cazul unei legături covalente reprezintă puterea lor de atracție pentru electroni. Pe baza acestor numere poate fi evaluat caracterul parțial ionic al legăturii. Această măsură relativă a puterii de atracție a electronilor într-o legătură covalentă este denumită *electronegativitatea elementului* fiind notată cu simbolul literal χ . Deoarece această notație este utilizată în subcapitolul următor la notarea unor indici topologici, în această lucrare, pentru electronegativitate s-a utilizat notația **E**.

Scara electronegativității se întinde de la cesiu (0.7) până la fluor (3.98) care este de departe cel mai electronegativ element din sistemul periodic.

Atomii de hidrogen și iod deși foarte diferiți, au o electronegativitate aproximativ egală. În legătura H-I, cei doi atomi exercită aproximativ aceeași atracție asupra perechii de electroni puși în comun care constituie de fapt legătura covalentă dintre ei. Deci, această legătură este foarte asemănătoare cu a legăturii total covalente din moleculele elementare H-H și I-I. Prin simplă deducție, energia legăturii H-I este foarte apropiată de energia medie a legăturilor H-H și I-I. Considerând reacțiile:



cu valorile căldurilor de formare corespunzătoare, se observă că sunt într-o proporționalitate directă cu diferențele de electronegativitate dintre moleculele halogenurilor de hidrogen (0.7 - HBr, 0.9-HCl, 1.9-HF) ce descriu caracterul ionic al legăturilor (12% - HBr, 17%-HCl, 45%-HF).

Aceasta înseamnă că legăturile din moleculele ultimelor 3 halogenuri de hidrogen sunt mai puternice decât media legăturilor în moleculele substanțelor elementare respective și că energia mai mare a legăturii este determinată de diferența dintre electronegativitățile celor doi atomi. De fapt, cu cât diferența de electronegativitate dintre două elemente este mai mare, cu atât legătura dintre ele este mai puternică. Practic, proprietățile elementare sunt corelate reciproc descriind și comportamentul chimic al elementelor.

Diferența dintre valorile electronegativităților a două elemente caracterizează într-o bună măsură natura legăturii compusului pe care îl formează fiind și o caracteristică a stabilității acestuia.

Concluziv, o bună predicție a electronegativității pentru elementele pentru care această proprietate nu este cunoscută, prezintă importanța estimării caracteristicilor compușilor ipotetici pe care acestea i-ar putea forma.

Scopul aplicației este crearea unui model neural de corelare a structurii electronice cu valoarea electronegativității elementelor și de predicție a valorii electronegativității elementelor artificiale care este necunoscută.

II.1.3 Aplicații existente în domeniu

Sigman&Rives¹⁰⁸ descriu o aplicație similară de predicție a potențialelor de ionizare I-III utilizând o rețea neurală artificială cu propagare înapoi. Pentru 222 tipare cu valori cuprinse între

3,984 V (Cs)	- 24,587 V (He)	- pentru primul potențial de ionizare
10,004 V (Ba)	- 76,638 V (Li)	- pentru al doilea potențial de ionizare
19,174 V (La)	- 153,893 V (Be)	- pentru al treilea potențial de ionizare

utilizând o rețea neurală cu 7 unități pe stratul de intrare, 6 unități pe stratul ascuns, și 1 unitate pe stratul de ieșire, au fost necesare 96 000 epoci pentru parcurgerea etapei de învățare.

Pentru validarea modelului a fost utilizată o procedură de validare denumită generos “validare încrucișată” prin care s-au omis 10 tipare selectate arbitrar din setul de antrenare: Se(34-I), Ac(89-I), Al(13-II), Sr(38-II), Pd(46-II), Yb(70-II), B(5-III), Mg(12-III), Ce(58-III), Th(90-III) unde prima cifră reprezintă numărul atomic iar cea de-a doua primul, al doilea, respectiv al treilea potențial de ionizare.

Această corelare între configurația electronică și potențialele de ionizare este logistic simplu de intuit având în vedere că valoarea potențialului de ionizare depinde în mod direct de stratul de pe care se cedează electronul. Intre valorile calculate și cele experimentale aceștia obțin un coeficient de corelare liniară de 0,976, care denotă într-adevăr corelarea scontată.

Având în vedere domeniul larg al potențialelor de ionizare 3,984 - 153,983V variabilelor de ieșire, Sigman&Rives utilizează pentru obținerea unui coeficient de corelare bun nu o normalizare a acestor valori ci o logaritmare. Această intervenție de tip logaritmă este ușor forțată, singurul scop fiind de a îmbunătăți parametrii de corelare liniară dintre valorile calculate și cele experimentale prin reducerea domeniului de împrăștiere a valorilor variabilelor de ieșire.

II.2 Corelare structură electronică - căldură specifică

Scopul aplicației este predicția căldurii specifice a elementelor artificiale utilizând o rețea neurală antrenată cu valori argumentate ale vectorilor de intrare și ieșire pentru elementele pentru care această proprietate este cunoscută.

II.2.1 Alegerea și justificarea variabilelor de corelare

Pentru corelarea structură electronică - căldură specifică, s-au luat în considerare elementele pentru care această proprietate este cunoscută, folosind ca variabile independente valori ce reflectă configurația electronică a elementului iar ca variabilă independentă - valoarea proprie a căldurii specifice.

Pentru a descrie cât mai fidel schimbările configurației electronice în sistemul periodic al elementelor, **pentru variabilele independente** s-a utilizat următoarea mapare:

-**prima variabilă independentă** a fost dată de perioada din care face parte elementul, făcându-se astfel diferențierea între orbitalii de completare și ceilalți orbitali de același gen

-**următoarele patru variabile independente** au fost date de numărul electronilor **s, p, d, f** de pe orbitalul în completare.

-**variabila independentă**, s-a ales valoarea proprie a căldurii specifice a elementului.

Astfel, pentru corelare, s-au avut în vedere cinci variabile independente. În tabelul 1 sunt redată spre exemplificare variabilele pentru trei elemente selectate aleator.

Tabelul 1 Exemplificarea variabilelor pentru corelarea structură electronică - căldură specifică.

Număr atomic	Element	Variabile independente	Variabila dependentă Cp [kJ/kg·K]
21	Sc	4,2,0,1,0	0,6
50	Sn	5,2,2,10,0	0,227
87	Fr	7,1,0,0,0	nedeterminat exp.

II.2.2 Justificarea utilizării modelului neural

Intr-un prim stadiu, a fost testată o regresie multiliniară a valorilor variabilelor independente cu variabilă dependentă - căldura specifică.

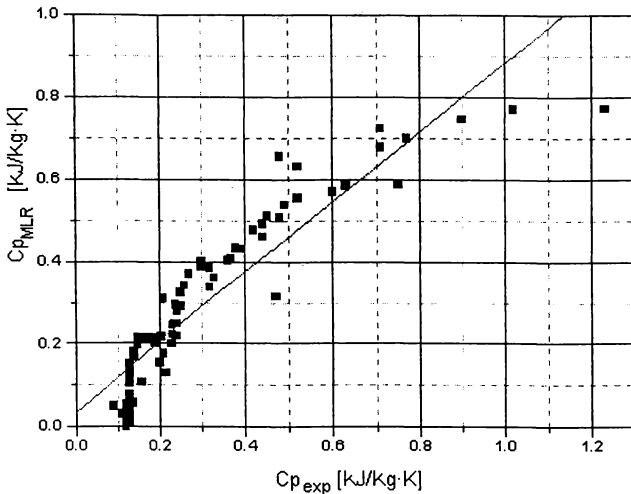


Fig.1 Reprezentarea grafică a relației liniare dintre căldura specifică calculată prin regresie multiliniară $C_{p_{MLR}}$ și cea experimentală $C_{p_{exp}}$.

În figura 1 este redată reprezentarea grafică a relației liniare dintre căldura specifică calculată prin regresie multiliniară $C_{p_{MLR}}$ și valoarea experimentală a căldurii specifice $C_{p_{exp}}$, iar relația 5 redă algoritmic această dependență:

$$C_{p_{MLR}} = 1,3253(\pm 0,0644) - 0,1841(\pm 0,0116) \cdot P + 0,0021(\pm 0,0160) \cdot s - 0,0230(\pm 0,0072) \cdot p - 0,0155(\pm 0,0025) \cdot d + 0,0002(\pm 0,0023) \cdot f \quad (5)$$

unde $C_{p_{MLR}}$ este valoarea calculată prin regresie multiliniară a căldurii specifice, P este perioada iar s, p, d, f numărul de electroni de pe ultimul strat de tip s, p, d , respectiv f .

Parametrii de corelare sunt:

Volumul eșantionului	$N = 82$
Numărul gradelor de libertate	$NL = 76$
Coeficientul de corelare	$R = 0,918$
Coeficientul de determinare	$R^2 = 0,842$
Deviația standard	$SD = 0,091$
Testul Fisher	$F = 405$

După cum se observă, valoarea coeficientului de corelare este mică, iar valorile de deviație ale coeficienților au valori comparabile cu ale acestora. Corelarea prin regresie multiliniară este deci nesatisfăcătoare, în cele ce urmează demonstrându-se superioritatea modelului neural pentru corelarea căldurii specifice cu configurația electronică a elementelor.

II.2.3 Modelul neural - Caracterizare

II.2.3.1 Numărul straturilor ascunse și al unităților ascunse

În cazul rețelelor cu propagare înapoi cu funcții de activare neliniare și unități binare s-a demonstrat că un singur strat rezolvă problema asimilării oricărei funcții continue. În cazul rețelelor cu propagare înapoi cu unități reale există lucrări^{65-67,77} care utilizează cel puțin două straturi ascunse. Pe de o parte nu există nici o argumentare logică a utilizării mai multor straturi ascunse iar pe de altă parte nu s-a reclamat în domeniu un randament mai bun al acestora în detrimentul rețelelor cu un singur strat ascuns. Deși această opțiune pentru un singur strat ascuns este susținută în ultimii 2 ani pe bază experimentală, utilizarea mai multor straturi de unități ascunse nu este principial greșită. Evitarea utilizării exclusive a unui singur strat ascuns poate fi interpretată ca o măsură de prevedere cu scopul de a evita un impact care ar surveni în măsura în care cercetări ulterioare ar demonstra limitele rețelelor cu un singur strat ascuns.

Prin urmare, rețeaua a fost dotată cu un singur strat de unități ascunse¹⁰⁹, singura problemă fiind cea a numărului de unități de pe stratul ascuns. Într-adevăr, numărul de unități de pe stratul ascuns a generat probleme a căror rezolvare s-a efectuat local doar pentru cazuri specifice, neexistând încă o generalizare a prescripțiilor emise de autori.

Pentru determinarea numărului de unități de pe stratul ascuns¹¹⁰, rețeaua a fost antrenată cu un număr diferit de unități pe stratul ascuns, menținând constante celelalte caracteristici - funcția de activare, rata de învățare, termenul de moment, algoritmul aleator de inițializare a ponderilor, algoritmul de corecție al acestora și numărul de epoci de rulare.

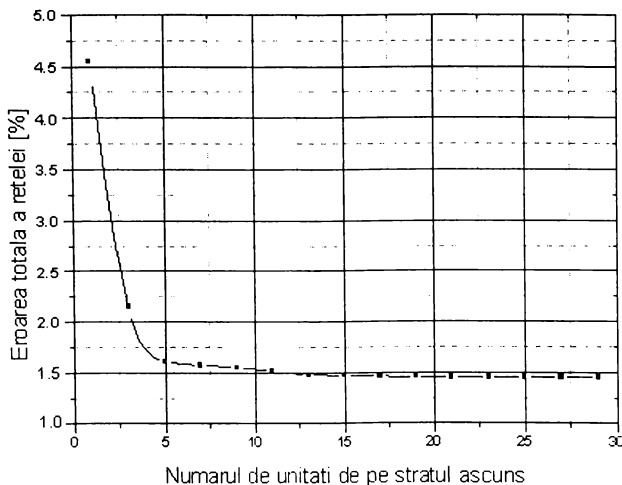


Fig. 2 Reprezentarea grafică a variației erorii totale a rețelei funcție de numărul de unități de pe stratul ascuns

Având în vedere algoritmul de inițializare aleatoare a matricilor de ponderi, fiecare punct al valorii erorii totale a rețelei a fost obținut prin **rularea rețelei până la confirmarea valorii celei mai mici** pentru același număr de unități de pe stratul ascuns.

Din figura 2, se observă că valoarea minimă a erorii totale a rețelei se obține pentru 22 unități pe stratul ascuns, valoare care a fost utilizată pentru rezolvarea aplicației propuse spre rezolvare.

Creșterea numărului de unități de pe stratul ascuns peste această valoare nu împieteează asupra valorii erorii totale, rețeaua evoluând în aceeași direcție favorabilă fără un consum semnificativ de timp suplimentar.

De fapt, unicul motiv de reducere a numărului de unități de pe stratul ascuns este consumul suplimentar de timp, consum care în contextul predicției valorilor căldurii specifice pentru elementele artificiale este preferabil decât o predicție de o precizie mai scăzută.

Utilizând un program realizat în Turbo Pascal 7 care salvează pe hard-disk valorile ponderilor la fiecare 10 epoci nu au fost necesare mai mult de 15 de minute pentru parcurgerea a 2000 de epoci.

În cazul utilizării unui număr redus de unități pe stratul ascuns concomitent cu rularea unui număr de epoci mai mare, există posibilitatea supraînvățării rețelei cu rezultate notabile în ce privește reproducerea tiparelor învățate dar cu capacități scăzute de predicție pentru tiparele noi.

II.2.3.2 Selecția tiparelor

Având în vedere că scopul aplicației îl constituie predicția căldurii specifice pentru elementele artificiale, selecția tiparelor în vederea realizării scopului propus este o etapă de o importanță majoră.

În figura 3 este redată variația valorii căldurii specifice cu numărul atomic al elementelor din tabelul periodic al elementelor¹¹¹.

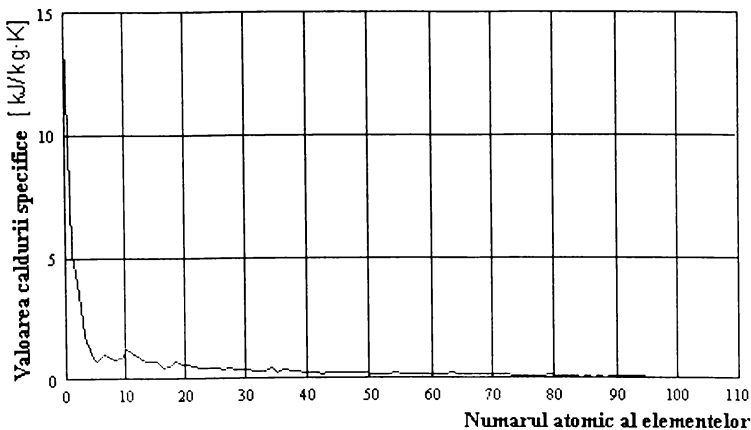


Fig 3. Reprezentarea grafică a variației căldurii specifice în tabelul periodic al elementelor.

Se observă o diferență semnificativă între valorile căldurii specifice ale primelor 10 elemente în comparație cu valorile celorlalte elemente.

S-a constatat experimental că există doi factori care conduc la rezultate slabe ale procesului de învățare:

1 - un domeniu larg de împrăștiere al variabilelor de ieșire

2 - o acoperire inegală a domeniului variabilelor de ieșire.

Pentru a caracteriza primul factor, se va defini un **factor de împrăștiere SF** (spreading factor) descris de relația¹⁰⁹:

$$SF = \frac{MAX_{out}}{MIN_{out}} \quad (6)$$

unde: - MAX_{out} este valoarea maximă a variabilei de ieșire

- MIN_{out} este valoarea minimă a variabilei de ieșire

De reținut că scăderea influenței primului factor nu se poate face prin normalizarea valorilor variabilelor de ieșire, valoarea factorului de împrăștiere rămânând din definiție aceeași. Sensul acestui factor este de a decree largimea domeniului variabilei de ieșire, experimental dovedindu-se că valoarea factorului de împrăștiere este de dorit a fi cât mai mică.

Selecția inițială a tiparelor pentru etapa de antrenare este importantă deoarece rețeaua va învăța mai bine mecanismul de corelare a variabilelor de intrare cu cele de ieșire pentru subdomeniile cele mai bine reprezentate în domeniul total de variație.

Având în vedere scopul principal al aplicației și faptul că valoarea căldurii specifice pentru hidrogen este de peste 150 de ori mai mare decât cea a radonului, nu este recomandabil a folosi tiparele primelor elemente pentru etapa de învățare.

Pentru această aplicație, în cazul în care se iau în considerare tiparele corespunzătoare pentru toate elementele, factorul de împrăștiere are valoarea:

$$SF = 14.3(H)/0.09(Rn) = 159 \quad (7)$$

Omitând tiparele corespunzătoare primelor 2 elemente, valoarea factorului de împrăștiere va ajunge la:

$$SF = 3.6/0.09(Rn) = 40 \quad (8)$$

iar omitând tiparele corespunzătoare primelor 10 elemente, va ajunge la:

$$SF = 1.23(Na)/0.09(Rn) = 13.7 \quad (9)$$

Antrenând rețeaua fără tiparele corespunzătoare primelor 2 elemente și reprezentând grafic valoarea calculată a căldurii specifice în funcție de valoarea experimentală, se obține o dependență liniară de tipul celei din figura 4.

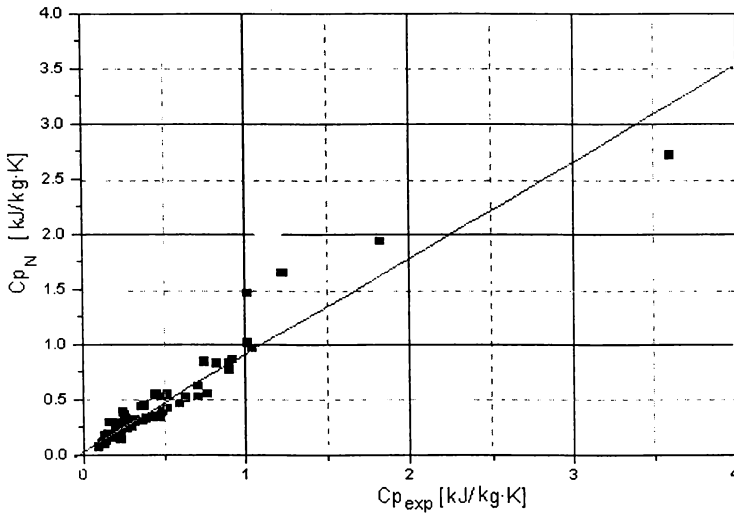


Fig. 4 Dependența liniară dintre valorile de predicție ale căldurii specifice C_{p_N} și cele experimentale $C_{p_{exp}}$ în cazul omiterii din etapa de învățare a primelor 2 tipare

După cum se observă, acuratețea predicției lasă de dorit pentru cele câteva valori care se constituie ca excepție 3,5 (Li), 1.82 (Be), 1.23(Na), 1.04(N), 1.02(B) .

Motivul diferenței între valoarea de predicție și cea experimentală este evident: insuficienta reprezentare a valorilor de ieșire în domeniul 1 - 3,6 kJ/kgK.

Concluzia care derivă din această constatare este că omițând tiparele corespunzătoare primelor 10 elemente, rețeaua va putea fi antrenată în vederea predicției valorilor căldurii specifice pentru elementele pentru care această proprietate este încă necunoscută.

II.2.3.3 Descrierea rețelei neurale utilizate

Schema bloc a rețelei neurale utilizate este redată în figura 5. Notațiile din schemă corespund:

- W1 - matricea de ponderi dintre stratul de intrare și stratul ascuns
- W2 - matricea de ponderi dintre stratul ascuns și stratul de ieșire

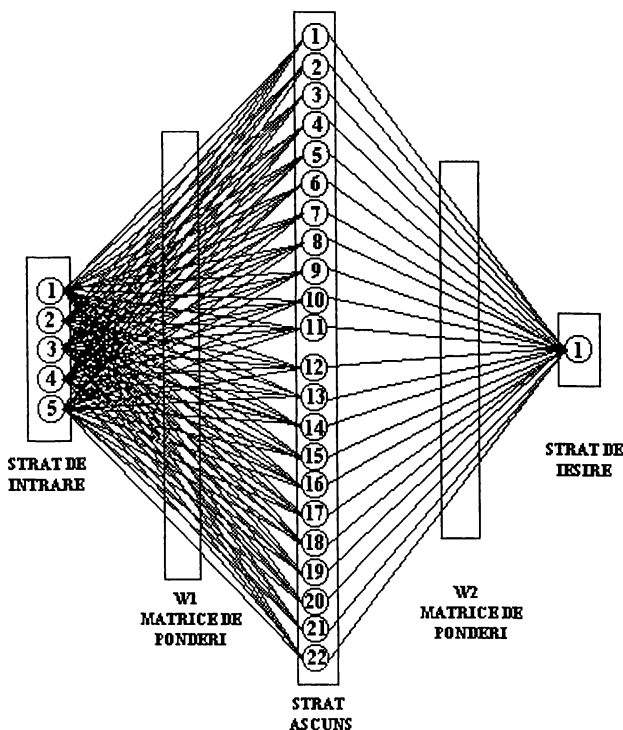


Fig.5. Schema bloc a rețelei neurale utilizate

Principalele variabile ale rețelei neurale utilizate sunt:

- Funcția de activare - funcția sigmoidă $F(x) = \frac{1}{1 + e^{-\beta \cdot x}}$ cu $\beta=3$
- Rata de învățare $\eta = \text{intre } 1 \text{ și } 0.001$
- Termenul **moment** $\gamma = 0.1$
- Nu s-au utilizat unități **bias**¹¹⁰
- Numărul de epoci $Nep = 2000$
- Numărul de unități de pe stratul de intrare $Ni = 5$
- Numărul de unități de pe stratul ascuns $Nhu = 22$
- Numărul de unități de pe stratul de ieșire $No = 1$
- Numărul de tipare $Ntip = 82$

II.2.4 Rezultate experimentale

Deci, pentru această aplicație, în etapa de învățare s-au utilizat tiparele elementelor cu număr atomic mai mare de 10, pentru care valoarea căldurii specifice este cunoscută.

În tabelul 3 sunt redată tiparele elementelor utilizate în etapa de antrenare a rețelei neurale în vederea predicției căldurii specifice a elementelor pentru care valoarea acestuia este nedeterminată experimental.

Valorile din tabel corespund:

In coloana 1 este dat numărul atomic al elementului în vederea identificării acestuia.

In coloanele 2-5 sunt trecute elementele vectorului de intrare:

- P - perioada din care face parte elementul

-s, p, d, f - numărul electronilor s, p, d, f de pe orbitalul în completare.

In coloana 6 este redată valoarea experimentală a căldurii specifice a elementului¹¹¹.

In coloana 7 este redată valoarea calculată cu rețeaua neurală a căldurii specifice a elementului.

In coloana 8 este redată eroarea procentuală a valorii calculate față de cea experimentală.

Tabelul 3

Tipare experimentale; valori experimentale și valori calculate ale căldurii specifice

Nr. atomic	Perioada	electroni s pe stratul de completare	electroni p pe stratul de completare	electroni d pe stratul de completare	electroni f pe stratul de completare	$C_{p,exp}$ [kJ/kgK]	$C_{p,N} \cdot 10^{-3}$ [kJ/kgK]	Eroare [%]
0	1	2	3	4	5	6	7	8
11.	3	1	0	0	0	1,23	1,16	-5,66
12.	3	2	0	0	0	1,02	1,046	2,57
13.	3	2	1	0	0	0,9	0,899	-0,07
14.	3	2	2	0	0	0,71	0,713	0,48
15.	3	2	3	0	0	0,77	0,769	-0,19
16.	3	2	4	0	0	0,71	0,654	-7,9
17.	3	2	5	0	0	0,48	0,529	10,11
18.	3	2	6	0	0	0,52	0,477	-8,35
19.	4	1	0	0	0	0,75	0,789	5,2
20.	4	2	0	0	0	0,63	0,62	-1,54
21.	4	2	0	1	0	0,6	0,584	-2,65
22.	4	2	0	2	0	0,52	0,533	2,47
23.	4	2	0	3	0	0,49	0,495	0,98
24.	4	1	0	5	0	0,45	0,451	0,29
25.	4	2	0	5	0	0,48	0,468	-2,58
26.	4	2	0	6	0	0,44	0,454	3,26
27.	4	2	0	7	0	0,42	0,434	3,3
28.	4	2	0	8	0	0,44	0,411	-6,59
29.	4	1	0	10	0	0,38	0,378	-0,58
30.	4	2	0	10	0	0,39	0,404	3,54
31.	4	2	1	10	0	0,37	0,356	-3,7
32.	4	2	2	10	0	0,32	0,326	1,79
33.	4	2	3	10	0	0,33	0,321	-2,64
34.	4	2	4	10	0	0,32	0,357	11,43
35.	4	2	5	10	0	0,47	0,461	-2
36.	4	2	6	10	0	0,25	0,245	-2,01
37.	5	1	0	0	0	0,363	0,364	0,15
38.	5	2	0	0	0	0,3	0,309	3,06
39.	5	2	0	1	0	0,3	0,289	-3,82

Tabelul 3 (Continuare)

0	1	2	3	4	5	6	7	8
40.	5	2	0	2	0	0,27	0,275	1,91
41.	5	1	0	4	0	0,26	0,243	-6,37
42.	5	1	0	5	0	0,25	0,241	-3,45
43.	5	1	0	6	0	0,21	0,239	13,98
44.	5	1	0	7	0	0,238	0,236	-0,87
45.	5	1	0	8	0	0,242	0,232	-4,28
46.	5	0	0	10	0	0,24	0,243	1,27
47.	5	1	0	10	0	0,235	0,232	-1,08
48.	5	2	0	10	0	0,23	0,236	2,52
49.	5	2	1	10	0	0,23	0,222	-3,31
50.	5	2	2	10	0	0,227	0,215	-5,4
51.	5	2	3	10	0	0,21	0,202	-3,93
52.	5	2	4	10	0	0,2	0,192	-4,04
53.	5	2	5	10	0	0,214	0,196	-8,47
54.	5	2	6	10	0	0,158	0,199	25,9
55.	6	1	0	0	0	0,24	0,205	-14,76
56.	6	2	0	0	0	0,204	0,194	-4,8
57.	6	2	0	1	0	0,19	0,182	-4,01
58.	6	2	0	0	2	0,19	0,175	-7,89
59.	6	2	0	0	3	0,19	0,18	-5,05
60.	6	2	0	0	4	0,19	0,191	0,35
61.	6	2	0	0	5	0,18	0,182	1,04
62.	6	2	0	0	6	0,2	0,176	-11,79
63.	6	2	0	0	7	0,18	0,203	12,74
64.	6	2	0	1	7	0,23	0,205	-10,85
65.	6	2	0	0	9	0,18	0,186	3,07
66.	6	2	0	0	10	0,17	0,173	1,87
67.	6	2	0	0	11	0,16	0,165	3,41
68.	6	2	0	0	12	0,17	0,16	-5,89
69.	6	2	0	0	13	0,16	0,155	-2,9
70.	6	2	0	0	14	0,15	0,151	0,78
71.	6	2	0	1	14	0,15	0,147	-2,01
72.	6	2	0	2	14	0,14	0,143	1,97
73.	6	2	0	3	14	0,14	0,139	-0,5
74.	6	2	0	4	14	0,13	0,137	5,13
75.	6	2	0	5	14	0,13	0,135	3,63
76.	6	2	0	6	14	0,13	0,133	2,51
77.	6	2	0	7	14	0,13	0,132	1,63
78.	6	1	0	9	14	0,13	0,132	1,72
79.	6	1	0	10	14	0,128	0,131	2,41
80.	6	2	0	10	14	0,139	0,129	-7,46
81.	6	2	1	10	14	0,13	0,13	-0,22
82.	6	2	2	10	14	0,13	0,13	0,12
83.	6	3	0	10	14	0,12	0,123	2,76
84.	6	4	0	10	14	0,12	0,115	-3,86

Tabelul 3 (Continuare)

0	1	2	3	4	5	6	7	8
85.	6	5	0	10	14	*	*	*
86.	6	6	0	10	14	0,09	0,102	13,47
87.	7	1	0	0	0	*	*	*
88.	7	2	0	0	0	0,12	0,153	27,21
89.	7	2	0	1	0	*	*	*
90.	7	2	0	2	0	0,12	0,137	13,95
91.	7	2	0	1	2	0,12	0,129	7,45
92.	7	2	0	1	3	0,12	0,129	7,87
93.	7	2	0	1	4	0,12	0,135	12,27
94.	7	2	0	0	6	0,13	0,138	6,49
95.	7	2	0	0	7	0,11	0,126	14,50

* Valoarea căldurii specifice nu este încă determinată experimental¹¹¹

Rețeaua antrenată cu tiparele stabilite a fost utilizată pentru reproducerea tiparelor învățate. In figura 6 este redată dependența liniară dintre valorile calculate și cele experimentale ale căldurii specifice pentru toate tiparele utilizate în etapa de învățare.

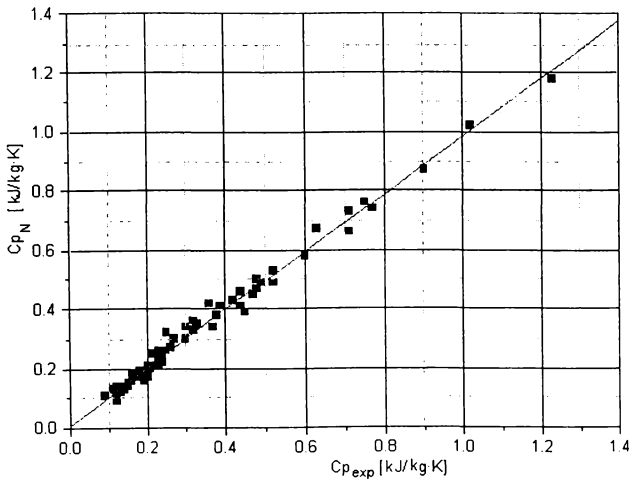


Fig. 6. Dependența liniară dintre valorile căldurii specifice calculate cu rețeaua neurală C_{p_N} și cele experimentale $C_{p_{exp}}$

Ecuția de corelare liniară dintre valorile calculate cu rețeaua neurală C_{p_N} și cele experimentale $C_{p_{exp}}$ este redată mai jos:

$$C_{p_N} = 0,0066(\pm 0,0041) + 0,9887(\pm 0,0111)C_{p_{exp}} \quad (10)$$

Parametrii de corelare sunt:

Volumul eșantionului	$N = 82$
Numărul gradelor de libertate	$NL = 80$
Coefficientul de corelare	$R = 0,995$
Coefficientul de determinare	$R^2 = 0,989(7)$
Deviația standard	$SD = 0,022$
Testul Fisher	$F = 7648$

Din valorile parametrilor de corelare se observă o foarte bună corelare prin regresie liniară a valorile calculate ale căldurii specifice cu cele experimentale.

II.2.5 Validarea modelului neural

Pentru verificarea și acceptarea modelului de corelare, s-a utilizat metoda validării încrucișate¹¹² (cross-validation procedure). Pentru o încărcare egală a domeniului de valori al variabilei de ieșire, întregul set de date a fost divizat în două seturi, criteriul de apartenență fiind dat de paritatea numărului curent al tiparului.

Având în vedere numărul înjumătățit de tipare, rețelei descrise anterior i s-au adus modificări în ce privește numărul de unități de pe stratul ascuns care a fost redus la 11 și numărul de epoci de antrenare 1000.

Intr-un prim stadiu, rețeaua a fost antrenată cu setul **impar** de date pentru predicția setului **par**, ulterior inversându-se procedura.

Ulterior etapei de învățare, rețeaua a fost alimentată doar cu vectorii de intrare ai celuilalt subset, și pe baza ponderilor calibrate, în cadrul unei semiepoci de propagare înainte s-a efectuat o predicție a valorilor vectorilor de ieșire.

În figura 7, este redată variația liniară dintre valoarea calculată a căldurii specifice pentru setul de date **par** utilizând rețeaua antrenată cu setul de date **impar** și valoarea experimentală a acesteia pentru setul de date **par**.

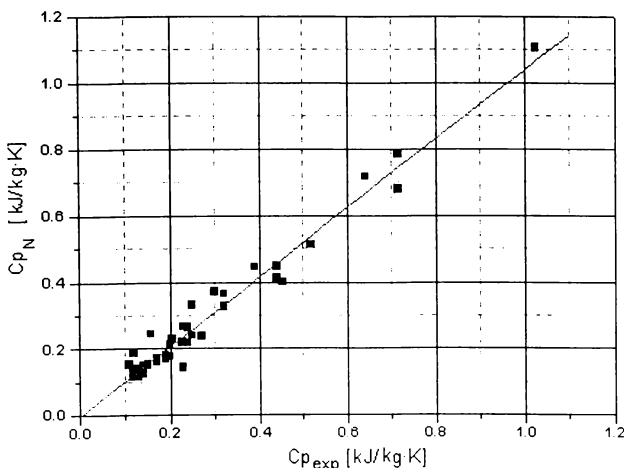


Fig. 7 Reprezentarea grafică a dependenței liniare dintre căldura specifică a **tiparelor pare** calculată pe baza tiparelor **impar** C_{pN} și cea experimentală pentru tiparele **pare** $C_{p_{exp}}$

Ecuația de corelare pentru regresia liniară dintre valorile calculate ale căldurii specifice pentru setul de date par calculate pe baza setului de date impar C_{pN} și cele experimentale $C_{p_{exp}}$ este redată mai jos:

$$C_{pN} = -0,0018(\pm 0,0148) + 1,0405(\pm 0,0426)C_{p_{exp}} \quad (11)$$

Parametrii de corelare sunt:

Volumul eșantionului	N = 41
Numărul gradelor de libertate	NL = 39
Coeficientul de corelare	R = 0,981
Coeficientul de determinare	R ² = 0,962
Deviația standard	SD = 0,041
Testului Fisher	F = 896

În figura 8, este redată variația liniară dintre valoarea calculată a căldurii specifice pentru setul de date **impar** utilizând rețeaua antrenată cu setul de date par și valoarea experimentală a acesteia pentru setul de date **impar**.

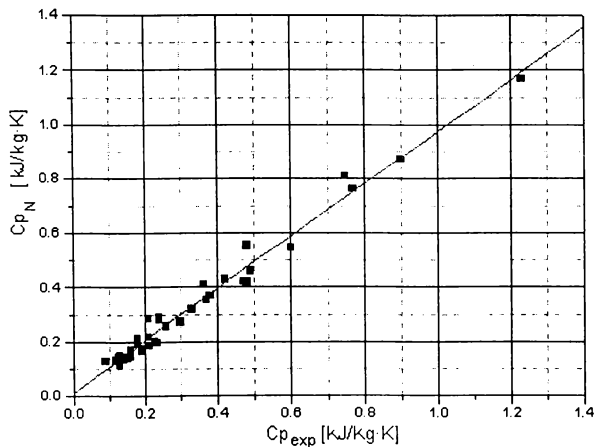


Fig. 8 Dependența liniară dintre căldura specifică a **tiparelor impare** calculată pe baza tiparelor pare C_{pN} funcție de cea experimentală pentru tiparele impare $C_{p_{exp}}$

Ecuația de corelare prin regresie liniară a valorilor căldurii specifice pentru setul de date impar C_{pN} calculate pe baza setului de date par și cele experimentale $C_{p_{exp}}$ este redată mai jos:

$$C_{pN} = 0,0142(\pm 0,0092) + 0,9592(\pm 0,023)C_{p_{exp}} \quad (12)$$

Parametrii de corelare sunt:

Volumul eșantionului	N = 41
Numărul gradelor de libertate	NL = 39
Coeficientul de corelare	R = 0,989
Coeficientul de determinare	R ² = 0,988(7)
Deviația standard	SD = 0,036
Testul Fisher	F = 1684

Parametrii de corelare pentru cele două subseturi de date calculate fiecare pe baza celui alt argumentează apartenența celor două subseturi de date la aceeași populație, precum o stabilitate și o capacitate de predicție foarte bună a rețelei neurale utilizate.

II.2.6 Predicții realizate pe baza modelului neural

Ulterior validării modelului de corelare, s-a utilizat rețeaua antrenată cu cele 82 de tipare în scopul predicției căldurii specifice pentru elementele la care valoarea acesteia este nedeterminată experimental. În acest scop, s-au prezentat succesiv rețelei vectorii de intrare care reflectă structura electronică a elementelor, pe baza ponderilor create în etapa de învățare calculându-se valoarea unității de ieșire ca valoare probabilă a căldurii specifice.

În tabelul 4, sunt redate valorile căldurilor specifice ale elementelor pentru care această proprietate este necunoscută

Tabelul 4 Valoarea de predicție a căldurii specifice pentru elemente la care aceasta este încă nedeterminată experimental

Număr atomic	Simbol	Căldură specifică [kJ/kgK]	Număr atomic	Simbol	Căldură specifică [kJ/kgK]
85	At	0,119	103	Lr	0,141
87	Fr	0,156	104	Rf	0,137
89	Ac	0,123	105	Ha	0,134
96	Cm	0,138	106	Sg	0,132
97	Bk	0,152	107	Ns	0,132
98	Cf	0,159	108	Hs	0,135
99	Es	0,154	109	Mt	0,139
100	Fm	0,147	110	Unn	0,145
101	Md	0,145	111	Unu	0,141
102	No	0,145			

Ținând cont de variația proprietăților atomice de-a lungul perioadelor și grupelor din sistemul periodic al elementelor, valorile prezise ale căldurii specifice pentru elementele la care această proprietate nu este cunoscută sunt pertinente.

II.2.7 Concluzii

Această aplicație s-a constituit în utilizarea unui model neural care să coreleze variația configurației electronice a elementelor de-a lungul tabelului periodic al elementelor cu căldura specifică elementară.

Inițial s-a testat corelarea prin regresie multiliniară a variabilelor care descriu configurația electronică a elementelor cu valoarea căldurii specifice a acestora. Prin valorile parametrilor de corelare prin regresie liniară dintre valorile calculate și cele experimentale, regresia multiliniară s-a dovedit necorespunzătoare.

Având în vedere utilizarea unei validări severe prin metoda validării încrucișate și parametrii de corelare obținuți la corelarea liniară dintre valorile calculate și cele reale, se poate constata o bună stabilitate și capacitate de corelare a modelului neural utilizat.

Odată realizat, modelul a fost utilizat pentru predicția căldurii specifice a elementelor pentru care această proprietate este necunoscută. Pentru aceasta, rețelei neurale antrenate cu tiparele elementelor pentru care valoarea căldurii specifice este cunoscută, i-au fost introduse doar valorile vectorilor de intrare ce caracterizează configurația electronică a acestor elemente, valorile de ieșire constituindu-se în valori probabile ale căldurii specifice pentru aceste elemente.

Pentru exemplificare, în figura 9, este redată variația căldurii specifice pentru elementele perioadelor 6 și 7 pentru a observa evoluția în context a acestei proprietăți pentru ultimele elemente. Valorile scrise în chenar accentuat sunt valori de predicție ale modelului neural anterior descris.

P	55	56	57	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
6	Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
	0.24	0.204	0.19	0.14	0.14	0.13	0.13	0.13	0.130	0.13	0.128	0.139	0.13	0.13	0.12	0.12	0.119	0.09
7	Fr	Ra	Ac	Rf	Ha	Sg	Ns	Hs	Mt	Unn	Unu							
	0.156	0.12	0.123	0.137	0.134	0.132	0.132	0.135	0.139	0.145	0.154							

Lantanide	58	59	60	61	62	63	64	65	66	67	68	69	70	71
	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
	0.19	0.19	0.19	0.18	0.20	0.18	0.23	0.18	0.17	0.16	0.17	0.16	0.15	0.15
Actinide	90	91	92	93	94	95	96	97	98	99	100	101	102	103
	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr
	0.12	0.12	0.12	0.12	0.13	0.11	0.138	0.152	0.159	0.154	0.147	0.145	0.145	0.141

Fig. 9 Variația căldurii specifice pentru elementele perioadelor 6 și 7

Având în vedere variația proprietăților în tabelul periodic al elementelor, valorile de predicție sunt pertinente, având o bună probabilitate de a fi confirmate experimental.

II.3 Corelări structură electronică - electronegativitate

Scopul aplicației este predicția electronegativității elementelor artificiale utilizând o rețea neurală antrenată cu valori argumentate ale vectorilor de intrare și ieșire proprii elementelor pentru care această proprietate este cunoscută¹¹¹.

II.3.1 Alegerea și justificarea variabilelor de corelare

Pentru corelarea structură electronică - electronegativitate, s-au luat în considerare elementele pentru care această proprietate este cunoscută folosind ca variabile independente valori ce reflectă configurația electronică a elementului iar ca variabilă dependentă - valoarea proprie a electronegativității.

Pentru a descrie variația configurației electronice în sistemul periodic al elementelor, pentru variabilele independente s-a utilizat maparea din aplicația anterioară:

-prima variabilă independentă **P** a fost dată de perioada din care face parte elementul, făcându-se astfel diferențierea între orbitalii de completare și ceilalți orbitali de același gen

-următoarele patru variabile independente au fost date de numărul electronilor **s, p, d, f** de pe orbitalul în completare.

-variabila independentă, s-a ales valoarea proprie a electronegativității elementului.

În tabelul 5 sunt redată spre exemplificare variabilele pentru câteva elemente selectate aleator:

Tabelul 5. Exemplificarea selecției variabilelor pentru corelarea structură electronică - electronegativitate

Număr atomic	Element	Variabile independente	Variabila dependentă E
20	Ca	4,2,0,0,0	1,00
51	Sb	5,2,3,10,0	2,05
88	Ra	7,2,0,0,0	0,90

II.3.2 Justificarea utilizării modelului neural

Într-un prim stadiu, a fost testată o regresie multiliniară a valorilor variabilelor independente cu variabila dependentă - electronegativitatea.

Relația obținută este redată mai jos:

$$E_{\text{MLR}} = 2,238(\pm 0,194) - 0,161(\pm 0,032) \cdot P - 0,117(\pm 0,080) \cdot s + 0,263(\pm 0,031) \cdot p + 0,048(\pm 0,009) \cdot d + 0,038(\pm 0,008) \cdot f \quad (13)$$

unde E_{MLR} este valoarea calculată a electronegativității, **P** - perioada în care este situat elementul iar **s,p,d,f** - numărul electronilor **s, p, d, f** de pe orbitalul în completare.

Parametrii de corelare sunt:

Volumul eșantionului	N = 96
Numărul gradelor de libertate	NL= 90
Coeficientul de corelare	R = 0,858
Coeficientul de determinare	R ² = 0,731
Deviația standard	SD= 0,336
Testul Fisher	F = 216

În figura 10 este redată dependența liniară dintre valoarea calculată prin regresie multiliniară a electronegativității și valoarea experimentală a acesteia pentru primele 96 de elemente.

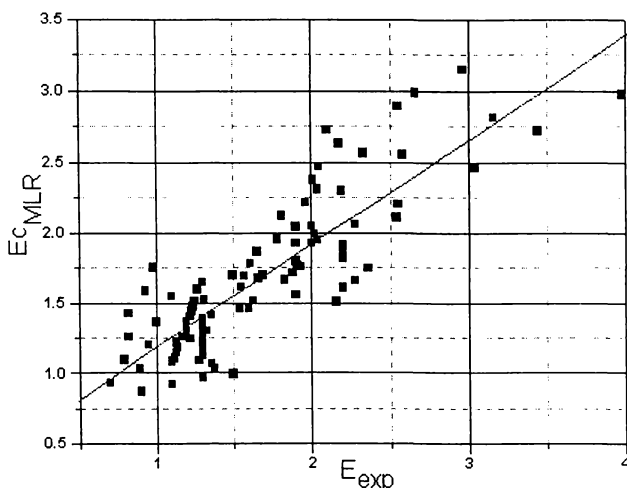


Fig. 10. Dependența liniară dintre valorile electronegativității calculate prin regresie multiliniară E_{C_MLR} și cele experimentale E_{exp}

După cum se observă, valoarea coeficientului de corelare este mică, iar valorile de deviație ale coeficienților au valori comparabile cu ale acestora, fapt ce denotă o slabă corelare în cazul regresiei multiliniare.

II.3.3 Caracterizarea modelului neural

II.3.3.1 Numărul straturilor ascunse și al unităților ascunse

Având în vedere argumentația aplicației precedente, s-a utilizat o rețea neurală cu un singur strat ascuns.

Pentru a determina numărul optim de unități de pe stratul ascuns, s-a utilizat aceeași procedură din aplicația anterioară: rețeaua a fost antrenată cu un număr diferit de unități pe stratul ascuns, menținând constante celelalte caracteristici - funcția de activare, rata de învățare, termenul de moment, algoritmul aleator de inițializare a ponderilor și algoritmul de corecție al acestora și numărul de epoci de rulare.

Având în vedere algoritmul de inițializare aleatoare a matricilor de ponderi, fiecare punct al valorii erorii totale a rețelei a fost obținut prin rularea rețelei *până la confirmarea valorii celei mai mici* pentru același număr de unități de pe stratul ascuns.

În figura 11 este redată variația erorii totale a rețelei funcție de numărul de unități de pe stratul ascuns.

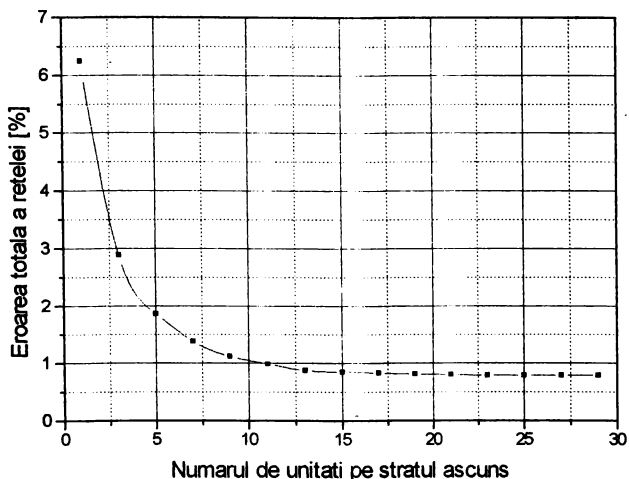


Fig. 11. Variația erorii totale a rețelei funcție de numărul de unități de pe stratul ascuns

Se observă că valoarea minimă a erorii totale a rețelei se obține pentru 24 de unități pe stratul ascuns, valoare care a fost utilizată pentru rezolvarea aplicației propuse spre rezolvare.

Și pentru această aplicație, creșterea numărului de unități de pe stratul ascuns peste această valoare nu împietează asupra valorii erorii totale, rețeaua evoluând în aceeași direcție favorabilă fără un consum semnificativ de timp suplimentar.

S-au utilizat 24 de unități pe stratul ascuns, urmărind ideea că în contextul predicției valorilor electronegativității pentru elementele artificiale, este preferabil un consum de timp suplimentar decât o predicție de o precizie mai scăzută.

Utilizând un program realizat în Turbo Pascal 7 care salvează pe hard-disk valorile ponderilor la fiecare 10 de epoci, pentru parcurgerea a 2000 de epoci au fost necesare aproximativ 20 de minute

II.3.3.2 Selecția tiparelor

În figura 12 este redată variația valorii electronegativității cu numărul atomic al elementelor din tabelul periodic al elementelor^[11].

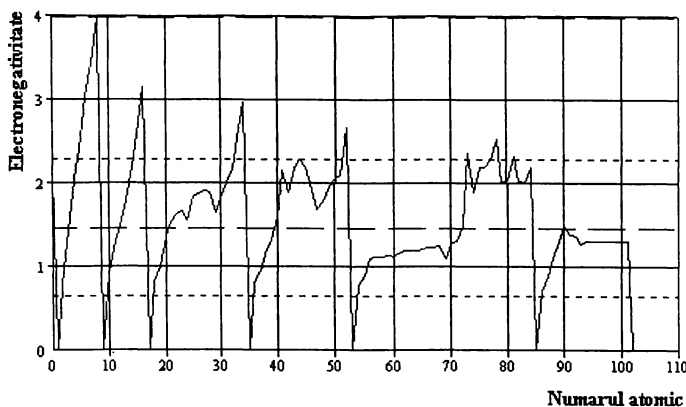


Fig 12. Variația electronegativității în tabelul periodic al elementelor.

Având în vedere că valoarea cea mai mică a electronegativității unui element este de 0,7(Fr) iar cea mai mare este 3,98(F), conform relației (14), factorul de împrăștiere SF are o valoare acceptabilă:

$$SF = \frac{MAX_{out}}{MIN_{out}} \cdot \frac{3,98}{0,7} = 5,69 \quad (14)$$

Având în vedere că factorul de împrăștiere are o valoare suficient de mică iar valorile acoperă într-un mod relativ egal domeniul de ieșire, nu s-a renunțat la nici un tipar excepție făcând tiparele corespunzătoare elementelor din grupa a VIII-a pentru care această proprietate nu își are sensul.

Pentru caracterizarea efortului de învățare al rețelei, s-a introdus noțiunea de **factor specific de antrenare** (specific training factor¹⁰⁹), calculat pe baza relației:

$$STF = \frac{N_i \cdot \left(\sum_1^{N_h} N_{uh} \right) \cdot N_{ep}}{N_{tip} \cdot N_o} = (5 \cdot 24 \cdot 2000) / (96 \cdot 1) = 2500 \quad (15)$$

unde:

- N_i - numărul elementelor din vectorul de intrare
- N_h - numărul straturilor ascunse
- N_{uh} - numărul unităților din stratul ascuns
- N_{ep} - numărul epocilor de antrenare a rețelei
- N_{tip} - numărul de tipare din setul de date utilizat în etapa de învățare
- N_o - numărul elementelor din vectorul de ieșire

Semnificația factorului de antrenare este de a cuantifica efortul mediu depus de rețeaua neurală în etapa de antrenare pentru învățarea unui singur tipar.

Necesitatea unui asemenea factor rezultă din dorința de comparare a volumului de calcul la aplicații similare în vederea optimizării configurației modelelor neurale utilizate. Pentru o aplicație similară de corelare a potențialelor de

ionizare cu structura electronică a elementelor, Sigman&Reeves⁴ utilizează la învățare o rețea neurală cu următoarele caracteristici:

- numărul unităților de pe stratul de intrare $N_i = 7$
- numărul straturilor ascunse $N_h = 1$
- numărul unităților din stratul ascuns $N_{uh} = 6$
- numărul de epoci parcurse $N_{ep} = 96\ 000$
- numărul unităților de pe stratul de ieșire $N_o = 1$
- numărul de tipare $N_{tip} = 222$

$$STF = \frac{N_i \cdot N_a \cdot N_{ep}}{N_{tip} \cdot N_o} = (7 \cdot 6 \cdot 96000) / (222 \cdot 1) = 18162 \quad (16)$$

Comparativ, din valoarea factorilor specifici de antrenare pentru cele două aplicații, rezultă că efortul la învățare pentru cazul de față este considerabil mai mic, fapt ce denotă o optimizare mai performantă a configurației rețelei neurale utilizate în lucrarea de față.

II.3.3.3 Descrierea rețelei neurale utilizate

Schema bloc a rețelei neurale utilizate este redată în figura 13

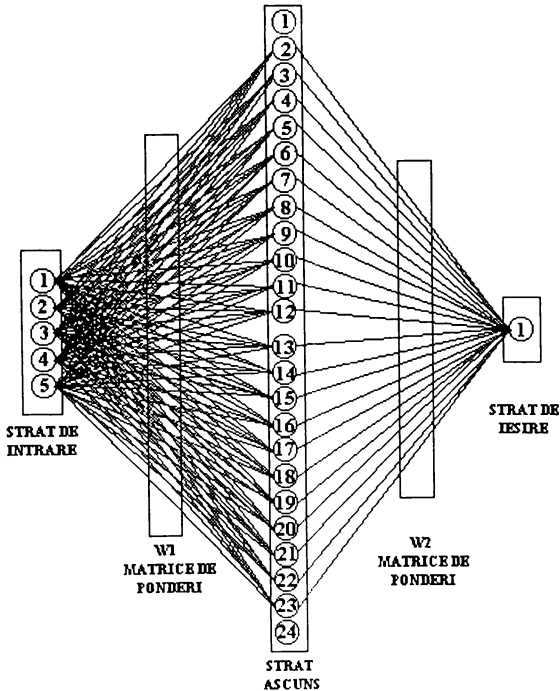


Fig. 13. Schema bloc a rețelei neurale utilizate pentru rezolvarea aplicației

Notațiile din figură corespund:

- W1 - matricea de ponderi dintre stratul de intrare și stratul ascuns
- W2 - matricea de ponderi dintre stratul ascuns și stratul de ieșire

Principalele variabile ale rețelei neurale utilizate sunt:

- Funcția de activare - funcția sigmoidă $F(x) = \frac{1}{1 + e^{-\beta \cdot x}}$ cu $\beta=3$
- Rata de învățare $\eta =$ între 1 și 0.01
- Termenul moment $\gamma = 0.1$
- Numărul de epoci $N_{ep} = 2000$
- Numărul de unități de pe stratul de intrare $N_i = 5$
- Numărul de unități de pe stratul ascuns $N_h = 24$
- Numărul de unități de pe stratul de ieșire $N_o = 1$
- Numărul de tipare $N_{tip} = 96$
- Nu s-au utilizat unități BIAS¹¹⁰

II.3.4 Rezultate experimentale

Pentru rezolvarea acestei aplicații au fost utilizate ca bază de învățare tiparele primelor 96 de elemente. Astfel, în tabelul 6, sunt redată tiparele utilizate în etapa de antrenare a rețelei.

Tabelul6. Tipare experimentale; valori experimentale și calculate ale electronegativității

Nr. atomic	Perioada	electroni s pe stratul de completare	electroni p pe stratul de completare	electroni d pe stratul de completare	electroni f pe stratul de completare	Valoarea experiment. E_{exp}	Valoarea calculată E_{CN}	Eroare [%]
0	1	2	3	4	5	6	7	8
1	1	1	0	0	0	2,2	2,199	-0,05
2	2	1	0	0	0	0,98	0,994	1,46
4	2	2	0	0	0	1,57	1,571	0,09
5	2	2	1	0	0	2,04	2,049	0,42
6	2	2	2	0	0	2,55	2,574	0,94
7	2	2	3	0	0	3,04	3,044	0,12
8	2	2	4	0	0	3,44	3,489	1,43
9	2	2	5	0	0	3,98	3,792	-4,73
11	3	1	0	0	0	0,93	0,859	-7,68
12	3	2	0	0	0	1,31	1,283	-2,09
13	3	2	1	0	0	1,61	1,563	-2,89
14	3	2	2	0	0	1,9	1,861	-2,07
15	3	2	3	0	0	2,19	2,186	-0,17
16	3	2	4	0	0	2,58	2,632	2,01
17	3	2	5	0	0	3,16	3,158	-0,07
19	4	1	0	0	0	0,82	0,82	0,05
20	4	2	0	0	0	1	1,095	9,55
21	4	2	0	1	0	1,36	1,409	3,57
22	4	2	0	2	0	1,54	1,478	-4,01
23	4	2	0	3	0	1,63	1,559	-4,37
24	4	1	0	5	0	1,66	1,798	8,33
25	4	2	0	5	0	1,55	1,737	12,04
26	4	2	0	6	0	1,83	1,77	-3,27

Tabelul 6. (Continuare)

0	1	2	3	4	5	6	7	8
27	4	2	0	7	0	1,88	1,773	-5,72
28	4	2	0	8	0	1,91	1,755	-8,13
39	4	1	0	10	0	1,9	1,884	-0,84
30	4	2	0	10	0	1,65	1,693	2,62
31	4	2	1	10	0	1,81	1,803	-0,38
32	4	2	2	10	0	2,01	1,967	-2,12
33	4	2	3	10	0	2,18	2,196	0,73
34	4	2	4	10	0	2,55	2,522	-1,09
35	4	2	5	10	0	2,96	2,999	1,33
37	5	1	0	0	0	0,82	0,792	-3,46
38	5	2	0	0	0	0,95	0,959	0,99
39	5	2	0	1	0	1,22	1,21	-0,79
40	5	2	0	2	0	1,33	1,372	3,16
41	5	1	0	4	0	1,6	1,65	3,1
42	5	1	0	5	0	2,16	1,94	-10,2
43	5	1	0	6	0	1,9	2,098	10,4
44	5	1	0	7	0	2,2	2,146	-2,46
45	5	1	0	8	0	2,28	2,127	-6,69
46	5	0	0	10	0	2,2	2,227	1,22
47	5	1	0	10	0	1,93	1,999	3,58
48	5	2	0	10	0	1,69	1,81	7,12
49	5	2	1	10	0	1,78	1,843	3,56
50	5	2	2	10	0	1,96	1,906	-2,74
51	5	2	3	10	0	2,05	2,016	-1,67
52	5	2	4	10	0	2,1	2,216	5,5
53	5	2	5	10	0	2,66	2,601	-2,23
55	6	1	0	0	0	0,79	0,776	-1,74
56	6	2	0	0	0	0,89	0,885	-0,57
57	6	2	0	1	0	1,1	1,125	2,23
58	6	2	0	0	2	1,12	1,08	-3,59
59	6	2	0	0	3	1,13	1,129	-0,05
60	6	2	0	0	4	1,14	1,162	1,92
61	6	2	0	0	5	1,13	1,187	5,01
62	6	2	0	0	6	1,17	1,205	3,03
63	6	2	0	0	7	1,2	1,219	1,55
64	6	2	0	1	7	1,2	1,237	3,1
65	6	2	0	0	9	1,2	1,23	2,53
66	6	2	0	0	10	1,22	1,231	0,89
67	6	2	0	0	11	1,23	1,229	-0,07
68	6	2	0	0	12	1,24	1,226	-1,09
69	6	2	0	0	13	1,25	1,224	-2,09
70	6	2	0	0	14	1,1	1,222	11,13
71	6	2	0	1	14	1,27	1,241	-2,29
72	6	2	0	2	14	1,3	1,262	-2,95
73	6	2	0	3	14	1,5	1,574	4,94
74	6	2	0	4	14	2,36	2,284	-3,22
75	6	2	0	5	14	1,9	2,012	14,56
76	6	2	0	6	14	2,2	2,068	-5,98
77	6	2	0	7	14	2,2	2,039	-7,31

Tabelul6.(Continuare)

0	1	2	3	4	5	6	7	8
78	6	1	0	9	14	2,28	2,374	4,12
79	6	1	0	10	14	2,54	2,432	-4,25
80	6	2	0	10	14	2	2,102	5,08
81	6	2	1	10	14	2,04	2,133	4,54
82	6	2	2	10	14	2,33	2,291	-1,67
83	6	3	0	10	14	2,02	1,996	-1,2
84	6	4	0	10	14	2	2,019	0,97
85	6	5	0	10	14	2,2	2,185	-0,67
87	7	1	0	0	0	0,7	0,774	10,56
88	7	2	0	0	0	0,9	0,854	-5,09
89	7	2	0	1	0	1,1	1,108	0,74
90	7	2	0	2	0	1,3	1,294	-0,47
91	7	2	0	1	2	1,5	1,447	-3,55
92	7	2	0	1	3	1,38	1,42	2,92
93	7	2	0	1	4	1,36	1,348	-0,85
94	7	2	0	0	6	1,28	1,241	-3,08
95	7	2	0	0	7	1,3	1,264	-2,77
96	7	2	0	1	7	1,3	1,259	-3,15
97	7	2	0	0	9	1,3	1,288	-0,92
98	7	2	0	0	10	1,3	1,291	-0,7
99	7	2	0	0	11	1,3	1,29	-0,81
100	7	2	0	0	12	1,3	1,285	-1,15
101	7	2	0	0	13	1,3	1,278	-1,66
102	7	2	0	0	14	1,3	1,271	-2,26

In figura 14 este redată variația erorii totale a rețelei pe parcursul procesului de antrenare pe parcursul celor 2000 de epoci.

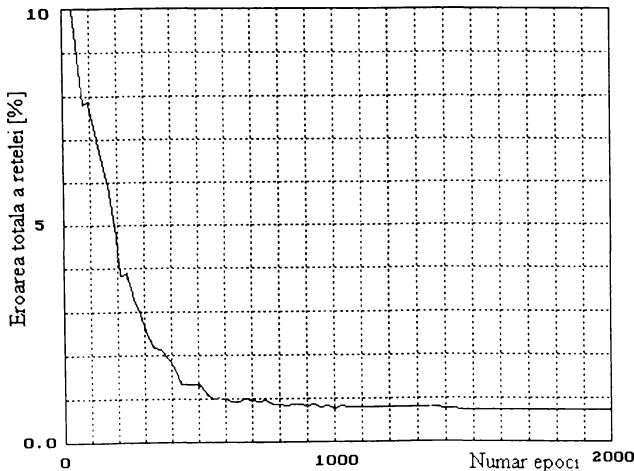


Fig. 14 Variația erorii totale a rețelei funcție de numărul de epoci parcurse

Din variația erorii cu numărul de epoci parcurse, efortul rețelei putea fi redus cu $\cong 50\%$, dar având în vedere specificul aplicației, - predicția electronegativității pentru elementele pentru care această proprietate este încă necunoscută - s-a preferat valoarea de 2000.

În figura 15 este redată dependența liniară dintre valorile calculate cu rețeaua neurală E_{CN} și cele experimentale E_{exp} ale electronegativității pentru toate tiparele utilizate în etapa de învățare.

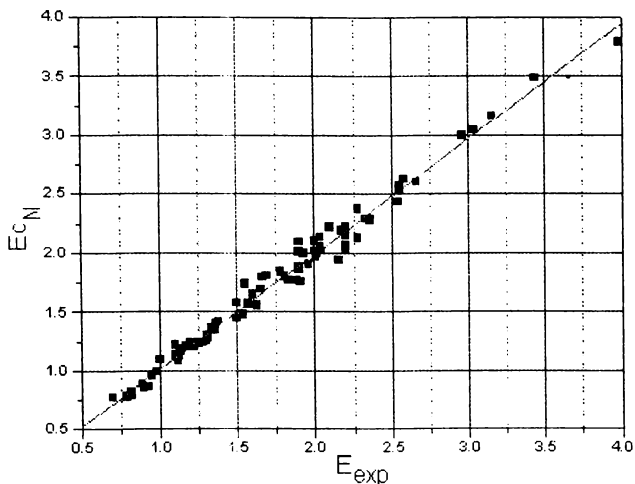


Fig. 15 Dependența liniară între valorile calculate cu rețeaua neurală E_{CN} și cele experimentale E_{exp}

Ecuția de corelare liniară dintre valorile calculate ale electronegativității E_{CN} și cele experimentale E_{exp} este redată mai jos:

$$E_{CN} = 0,037(\pm 0,021) + 0,977(\pm 0,012) \cdot E_{exp} \quad (17)$$

Parametrii de corelare liniară sunt:

Volumul eșantionului	N = 96
Numărul gradelor de libertate	NL = 94
Coefficientul de corelare total	R = 0,993
Coefficientul de determinare	R ² = 0,987
Deviația standard	SD = 0,071
Testul Fisher	F = 6971

Din valorile parametrilor de corelare se observă o foarte bună corelare liniară între valorile calculate și cele experimentale ale electronegativității.

II.3.5 Validarea modelului

Pentru validarea modelului s-a utilizat procedura validării încrucișate (cross-validation procedure). Pentru o încărcare egală a domeniului de valori al variabilei

de ieșire, întregul set de date a fost divizat în două seturi, criteriul de apartenență fiind dat de numărul par sau impar al tiparului.

Având în vedere numărul înjumătățit de tipare, rețelei descrise anterior i s-au adus modificări în ce privește numărul de unități de pe stratul ascuns care a fost redus la 12 și numărul de epoci de antrenare care a fost redus la 1000.

Intr-un prim stadiu, rețeaua a fost antrenată cu setul impar de date pentru predicția setului par, ulterior inversându-se procedura. Ulterior etapei de învățare, rețelei i s-a prezentat doar vectorii de intrare ai celuilalt subset, și pe baza ponderilor calibrate, în cadrul unei semiepoci de propagare înainte s-a efectuat o predicție a valorilor vectorilor de ieșire.

În figura 16 este redată dependența liniară dintre valorile electronegativității tiparelor impare calculate pe baza tiparelor pare, E_{CN} și cele experimentale E_{exp} .

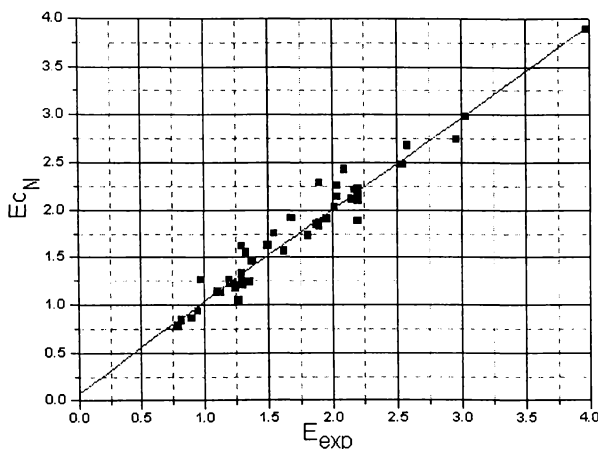


Fig. 16 Dependența liniară dintre electronegativitatea **tiparelor impare** calculată pe baza tiparelor pare E_{CN} și valorile experimentale E_{exp}

Pentru valorile electronegativității tiparelor impare calculate pe baza tiparelor pare, E_{CN} și cele experimentale E_{exp} există o dependență liniară exprimată de ecuația:

$$E_{CN} = 0,076(\pm 0,059) + 0,964(\pm 0,032) \cdot E_{exp} \quad (18)$$

Parametrii de corelare sunt:

Volumul eșantionului	$N = 48$
Numărul gradelor de libertate	$NL = 46$
Coeficientul de corelare	$R = 0,975$
Coeficientul de determinare	$R^2 = 0,951$
Deviația standard	$SD = 0,020$
Testul Fisher	$F = 887$

În figura 17, este redată variația liniară dintre valoarea calculată a electronegativității pentru setul de date par utilizând rețeaua antrenată cu setul de date impar și valoarea experimentală a acesteia.

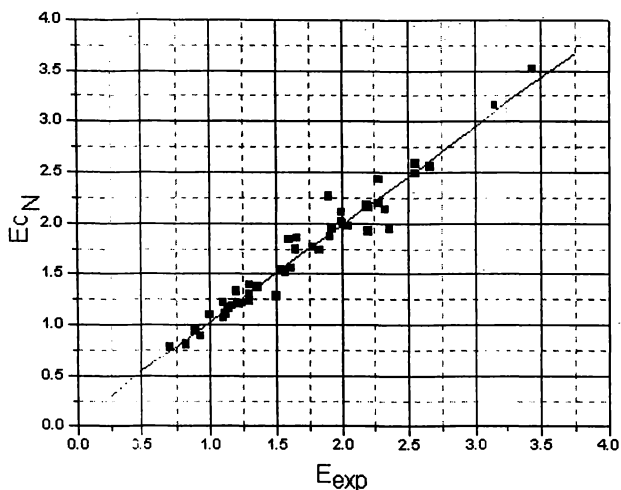


Fig. 17. Dependența liniară dintre electronegativitatea tiparelor pare calculată pe baza tiparelor impare E_{cN} și valorile experimentale E_{exp}

Ecuția de corelare liniară dintre valorile electronegativității tiparelor pare calculate pe baza tiparelor impare este:

$$E_{cN} = 0,063(\pm 0,053) + 0,962(\pm 0,029) \cdot E_{exp} \quad (19)$$

Parametrii de corelare sunt:

Volumul eșantionului	$N = 48$
Numărul gradelor de libertate	$NL = 46$
Coefficientul de corelare	$R = 0,979$
Coefficientul de determinare	$R^2 = 0,959$
Deviația standard	$SD = 0,123$
Testul Fisher	$F = 1069$

Parametrii de corelare pentru cele două subseturi de date calculate fiecare pe baza celuilalt argumentează o stabilitate și o capacitate de predicție foarte bună a rețelei neurale utilizate.

II.3.6 Predicții realizate pe baza modelului neural

Ulterior validării modelului de corelare, s-a utilizat rețeaua antrenată cu cele 96 de tipare în scopul predicției valorilor electronegativității elementelor la care valoarea acesteia nu este cunoscută. În acest scop, s-au prezentat succesiv rețelei vectorii de intrare care reflectă structura electronică a elementelor, pe baza ponderilor create în etapa de învățare calculându-se valoarea unității de ieșire ca valoare probabilă a electronegativității.

În tabelul 7, sunt redată valorile electronegativității elementelor cu numărul atomic specificat.

Tabelul 7 Valoarea de predicție a electronegativității pentru elemente pentru care valoarea experimentală este încă nedeterminată

Nr. atomic	103	104	105	106	107	108	109	110	111
Simbol	Lr	Rf	Ha	Sg	Ns	Hs	Mt	Unn	Unu
E	1,27	1,28	1,28	1,54	2,21	2,12	2,02	1,97	1,95

Ținând cont de variația proprietăților atomice de-a lungul perioadelor și grupelor din sistemul periodic al elementelor, valorile prezise ale electronegativității pentru elementele la care această proprietate nu este cunoscută sunt pertinente.

II.3.7 Concluzii

Această aplicație a urmărit utilizarea unui model neural pentru a stabili o corelare între configurația electronică a elementelor și electronegativitatea acestora.

Inițial, s-a încercat o corelare prin regresie multiliniară având ca variabile independente valori ce descriu configurația electronică iar ca variabilă dependentă valoarea electronegativității elementului, dar valorile parametrilor de corelare s-au dovedit a fi total necorespunzători. În cazul utilizării modelului neural, valorile parametrilor de corelare liniară dintre valorile calculate și cele experimentale și testarea prin validare încrucișată a acestui model, indică cu certitudine o corelare între această proprietate și configurația electronică a elementelor.

Modelul neural realizat și antrenat cu tiparele a 96 de elemente a fost utilizat pentru predicția electronegativității elementelor pentru care această proprietate este necunoscută. În acest scop, rețelei i s-au prezentat vectorii de intrare ai elementelor pentru care valoarea electronegativității este necunoscută, aceasta calculând pe baza ponderilor calibrate în procesul de învățare valoarea de ieșire care s-a constituit ca valoare probabilă a electronegativității.

Pentru exemplificare, în figura 18, este redată variația electronegativității pentru elementele perioadelor 6 și 7 pentru a fi observată evoluția în context a acestei proprietăți pentru ultimele elemente. Valorile trecute în chenar îngroșat sunt valori de predicție ale modelului neural anterior descris.

6	55	56	57	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
	Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
	0.79	0.89	1.10	1.3	1.5	2.38	1.9	2.2	2.20	2.28	2.54	2.00	2.04	2.33	2.02	2.0	2.2	0
7	87	88	89	104	105	106	107	108	109	110	111							
	Fr	Ra	Ac	Rf	Ha	Sg	Ns	Hs	Mt	Unn	Unu							
	0.7	0.9	1.1	1.28	1.28	1.54	2.21	2.12	2.02	1.97	1.95							

Lantanide	58	59	60	61	62	63	64	65	66	67	68	69	70	71
	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
	1.12	1.13	1.14	1.13	1.17	1.2	1.20	1.2	1.22	1.23	1.24	1.25	1.1	1.27
Actinide	90	91	92	93	94	95	96	97	98	99	100	101	102	103
	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr
	1.3	1.5	1.38	1.38	1.28	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.27

Fig. 18 Variația electronegativității pentru elementele perioadelor 6 și 7

Având în vedere variația acestei proprietăți în tabelul periodic al elementelor, valorile prezise sunt pertinente, având o bună probabilitate de a fi confirmate experimental.

III. CORELĂRI STRUCTURĂ CHIMICĂ - PROPRIETĂȚI FIZICE

III.1 Introducere

Structura moleculară reprezintă elementul fundamental ce condiționează proprietățile substanțelor chimice¹¹³. De aceea, cunoașterea structurii și mai ales cuantificarea ei, reprezintă un pas important, absolut necesar pentru înțelegerea comportamentului fizico-chimic al moleculelor¹¹⁴.

Modalitățile de abordare a caracterizării cantitative a structurii moleculare sunt de o mare diversitate. Pot fi, însă, distinse două direcții principale:

1 - descrierea stării nivelelor energetice pe care se află electronii de valență prin ecuațiile diferențiale ale chimiei cuantice^{115,116,117,118}

2 - descrierea topologiei moleculare (vecinătăți, ramificare, legături, numărul și tipul atomilor, modul de legare a lor, număr de legături)¹¹⁹

În cazul abordării topologice, moleculele sunt reprezentate prin grafuri, relațiile de vecinătate dintre atomi determinate de starea legat-nelegat directă²⁵ fiind relații de importanță majoră ce definesc conectivitatea¹²⁰. Diferențierea structurilor izomere în cadrul modelului topologic se face prin dezvoltarea unor invarianți topologici și concretizarea acestora în măsurile cantitative adecvate capabile să rețină suficientă informație structurală.

Pentru a putea corela proprietățile moleculare redade cantitativ prin numere - cu structura, s-a pus problema cuantificării acestora prin valori numerice ce se constituie ca descriptori moleculari. Aceștia, în funcție de modul de determinare se pot clasifica în descriptori cuantici (indici cuantici de reactivitate rezultați prin calcule de chimie cuantică) și descriptori topologici (indici topologici calculați pe baza grafurilor chimice).

Având în vedere că reprezentarea unei structuri presupune redarea ei printr-o mulțime de componente între care există diferite relații de interdependență, Randić propune două posibilități de reprezentare¹²¹:

1 - Codurile moleculare obținute din enumerarea atomilor prin numărul de marcarea, apar sub forma unei liste a componentelor.

Utilizate ca date de intrare, codurile moleculare permit manipularea computațională a structurilor, în așa fel încât există posibilitatea reconstituirii integrale a acestora. Acest fapt denotă o corespondență biunivocă între codurile moleculare și structura moleculelor.

2 - Invarianții structurali, care redau practic anumite proprietăți matematice ale structurii.

Invarianții structurali sunt fie descrieri matematice ale grafurilor chimice, fie sunt obținuți în urma unor calcule aplicate componentelor cuprinse în codurile moleculare. Practic, este nevoie de o colecție largă de invarianți structurali pentru caracterizarea cât mai fidelă a structurii. Condensarea acestora sub forma indicilor topologici elimină posibilitatea existenței unei relații biunivoce structură-indici topologici.

Indicii topologici¹²² cuantifică diverse aspecte ale structurii grafului molecular fără hidrogeni permițând stabilirea unor relații structură-proprietăți fizico-

chimice prin căutarea legăturilor între cele două tipuri de proprietăți, matematică și fizico-chimică. Ambele tipuri de proprietăți depind de caracteristicile structurale, însă cele matematice permit o interpretare directă a factorilor care determină caracteristicile fizice, chimice sau biologice ale moleculelor. Deoarece nu există un indice topologic care să poată descrie perfect o anumită proprietate, există posibilitatea ca structuri diferite să genereze aceleași valori pentru indicii topologici. Acest fenomen poartă numele de **degenerare** și, uneori, afectează rezultatele corelărilor.

III.2 Aplicații existente în domeniu

Teoria grafurilor și indicii topologici și-au găsit aplicații deosebite în foarte multe domenii ale cercetării chimice dintre care amintim:

- corelarea structurii chimice cu activitatea biologică pentru proiectarea unor medicamente cu anumite proprietăți^{123,124};
- enumerarea compușilor aciclici și caracterizarea ramificării^{125,126,127};
- enumerarea structurilor ciclice^{128,129};
- codificare unică a structurilor moleculare¹³⁰ și elaborarea unei nomenclaturi raționale^{131,132};
- enumerarea și clasificarea sistematică a moleculelor cu simetrie înaltă^{133,134};
- stabilirea legăturilor între matricea Huckel și matricea de adiacență;
- aplicații în cinetica chimică¹³⁵, și în reprezentarea reacțiilor chimice¹³⁶.

Pentru detalii privitoare la indicii topologici și la aplicațiile lor se pot consulta o serie de articole de sinteză dedicate acestui subiect^{119,131,137-141}.

În continuare, se va face o scurtă prezentare a modului de calcul al principalilor indici topologici și o clasificare a acestora.

III.3 Relații numerice în teoria grafurilor

Stabilirea unor relații cantitative între structură și proprietățile fizice ale compușilor chimici reprezintă unul din dezideratele majore ale cercetărilor actuale. De aceea, în acest capitol sunt studiate cu ajutorul analizei corelaționale relațiile dintre punctele de fierbere a 100 alcoolii și caracteristicile lor structurale, descrise topologic. Având în vedere că reprezentarea moleculelor este făcută cu ajutorul grafurilor, s-a considerat necesară abordarea unor relații numerice elementare din teoria grafurilor, utilizate în descrierea indicilor topologici.

Cele mai utilizate noțiuni topologice în chimie sunt cele legate de numărul de izomeri posibili pentru un număr dat de atomi, numărul de atomi, numărul de legături, numărul de cicluri, etc.

O teoremă de bază a teoriei grafurilor¹⁴² este cea referitoare la relația existentă între numărul de muchii și gradele vârfurilor. Astfel, dacă un graf G conține n vârfuri v_i ($i=1, \dots, n$), fiecare având gradul (valența) δ_i și un număr m de muchii, atunci între acestea există relația :

$$\sum_{i=1}^n \delta_i = 2 \cdot m \quad (1)$$

De aici rezultă că suma gradelor vârfurilor este un număr par.
Pentru un arbore cu n vârfuri și m muchii este adevărată relația:

$$m = n - 1 \quad (2)$$

iar pentru grafurile care descriu sisteme spațiale este valabilă relația:

$$m = n + c - 1 \quad (3)$$

unde c reprezintă numărul de fețe ale figurii spațiale.

Stocarea numerică a unui graf se poate face pe baza legăturilor dintre vârfuri¹⁴³, adică a **adiacențelor**. Această modalitate de reprezentare permite refacerea integrală a grafului. Pentru un graf, numerotarea vârfurilor poate fi făcută în orice ordine fără a afecta univocitatea reprezentării. Pentru grafurile care descriu ansambluri moleculare, este de preferat ca numerotarea să fie cât mai apropiată de normele IUPAC. Adiacențele se regăsesc în matricea corespunzătoare A asociată grafului, denumită **matrice de adiacență** (sau de conectivitate).

Această matrice este o matrice pătrată de $n \times n$ elemente, simetrică față de diagonală principală, având valorile:

$$\begin{cases} - 0 \text{ pentru perechile de atomi nelegați între ei} \\ - 1 \text{ pentru perechile de atomi legați} \end{cases}$$

Astfel, pentru 3-metilpentan, matricea de adiacențe este redată în figura 1:

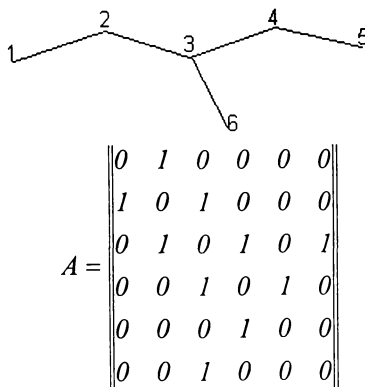


Fig. 1. Matricea de adiacență pentru 3-metilpentan

Dacă se efectuează suma elementelor unei linii sau a unei coloane a matricii de adiacență se obține chiar gradul aceluși vârf.

$$\delta_i = \sum_{j=1}^n a_{ij} \quad (4)$$

Un alt invariant fundamental pentru modelarea topologică este *matricea distanțelor*¹³⁸ notată cu **D**. Elementele ei reprezintă distanțele topologice dintre vârfurile *i* și *j* considerate pe drumul cel mai scurt. Este tot o matrice pătrată de $n \times n$ elemente, simetrică față de diagonala principală. Pentru 3-metilpentan, matricea distanțelor este prezentată în figura 2.

$$D = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 3 \\ 1 & 0 & 1 & 2 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 1 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 4 & 3 & 2 & 1 & 0 & 3 \\ 3 & 2 & 1 & 2 & 3 & 0 \end{pmatrix}$$

Fig. 2. Matricea distanțelor pentru 3-metilpentan

Dacă se efectuează suma elementelor unei linii sau a unei coloane a matricii de distanțe se obține suma distanțelor (s_i) de la vârful *i* la toate celelalte vârfuri:

$$s_i = \sum_{j=1}^n d_{ij} \quad (5)$$

Sumele δ_i și s_i (relațiile (4) și (5)) reprezintă cei mai simpli invarianti locali ai grafurilor corespunzătoare fiind, în fapt, numere asociate vârfurilor grafurilor. Pentru acești invarianti, nodurile grafului molecular (atomi) care sunt topologic echivalente au valori numerice egale, indiferent de modul în care s-a făcut numerotarea. Grafurile care nu conțin asemenea puncte sunt numite *grafuri identitate* și nu li se poate aplica nici o operație de simetrie cu excepția celei de identitate.

Pe lângă matricile prezentate anterior, în studiile de topologie moleculară sunt utilizate și alte matrici, dintre care amintim:

- **AM - Matricea de adiacență a muchiilor**. Ea este o matrice pătrată, simetrică față de diagonala principală, compusă din $m \times m$ elemente (*m* fiind numărul de muchii). Valorile elementelor acestei matrici sunt determinate astfel:

$$(i,j) \begin{cases} 1 & \text{dacă muchiile au un vârf comun} \\ 0 & \text{în caz contrar.} \end{cases}$$

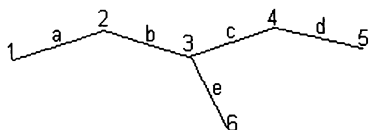
- **IM - Matricea de incidență muchii-noduri**¹⁴⁵. Se poate construi pentru orice graf cu *n* vârfuri și *m* muchii, astfel de matrice formată din $n \times m$ elemente. Liniile corespund vârfurilor, iar coloanele muchiilor. Valorile elementelor acestei matrici sunt prezentate astfel:

$$(i,j) = \begin{cases} 1 & \text{dacă muchia } j \text{ este incidentă vârfului } i \\ 0 & \text{în caz contrar} \end{cases}$$

Această matrice nu este simetrică față de diagonala principală. În Figura 3 sunt prezentate cele două tipuri de matrici pentru 3-metilpentan.

Dacă se efectuează suma elementelor pe linii în matricea de adiacență a muchiilor (AM) se obține "gradul" muchiei respective (numărul de muchii de care se leagă).

Modificări ale elementelor unor matrici conduc la construcția altora, care deasemenea pot fi utilizate pentru generarea unor invarianți structurali^{146,147,148}.



$$AM = \begin{array}{c} \begin{array}{ccccc} a & b & c & d & e \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 & 1 \\ \hline 0 & 1 & 0 & 1 & 1 \\ \hline 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 & 0 \\ \hline \end{array} \begin{array}{l} a \\ b \\ c \\ d \\ e \end{array} \end{array} \quad IM = \begin{array}{c} \begin{array}{ccccc} a & b & c & d & e \\ \hline 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 & 1 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ \hline \end{array} \end{array}$$

Fig. 3. AM - matricea de adiacență a muchiilor pentru 3-metilpentan
 IM - matricea de incidență muchii - noduri pentru 3-metilpentan

III.4. Indici topologici

Indicii topologici sunt numere asociate structurilor moleculare în scopul determinării unor relații cantitative structură-proprietăți. Aceștia se obțin prin diverse operații matematice efectuate asupra invarianților locali ai grafurilor asociate moleculelor. Invarianții locali sunt după cum am mai amintit, numere reprezentând gradul vârfurilor, suma distantelor, elemente ale primului vector propriu¹⁴⁹ sau elementele descrise mai sus ponderate cu diverse mărimi (electronegativități, raze covalente, raze van der Waals¹⁵⁰, etc.) De remarcat că nu există un indice topologic capabil să diferențieze toți compușii organici, cu aplicabilitate generală în corelări sau codificări¹⁵¹.

Având în vedere numărul foarte mare de indici topologici, aprecierea calității acestora este foarte dificilă, fiecare fiind de regula proiectat pentru o clasă de compuși și apreciat în funcție de capacitatea de corelare a uneia sau a mai multor

proprietăți ale acestora. Nici măcar gradul de degenerare (aceleași valori ale indicilor topologici pentru structuri diferite) nu este elocvent pentru aprecierea calității indicelui deoarece uneori, izomerii pentru care valoarea indicelui este aceeași, prezintă o variație nesemnificativă a proprietății cu care este testată calitatea indicelui.

La definirea fiecărui indice, pierderea de informație este inevitabilă, fapt care duce încă din start la imposibilitatea refacerii structurii.

III.4.1. Indicele de adiacență totală

Cel mai simplu indice este A' care se calculează ca sumă a tuturor elementelor matricii de adiacență

$$A' = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} \quad (6)$$

Un dezavantaj al acestui indice este redundanța sa cauzată de simetria matricii de adiacență. Acest indice poate fi redus la suma elementelor submatricii triunghiulare superioare, A , și este numit **adiacența totală a moleculei**¹³¹:

$$A = \frac{1}{2} \cdot \sum_{i=1}^n \sum_{j=1}^n a_{i,j} \quad (7)$$

Pentru 3-metilpentan se obțin valorile $A'=10$ și $A=5$

III.4.2. Indicii Wiener, Rouvray-Crawford și de polaritate

În cazul în care suma se aplică elementelor matricii distanțelor, rezultă un nou indice - **indicele Wiener**^{153,154,155}, w , primul indice structural de natură topologică. Denumit inițial "indice de drumuri" (path number), el reprezintă numărul total de legături care separă toate perechile de atomi din hidrocarburile saturate. Hosoya a extins definiția la grafurile ciclice, utilizând distanțele în locul drumurilor (în cazul grafurilor aciclice aceste două noțiuni coincid). Se poate vedea foarte ușor că w reprezintă jumătatea sumei tuturor elementelor ne-diagonale din matricea de distanțe. Acesta este dat de suma elementelor matricii triunghiulare superioare din matricea D :

$$w = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n d_{i,j} \quad (8)$$

S-a demonstrat¹⁵⁶ că între numărul n de atomi de carbon ai unei hidrocarburi liniare și indicele Wiener există relația :

$$w = \frac{n \cdot (n^2 - 1)}{6} \quad (9)$$

Rouvray a introdus¹⁵⁷ **indicele R** derivat din indicele Wiener și definit conform relației:

$$R = 2 \cdot w \quad (10)$$

Acesta reprezintă suma tuturor elementelor matricii de distanțe. R a fost utilizat cu succes la corelarea unor proprietăți fizice ale alcanilor, alchenelor, alchinilor și arenelor.

Tot Wiener¹⁵² a definit și **indicele de polaritate p** care se calculează însumând numărul de perechi de vârfuri separate prin câte trei muchii:

$$p = \frac{1}{2} \cdot \sum_i d_{3,i} \quad (11)$$

Factorul 1/2 din fața sumei anulează parcurgerea în dublu sens a distanțelor de 3 unități ($d_{3,i}$).

Pentru 3-metilpentan valorile indicilor prezentați mai sus sunt:

$$w=31; \quad R=62; \quad p=2$$

III.4.3. Indicele Platt

Platt a definit¹⁵⁸ **indicele F** ca sumă a numărului de muchii adiacente fiecărei muchii în parte. Acest indice se poate calcula:

- din suma tuturor gradelor muchiilor din structura moleculei

sau

- din matricea adiacenței muchiilor **AM**, însumând toate elementele.

Pentru 3-metilpentan se obține $F = 10$.

III.4.4. Indicele polinomial Altenburg

Indicele polinomial y, introdus de Altenburg^{159,160}, se obține prin exprimarea indicelui Wiener ca sumă de termeni care depind de o variabilă indexată q_i , ce reprezintă numărul de perechi de atomi separate prin legături de lungime d_i .

Polinomul propus de Altenburg pentru caracterizarea cantitativă a unui graf are forma:

$$y(d) = \sum_i q_i \cdot d_i, \quad i=1,2,\dots,d_{\max} \quad (12)$$

Acesta a fost obținut plecând de la următoarea relație de calcul a indicelui Wiener:

$$w = \sum_i q_i \cdot d_i \quad (13)$$

Astfel, expresia polinomială a acestui indice în cazul 3-metilpentanului este, conform relației (12), următoarea:

$$y = 5d_1 + 5d_2 + 4d_3 + 1d_4 \quad (14)$$

Din polinomul y prin înlocuirea variabilelor d_i cu valoarea distanței pe care o reprezintă, conform relației (13) se obține valoarea indicelui Wiener:

$$y = 5d_1 - 5d_2 + 4d_3 + 1d_4 = 5 \cdot 1 + 5 \cdot 2 + 4 \cdot 3 + 1 \cdot 4 = 31$$

III.4.5. Indicele Hosoya

Pornind de la faptul că proprietățile fizice și termodinamice sunt determinate de structura topologică a moleculelor, Hosoya a propus¹⁶¹⁻¹⁶³ indicele Z definit prin relația:

$$Z = \sum_{k=0}^{n/2} p(G, k) \quad (15)$$

unde $p(G, k)$ reprezintă numărul de noduri în care pot fi alese k muchii dintr-un graf aciclic în așa fel încât: să nu fie adiacente două câte două. Prin definiție $p(G, 0) = 1$, iar $p(G, 1)$ reprezintă chiar numărul de muchii din graful molecular. Alternativ, pentru grafurile aciclice Z poate fi definit drept sumă a valorilor absolute ale coeficienților polinomului caracteristic $P_H(G, x)$:

$$P_H(G, x) = \sum_{k=0}^n (-1)^k p(G, k) x^{n-2k} = (-1)^n \det|A - xE| \quad (16)$$

unde n este numărul de atomi, iar A și E sunt matricea de adiacență și respectiv, matricea unitară; s reprezintă cel mai mare număr de muchii nelegate una de alta din graful aciclic.

Indicele Z a fost primul cunoscut sub numele de "indice topologic"¹⁴⁶, fiind aplicat pentru corelarea punctelor de fierbere, entropiilor, ordinilor de legătură calculate, precum și pentru codificarea structurilor chimice.

III.4.6. Indicii M_1 și M_2

Indicii prezentați până acum au fost obținuți prin :

- însumarea matricilor de adiacență A și de distanțe D , pe linii sau coloane, rezolvarea ecuației caracteristice asociate unei matrici, însumarea coeficienților polinomului caracteristic și enumerarea unor secvențe de subgrafuri din graful molecular.

Primii indici topologici bazați pe matricea de adiacență au fost introduși de Gutman și colaboratorii^{164,165}:

$$M_1 = \sum_{i=1}^n (\delta_i)^2 \quad (17)$$

$$M_2 = \sum_{i \text{ legat de } j} (\delta_i \cdot \delta_j) \quad (18)$$

unde δ_i reprezintă gradul (valența) nodului i din graful molecular fără hidrogeni.

Acești indici au fost utilizați pentru studii de ramificare, compararea unor grafuri și recunoașterea similarității moleculare^{166,167}, corelări cu energia electronilor în sistemele conjugate¹⁶⁸ și cu energia de rezonanță¹⁶⁹.

III.4.7. Indici de tip Randić

În încercarea de a caracteriza gradul de ramificare moleculară printr-un singur descriptor, Randić a propus¹⁶⁹ un indice calculat conform relației (17):

$${}^1\chi = \sum_{i \text{ legat de } j} (\delta_i \cdot \delta_j)^{-\frac{1}{2}} \quad (19)$$

unde δ_i și δ_j reprezintă gradele vârfurilor adiacente i și j .

Indicele Randić prezintă o bună corelare cu punctele de fierbere ale alcanilor, dar degenerază destul de rapid¹⁷⁰ (la $n = 8$ atomi de carbon).

III.4.8. Indicele J

Pentru a elimina principalul dezavantaj al indicelui Randić (degenerarea destul de rapidă începând cu $n = 8$ atomi de carbon), Balaban¹⁷¹ a aplicat relația (20) introducând astfel un nou indice al sumei distanțelor de la un vârf i la toate celelalte vârfuri din graf (s_i), definit astfel:

$$J = \frac{q}{\mu + 1} \cdot \sum (s_i \cdot s_j)^{-\frac{1}{2}} \quad (20)$$

În relația (20), q reprezintă numărul de muchii, μ este numărul ciclomatic ($\mu = q - n + 1$, cu $n =$ numărul de vârfuri), iar suma se efectuează doar pentru perechile "i legat de j".

Indicele J prezintă o degenerare mult mai mică¹⁷²; ea apare la $n = 12$ atomi de carbon. Comportarea indicelui J în corelări liniare monoparametrice cu proprietăți de genul punct de fierbere, entalpii, timpi de retenție, etc., este nesatisfăcătoare. Totuși, aceasta se îmbunătățește spectaculos atunci când indicele J este folosit în corelări multiparametrice în special în acelea în care unul dintre parametri este numărul de vârfuri ale grafului.

III.4.9. Indici derivați din indicii J și χ ^{174,175}

Prin înlocuirea δ_i și δ_j (gradele vârfurilor adiacente i și j) cu produsul $\delta_i \cdot E$ respectiv $\delta_j \cdot E$, unde E este electronegativitatea elementelor din vârfurile i și j, în relația 19, rezultă un indice derivat χ^{el} .

$$\chi^{el} = \sum_{i \text{ legat de } j} (\delta_i E_i \cdot \delta_j E_j)^{-\frac{1}{2}} \quad (21)$$

Prin înlocuirea s_i și s_j (gradele vârfurilor adiacente i și j) cu produsul $s_i \cdot E$ respectiv $s_j \cdot E$, unde E este electronegativitatea elementelor din vârfurile i și j în relația 20, rezultă un indice derivat J^{el} .

$$J^{el} = \frac{q}{\mu + 1} \cdot \sum (s_i E_i \cdot s_j E_j)^{-\frac{1}{2}} \quad (22)$$

Intr-un mod similar, prin înlocuirea δ_i , δ_j cu $\delta_i \cdot r_{c_i}$, $\delta_j \cdot r_{c_j}$ și a s_i , s_j cu $s_i \cdot r_{c_i}$, $s_j \cdot r_{c_j}$ în relația 19 și respectiv 20 se obțin doi indici derivați care țin seama de raza covalentă a legăturii dintre vârfurile i și j descriși de relațiile 23 și 24.

$$\chi^{rc} = \sum_{i \text{ legat de } j} (\delta_i r_{c_i} \cdot \delta_j r_{c_j})^{-\frac{1}{2}} \quad (23)$$

$$J^{rc} = \frac{q}{\mu + 1} \cdot \sum (s_i r_{c_i} \cdot s_j r_{c_j})^{-\frac{1}{2}} \quad (24)$$

Deasemenea, în cazul în care în locul razei covalente se utilizează raza van der Waals a legăturii dintre vârfurile i și j se obțin indicii descriși de relațiile 25 și 26:

$$\chi^w = \sum_{i \text{ legat de } j} (\delta_i r_{w_i} \cdot \delta_j r_{w_j})^{-\frac{1}{2}} \quad (25)$$

$$J^w = \frac{q}{\mu + 1} \cdot \sum (s_i r_{w_i} \cdot s_j r_{w_j})^{-\frac{1}{2}} \quad (26)$$

O altă categorie derivată din indicii χ , J și derivații de mai sus se obține¹⁷⁶ prin diferențe conform relațiilor (27,28,29,30):

$$\chi d_1 = \chi^{rc} - \chi \quad (27)$$

$$\chi d_2 = \chi^{el} - \chi \quad (28)$$

$$Jd_1 = J^{rc} - J \quad (29)$$

$$Jd_2 = J^{el} - J \quad (30)$$

III.4.10. Indicii Kier și Hall

Utilizând ca bază indicele Randiç, Kier și Hall au pornit de la indicele Randiç, încercând să elimine o serie de deficiențe care au fost observate. Pentru eliminarea acestor deficiențe Kier și Hall au definit¹⁷⁷ un indice de conectivitate generalizat:

$${}^m \chi = \sum_r (\delta_1 \cdot \delta_2 \cdot \dots \cdot \delta_{m+1})^{-\frac{1}{2}} \quad (31)$$

unde r reprezintă toate drumurile posibile de lungime m . Se observă că pentru $m=1$ se obține chiar indicele Randiç.

III.5. Concluzii

Domeniul de cercetare reprezentat de aplicarea modelului topologic în studiul structurii moleculelor este foarte vast și se află într-o continuă dezvoltare. Deoarece posibilitățile de combinare ale invariantilor locali în scopul găsirii unor indici topologici adecvați pentru corelări cantitative structură chimică-proprietăți moleculare sunt practic nelimitate, generarea de noi descriptori moleculari constituie un domeniu extrem de studiat și într-o amplă dezvoltare.

În utilizarea descriptorilor topologici pentru corelarea cu diverse proprietăți, au fost propuse două criterii de bază de către Randiç¹⁴⁶:

- (1) Utilizarea unui număr cât mai restrâns de descriptori.
- (2) Utilizarea unor descriptori care să fie cât mai ușor interpretabili din punct de vedere al semnificației fizice.

Modul de selecție al acestor descriptori topologici precum și performanțele lor pentru explicarea variației unor proprietăți fizice în cadrul unei serii de izomeri vor fi prezentate în secțiunea următoare, dedicată corelării structurii topologice cu punctele de fierbere ale alcoolilor alifatici.

III.6.1. Introducere

Importanța stabilirii unor relații cantitative structură chimică-proprietăți fizice este evidentă. Aceste relații pot fi utilizate în sens explicativ, permițând decelarea elementelor structurale responsabile pentru fenomenul fizic observat, precum și calculul predictiv, în acest caz ecuațiile corelaționale dacă sunt validate statistic corespunzător, pot servi pentru evaluarea respectivei proprietăți fizice pentru o serie de compuși, anterior determinării ei sau în situația în care ea nu poate fi evaluată din diferite motive.

Obiectivul acestui capitol este stabilirea legăturii dintre structura și punctele de fierbere pentru un set de 100 alcooli alifatici. Pentru cuantificarea structurii s-au utilizat o serie de indici topologici adecvați precum și indici derivați din maparea logică a conformației moleculare.

Având în vedere datele din literatură sistematizate anterior s-au calculat o serie de indici topologici utilizați pentru stabilirea unor astfel de relații.

În tabelul 1 sunt redate denumirea, punctul de fierbere și valorile principalilor indici topologici pentru setul de alcooli testat.

Modalitatea de calcul a indicilor topologici utilizați este redată în acest capitol de relațiile:

χ	-	relația (19)
J	-	relația (20)
χ^{el}, χ^{rc}	-	relațiile (21),(23)
J^{el}, J^{rc}	-	relațiile (22),(24)
$\chi d_1, \chi d_2$	-	relațiile (27),(28)
$J d_1, J d_2$	-	relațiile (29),(30)
χ^w, J^w	-	relațiile (25),(26)

Pentru corelarea indicilor topologici cu punctele de fierbere ale alcoolilor s-au utilizat atât rețele neurale cât și regresii multiliniare, rezultatele obținute fiind analizate comparativ.

Tabelul 1. Punctele de fierbere și indicii topologici corespunzătorii setului de alcooli utilizați

Nr	Alcool	Pf°C	γ	J	χ^{el}	γ^c	J^{el}	J^c	χd_1	χd_2	Jd ₁	Jd ₂	χ^w	J^w
1.	Etanol	78,0	1,414	1,633	0,155	1,918	0,178	2,214	0,504	-1,260	0,581	-1,454	0,856	0,988
2.	Propanol	97,1	1,914	1,975	0,217	2,567	0,228	2,635	0,653	-1,697	0,660	-1,747	1,150	1,182
3.	Izopropilic	82,4	1,732	2,324	0,199	2,316	0,266	3,107	0,584	-1,533	0,783	-2,057	1,038	1,393
4.	Butanol	111,6	2,414	2,191	0,280	3,217	0,259	2,900	0,802	-2,134	0,709	-1,931	1,444	1,305
5.	2M-1Propanol	108,1	2,270	2,539	0,262	3,029	0,301	3,361	0,759	-2,008	0,821	-2,239	1,359	1,512
6.	2Butanol	99,5	2,270	2,539	0,266	3,014	0,298	3,371	0,744	-2,004	0,831	-2,241	1,355	1,515
7.	2M-2Propanol	82,4	2,000	3,024	0,235	2,655	0,355	4,014	0,655	-1,765	0,990	-2,669	1,193	1,804
8.	Pentanol	138,0	2,914	2,339	0,343	3,866	0,281	3,082	0,952	-2,572	0,743	-2,058	1,738	1,389
9.	3M-1Butanol	131,0	2,770	2,627	0,324	3,679	0,316	3,461	0,909	-2,446	0,833	-2,312	1,653	1,560
10.	2M-1Butanol	128,0	2,808	2,754	0,329	3,728	0,330	3,630	0,920	-2,479	0,876	-2,424	1,676	1,636
11.	2Pentanol	119,3	2,770	2,627	0,329	3,664	0,313	3,471	0,894	-2,441	0,843	-2,314	1,649	1,563
12.	3Pentanol	116,2	2,808	2,754	0,333	3,713	0,327	3,642	0,905	-2,475	0,888	-2,427	1,671	1,639
13.	3M-2Butanol	112,9	2,643	2,993	0,313	3,498	0,357	3,953	0,856	-2,330	0,960	-2,637	1,574	1,780
14.	2M-2Butanol	102,3	2,561	3,168	0,305	3,383	0,376	4,189	0,822	-2,256	1,020	-2,792	1,523	1,886
15.	Hexanol	157,6	3,414	2,447	0,405	4,515	0,296	3,216	1,101	-3,009	0,769	-2,151	2,032	1,451
16.	3M-1Pentanol	153,0	3,308	2,832	0,392	4,377	0,343	3,720	1,069	-2,916	0,888	-2,489	1,970	1,678
17.	4M-1pentanol	151,9	3,270	2,678	0,387	4,328	0,324	3,518	1,058	-2,883	0,840	-2,354	1,947	1,587
18.	2M-1Pentanol	149,0	3,308	2,832	0,392	4,377	0,342	3,723	1,069	-2,916	0,891	-2,490	1,970	1,679
19.	2E-1Butanol	147,0	3,346	2,992	0,397	4,427	0,361	3,934	1,081	-2,949	0,942	-2,631	1,992	1,774
20.	2,3-diM-1Butanol	144,5	3,181	3,144	0,376	4,212	0,381	4,132	1,031	-2,805	0,987	-2,764	1,895	1,864
21.	3,3-diM-1Butanol	143,0	3,061	3,155	0,361	4,056	0,383	4,142	0,995	-2,700	0,987	-2,772	1,824	1,869
22.	2Hexanol	140,0	3,270	2,678	0,391	4,313	0,322	3,526	1,043	-2,879	0,848	-2,356	1,943	1,590
23.	2,2-diM-1Butanol	136,5	3,121	3,360	0,368	4,135	0,407	4,416	1,013	-2,753	1,056	-2,954	1,860	1,992
24.	3Hexanol	135,0	3,308	2,832	0,396	4,363	0,339	3,734	1,054	-2,912	0,902	-2,493	1,965	1,682
25.	3M-2Pentanol	134,3	3,181	3,144	0,380	4,197	0,378	4,140	1,016	-2,801	0,995	-2,766	1,891	1,866

Nr	Alcool	Pf °C	γ	J	γ^{el}	γ^{re}	J^{el}	J^{re}	χ_{d1}	χ_{d2}	χ^w	J^w		
26.	4M-2Pentanol	131,6	3,126	2,953	0,373	4,126	0,356	3,887	1,000	-2,753	0,934	-2,598	1,858	1,753
27.	2M-3Pentanol	126,5	3,181	3,144	0,380	4,197	0,377	4,145	1,016	-2,801	1,000	-2,767	1,891	1,868
28.	3M-3Pentanol	122,4	3,121	3,360	0,375	4,111	0,402	4,432	0,990	-2,746	1,071	-2,958	1,853	1,997
29.	2M-2Pentanol	121,1	3,061	3,155	0,368	4,032	0,378	4,158	0,972	-2,693	1,004	-2,776	1,817	1,874
30.	3,3-diM-2Butanol	120,4	2,943	3,541	0,350	3,889	0,427	4,660	0,946	-2,593	1,119	-3,115	1,751	2,101
31.	2,3-diM-2Butanol	118,4	2,943	3,541	0,353	3,880	0,425	4,667	0,937	-2,591	1,125	-3,116	1,748	2,103
32.	Heptanol	176,4	3,914	2,530	0,468	5,165	0,308	3,318	1,250	-3,446	0,788	-2,222	2,326	1,498
33.	4M-1Hexanol	173,0	3,808	2,862	0,455	5,027	0,349	3,752	1,219	-3,353	0,890	-2,513	2,264	1,694
34.	5M-1Hexanol	170,0	3,770	2,716	0,450	4,977	0,331	3,561	1,207	-3,320	0,845	-2,385	2,242	1,607
35.	3M-1Hexanol	169,0	3,808	2,920	0,455	5,027	0,356	3,829	1,219	-3,353	0,909	-2,564	2,264	1,728
36.	2M-1Hexanol	164,0	3,808	2,862	0,455	5,027	0,348	3,756	1,219	-3,353	0,894	-2,514	2,264	1,695
37.	2Heptanol	159,0	3,770	2,716	0,454	4,963	0,329	3,568	1,192	-3,316	0,852	-2,387	2,237	1,610
38.	2,4-diM-1Pentanol	159,0	3,664	3,099	0,437	4,839	0,377	4,065	1,176	-3,227	0,966	-2,722	2,179	1,835
39.	3Heptanol	157,0	3,808	2,862	0,459	5,012	0,345	3,765	1,204	-3,349	0,903	-2,517	2,260	1,698
40.	4Heptanol	156,0	3,808	2,920	0,459	5,012	0,352	3,843	1,204	-3,349	0,923	-2,568	2,260	1,732
41.	5M-2Hexanol	151,0	3,626	2,928	0,436	4,775	0,355	3,845	1,149	-3,190	0,917	-2,573	2,152	1,735
42.	5M-3Hexanol	148,0	3,664	3,099	0,441	4,825	0,374	4,076	1,161	-3,223	0,977	-2,725	2,175	1,838
43.	2M-2Hexanol	143,0	3,561	3,112	0,430	4,682	0,376	4,092	1,121	-3,131	0,981	-2,736	2,111	1,845
44.	2M-3Hexanol	143,0	3,681	3,171	0,443	4,846	0,382	4,173	1,166	-3,238	1,002	-2,789	2,185	1,881
45.	3M-3Hexanol	143,0	3,621	3,373	0,438	4,760	0,406	4,440	1,139	-3,184	1,066	-2,967	2,147	2,002
46.	3E-3Pentanol	142,0	3,682	3,583	0,445	4,839	0,431	4,717	1,157	-3,237	1,134	-3,152	2,183	2,126
47.	2,3-diM-3Pentanol	139,7	3,504	3,708	0,423	4,608	0,447	4,879	1,104	-3,081	1,171	-3,261	2,078	2,200
48.	2,4-diM-3Pentanol	138,7	3,553	3,464	0,427	4,681	0,418	4,558	1,128	-3,127	1,094	-3,046	2,110	2,055
49.	2,2-diM-3Pentanol	135,0	3,481	3,623	0,418	4,588	0,438	4,764	1,106	-3,064	1,141	-3,186	2,067	2,149
50.	2,4-diM-2Pentanol	133,1	3,417	3,389	0,412	4,494	0,409	4,456	1,078	-3,004	1,067	-2,980	2,027	2,010

Nr	Alcool	Pf°C	χ	J	χ^{el}	χ^c	J ^{el}	J ^c	χ_{d1}	χ_{d5}	J _{d1}	J _{d2}	χ^w	J ^w
51.	2,3,3-triM-2Butanol	131,0	3,250	4,020	0,391	4,278	0,486	5,285	1,028	-2,859	1,264	-3,534	1,929	2,384
52.	Octanol	195,1	4,414	2,595	0,530	5,814	0,317	3,399	1,400	-3,884	0,804	-2,278	2,621	1,535
53.	6M-1Heptanol	188,6	4,270	2,747	0,512	5,627	0,336	3,596	1,357	-3,758	0,850	-2,411	2,536	1,624
54.	4M-1Heptanol	188,0	4,308	2,955	0,517	5,676	0,361	3,869	1,368	-3,791	0,914	-2,593	2,558	1,747
55.	2Octanol	180,0	4,270	2,747	0,517	5,612	0,335	3,602	1,342	-3,753	0,855	-2,412	2,531	1,626
56.	25-diM-1Hexanol	179,5	4,164	3,061	0,499	5,489	0,374	4,010	1,325	-3,665	0,949	-2,687	2,473	1,811
57.	4Octanol	176,3	4,308	2,955	0,521	5,661	0,358	3,883	1,353	-3,787	0,928	-2,597	2,554	1,752
58.	6M-3Heptanol	174,0	4,164	3,061	0,503	5,474	0,371	4,018	1,310	-3,661	0,958	-2,689	2,469	1,813
59.	5M-3Heptanol	172,0	4,202	3,223	0,508	5,523	0,391	4,231	1,321	-3,694	1,008	-2,832	2,491	1,909
60.	3Octanol	171,0	4,308	2,877	0,521	5,661	0,349	3,777	1,353	-3,787	0,901	-2,528	2,554	1,704
61.	5M-2Heptanol	170,0	4,164	3,061	0,503	5,474	0,373	4,013	1,310	-3,661	0,952	-2,688	2,469	1,812
62.	4M-3Heptanol	170,0	4,219	3,325	0,510	5,545	0,403	4,366	1,326	-3,709	1,041	-2,921	2,501	1,970
63.	2,4,4-triM-1Pentanol	168,5	3,954	3,467	0,473	5,217	0,424	4,541	1,262	-3,482	1,074	-3,043	2,350	2,051
64.	2M-3Heptanol	167,5	4,181	3,155	0,505	5,496	0,382	4,146	1,315	-3,675	0,990	-2,773	2,479	1,870
65.	3M-2Heptanol	166,1	4,181	3,155	0,505	5,496	0,384	4,139	1,315	-3,675	0,984	-2,771	2,479	1,868
66.	3,4-diM-2Hexanol	165,5	4,091	3,576	0,494	5,380	0,435	4,689	1,288	-3,597	1,114	-3,140	2,426	2,117
67.	2M-4Heptanol	164,0	4,164	3,151	0,503	5,474	0,382	4,140	1,310	-3,661	0,989	-2,770	2,469	1,868
68.	3M-3Heptanol	163,0	4,121	3,330	0,501	5,410	0,403	4,375	1,288	-3,621	1,045	-2,927	2,441	1,974
69.	3M-4Heptanol	162,0	4,219	3,325	0,510	5,545	0,403	4,368	1,326	-3,709	1,044	-2,922	2,501	1,971
70.	4M-4Heptanol	161,0	4,121	3,431	0,501	5,410	0,415	4,510	1,288	-3,621	1,078	-3,016	2,441	2,034
71.	2M-3E-3Pentanol	160,0	4,065	3,919	0,493	5,336	0,474	5,151	1,271	-3,571	1,232	-3,445	2,408	2,323
72.	2,3-diM-2Hexanol	160,0	3,981	3,589	0,483	5,228	0,436	4,711	1,247	-3,499	1,122	-3,153	2,359	2,126
73.	2,3,4-triM-3Pentanol	156,5	3,887	4,014	0,471	5,105	0,486	5,273	1,218	-3,416	1,260	-3,528	2,303	2,379
74.	2M-3E-2Pentanol	156,0	4,019	3,793	0,488	5,277	0,460	4,979	1,258	-3,532	1,186	-3,332	2,381	2,247
75.	2M-2Heptanol	156,0	4,061	3,073	0,493	5,331	0,373	4,034	1,270	-3,568	0,961	-2,700	2,406	1,820

Nr	Alcool	Pf °C	χ	J	χ^{el}	χ^{rc}	J^{el}	J^{rc}	χd_1	χd_2	Jd ₁	Jd ₂	χ^w	J^w
76.	2,5-diM-2Hexanol	154,5	3,917	3,281	0,475	5,144	0,398	4,306	1,227	-3,442	1,025	-2,882	2,321	1,943
77.	2,2,4-triM-3Pentanol	150,5	3,854	3,878	0,465	5,072	0,470	5,093	1,217	-3,390	1,215	-3,408	2,287	2,298
78.	2,4,4-triM-2Pentanol	147,5	3,707	3,746	0,449	4,872	0,455	4,916	1,165	-3,259	1,170	-3,291	2,198	2,219
79.	Nonanol	213,3	4,914	2,648	0,593	6,463	0,325	3,464	1,549	-4,321	0,816	-2,323	2,915	1,565
80.	7M-1Octanol	206,0	4,770	2,773	0,575	6,276	0,340	3,628	1,506	-4,195	0,854	-2,433	2,830	1,639
81.	3Nonanol	195,0	4,808	2,886	0,584	6,311	0,352	3,784	1,503	-4,224	0,898	-2,534	2,848	1,708
82.	2Nonanol	193,5	4,770	2,773	0,579	6,261	0,339	3,632	1,491	-4,191	0,859	-2,434	2,826	1,640
83.	5Nonanol	193,0	4,808	2,998	0,584	6,311	0,364	3,936	1,503	-4,224	0,938	-2,634	2,848	1,776
84.	4Nonanol	192,5	4,808	2,968	0,584	6,311	0,361	3,895	1,503	-4,224	0,927	-2,607	2,848	1,758
85.	4E-4Heptanol	182,0	4,682	3,690	0,571	6,138	0,448	4,845	1,456	-4,111	1,155	-3,243	2,771	2,186
86.	2M-2Octanol	178,0	4,561	3,044	0,555	5,980	0,371	3,990	1,420	-4,005	0,947	-2,673	2,700	1,801
87.	2,6-diM-3Heptanol	175,0	4,537	3,301	0,550	5,958	0,401	4,331	1,421	-3,987	1,030	-2,900	2,688	1,955
88.	2,6-diM-4Heptanol	174,5	4,520	3,337	0,548	5,936	0,405	4,380	1,416	-3,972	1,043	-2,932	2,678	1,977
89.	2,6-diM-2Heptanol	173,0	4,417	3,205	0,537	5,793	0,393	4,202	1,377	-3,879	0,996	-2,815	2,615	1,897
90.	3,6-diM-3Heptanol	173,0	4,477	3,465	0,545	5,872	0,421	4,545	1,395	-3,932	1,081	-3,043	2,651	2,052
91.	2,2,3-triM-3Hexanol	156,0	4,311	4,102	0,524	5,656	0,498	5,383	1,345	-3,786	1,281	-3,603	2,553	2,429
92.	Decanol	231,1	5,414	2,691	0,656	7,113	0,331	3,517	1,698	-4,758	0,827	-2,360	3,209	1,590
93.	3,7-diM-1Octanol	212,5	5,164	3,100	0,624	6,787	0,381	4,052	1,624	-4,539	0,952	-2,719	3,062	1,831
94.	2Decanol	211,0	5,270	2,796	0,642	6,911	0,343	3,659	1,640	-4,628	0,863	-2,454	3,120	1,653
95.	4Decanol	210,5	5,308	2,973	0,647	6,960	0,363	3,896	1,652	-4,661	0,924	-2,610	3,142	1,759
96.	3,6-diM-3octanol	202,2	4,888	3,718	0,598	6,400	0,453	4,874	1,512	-4,290	1,156	-3,264	2,891	2,200
97.	3E-3Octanol	199,0	5,249	3,626	0,633	6,905	0,432	4,791	1,655	-4,616	1,165	-3,195	3,114	2,157
98.	2,6-diM-4octanol	195,0	5,058	3,428	0,615	6,635	0,418	4,495	1,577	-4,442	1,066	-3,011	2,995	2,029
99.	2,7-diM-3Octanol	193,5	5,037	3,242	0,613	6,607	0,396	4,248	1,571	-4,424	1,006	-2,846	2,982	1,918
100.	3E-2M-3Heptanol	193,0	5,065	3,885	0,619	6,635	0,473	5,095	1,570	-4,446	1,210	-3,412	2,996	2,300

III.6.2. Considerații privind alegerea variabilelor utilizate în corelare

Pentru alegerea variabilelor utilizate în corelare, s-a testat cu ajutorul regresiei liniare fiecare indice topologic calculat pentru a stabili cantitatea de informație structurală conținută și pentru a selecta variantele optime pentru o corelare multiplă.

În Tabelul 2 sunt date valorile coeficienților de corelare și deviațiile standard pentru regresii liniare dintre fiecare din indicii topologici calculați și punctele de fierbere a alcoolilor.

Tabelul 2 Parametrii statistici corespunzători ecuațiilor corelaționale liniare dintre punctele de fierbere și indici topologici

Indice topologic	Coeficient de corelare	Deviația standard	Indice topologic	Coeficient de corelare	Deviația standard
χ	0,9611	8,527	J	0,1816	30,364
χ^{el}	0,9552	9,138	J^{el}	0,2388	29,984
χ^{rc}	0,9630	8,319	J^{rc}	0,1617	30,471
χd_1	0,9689	7,639	$J d_1$	0,0957	30,736
χd_2	-0,9619	8,441	$J d_2$	-0,1733	30,410
χ^w	0,9604	8,391	J^w	0,1686	30,435

Se observă din tabelul 2 o foarte bună corelare pentru indicii de tip χ rezultatele cele mai bune fiind obținute pentru χ și χd_1 . Pentru indicii de tip J corelările obținute sunt nesemnificative. Rezultatele obținute nu sunt neașteptate.

Din literatură¹⁷² se cunoaște că indicii de tip J își îmbunătățesc capacitatea de corelare prin asociere cu numărul atomilor din moleculă, N.

Pe baza rezultatelor sistematizate în tabelul 2, pentru analiza relației structură-puncte de fierbere s-au selectat următoarele perechi de indici:

- 1 - $\chi + \chi d_1$
- 2 - $\chi + J$
- 3 - J+N

III.6.3. Descrierea modelului neural

Modelul neural utilizat pentru modelarea structurală a punctelor de fierbere diferă de cel utilizat în capitolul anterior, prin algoritmul de corectare a ponderilor.

Algoritmul cel mai utilizat pentru corecția ponderilor este:

$$\Delta\omega_{ij}(t+1) = \eta\delta_i^p a_j^p + \gamma\Delta\omega_{ij}(t) \quad \gamma \in (0,1) \quad (32)$$

unde:

- t indexează epoca
- η este rata de învățare

- γ este termenul moment care presupune o corecție a ponderilor la momentul $t+1$ afectată direct proporțional cu valoarea γ de corecția ponderilor efectuată într-o epocă anterioară.

Dacă termenul **moment**, γ , este discutabil în ceea ce privește oportunitatea și valoarea sa, valoarea ratei de învățare η influențează viteza de înaintare pe direcția de convergență spre minim a erorii totale a rețelei.

Cea mai utilizată modalitate de variație a ratei de învățare la antrenarea rețelei este cea descendentă, procesul evoluând cu valori descrescătoare ale acesteia pe măsura apropierii de finalul învățării. Nu există practic o regulă de variație a ratei de învățare, schimbarea valorii acesteia presupunând de la caz la caz o abordare personalizată.

În figura 4 este prezentată evoluția erorii rețelei cu numărul de epoci parcurse în etapa de învățare. Se observă că valoarea erorii la sfârșitul antrenării este de 1.67%.

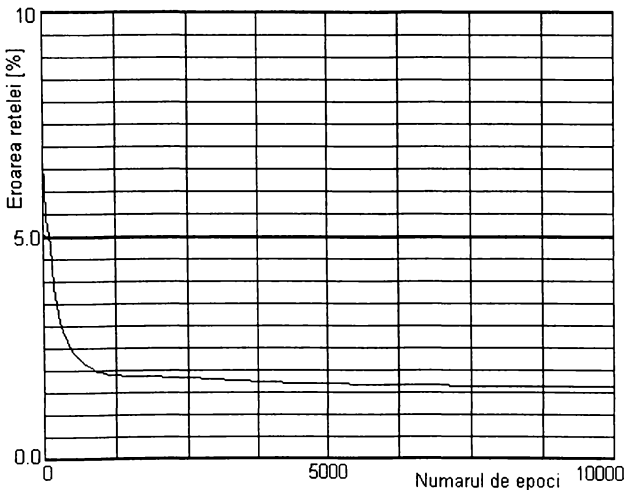


Fig. 4 Variația erorii rețelei pe parcursul procesului de învățare în cazul algoritmului uzual

Având în vedere situația analizată, în vederea optimizării convergenței spre valoarea de minim, este de dorit ca valoarea ratei de învățare să crească atunci când gradientul de eroare este negativ și să scadă atunci când valoarea acestuia este pozitivă. Exprimarea relațională a acestei corecții este:

$$\eta_t = \eta_{t-1} - m_g \frac{\partial E}{\partial \eta} \quad (33)$$

unde rata de învățare η_t corespunzătoare epocii t este corectată în funcție de evoluția de moment a erorii rețelei $\frac{\partial E}{\partial \eta}$ față de epoca anterioară. Valoarea coeficientului m_g este aleasă astfel încât variația ratei de învățare să fie suficient de semnificativă pentru a exploata avantajele algoritmului.

Un algoritm simplu de realizare a acestei reguli presupune luarea în considerare a următoarelor elemente:

- Startul procesului de antrenare trebuie realizat cu o valoare a ratei de învățare rezonabil de mică.
- Trebuie stabilită o valoare a coeficientului m_g .
- Se impune constatarea valorii erorii totale a rețelei la momentul t , și $t+1$.
- Este necesară compararea valorilor erorilor și adaptarea ratei de învățare conform regulilor următoare:

- Dacă $E\%_{t+1} > E\%_t \Rightarrow \eta = \eta + m_g \frac{\partial E}{\partial \alpha}$

- Dacă $E\%_{t+1} < E\%_t \Rightarrow \eta = \eta - m_g \frac{\partial E}{\partial \alpha}$

Se obține astfel o adaptare a procesului de învățare la gradientul de eroare momentan, fapt ce conduce la accelerarea convergenței spre minimum de eroare al rețelei. Spre deosebire de algoritmul normal de învățare, acest algoritm necesită parcurgerea unui număr mai mare de epoci, însă valoarea erorii totale a rețelei la sfârșitul procesului de învățare este mai redusă.

Acest algoritm este recomandat a fi utilizat la învățarea unei rețele neurale în cazul în care rețeaua antrenată urmează a fi utilizată în scopul predicției valorilor vectorului de ieșire pentru tipare la care se cunosc doar vectorii de intrare.

În Figura 5 este redată variația erorii rețelei pe parcursul procesului de învățare cu tiparele corespunzătoare aplicației de la subpunctul III.6.5 din acest capitol. Se observă că valoarea erorii la sfârșitul etapei de antrenare este de 0.89, fapt care confirmă în cazul concret studiat o comportare superioară a rețelei neurale bazate pe algoritmul adaptiv în corecția ratei de învățare.

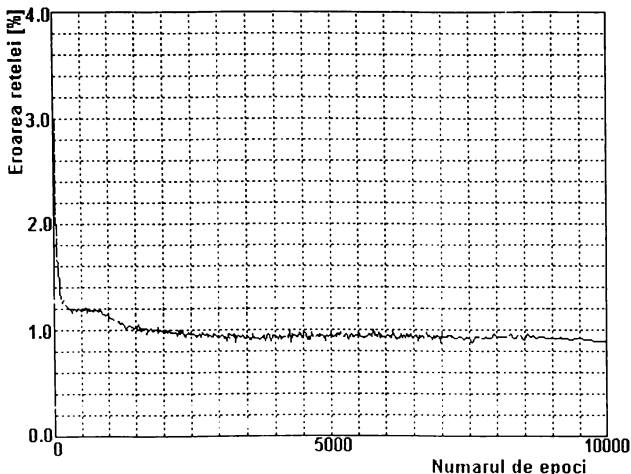


Fig. 5 Variația erorii rețelei pe parcursul procesului de învățare în cazul algoritmului de învățare adaptivă

În figura 6 este redată evoluția ratei de învățare a rețelei în cazul utilizării algoritmului adaptiv de corecție a ponderilor.

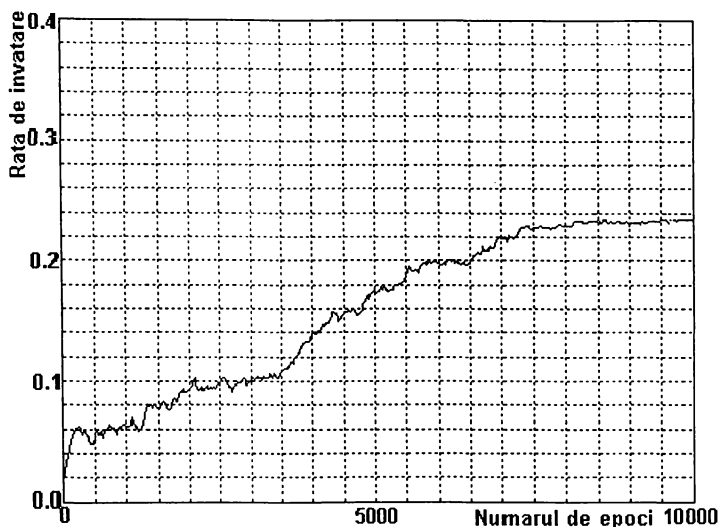


Fig.6 Variația ratei de învățare pe parcursul procesului de învățare în cazul algoritmului de învățare adaptivă

Se observă că, pe măsură ce se reduce eroarea totală a rețelei, deci în cazul situației pe un gradient descendent al erorii, convergența spre minim este accelerată prin valori crescătoare ale ratei de învățare η iar un gradient ascendent al erorii, fie și de extindere locală induce o valoare redusă a ratei de învățare, deci o decelerare a avansării pe direcția de minim a erorii rețelei.

Caracteristicile rețelelor neurale exploatate în acest capitol sunt:

- funcția de activare sigmoidă: $F(x) = \frac{1}{1 + e^{-\beta x}}$ cu $\beta=3$,
- algoritmul de corecție a ponderilor utilizează corecția adaptivă,
- termenul moment este $\gamma=0,2$,
- nu s-au utilizat unități BIAS¹¹⁰,
- numărul de epoci de antrenare este - 10 000,
- numărul de unități pe stratul ascuns este 24.

III.6.4. Corelarea punctelor de fierbere ale alcoolilor alifatici cu indici topologici

Conform variabilelor de intrare selectate, au fost efectuate corelări între punctele de fierbere a alcoolilor - Pf și următoarele seturi de indici topologici:

$$1 - \chi + \chi d_1$$

$$2 - \chi + J$$

$$3 - J+N$$

În consecință, drept variabilă dependentă s-a adoptat temperatura de fierbere a alcoolilor iar ca variabile independente - indicii topologici.

Pentru fiecare caz, s-au efectuat secvențial corelări multiliniare și s-au aplicat modele neurale, comparându-se indicatorii statistici corelaționali și evidențiindu-se avantajele și dezavantajele celor două modele.

Astfel, în figurile 7-8, 9-10, 11-12 sunt redată grafic dependențele liniare dintre punctele de fierbere calculate prin regresie multiliniară respectiv, cu rețele neurale și valorile experimentale ale punctelor de fierbere.

Ecuțiile 34-35, 36-37, 38-39 redau dependența analitică dintre punctele de fierbere calculate prin cele două modele și punctele de fierbere experimentale. În tabelele 3, 4, și 5 sunt sistematizați indicatorii statistici pentru cele două modele.

Performanța superioară a modelului neural în detrimentul regresiei multiliniare este prezentată comparativ prin variația procentuală [%] a indicatorilor statistici. În măsura în care variația indicatorilor statistici nu are sens (de exemplu pentru valorile volumului eșantionului N, sau pentru numărul gradelor de libertate NL la care compararea este în acest caz neconcludentă), s-a utilizat semnul “*”).

Variația procentuală [%] a indicatorilor statistici pentru modelul neural în comparație cu regresia multiliniară se concretizează în plus de performanță în cazul în care:

- Valorile *coeficientului de corelare*, *coeficientului de determinare* și *Testul Fisher* sunt mai mari;
- Valoarea *deviației standard* este mai mică.

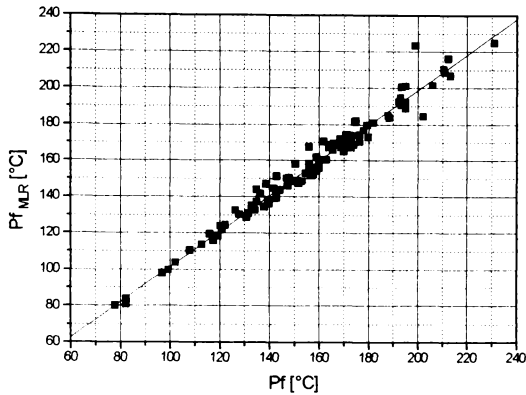


Fig.7

Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate prin regresie multiliniară Pf_{MLR} pe baza indicilor $\chi, \chi d_1$ și valoarea experimentală Pf a acestora

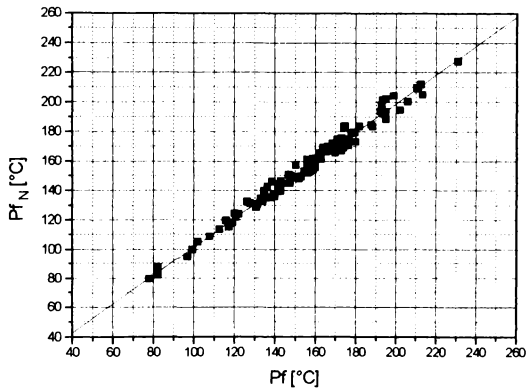


Fig.8

Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate cu rețeaua neurală Pf_N utilizând indicii $\chi, \chi d_1$ și valoarea experimentală Pf a acestora

Ecuția de regresie multiliniară $Pf=f(\chi, \chi d_1)$ este:

$$Pf_{MLR} = -31,369(\pm 4,543) - 182,669(\pm 15,836) \cdot \chi + 733,214(\pm 53,204) \cdot \chi d_1 \quad (34)$$

Ecuția de regresie liniară $Pf_N=f(Pf)$ este:

$$Pf_N = -1,247(\pm 2,007) + 1,009(\pm 0,013) \cdot Pf \quad (35)$$

Tabelul 3. Compararea calității celor două tipuri de modele:

	REGRESIE MULTILINIARA	RETEA NEURALA	COMPARARE %
-Volumul eșantionului :	N = 100	N = 100	*
-Numărul gradelor de libertate :	NL= 97	NL= 98	*
-Coeficientul de corelare	R = 0,987	R = 0,992	+0,55
-Coeficientul de determinare	R ² = 0,974	R ² = 0,985	+1,10
-Deviația standard	SD = 4,985	SD = 3,090	-8,01
-Testul Fisher	F = 1831	F = 6406	+249

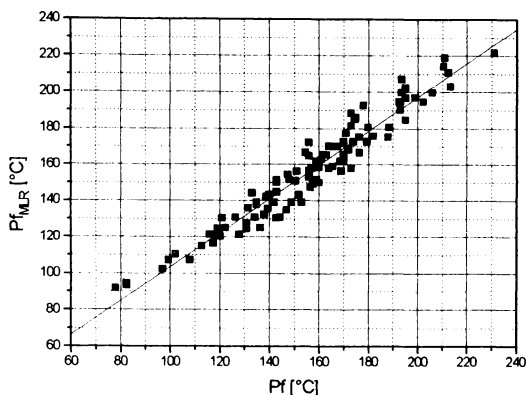


Fig. 9.

Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate prin regresie multiliniară Pf_{MLR} pe baza indicilor N,J și valoarea experimentală Pf a acestora

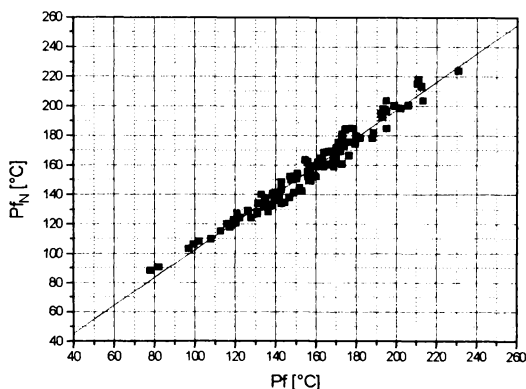


Fig. 10.

Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate cu rețeaua neurală Pf_N utilizând indicii N,J și valoarea experimentală Pf a acestora

Ecuția de regresie multiliniară $Pf=f(N,J)$ este:

$$Pf_{MLR} = 75,185(\pm 5,598) + 19,757(\pm 0,527) \cdot N - 26,332(\pm 2,038) \cdot J \quad (36)$$

Ecuția de regresie liniară $Pf_N=f(Pf)$ este:

$$Pf_N = -2,024(\pm 3,044) + 1,016(\pm 0,192) \cdot Pf \quad (37)$$

Tabelul 4. Compararea calității celor două tipuri de modele:

	REGRESIE MULTILINIARA	RETEA NEURALA	COMPARARE %
-Volumul eșantionului :	N = 100	N = 100	*
-Numărul gradelor de libertate :	NL= 97	NL = 98	*
-Coeficientul de corelare	R = 0,968	R = 0,983	+1,52
-Coeficientul de determinare	R ² = 0,938	R ² = 0,966	+3,05
-Deviatia standard	SD = 7,751	SD = 5,668	-6,88
-Testul Fisher	F = 729	F = 2810	285

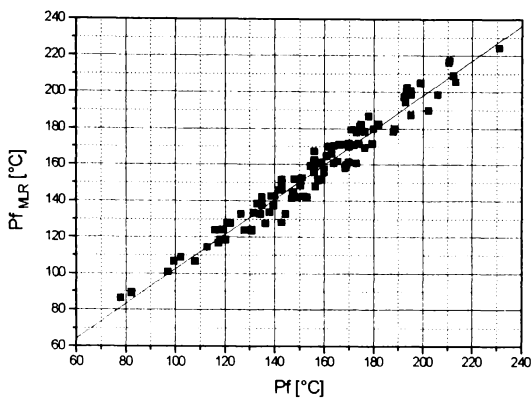


Fig. 11

Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate prin regresie multiliniară $Pf_{ML,R}$ pe baza indicilor χ, J și valoarea experimentală Pf a acestora

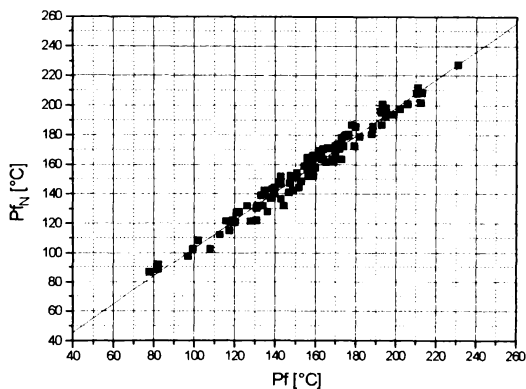


Fig. 12

Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate cu rețeaua neurală Pf_N utilizând indicii χ, J și valoarea experimentală Pf a acestora

Ecuția de regresie multiliniară $Pf_{ML,R}=f(\chi, J)$ este:

$$Pf_{ML,R} = 54,766(\pm 4,699) + 38,118(\pm 0,825)\chi - 13,780(\pm 1,548) \cdot J \quad (38)$$

Ecuția de regresie liniară $Pf_N=f(Pf)$ este:

$$Pf_N = 7,465(\pm 2,616) + 0,954(\pm 0,016) \cdot Pf \quad (39)$$

Tabelul 5. Compararea calității celor două tipuri de modele:

	REGRESIE MULTILINIARA	RETEA NEURALA	COMPARARE %
-Volumul eșantionului :	N = 100	N = 100	*
-Numărul gradelor de libertate :	NL= 97	NL= 98	*
-Coeficientul de corelare	R = 0,979	R = 0,986	+0.72
-Coeficientul de determinare	R ² = 0,958	R ² = 0,972	+1.44
-Deviatia standard	SD = 6,360	SD = 5,014	-21.16
-Testul Fisher	F = 1106	F = 3381	+205

Din valorile comparative ale parametrilor statistici pentru cele două tipuri de corelări, se observă că, indiferent de indicii topologici utilizați, modelul neural este superior regresiei multilinare din următoarele motive:

- Valorile coeficienților de corelare și determinare superioare cu $\cong 1\%$
- Deviația standard inferioară cu 21 - 38 %
- Valori pentru testul Fisher superioare cu 205 - 249%

Cea mai reprezentativă îmbunătățire obținută la utilizarea modelului neural o reprezintă reducerea deviației standard SD, fapt care impune modelul neural mai ales în cazul în care acesta este utilizat în scop **predictiv**.

Corelarea punctelor de fierbere ale alcoolilor cu indicii topologici conferă o bună reproductibilitate a valorilor experimentale pentru indicii topologici de tip χ , fără contribuția altor indici iar pentru indicii de tip **J**, cu contribuția numărului de atomi din moleculă **N** sau a unui indice de tip χ .

III.6.5 Studiu corelativ între discriminanți structurali și punctele de fierbere ale alcoolilor alifatici

Testele de corelare a punctelor de fierbere ale alcoolilor alifatici cu indicii topologici au dat rezultate semnificative, mai ales în cazul utilizării rețelelor neurale. Caracterizarea structurală prin indici topologici s-a dovedit a fi suficient de completă, însă valorile indicilor topologici sunt lipsite de o semnificație structurală directă.

Pentru a elimina acest inconvenient, s-au luat în considerare principalii factori care induc o variație semnificativă a temperaturii de fierbere, astfel considerați încât influența lor asupra temperaturii de fierbere să fie direct proporțională cu valoarea proprie. În continuare va fi prezentată transpunerea cantitativă a aspectelor structurale implicate, în discriminanți a căror valoare să fie direct proporțională cu temperatura de fierbere.

1. Temperatura de fierbere crește cu creșterea numărului de atomi de carbon din moleculă, **Nc**.

2. Temperatura de fierbere crește cu poziționarea extremă a grupării -OH, aspect care poate fi cuantificat prin **Ia**, denumit *indice de acentricitate*.

$$Ia = \frac{Lc_{max}}{C_{OH}} \quad (40)$$

unde:

- Lc_{max} este numărul de atomi de carbon din catena de lungime maximă a moleculei
- C_{OH} - reprezintă numărul atomului de carbon la care este legată gruparea -OH

3. Temperatura de fierbere scade cu numărul și mărimea substituenților și crește cu distanța față de gruparea -OH. Se definește un indice incremental de substituție **Ii**, care ține seama de tipul, numărul substituenților și depărtarea lor față de gruparea OH. Acesta se calculează conform relației:

$$I_i = \frac{I}{1 + \sum \frac{A_i}{1 + N_c - c}} \quad (41)$$

în care $A_i = \sqrt{N_c}$, este un coeficient al efectului inductiv al substituentului, N_c fiind numărul atomilor de carbon din care este format substituentul, iar $N_c - c$ reprezintă numărul legăturilor C-C care despart gruparea OH de substituent.

De menționat că maparea prin discriminanți structurali utilizată în acest capitol este originală, ea având un caracter intuitiv.

În tabelul 6 sunt redată valorile acestor descriptori pentru fiecare din cei 100 alcooli, aranjați pe liniile tabelului sub forma tiparelor utilizate la antrenarea rețelei neurale. Astfel în coloanele 1-3 sunt trecute valorile care compun vectorul de intrare al rețelei, iar în coloana 4 valoarea variabilei de ieșire a acesteia.

Semnificațiile valorilor din tabel sunt:

- coloana 0 - numărul curent care este identic cu cel din tabelul 1, fiind utilizat la identificarea alcoolilor.
- coloana 1 - N_c - numărul atomilor de carbon din molecula alcoolilor
- coloana 2 - I_a - valoarea indicelui de acentricitate
- coloana 3 - I_i - valoarea indicelui incremental de substituție
- coloana 4 - P_f - punctele de fierbere experimentale a alcoolilor
- coloana 5 - P_{f_N} - punctele de fierbere a alcoolilor calculate cu modelul neural
- coloana 6 - Err - coeficientul procentual de variație a valorii calculate față de cea experimentală a punctelor de fierbere

Tabelul 6 Valorile descriptorilor utilizați pentru setul de 100 alcooli

Nr. crt.	N_c	I_a	I_i	P_f [°C]	P_{f_N} [°C]	Err [%]
0	1	2	3	4	5	6
1	2	2	1	78	81,4	4,35
2	3	3	1	97,1	96,6	-0,51
3	3	1,5	1	82,4	84,6	2,66
4	4	4	1	117,6	116,4	-1,02
5	4	3	0,667	108,1	105,2	-2,68
6	4	2	1	99,5	99,9	0,4
7	4	1,5	0,5	82,4	86	4,36
8	5	5	1	138	137,7	-0,21
9	5	4	0,75	131	128,8	-1,67
10	5	4	0,667	128	126	-1,56
11	5	2,5	1	119,3	120,1	0,67
12	5	1,667	1	116,2	113,7	-2,15
13	5	2	0,667	112,9	108,8	-3,63
14	5	2	0,5	102,3	105,1	2,73
15	6	6	1	157,6	159,5	1,2
16	6	5	0,75	153	150,8	-1,43
17	6	5	0,8	151,9	152,8	0,59
18	6	5	0,667	149	147,5	-1

III. Corelări structură chimică - proprietăți fizice

0	1	2	3	4	5	6
19	6	4	0,586	147	145,2	-1,22
20	6	4	0,545	144,5	138,8	-3,94
21	6	4	0,6	143	140,8	-1,53
22	6	3	1	140	141,4	1
23	6	4	0,5	136,5	137,1	0,43
24	6	2	1	135	136,2	0,88
25	6	2,5	0,667	134,3	127,9	-4,76
26	6	2,5	0,75	131,6	130,5	-0,83
27	6	1,667	0,667	126,5	125,2	-1,02
28	6	1,667	0,5	122,4	120,2	-1,79
29	6	2,5	0,5	121,1	122,4	1,07
30	6	2	0,5	120,4	121	0,49
31	6	2	0,4	118,4	117,9	-0,42
32	7	7	1	176,4	179,3	1,64
33	7	6	0,8	173	173,2	0,11
34	7	6	0,833	170	174,4	2,58
35	7	6	0,75	169	171,3	1,36
36	7	6	0,667	164	168	2,43
37	7	3,5	1	159	161,7	1,69
38	7	5	0,571	159	163,2	2,64
39	7	2,333	1	157	157,9	0,57
40	7	1,75	1	156	156,7	0,44
41	7	3	0,8	151	152,6	1,05
42	7	2	0,75	148	149,2	0,81
43	7	3	0,5	143	141,5	-1,04
44	7	2	0,667	143	146,3	2,3
45	7	2	0,5	143	140,2	-1,95
46	7	1,667	0,414	142	136,6	-3,8
47	7	1,667	0,4	139,7	136,1	-2,57
48	7	1,667	0,5	138,7	139,8	0,79
49	7	1,667	0,5	135	139,8	3,55
50	7	2,5	0,429	133,1	137,8	3,53
51	7	2	0,333	131	133,8	2,13
52	8	8	1	195,1	195,2	0,05
53	8	7	0,857	188,6	191,9	1,74
54	8	7	0,8	188	190,1	1,11
55	8	4	1	180	179,6	-0,22
56	8	6	0,588	179,5	183,2	2,06
57	8	2	1	176,3	176,2	-0,05
58	8	2,333	0,8	174	170,6	-1,95
59	8	2,333	0,75	172	169	-1,74
60	8	2,667	1	171	176,8	3,39
61	8	3,5	0,8	170	171,5	0,88
62	8	2,333	0,667	170	166,1	-2,29
63	8	5	0,5	168,5	175	3,85
64	8	2,333	0,667	167,5	166,1	-0,83
65	8	3,5	0,667	166,1	166,9	0,48
66	8	3	0,545	165,5	161,8	-2,23
67	8	1,75	0,75	164	168,6	2,8

0	1	2	3	4	5	6
68	8	2,333	0,5	163	160	-1,84
69	8	1,75	0,667	162	165,8	2,34
70	8	1,75	0,5	161	159,9	-0,68
71	8	1,667	0,343	160	153,8	-3,87
72	8	3	0,4	160	156	-2,5
73	8	1,667	0,333	156,5	153,4	-1,98
74	8	2,5	0,293	156	151,6	-2,82
75	8	3,5	0,5	156	160,6	2,94
76	8	3	0,444	154,5	157,8	2,13
77	8	1,667	0,333	150,5	153,4	1,92
78	8	2,5	0,375	147,5	151,1	2,44
79	9	9	1	213,3	206,8	-3,04
80	9	8	0,875	206	204,6	-0,67
81	9	3	1	195	191,9	-1,58
82	9	4,5	1	193,5	193,9	0,2
83	9	1,8	1	193	191,4	-0,82
84	9	2,25	1	192,5	191,7	-0,41
85	9	1,75	0,414	182	178,6	-1,86
86	9	4	0,5	178	178	0
87	9	2,333	0,588	175	180,7	3,25
88	9	1,75	0,6	174,5	181	3,72
89	9	3,5	0,455	173	176	1,73
90	9	2,333	0,444	173	175,9	1,67
91	9	2	0,333	156	162,1	3,91
92	10	10	1	231,1	227,5	-1,55
93	10	8	0,677	212,5	210,1	-1,12
94	10	5	1	211	204,7	-2,98
95	10	2,5	1	210,5	203,2	-3,46
96	10	2,667	0,444	202,2	197,1	-2,52
97	10	2,667	0,414	199	194,2	-2,41
98	10	2	0,6	195	195	0
99	10	2,667	0,588	193,5	194,7	0,62
100	10	2,333	0,343	193	187,9	-2,64

Pentru început s-a testat o corelare cu ajutorul regresiei multiliniare utilizând ca variabilă dependentă temperatura de fierbere a alcoolilor iar ca variabile independente valorile descriptorilor menționați anterior.

În figura 13 este redată dependența liniară dintre valorile calculate prin regresie multiliniară Pf_{MLR} utilizând descriptorii menționați mai sus și valorile experimentale ale punctelor de fierbere Pf , ecuația de regresie multiliniară fiind descrisă de relația 42.

Se observă că utilizarea acestor parametri permite obținerea unui coeficient de corelare foarte bun care depășește valorile obținute prin regresie multiliniară a punctelor de fierbere cu indici topologici; deasemenea celelalți indicatori statistici sunt superiori.

Acești parametri structurali au fost utilizați în scopul predicției punctelor de fierbere cu ajutorul modelului neural.

Rețeaua neurală utilizată este descrisă la subpunctul III.6.3. Aceasta utilizează ca date de intrare valorile descriptorilor iar ca date de ieșire valorile punctelor de fierbere ale alcoolilor corespunzători.

În figura 14 este prezentată dependența liniară dintre valorile calculate cu modelul neural Pf_N , și cele experimentale Pf . Ecuația de regresie multiliniară corespunzătoare este descrisă de relația 43.

În tabelul 7 sunt prezentate comparativ valorile parametrilor statistici ai corelării pentru cele două tipuri de modele utilizate. Din valorile acestora rezultă superioritatea modelului neural, atât din punctul de vedere al coeficienților de corelare și determinare, cât și din punctul de vedere al dispersiei, caracterizate prin valoarea deviației standard și prin Testul Fisher.

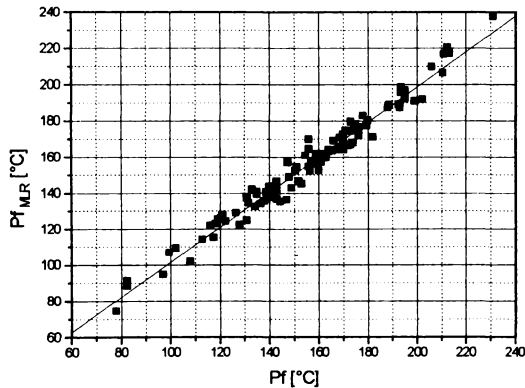


Fig. 13 Reprezentarea relației liniare dintre punctele de fierbere calculate prin regresie multiliniară Pf_{MLR} din indicii Nc, Ia, Ii și valoarea experimentală, Pf a acestora.

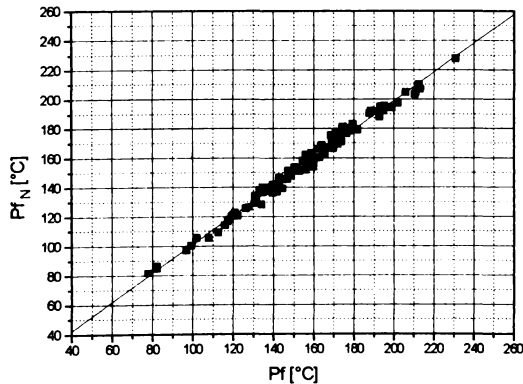


Fig. 14 Reprezentarea relației liniare dintre punctele de fierbere calculate cu rețeaua neurală Pf_N din indicii Nc, Ia, Ii și valoarea experimentală Pf a acestora

Ecuația de regresie multiliniară $Pf_{MLR}=f(Nc,L,C,St)$ este:

$$Pf_{MLR} = 6,087(\pm 3,063) + 16,248(\pm 0,321) \cdot Nc + 4,173(\pm 0,308) \cdot Ia + 27,349(\pm 2,517) \cdot Ii \quad (42)$$

Ecuția de regresie liniară $Pf_N=f(Pf)$ este:

$$Pf_N = 3,110(\pm 1,726) + 0,980(\pm 0,011) \cdot Pf \quad (43)$$

Tabelul 7. Indicatorii statistici aferenți celor două modele utilizate.

	REGRESIE MULTILINIARA	RETEA NEURALA	COMPARARE %
-Volumul eșantionului :	N = 100	N = 100	-
-Numărul gradelor de libertate :	NL= 96	NL= 98	-
-Coeficientul de corelare	R = 0,986	R = 0,994	+0,81
-Coeficientul de determinare	R ² = 0,972	R ² = 0,988	+1,62
-Deviatia standard	SD = 5,183	SD = 3,108	-0,03
-Testul Fisher	F = 1127	F = 8193	+627

III.6.6 Validarea modelului

Pentru testarea modelului, s-a utilizat procedura validării încrucișate. Pentru o încărcare cât mai uniformă a domeniului variabilelor de ieșire, întreg setul de date a fost divizat în două subseturi, folosind ca unic criteriu de selecție numărul curent (par sau impar) al întregului set de date.

Într-un prim stadiu, rețeaua a fost antrenată cu setul de date impar, pe baza ponderilor obținute fiind precise valorile setului de date par. În figura 15 este redată dependența liniară dintre punctele de fierbere astfel calculate și valorile experimentale pentru setul de date par. Relația 44 reprezintă ecuația de regresie liniară dintre valorile calculate și cele experimentale pentru setul de date par.

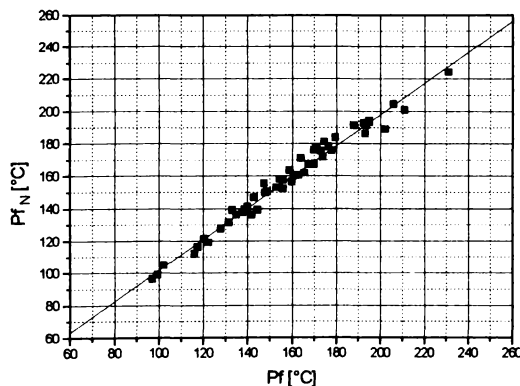


Fig. 15 Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate cu rețeaua neurală Pf_N și valoarea experimentală Pf a acestora pentru seturile de date pare

Ecuția de regresie liniară $Pf_N=f(Pf)$ este:

$$Pf_N = 5,566(\pm 3,277) + 0,962(\pm 0,020) \cdot Pf \quad (44)$$

Ulterior, rețeaua a fost antrenată cu setul de date par, pe baza ponderilor obținute fiind precise valorile setului de date impar.

In figura 16 este redată dependența liniară dintre temperatura de fierbere astfel calculată și cea experimentală pentru setul de date impar. Relația 45 reprezintă ecuația de regresie liniară dintre valorile calculate și cele experimentale pentru setul de date impar.

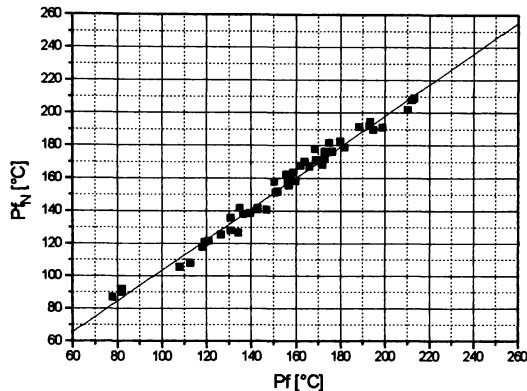


Fig.16 Reprezentarea grafică a relației liniare dintre punctele de fierbere calculate cu rețeaua neurală Pf_N și valoarea experimentală Pf a acestora pentru seturile de date impare

Ecuatia de regresie liniară $Pf_N=f(Pf)$ este:

$$Pf_N = 8,288(\pm 3,118) + 0,949(\pm 0,020) \cdot Pf \quad (45)$$

Indicatorii statistici caracteristici ecuațiilor corelaționale corespunzătoare celor două subseturi de date calculate fiecare pe baza celui alt (tabelul 8) argumentează apartenența celor două subseturi de date la aceeași populație din care provin și seturile de bază utilizate la antrenarea rețelei. Valorile acestor parametrii statistici argumentează stabilitatea și capacitatea de predicție foarte bună a rețelei neurale utilizate.

Tabelul 8

Parametrii de corelare ai regresiei liniare dintre valorile calculate și cele experimentale

	Setul de date par	Setul de date impar
-Volumul eșantionului :	N = 50	N = 50
-Numărul gradelor de libertate	NL= 48	NL= 48
-Coeficientul de corelare	R = 0,989	R = 0,989(7)
-Coeficientul de determinare	$R^2 = 0,979$	$R^2 = 0,979(6)$
-Deviatia standard	SD = 4,237	SD = 4,401
-Testul Fisher	F = 2243	F = 2305

III.7. Concluzii

Pentru corelarea punctelor de fierbere corespunzătoare a unui set de 100 alcoolii alifatici cu indicii topologici care cuantifică structura acestora, s-a testat inițial capacitatea de corelare liniară a fiecărui indice topologic. Cele mai bune rezultate s-au obținut cu indicii de tip χ . Este știut faptul că indicii de tip J nu

corelează corespunzător proprietățile fizice ale substanțelor organice utilizate, ci doar împreună cu numărul de atomi ai moleculei. De aceea s-au selectat indicii de tip χ care dau cei mai buni coeficienți de corelare liniară (χ , χd_1), precum și indicele J împreună cu numărul de atomi din moleculă pentru a se testa capacitatea de corelare combinată a acestuia cu temperatura de fierbere a alcoolilor. Pentru corelare s-au utilizat două modalități de corelare: corelarea multiliniară și cea bazată pe un model neural.

În toate cazurile, modelul neural s-a dovedit a fi superior regresiei multiliniare, atât prin valorile coeficienților de corelare și determinare cât și prin efectul de reducere a dispersiei.

Având în vedere absența unei semnificații fizice directe a indicilor topologici, s-a realizat o cuantificare originală a structurii alcoolilor din setul de date prin descriptorii derivați din considerente structurale ce presupun o dependență directă între valoarea acestora și variația punctelor de fierbere.

Corelarea acestor descriptorii structurali cu punctele de fierbere s-a realizat atât prin regresie multiliniară cât și utilizând un model neural. Modelul neural a fost testat printr-o procedură de validare încrucișată, parametrii de corelare fiind superiori celor obținuți prin regresie multiliniară.

Ambele modalități s-au dovedit a fi superioare indicatorilor statistici modelului utilizat de Hall¹⁷⁸ pentru corelarea punctelor de fierbere ale alcoolilor cu o combinație de indici electronici, topologici și de stare de valență.

IV UTILIZAREA REȚELELOR NEURALE ARTIFICIALE LA MODELAREA UNUI PROCES ELECTROCHIMIC

A fost luat în studiu cazul particular al reducerii nitrobenzenului la anilină, proces studiat în detaliu în literatura de specialitate¹⁷⁹⁻¹⁸². Realizarea practică a acestui proces presupune reactoare electrochimice cu spațiile electrodice separate prin membrane sau diafragme în scopul evitării degradării prin oxidare a produsului obținut la catod¹⁸³⁻¹⁸⁵.

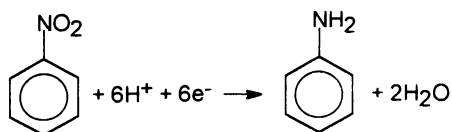
Recent, a fost demonstrată posibilitatea realizării cu randamente de curent mari a procesului într-un reactor electrochimic necompartimentat utilizând în calitate de catod un strat fix de granule de cupru Raney¹⁸⁶, ca electrolit suport o soluție apoasă de sulfat de sodiu la pH=5,5, substratul organic fiind emulsionat prin agitare.

IV.1 Chimismul procesului

Mecanismul procesului de reducere catodică a nitrobenzenului depinde de condițiile în care are loc electroliza. Natura materialului din care este confecționat catodul și pH-ul soluției de electrolit sunt factorii principali care influențează desfășurarea procesului. În cazul reducerii nitrobenzenului într-un reactor electrochimic necompartimentat, echipat cu catod din cupru scheletat și anod din platină, în soluție de sulfat de sodiu la pH=5, procesele principale care au loc la electrozi sunt următoarele:

La catod:

Reducerea nitrobenzenului la anilină:

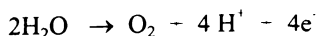


- degajarea hidrogenului:



La anod: - Pt

- degajarea oxigenului:



- reacții de oxidare a anilinei.

În masa soluției de electrolit au loc și alte procese la care participă intermediarii formați în procesele de electrod sau reacții de formare a unor produși de condensare.

Pentru obținerea unor randamente de curent mari în reactoarele electrochimice necompartimentate, este necesar a stimula selectiv reducerea nitrobenzenului la anilină la catod și degajarea oxigenului la anod. În ce privește electrolitul suport utilizat, atât reducerea nitrobenzenului la catod cât și reacția de degradare prin oxidare a anilinei depind de etapa de transport a moleculelor organice din electrolit la electrod. Deci, o stimulare selectivă a reacției de electrod utilă devine posibilă dacă procesul este condus astfel încât:

- densitatea de curent catodică să fie scăzută la o valoare care să tindă spre valoarea curentului limită de reducere a nitrobenzenului fără a o depăși.
- densitatea de curent anodică să fie cât mai mare posibilă la o valoare care să depășească curentul limită de oxidare al anilinei

O mod simplu de realizare a acestor considerente în cazul unui reactor electrochimic necompartimentat este de tip constructiv prin asigurarea unui raport suprafață catodică - suprafață anodică S_c/S_a mare.

La reducerea electrochimică a nitrobenzenului la anilină, indiferent de sistemul solvent-electrolit suport utilizat, procesul catodic este controlat de etapa de **transport a substratului organic din masa soluției de electrolit la catod**. Depășirea curentului limită de reducere a nitrobenzenului va provoca diminuarea randamentului de curent pentru obținerea anilinei, cu atât mai mult cu cât densitatea de curent este mai mare. Ca urmare, la densități mari de curent, în mediu apos, este favorizat procesul de degajare a hidrogenului.

Într-un reactor electrochimic necompartimentat, anilina formată la catod poate accede la anod, prin difuziune sau convecție, unde este degradată prin oxidare, fapt care contribuie atât la scăderea randamentului de curent, cât și la scăderea conversiei utile.

Și la anod procesul de oxidare a anilinei este controlat de etapa de transport a substratului organic, dar aici depășirea curentului limită este favorabilă, întrucât stimulează procesul de degajare a oxigenului.

IV.2 Descrierea instalației experimentale și selecția parametrilor de proces

Pentru realizarea determinărilor experimentale s-a utilizat un reactor electrochimic dintr-un vas cilindric de sticlă (dimensiuni $H=0.08$ m , $\Phi=0.04$ m) prevăzut cu manta de răcire (figura 1).

Semnificațiile notațiilor din figură sunt:

- | | |
|---------------------------|-----------------------|
| 1 - Baie termostată | S - Sursă |
| 2 - Agitator | A - Ampermetru |
| 3 - Celulă electrochimică | R - Rezistență etalon |
| 4 - Anod platină | C - Coulombmetru |
| 5 - Catod cupru scheletat | V - Voltmetru |

Catodul cilindric (dimensiuni $H=0.05$ m , $\Phi=0.038$ m) a fost confecționat dintr-o folie de cupru 0.0002 m pe care s-a depus un strat de cupru scheletat cu suprafață specifică mare obținut prin tescompunerea termică a glioxilului

heteropolinuclear de cupru(II) și aluminiiu (III)¹⁸⁷. Drept contraelectrod s-a utilizat un fir de platină ($\Phi=0.001\text{ m}$) plasat central în interiorul catodului, realizându-se astfel un raport constructiv al suprafețelor S_c/S_a de aproximativ 40.

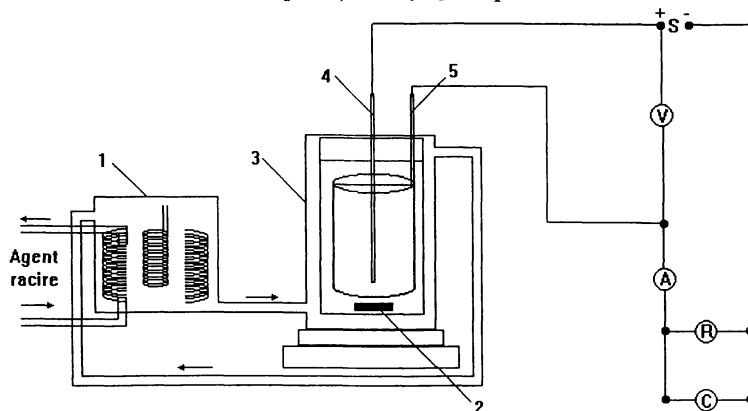


Fig. 1. Instalația experimentală

Amestecul inițial de reacție a fost constituit din 100 cm^3 soluție de Na_2SO_4 10% la $\text{pH} = 5.5$ utilizat ca electrolit, în care s-a efectuat inițial o electroliză în scopul formării electrozilor, apoi s-au introdus 20 cm^3 nitrobenzen și s-a realizat experimentul în condițiile prestabilite. Amestecul final a fost colectat, faza organică a fost separată de cea anorganică, concentrația de anilină din faza organică fiind determinată printr-o metodă cromatografică gaz-lichid în următoarele condiții:

Coloană: oțel inox, 1×0.004 umplută cu OV-17 (5%) pe suport Chromosorb W-AW, DMCS 80-100 mesh

Gaz purtător: azot - 100ml/min

Detector: FID

Temperatură injector : 240°C

Temperatură detector : 240°C

Standard intern: n- decan

Întrucât solubilitatea anilinei în soluții apoase este scăzută, **în calcule nu s-a ținut cont de anilina dizolvată în faza apoasă.**

S-au utilizat reactivi Fluka:

Nitrobenzen puriss p.a.

Sulfat de sodiu decahidrat puriss p.a. cristalizat

Acid sulfuric 97% puriss p.a.

n- decan puriss - standard pentru cromatografie gaz-lichid

Instalația experimentală permite variația principalilor parametri de care depinde randamentul de curent în anilină pentru procesul studiat. Astfel, plaja de variație a acestora a fost:

- densitatea de curent	- i -	100 - 500 A/m^2
- temperatura	- t -	25 - 60°C
- cantitatea de electricitate trecută prin reactorul electrochimic	- Q -	2000 - 6000 C

Designul electrozilor permite operarea procesului la un raport suprafață catodică/suprafață anodică foarte mare, fapt ce induce pe de o parte stimularea procesului catodic util de reducere a anilinei iar pe de altă parte, prin degajarea oxigenului la anod - inhibarea procesului anodic de degradare a anilinei formate.

Menținând constanți pe parcursul desfășurării procesului, raportul suprafață catodică/suprafață anodică, concentrația electrolitului suport și regimul hidrodinamic, principalii parametri care influențează eficiența procesului desfășurat într-un reactor necompartimentat sunt considerați:

- densitatea de curent i [A/m^2]
- temperatura t [$^{\circ}C$]
- cantitatea de electricitate trecută prin reactorul electrochimic Q_{exp} [C].

IV.3 Rezultate experimentale

În tabelul 1 sunt redate datele experimentale ce presupun condițiile de lucru (coloanele 1-3) în care au fost obținute randamentele de curent experimentale (coloana 4) - η_{exp} exprimate în procente.

Tabelul 1. Condițiile experimentale și rezultatele obținute

Nr. tipar	$Q_{exp} \cdot 10^{-2}$ [C]	t [$^{\circ}C$]	$i \cdot 10^{-1}$ [A/m^2]	η_{exp} [%]	η_{CN} [%]	Eroare [%]
0	1	2	3	4	5	6
1.	40	25	10	65,4	65,37	-0,05
2.	20	35	30	76,1	75,01	-1,43
3.	20	25	10	71,7	71,11	-0,82
4.	20	35	40	76,4	75,67	-0,95
5.	42	25	20	60,1	59,36	-1,24
6.	20	35	50	76,8	75,42	-1,79
7.	20	25	25	66,8	66,56	-0,37
8.	40	35	30	66,2	66,92	1,08
9.	41.5	25	40	54,8	54,38	-0,77
10.	40	35	40	64,1	63,41	-1,08
11.	20	25	40	59,2	60,37	1,97
12.	40	35	50	60,3	59,53	-1,27
13.	40	45	10	69,3	69,35	0,07
14.	60	35	30	38,2	37,91	-0,76
15.	20	45	10	76,0	77,59	2,10
16.	60	35	40	37,1	36,12	-2,64
17.	40	45	25	67,2	68,41	1,81
18.	60	35	50	35,8	35,25	-1,53
19.	20.86	45	25	81,2	79,44	-2,17
20.	20	45	30	80,1	79,88	-0,27
21.	40	45	40	67,7	67,26	-0,65
22.	20	45	40	79,9	80,55	0,81
23.	20	45	40	82,6	80,55	-2,48
24.	20	45	50	79,8	80,83	1,29
25.	40	60	10	60,6	61,20	0,99

Tabelul 1. (Continuare)

0	1	2	3	4	5	6
1.	40	45	30	68	68,16	0,23
2.	20	60	10	75,9	74,68	-1,60
3.	40	45	40	67,4	67,26	-0,21
4.	41,86	60	25	62,8	61,72	-1,72
5.	40	45	50	66,8	65,41	-2,08
6.	20	60	25	72,8	74,99	3,00
7.	60	45	30	47,3	46,85	-0,94
8.	40	60	40	64,7	64,46	-0,37
9.	60	45	40	43,4	43,96	1,29
10.	20	60	40	74,7	75,57	1,17
11.	60	45	50	42,1	41,05	-2,49
12.	60	60	40	54,7	54,71	0,02
13.	20	55	30	77,4	77,73	0,43
14.	62,63	60	25	53,3	53,56	0,49
15.	20	55	40	77,1	78,16	1,38
16.	60	60	10	55,1	53,48	-2,94
17.	20	55	50	77,5	78,63	1,46
18.	60	45	40	44,2	43,96	-0,54
44.	40	55	30	64,6	65,14	0,84
45.	60	45	25	48,6	48,14	-0,94
46.	40	55	40	65,1	65,95	1,30
47.	60	45	10	50,6	51,14	1,07
48.	40	55	50	67,2	66,37	-1,23
49.	60	25	40	34,6	34,87	0,78
50.	60	55	30	54,1	53,30	-1,47
51.	60	25	25	35,3	35,27	-0,07
52.	60	55	40	52,1	52,59	0,94
53.	60	25	10	38,4	39,01	1,59
54.	60	55	50	50,2	51,14	1,88
55.	60	25	25	35,6	35,27	-0,92

Așa cum se poate observa din datele prezentate, în anumite condiții, randamentul de curent pentru formarea anilinei ajunge la valori apropiate de cele raportate în literatura de specialitate pentru cazul folosirii reactoarelor electrochimice cu spațiile electrodice separate^{184,185}

IV.4 Modelarea procesului

Pentru justificarea utilizării rețelelor neurale la modelarea procesului, s-a încercat inițial o corelare printr-o regresie liniară multivariabilă având ca variabilă dependentă randamentul de curent experimental iar ca variabile independente densitatea de curent - i , temperatura - t și cantitatea de electricitate trecută prin reactorul electrochimic - Q_{exp} .

În figura 2 s-a reprezentat grafic variația liniară dintre valorile randamentului calculat prin regresie multiliniară și valorile experimentale.

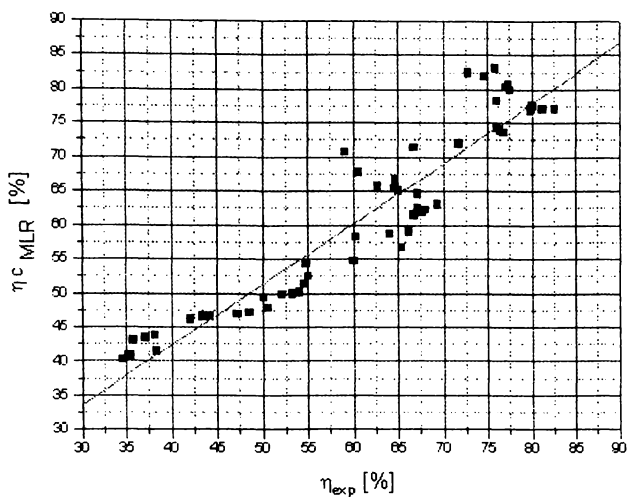


Fig. 2. Reprezentarea grafică a dependenței liniare dintre randamentele de curent calculate prin regresie multiliniară η_{MLR} și cele experimentale η_{exp}

Relația obținută pentru valorile randamentului calculat prin regresie multiliniară funcție de variabilele de proces este:

$$\eta_{\text{MLR}} = 79,697 (\pm 3,432) - 0,764 (\pm 0,040) \cdot Q_{\text{exp}} + 0,317 (\pm 0,056) \cdot t - 0,039 (\pm 0,052) \cdot i \quad (1)$$

Parametrii de corelare sunt:

Volumul esantionului	N = 55
Numărul gradelor de libertate	NL = 51
Coefficientul de corelare	R = 0,941
Coefficientul de determinare	R ² = 0,886
Deviația standard	SD = 4,928
Testul Fisher.	F = 396

Se observă că deși valoarea coeficientului de corelare este rezonabilă, valorile coeficientului de determinare și mai ales valoarea deviației standard fac regresia multiliniară improprie la modelarea procesului.

Ulterior realizării experimentelor, s-a utilizat o rețea neurală artificială cu propagare înapoi (Backpropagation Neural Network) în scopul validării datelor experimentale⁹⁸ și a furnării ponderilor ce corelează variabilele de intrare cu cele de ieșire ale procesului studiat¹. Sursa programului a fost realizată în Turbo Pascal, fiind proiectată pentru cele 55 de tipare experimentale și având următoarele caracteristici⁵⁴:

- vector de intrare tridimensional: Q_{exp}, t, i
- vector de ieșire unidimensional η_{exp}
- vectorul de pe stratul ascuns constituit din 14 noduri

- funcție de activare sigmoidă cu $\beta = 3$
- rata de învățare $\eta = (0,5 \dots 0,1)$
- termenul *momentum* = 0,1
- număr de epoci rulate = 1000

În figura 3 este redată variația erorii totale a rețelei cu numărul de epoci parcurse.

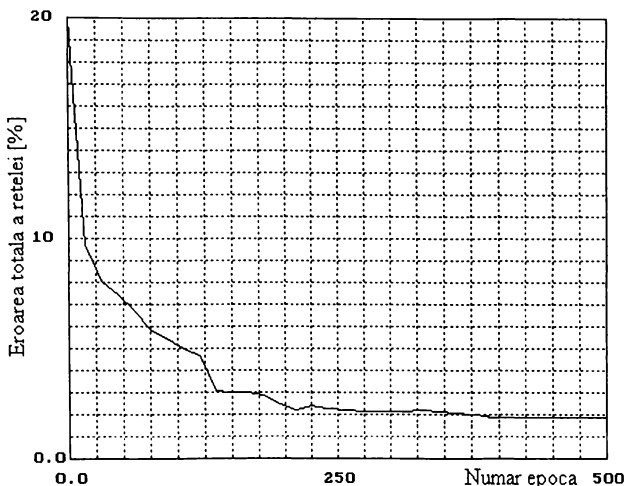


Fig. 3. Variația erorii totale a rețelei cu numărul de epoci parcurse

În urma rulării programului, valorile obținute ale randamentului de curent în anilină η_{CN} sunt trecute în tabelul 1 coloana 5, iar erorile procentuale ale valorilor randamentului de curent calculat față de cel experimental, în coloana 6. În vederea comparării rezultatelor calculate cu cele experimentale, în figura 4 s-au reprezentat valorile calculate ale randamentului de curent în funcție de cele experimentale efectuându-se și o regresie liniară între cele două seturi de valori.

Dependența liniară dintre randamentul de curent calculat cu rețeaua neurală η_{CN} funcție de cel experimental η_{exp} , este dată de relația 2:

$$\eta_{CN} = -0,23 (\pm 0,55) + 1,00 (\pm 0,01) \cdot \eta_{exp} \quad (2)$$

Parametrii de corelare sunt:

Volumul esantionului	N	= 55
Numărul gradelor de libertate	NL	= 51
Coefficientul de corelare	R	= 0,998
Coefficientul de corelare total	R ²	= 0,998
Deviația standard	DS	= 0,906
Testul Fischer	F	= 13275

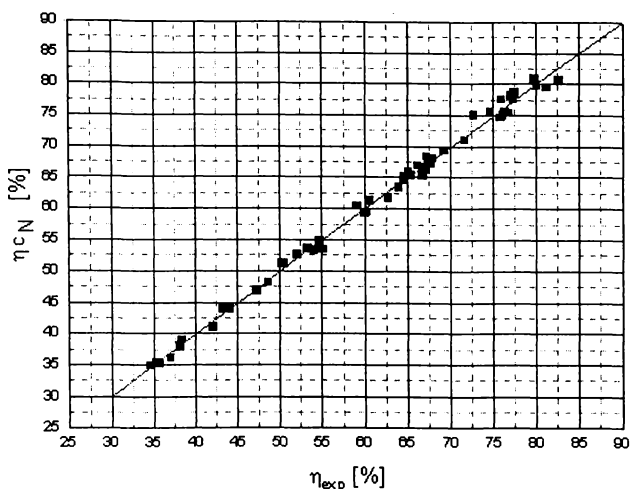


Fig. 4. Reprezentarea grafică a dependenței liniare dintre randamentele de curent calculate cu rețeaua neurală η_{cN} și cele experimentale η_{exp}

Prin compararea tabelară și grafică a randamentului de curent calculat η_{cN} cu cel experimental η_{exp} , se observă o foarte bună reproductibilitate a valorii experimentale prin calcul, deci o intuire fidelă a mecanismului care guvernează procesul.

IV.5 Validarea modelului

Pentru testarea modelului s-a utilizat metoda validării încrucișate (cross validation method¹¹²), în care cele 55 seturi de date experimentale au fost divizate în două subseturi. Regula de divizare a fost de natură arbitrară din punctul de vedere al procesului, selecția făcându-se după paritatea numărului de curent al tuturor tiparelor experimentale. Ulterior divizării setului de date în cele două subseturi, fiecare din ele a fost utilizat la antrenarea unei rețele neurale în scopul predicției valorilor celuilalt.

Relația 3 exprimă analitic dependența liniară dintre valorile de predicție ale randamentului de curent calculat η_{cN} și cele experimentale η_{exp} , pentru tiparele cu număr de curent impar, în figura 5 această dependență fiind redată grafic.

$$\eta_{cN} = 1,093 (\pm 1,052) + 0,979(\pm 0,017) \cdot \eta_{exp} \quad (3)$$

Parametrii de corelare sunt:

Volumul esantionului	N	= 28
Numărul gradelor de libertate	NL	= 26
Coefficientul de corelare	R	= 0,996
Coefficientul de determinare	R ²	= 0,992
Deviatia standard	SD	= 1,241
Testul Fischer	F	= 3307

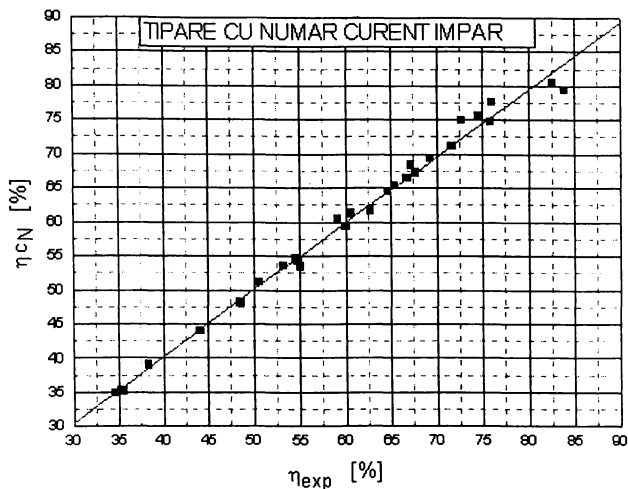


Fig. 5. Reprezentarea grafică a dependenței liniare dintre valorile de predicție ale randamentului de curent calculat η_{CN} și cele experimentale η_{exp}

Relația 4 exprimă analitic dependența liniară dintre valorile de predicție ale randamentului de curent calculat η_{CN} și cele experimentale η_{exp} , pentru tiparele cu număr curent par, în figura 6 această dependență fiind redată grafic.

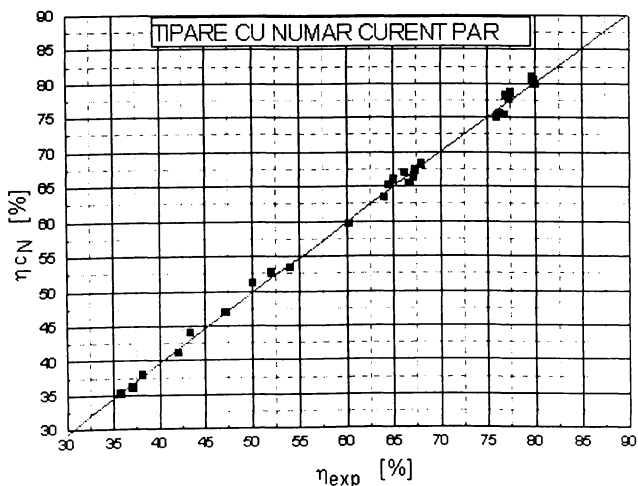


Figura 6. Reprezentarea grafică a dependenței liniare dintre valorile de predicție ale randamentului de curent calculat η_{CN} și cele experimentale η_{exp}

$$\eta_{CN} = -0,886 (\pm 0,698) + 1,012 (\pm 0,011) \cdot \eta_{exp} \quad (4)$$

Parametrii de corelare sunt:

Volumul eșantionului	N = 27
Numărul gradelor de libertate	NL = 25
Coefficientul de corelare	R = 0,998(6)
Coefficientul de determinare	R ² = 0,997
Deviația standard	DS = 0,658
Testul Fischer	F = 8682

Se remarcă o foarte bună predicție a valorilor randamentului de curent pentru fiecare subset de către rețeaua antrenată cu contrasetul corespunzător, fapt ce presupune apartenența tuturor tiparelor la același set de date - setul tiparelor experimentale.

IV.6 Decelarea influenței parametrilor de proces în vederea optimizării procesului

Pentru a sesiza importanța fiecărui parametru în cadrul procesului, s-au utilizat trei seturi de date în care a fost variat doar un singur parametru de intrare, ceilalți menținându-se constanți. Utilizând rețeaua antrenată cu cele 55 de tipare experimentale, *pe baza ponderilor calibrate la antrenare, într-o semiepocă de propagare înainte a rețelei*, s-au calculat valorile de predicție ale randamentelor de curent corespunzătoare seturilor de date de intrare.

În tabelele 2, 3 și 4 sunt redate condițiile propuse ale experimentelor pentru care rețeaua antrenată cu tiparele experimentale calculează prin ponderile de conexiune datele de ieșire corespunzătoare datelor de intrare impuse în scopul decelării influenței fiecărui parametru de proces asupra randamentului de curent..

Astfel, în tabelul 2, sunt redați parametrii de proces impuși ca date de intrare pentru rețea: cantitatea de electricitate trecută prin reactor Q este variată între limitele 1500 și 6500 C temperatura și densitatea de curent fiind menținute la valori constatante 45°C și respectiv 500 A/m², iar ca date de ieșire ale rețelei - valorile de predicție ale randamentelor de curent.

Tabelul 2 Variația randamentului de curent calculat η_{CN} funcție de cantitatea de electricitate trecută prin reactor

Nr. tipar	Q·10 ⁻² [C]	t [°C]	i·10 ⁻¹ [A/m ²]	η_{CN} [%]
0	1	2	3	4
1	15	45	50	81,13
2	16	45	50	81,17
3	18	45	50	81,10
4	20	45	50	80,83
5	22	45	50	80,34
6	24	45	50	79,65
7	26	45	50	78,72
8	28	45	50	77,56
9	30	45	50	76,14
10	32	45	50	74,46
11	34	45	50	72,53
12	36	45	50	70,36

Tabelul 2 (continuare)

0	1	2	3	4
13	38	45	50	67,98
14	40	45	50	65,41
15	42	45	50	62,72
16	45	45	50	58,55
17	46	45	50	57,16
18	48	45	50	54,42
19	50	45	50	51,77
20	52	45	50	49,25
21	54	45	50	46,90
22	56	45	50	44,75
23	58	45	50	42,80
24	60	45	50	41,05
25	62	45	50	39,51
26	64	45	50	38,15
27	65	45	50	37,54

În figura 7 este reprezentată variația estimată predictiv a randamentului de curent η_{cN} funcție de cantitatea de electricitate trecută prin reactor.

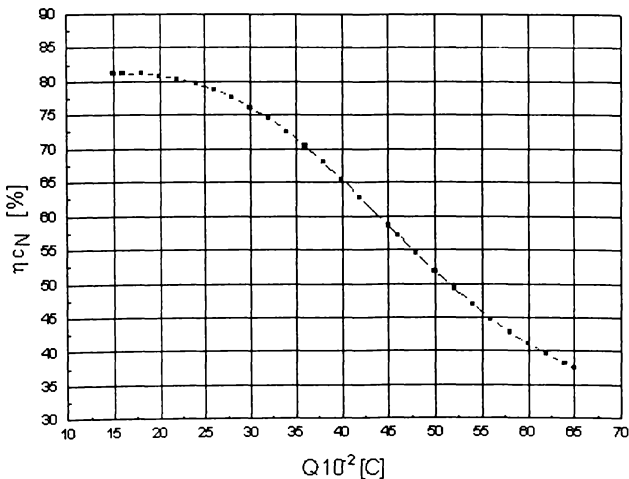


Fig. 7. Variația randamentului de curent η_{cN} funcție de cantitatea de electricitate trecută prin reactor Q

Se constată o dependență descrescătoare a randamentului de curent cu creșterea cantității de electricitate trecută prin reactor.

Fenomenul se datorează faptului că în cursul electrolizei crește concentrația anilinei în electrolit, având drept consecință intensificarea transportului produsului util din masa de reacție la anod, unde anilina este degradată. *Astfel, creșterea cantității de electricitate trecută prin reactorul electrochimic provoacă scăderea randamentului de curent în anilină.*

În tabelul 3, sunt redați parametrii de proces impuși ca date de intrare pentru rețea: temperatura este variată între limitele 20 și 65°C, cantitatea de electricitate trecută prin reactor și densitatea de curent fiind menținute la valori constante 2000 C și respectiv 500 A/m² iar ca date de ieșire ale rețelei - valorile de predicție ale randamentelor de curent η_{cN} .

Tabelul 3 Valorile randamentului de curent calculat η_{cN} funcție de temperatura de reacție t

Nr. tipar	Q·10 ⁻² [C]	t [°C]	i·10 ⁻¹ [A/m ²]	η_{cN} [%]
0	1	2	3	4
1	20	20	50	52,44
2	20	22	50	53,16
3	20	24	50	54,93
4	20	26	50	57,80
5	20	28	50	61,62
6	20	30	50	65,99
7	20	32	50	70,28
8	20	34	50	73,93
9	20	36	50	76,69
10	20	38	50	78,58
11	20	40	50	79,77
12	20	41	50	80,17
13	20	42	50	80,46
14	20	43	50	80,66
15	20	44	50	80,78
16	20	45	50	80,83
17	20	46	50	80,82
18	20	47	50	80,75
19	20	48	50	80,64
20	20	49	50	80,47
21	20	50	50	80,26
22	20	52	50	79,72
23	20	54	50	79,03
24	20	56	50	78,20
25	20	58	50	77,24
26	20	60	50	76,19
27	20	62	50	75,06
28	20	64	50	73,88
29	20	65	50	73,28

În figura 8 este reprezentată variația estimată predictiv a randamentului de curent η_{cN} funcție de temperatura masei de reacție t.

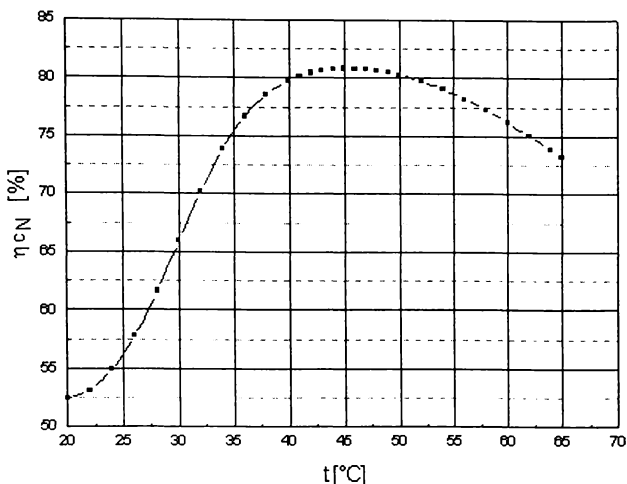


Fig. 8. Variația randamentului de curent η_{cN} funcție de temperatura masei de reacție t

Influența temperaturii de reacție asupra procesului este complexă, presupunând:

- pe de o parte influența benefică a temperaturii asupra transferului substratului organic din faza organică în faza apoasă în care are loc reacția principală datorită creșterii solubilității nitrobenzenului în faza apoasă

- pe de altă parte influența defavorabilă prin scăderea suprapotențialului reacției concurente de degajare a hidrogenului. De asemenea, conform datelor din literatura de specialitate¹⁸³, creșterea temperaturii duce la intensificarea reacțiilor secundare la care participă produșii de reacție formați la catod.

Conform dependenței randamentului de curent de temperatură din figura 6, se poate afirma că *până la temperatura de $\cong 45^\circ\text{C}$ procesul catodic util este determinat de transferul interfațic al nitrobenzenului, iar la valori mai mari ale temperaturii efectul benefic este supracompensat de scăderea suprapotențialului de degajare a hidrogenului și de creșterea ponderii proceselor secundare.*

În tabelul 4, sunt redați parametrii de proces impuși ca date de intrare pentru rețea: densitatea de curent este variată între limitele 50 și 600 A/m², temperatura și cantitatea de electricitate trecută prin reactor fiind menținute la valori constante 45°C și respectiv 500 A/m² iar ca date de ieșire ale rețelei - valorile de predicție ale randamentelor de curent.

Utilizând valorile din tabelul 4, s-au reprezentat grafic în figura 9 valorile de predicție ale randamentului de curent η_{cN} funcție de densitatea de curent.

Tabelul 4 Variația randamentului de curent calculat η_{CN}
funcție de densitatea de curent

Nr. tipar	$Q \cdot 10^{-2}$ [C]	t [°C]	$i \cdot 10^{-1}$ [A/m ²]	η_{CN} [%]
1	20	45	5	76,96
2	20	45	6	77,08
3	20	45	8	77,33
4	20	45	10	77,60
5	20	45	12	77,85
6	20	45	14	78,11
7	20	45	16	78,36
8	20	45	18	78,61
9	20	45	20	78,85
10	20	45	21	78,97
11	20	45	22	79,08
12	20	45	23	79,19
13	20	45	24	79,30
14	20	45	26	79,51
15	20	45	28	79,70
16	20	45	30	79,88
17	20	45	31	79,96
18	20	45	32	80,04
19	20	45	33	80,12
20	20	45	34	80,19
21	20	45	35	80,26
22	20	45	36	80,33
23	20	45	37	80,39
24	20	45	38	80,45
25	20	45	39	80,50
26	20	45	40	80,55
27	20	45	41	80,60
28	20	45	42	80,64
29	20	45	43	80,68
30	20	45	44	80,71
31	20	45	45	80,74
32	20	45	46	80,77
33	20	45	47	80,79
34	20	45	48	80,81
35	20	45	49	80,82
36	20	45	50	80,83
37	20	45	52	80,82
38	20	45	54	80,81
39	20	45	56	80,78
40	20	45	58	80,72
41	20	45	60	80,64

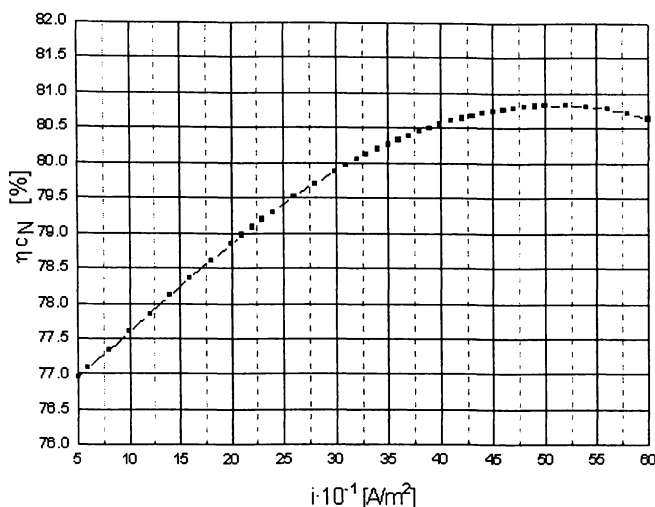


Fig.9 .Variația randamentului de curent η_{cN} funcție de densitatea de curent i

Conform așteptărilor, se constată o variație crescătoare a randamentului de curent odată cu creșterea densității de curent până la o valoare de 500 de A/m^2 după care valoarea randamentului de curent scade. Fenomenul este explicabil, deoarece creșterea densității de curent peste valoarea de $500 A/m^2$ va duce la scăderea randamentului de curent depășindu-se curentul limită de reducere a nitrobenzenului și rezultând implicit o creștere a ponderii reacției de degajare a H_2 .

Având în vedere că valoarea de $500 A/m^2$ este la capătul scării de valori a densității de curent, s-a considerat necesar a se confirma experimental scăderea randamentului de curent la depășirea acestei valori.

Tabelul 5 Valorile suplimentare ale randamentului de curent experimental η_{exp} și calculat η_{cN} funcție de variația densității de curent

Nr. tipar	$Q \cdot 10^{-2}$ [C]	t [°C]	$i \cdot 10^{-1}$ [A/m^2]	η_{exp} [%]	η_{cN} [%]	Eroare [%]
1.	20	45	50	80,40	80.83	0,53
2.	20	45	55	79,71	80.79	1,33
3.	20	45	60	79,12	80.64	1,88

Se observă că într-adevăr, valorile randamentului de curent experimental pentru variația densității peste valoarea de $500 A/m^2$ descresc pe măsura depărtării de această valoare.

Experimental se constată că la densități mici de curent creșterea densității de curent duce la creșterea randamentului de curent deoarece nu se depășește curentul limită iar la anod densitatea de curent nu este suficient de mare, astfel încât ponderea reacției de degradare a anilinei este neglijabilă.

La densitatea de curent la care randamentul este maxim, (500 A/m^2) - densitate de curent aplicabilă și la nivel industrial - randamentele de curent se află la valori ridicate, comparabile cu cele care se obțin în reactoare electrochimice compartimentate^{184,185}.

IV.7 Concluzii

În acest capitol s-a utilizat un model neural în scopul simulării, modelării și verificării capacității de predicție pentru un proces electrochimic multiparametric.

Aplicarea rețelelor neurale artificiale la modelarea proceselor este utilizată cu predilecție pentru *decelarea* influenței variabilelor de intrare asupra desfășurării procesului. Acest fapt permite atât relevarea influenței variabilelor de proces cât și determinarea condițiilor optime de desfășurare a procesului în ansamblu. Astfel, pe baza unui set de date experimentale, s-a antrenat o rețea neurală care pe baza ponderilor create în această etapă poate estima valori de ieșire - în acest caz randamentul de curent al procesului - pentru valori de intrare impuse. Din studiul influenței independente a fiecărei variabile de intrare pe întreg domeniul de variație experimental, se pot obține informații privitor la condițiile optime de funcționare ale procesului.

Concluziv, parametrii optimi ai procesului de reducere catodică a nitrobenzenului la anilină într-un reactor electrochimic necompartimentat, realizat în condițiile experimentale descrise în lucrare, sunt:

- cantitatea de electricitate trecută prin celula electrochimică - scăzută - limitată doar de posibilitățile de purificare ale produsului finit
- temperatura - 45°C
- densitate de curent - în jurul valorii de 500 A/m^2 .

Un factor de importanță deosebită în aplicarea rețelelor neurale artificiale la modelarea proceselor este cea a numărului relativ mic de date experimentale necesar pentru a obține modele ce aproximează fidel procesul. De aceea, această modalitate de modelare se pretează mai ales în cazul proceselor ce implică travaliu în culegerea datelor experimentale, fiind așa cum se demonstrează și în acest capitol net superioară metodelor clasice.

V. MODELAREA DINAMICĂ A PROCESULUI DE REȚINERE A Sr^{2+} PRIN SCHIMB IONIC

V.1 Noțiuni generale despre schimbul ionic

Schimbul ionic, ca fază apartine a procesului de adsorbție, reprezintă unul dintre cele mai complexe cazuri de transfer de masă din natură și guvernează o multitudine de manifestări ale materiei, la toate nivelele ei de structurare, inclusiv ale regnului viu.

Modelarea sa ridică probleme deosebite, datorită atât aspectului particular al numeroaselor fenomene care îl implică, dar mai ales caracterului fizico-chimic complex.

Cele mai multe modele de cinetică efectuate pentru cazul simplu cu contact unic (în regim static) sau pentru cel al stratului fix (regim dinamic) definite în literatură¹⁸⁹⁻¹⁹³ întâmpină dificultăți la verificarea practică deoarece concentrațiile componentelor din faza solidă nu se pot determina direct, astfel încât relațiile obținute, importante pentru descrierea fenomenului și unanim acceptate din punct de vedere teoretic, nu au fost, până în prezent, verificate practic.

La acestea se adaugă faptul că tendința accentuată către simplificare manifestată de numeroși autori, conduce fie la neglijarea aspectului hidrodinamic al procesului, fie invers, la reducerea către cazuri particulare de filtrare prin substraturi permeabile.

Mecanismul procesului de schimb ionic presupune trecerea unor ioni din soluție în rășină și a contraionilor din rășină în soluție, proces activ până la stabilirea echilibrului termodinamic. Ipotezele de modelare, deși simplificatoare, nu reușesc să surprindă într-o formă general valabilă schimbul ionic nici măcar pentru cazuistici restrânse.

Tematica acestui capitol face parte dintr-o preocupare mai vastă ce își propune studiul reținerii unor produși de fisiune ai uraniului $^{137}_{55}\text{Cs}$, $^{91}_{38}\text{Sr}$ care apar frecvent în circuitul secundar al reactorului nuclear CANDU sau în apele contaminate în cazul unui posibil accident nuclear¹⁹⁴⁻¹⁹⁶.

V.1.1 Echilibrul în schimbul ionic

Rășinile schimbătoare de ioni nu au afinități egale pentru toate tipurile de ioni. Variabilitatea preferințelor este adesea exprimată semicantitativ printr-o scară a afinităților pentru diferite tipuri de ioni. Cea mai utilizată exprimare este însă cea cantitativă printr-un factor de separare α_{ij} sau printr-un coeficient de selectivitate K_{ij} ambele corespunzând schimbului ionic binar.

1. Scări de selectivitate

O scară de selectivitate presupune ordinea preferințelor în cazul particular al unei anumite rășini. În tabelul 1 este redată o astfel de scară a selectivității¹⁹⁷ pentru un tip de rășină, cationii fiind listați în ordinea afinităților.

Tabelul 1 Scări de selectivitate pentru schimbătorii de ioni

Nr. crt	Rășini cationice puternic acide*		Rășini anionice puternic bazice**	
	Cation	α_i/Na^+	Anion	α_i/Cl^-
1	Ra^{2+}	13,0	CrO_4^{2-}	100,0
2	Ba^{2+}	5,8	SeO_4^{2-}	17,0
3	Pb^{2+}	5,0	SO_4^{2-}	9,1
4	Sr^{2+}	4,8	HSO_4^-	4,1
5	Cu^{2+}	2,6	NO_3^-	3,2
6	Ca^{2+}	1,9	Br^-	2,3
7	Zn^{2+}	1,8	$HAsO_4^{2-}$	1,5
8	Fe^{2+}	1,7	SeO_3^{2-}	1,3
9	Mg^{2+}	1,67	HSO_3^{3-}	1,2
10	K^+	1,67	NO_2^-	1,1
11	Mn^{2+}	1,6	Cl^-	1,0
12	NH_4^+	1,3	HCO_3^-	0,27
13	Na^+	1,0	CH_3COO^-	0,14
14	H^+	0,67	F^-	0,07

* Rășini cationice puternic acide cu grupări funcționale sulfonice grefate pe o matrice polistiren - divinilbenzen

** Rășini anionice puternic bazice cu grupări funcționale $-N(CH_3)_3$ grefate pe o matrice polistiren - divinilbenzen

De exemplu, valoarea $\alpha_{Sr/Na} = 4.8$ presupune că în cazul concentrațiilor egale în fază apoasă, stronțitul este preferat de rășină în proporție de 4.8/1.0 față de sodiu. În general, rășinile preferă ioni cu sarcina mare și cu grad mic de hidratare.

2. Coeficienți de selectivitate și factori de separare

Considerând de exemplu schimbul ionic



expresia constantei de echilibru este:

$$K = \frac{[\overline{NO_3^-}][Cl^-]}{[Cl^-][\overline{NO_3^-}]} \quad (2)$$

în care:

- bara superioară denotă faza legată de rășină
- K este constanta termodinamică de echilibru
- parantezele drepte semnifică activitatea ionică la echilibru.

Practic, se preferă exprimarea în concentrații, acestea fiind mai ușor măsurabile decât activitățile. Astfel ecuația 3 bazată pe valori ale concentrațiilor descrie prin valoarea K_{NO_3, Cl^-} schimbul ionic.

Valoarea $K_{NO_3^-/Cl^-}$ include coeficienții de activitate care sunt o funcție de tăria ionică, nefiind deci o constantă reală depinzând de tăria ionică.

$$K_{NO_3^-/Cl^-} = \frac{\overline{[NO_3^-]}[Cl^-]}{[Cl^-]\overline{[NO_3^-]}} = \frac{q_{NO_3^-} C_{Cl^-}}{q_{Cl^-} C_{NO_3^-}} \quad (3)$$

unde: - [] = concentrația, mol/l

- $q_{NO_3^-}$ = concentrația azotatului în faza rășină, echivalenți/l

- $C_{NO_3^-}$ = concentrația azotatului în fază apoasă, echivalenți/l

Factorul de separare binară $\alpha_{N/Cl}$ este utilizat în practica separării din ingineria chimică și datorită naturii simpliste și intuitive. Relația 4 descrie exprimarea matematică a acestui factor.

$$\alpha_{i,j} = \frac{\text{distributia ionului } i \text{ int re faze}}{\text{distributia ionului } j \text{ int re faze}} = \frac{y_i / x_i}{y_j / x_j} \quad (4)$$

Pentru exemplul anterior,

$$\alpha_{NO_3^-/Cl^-} = \frac{y_{NO_3^-} / x_{NO_3^-}}{y_{Cl^-} / x_{Cl^-}} = \frac{y_{NO_3^-} x_{Cl^-}}{x_{NO_3^-} y_{Cl^-}} = \frac{(q_{NO_3^-} / q)(C_{Cl^-} / C)}{(q_{Cl^-} / q)(C_{NO_3^-} / C)} = \frac{q_{NO_3^-} C_{Cl^-}}{q_{Cl^-} C_{NO_3^-}} \quad (5)$$

unde: y_{Cl^-} = fracția echivalentă a ionului Cl^- în rășină, q_{Cl^-}/q

$y_{NO_3^-}$ = fracția echivalentă a ionului NO_3^- în rășină, $q_{NO_3^-}/q$

x_{Cl^-} = fracția echivalentă a ionului Cl^- în apă, C_{Cl^-}/C

$x_{NO_3^-}$ = fracția echivalentă a ionului NO_3^- în apă, $C_{NO_3^-}/C$

$q_{NO_3^-}$ = concentrația ionului NO_3^- în rășină, echivalent/L

q_{Cl^-} = concentrația ionului Cl^- în rășină, echivalent/L

q = capacitatea totală de schimb a rășinii, echivalent/L

$C_{NO_3^-}$ = concentrația ionului NO_3^- în apă, echivalent/L

C_{Cl^-} = concentrația ionului Cl^- în apă, echivalent/L

C = concentrația ionică totală a apei, echivalent/L

Din ecuațiile (4) și (5) rezultă că pentru un schimb homovalent (monovalent/monovalent sau divalent/divalent), factorul de separare $\alpha_{i,j}$ și coeficientul de selectivitate $K_{i,j}$ sunt egale. In cazul schimbului ionic de valență inegală - schimb ionic heterovalent - factorul de separare nu este egal cu coeficientul de selectivitate. Considerând cazul schimbului ionic Sr^{2+} - R-H rezultă:



$$K_{Sr\ H} = \frac{(q_{Sr^{2+}} / C_{H^+}^2)}{(C_{Sr^{2+}} / q_{H^+}^2)} \quad (7)$$

Utilizând o combinație a relațiilor (4) și (7), se obține:

$$\alpha_{divalent/monovalent} = \alpha_{Sr/H} = K_{Sr/H} \frac{(q/C)}{(y_H/x_H)} \quad (8)$$

Semnificația acestor ecuații este că factorul de separare pentru schimbul ionic divalent/monovalent depinde de concentrația soluției C și de proporția de distribuție y_H/x_H între rășină și apă prin constanta q .

Cu cât concentrația soluției este mai mare, și cu cât factorul de separare divalent/monovalent este mai mic, selectivitatea tinde să favorizeze ionul monovalent pe măsura creșterii tăriei ionice.

V.1.2 Curbe izoterme

Valorile α_{ij} și K_{ij} pot fi determinate reprezentând curba de echilibru la temperatură constantă, adică concentrația adsorbită în faza solidă funcție de concentrația în fază apoasă. Din aspectul curbei se pot trage concluzii privitor la preferințele masei ionice față de specia reprezentată. Astfel:

- o izotermă favorabilă - convexă față de abscisă - presupune preferința rășinii pentru ionul respectiv din amestecul binar în detrimentul celuilalt (fig. 1 a)
- o izotermă nefavorabilă - concavă față de abscisă - presupune lipsa de preferință a rășinii pentru ionul respectiv din amestecul binar (fig. 1 b)

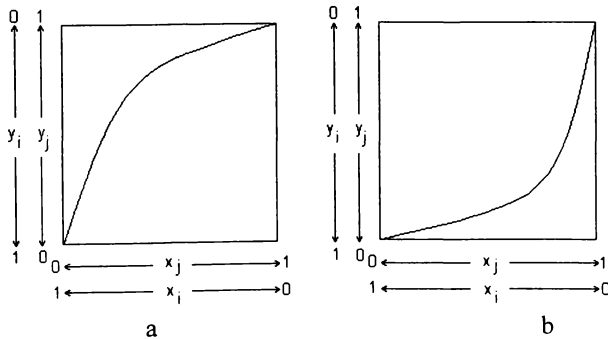


Fig. 1 Izoterme pentru un amestec binar ij
 a - izotermă favorabilă ionului j
 b - izotermă nefavorabilă ionului j

V.1.3 Schimbul ionic și cinetica de adsorbție

Cinetica procesului se realizează din punct de vedere intuitiv în patru etape distincte:

1. difuzia ionilor și a contraionilor prin filmul de lichid de la suprafața granulei de rășină;
2. impregnarea particulei de rășină cu soluția;
3. difuzia ionilor și a contraionilor prin porii particulei;
4. schimbul ionic propriu-zis la nivelul grupelor active ale rășinii.

În ansamblu, viteza procesului este controlată de etapa cea mai lentă de transfer. Date de literatură^{198,199} indică faptul că viteza procesului de schimb ionic este determinată în mod decisiv de difuzie și poate fi descrisă pe baza legii a II-a a lui Fick:

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \quad (9)$$

Participarea ionilor la schimb se supune acestei legi atât în ceea ce privește granula de rășină, cât și filmul de lichid.

În cazul schimbului ionic nestaționar, se poate accepta¹⁸⁹ a priori că deplasarea fazei lichide are loc simultan în tot spațiul astfel încât, la infinit, există un curent uniform, orientat după direcție axial simetrică (fig.2) față de elipsoidul de rotație S cu rază r , care împarte spațiul coloanei în domeniul interior D_1 și domeniul exterior D_2 .

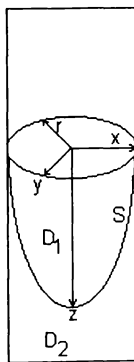


Fig. 2 Avansarea frontului pe coloană

Factorii determinați ai cineticii de schimb ionic sunt numeroși, dar cei mai importanți pot fi următorii: concentrația soluției, regimul hidrodinamic, dimensiunile particulelor de rășină, valoarea coeficientului de difuzie în particulă. Cinetica este determinată de caracteristicile ionilor participanți la schimb și de structura rășinii, prezentând valori mari pentru ioni puțin voluminoși și rășini cu porozitate ridicată.

Coefficienții de difuzie în rășină a ionilor monovalenți reprezintă $1/3$ până la $1/10$ din valorile de difuzie în apă, a ionilor bivalenți $1/5$ până la $1/100$, iar a celor trivalenți $1/10$ până la $1/100$.

Din analiza factorilor prezentați anterior se poate arăta că factorul cinetic al schimbului ionic poate fi controlat prin difuzia în filmul de lichid, în cazul ionilor puțin voluminoși (H^+ , Na^+) sau prin difuzia prin rășină, în cazul ionilor polivalenți. Pentru soluții care nu respectă condiția de diluție infinită, cinetica este în general controlată de etapa difuziunii în rășină, iar pentru concentrații scăzute factorul cinetic este tributari difuziunii în film²⁰⁰.

V.2 Determinări experimentale

Pentru cazul particular al reținerii ionilor de stronțiu în regim dinamic, pe coloană, instalația experimentală este redată în figura 3.

Practic, pentru fiecare caz în parte s-a lucrat la câte trei concentrații astfel alese încât scăderea ionică să apară cât mai repede și să reducă experimentul la timpi rezonabili din punct de vedere experimental. Coloanele au fost regenerate de fiecare dată în ciclul H-cationic cu o soluție 4% HCl, iar spălarea s-a efectuat până la reacția negativă a ionului Cl^- față de $AgNO_3$.

Instalația experimentală este formată din trei coloane identice ca dimensiuni, în care straturile de schimbători de ioni au fost dispuse prin multiplicarea lungimii primului (tehnica coloanelor multiple): 2:1 în cea de-a doua coloană, 3:1 în cea de-a treia coloană (1 fiind, în mod evident, lungimea coloanei de schimbător de ion din prima coloană). Masa ionică utilizată a fost un cationit puternic acid de tip PUROLITE C 100, destinată preferențial aplicațiilor speciale²⁰¹.

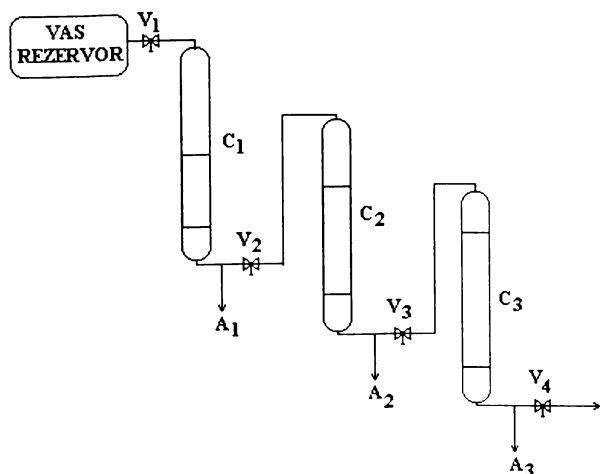


Figura 3. Instalația experimentală

Semnificațiile notațiilor din figura 3 sunt:

- C_1, C_2, C_3 - coloane cu schimbător de ioni
- V_1, V_2, V_3 - robinete de reglare a debitului
- A_1, A_2, A_3 - puncte de colectare

Concentrațiile de alimentare s-au selectat din condiții experimentale legate de:

1. Corelarea cantității de schimbător de ioni cu debitul de alimentare

În masa schimbătorului de ioni este posibilă apariția de canale preferențiale ce conduc la zone în care debitul de circulație a soluției este de ordin superior celui recomandat de adsorbție. Acest fapt induce o scăpare din coloană a unei concentrații de valoare mare datorită nestabilirii echilibrului termodinamic de schimb ionic.

Probabilitatea perturbării experimentului prin apariția de canale preferențiale este invers proporțională cu lungimea coloanei și direct proporțională cu debitul de alimentare și cu concentrația soluției. Deci, volumul de schimbător de ioni din prima coloană este determinant în ceea ce privește atât debitul cât și concentrația soluției de alimentare.

2. Considerente practice de realizare a experimentului.

Această problemă presupune încadrarea experimentului în anumite limite care vizează:

-o concentrație a efluentului care să nu vizeze limita de detectabilitate în analiza concentrației acestuia. Având în vedere precizia metodei de analiză este necesar ca o posibilă eroare să nu fie de un ordin apropiat variației dinamice a concentrației efluentului.

-obținerea de suficiente date experimentale pentru a caracteriza evoluția concentrației efluentului în timp

Astfel, pentru un debit de alimentare de 2l/h, un volum de 30 cm³ schimbător de ioni în prima coloană, s-au efectuat determinări cu soluții pornind de la o concentrație de 122·10⁻³N și continuând cu cele obținute prin înjumătățirea succesivă a acesteia până la obținerea valorii de concentrație care se încadrează în limitele experimentale propuse.

Primul nivel de concentrație acceptat a fost de 15,25·10⁻³N, stabilindu-se o plajă de variație a concentrației soluției de alimentare cuprinsă între 5 - 15,25·10⁻³N.

Volumul de schimbători de ioni pe cele trei coloane s-a stabilit într-un raport de 1 : 2 : 3. Ca și concentrații ale influentului de alimentare s-au ales : 5, 10, respectiv 15,25 mval/l. Astfel, s-a considerat suficient un număr de nouă seturi de date experimentale ca bază de date pentru simularea dinamică a reținerii Sr^{2+} pe schimbători de ioni.

Analiza concentrației Sr^{2+} s-a efectuat prin titrare complexometrică²⁰¹.

În tabelele 2, 3 și 4, în coloanele 1,2,3 sunt redate condițiile experimentale, iar în coloanele 4 și 5 valorile experimentale și respectiv calculate pe baza modelului neural pentru concentrațiile inițiale de 5·10⁻³, 10·10⁻³ și respectiv 15,25·10⁻³N. Notățiile din tabel corespund:

- C_i - concentrația soluției la alimentarea coloanei
- V_r - volumul redus al masei de schimbători de ioni
- C_{exp} - concentrația efluentului pentru colana corespunzătoare valorii V_r
- C_{cN} - concentrația de predicție (utilizând modelul neural) a efluentului

Tabelul 2. Condiții experimentale; valori experimentale și calculate pentru concentrația influentului coloanei 1 de $5 \cdot 10^{-3}N$

Nr. crt.	Timp [h]	$C_i \cdot 10^3 N$	V_r	$C_{exp} \cdot 10^3 N$	$C_{cN} \cdot 10^3 N$
0	1	2	3	4	5
1	1,5	5	0,167	0,31	0,13
2	2	5	0,167	0,52	0,34
3	2,5	5	0,167	0,9	0,77
4	2,85	5	0,167	1,2	1,26
5	3,15	5	0,167	1,6	1,74
6	3,5	5	0,167	2,2	2,32
7	3,85	5	0,167	2,91	2,89
8	4,1	5	0,167	3,4	3,31
9	4,25	5	0,167	3,6	3,56
10	4,5	5	0,167	4,2	4,01
11	4,65	5	0,167	4,6	4,28
12	4,9	5	0,167	4,8	4,76
13	5,2	5	0,167	5	5,00
14	8,5	5	0,5	0,2	0,23
15	9	5	0,5	0,5	0,45
16	9,5	5	0,5	0,9	0,79
17	10	5	0,5	1,3	1,28
18	10,5	5	0,5	1,8	1,95
19	10,85	5	0,5	2,5	2,53
20	11	5	0,5	2,9	2,8
21	11,25	5	0,5	3,4	3,28
22	11,5	5	0,5	3,8	3,79
23	11,75	5	0,5	4,5	4,33
24	12,25	5	0,5	5	4,98
25	16,5	5	1	0,1	0,29
26	17,5	5	1	0,7	0,83
27	18,5	5	1	1,8	1,83
28	19	5	1	2,2	2,53
29	19,25	5	1	3	2,92
30	19,5	5	1	3,6	3,35
31	19,75	5	1	4	3,8
32	20,15	5	1	4,6	4,55
33	20,25	5	1	4,8	4,75
34	20,5	5	1	5	4,97

Tabelul 3. Condiții experimentale; valori experimentale și calculate pentru concentrația influentului coloanei 1 de $10 \cdot 10^{-3}N$

Nr. crt.	Timp [h]	$C_i \cdot 10^3 N$	V_r	$C_{exp} \cdot 10^3 N$	$C_{cN} \cdot 10^3 N$
0	1	2	3	4	5
1.	0,33	10	0,167	0,53	0,72
2.	0,5	10	0,167	0,9	0,98
3.	0,67	10	0,167	1,38	1,43
4.	0,83	10	0,167	2,1	2,15
5.	1	10	0,167	3,78	3,48
6.	1,08	10	0,167	4,82	4,37
7.	1,17	10	0,167	5,92	5,6
8.	1,25	10	0,167	7,14	6,85
9.	1,33	10	0,167	8,22	8,19
10.	1,42	10	0,167	9,54	9,66
11.	1,5	10	0,167	9,76	9,98
12.	1,55	10	0,167	9,99	10,0
13.	3,67	10	0,5	0,3	0,31
14.	3,92	10	0,5	0,45	0,74
15.	4,17	10	0,5	0,59	1,18
16.	4,67	10	0,5	1,12	1,33
17.	4,92	10	0,5	1,4	1,62
18.	5,17	10	0,5	2,5	2,49
19.	5,42	10	0,5	4	4,06
20.	5,58	10	0,5	5,4	5,34
21.	5,67	10	0,5	6,2	6,08
22.	5,75	10	0,5	6,8	6,73
23.	5,83	10	0,5	7,6	7,34
24.	5,92	10	0,5	8	7,99
25.	6	10	0,5	8,86	8,51
26.	6,08	10	0,5	9,27	8,99
27.	6,17	10	0,5	9,5	9,47
28.	6,25	10	0,5	9,7	9,85
29.	6,33	10	0,5	9,92	9,93
30.	6,42	10	0,5	10	10,00
31.	10,5	10	1	0,6	0,30
32.	11	10	1	2	2,02
33.	11,25	10	1	3,4	3,40
34.	11,42	10	1	4,2	4,38
35.	11,58	10	1	5,2	5,26
36.	11,75	10	1	6	6,13
37.	11,92	10	1	7,2	6,93
38.	12,25	10	1	8,4	8,27
39.	12,4	10	1	9	8,79
40.	12,6	10	1	9,6	9,43
41.	12,8	10	1	10	10,00

Tabelul 4. Condiții experimentale; valori experimentale și calculate pentru concentrația influentului coloanei I de $15,25 \cdot 10^{-3}$ N

Nr. crt.	Timp [h]	$C_i \cdot 10^3$ N	V_r	$C_{exp} \cdot 10^3$ N	$C_{en} \cdot 10^3$ N
0	1	2	3	4	5
1.	0,42	15,25	0,167	0,8	0,67
2.	0,5	15,25	0,167	0,9	0,77
3.	0,58	15,25	0,167	1	0,9
4.	0,75	15,25	0,167	1,5	1,34
5.	0,92	15,25	0,167	2	2,16
6.	1,08	15,25	0,167	3,5	3,54
7.	1,17	15,25	0,167	4,5	4,69
8.	1,25	15,25	0,167	6	5,95
9.	1,33	15,25	0,167	7,5	7,41
10.	1,42	15,25	0,167	9	9,14
11.	1,5	15,25	0,167	10,75	10,61
12.	1,58	15,25	0,167	12,75	11,89
13.	1,67	15,25	0,167	13,75	13,02
14.	1,94	15,25	0,167	15,25	15,23
15.	3,25	15,25	0,5	1	1,03
16.	3,5	15,25	0,5	2	1,92
17.	3,75	15,25	0,5	4	3,9
18.	3,92	15,25	0,5	6	6,29
19.	4,08	15,25	0,5	9	9,13
20.	4,17	15,25	0,5	11	10,68
21.	4,25	15,25	0,5	12,75	11,88
22.	4,33	15,25	0,5	13	12,85
23.	4,5	15,25	0,5	14,6	14,16
24.	4,6	15,25	0,5	15,05	14,88
25.	4,75	15,25	0,5	15,25	15,20
26.	5,58	15,25	1	0,5	0,49
27.	6	15,25	1	0,75	0,99
28.	6,5	15,25	1	3	2,62
29.	6,67	15,25	1	4,5	3,87
30.	6,83	15,25	1	5,3	5,63
31.	7	15,25	1	7,5	8,13
32.	7,17	15,25	1	10,75	10,81
33.	7,4	15,25	1	14	13,46
34.	7,6	15,25	1	14,86	14,67
35.	7,75	15,25	1	15,25	15,21

V.3 Modelarea procesului

Inițial s-a testat o regresie multiliniară pentru a găsi o ecuație de regresie care să exprime variația concentrației la ieșirea din coloană în funcție de timp cele trei variabile independente - t , concentrația inițială a soluției C_i , și volumul de schimbători de ioni parcurs de soluție, exprimat prin volumul redus descris de relația:

$$V_r = V_t/V_u \quad (10)$$

în care:

- V_r - volumul redus
- V_t - volumul total de schimbători de ioni din cele trei coloane
- V_u - volumul de schimbători de ioni parcurs de soluție

Astfel, valorile volumului redus V_r de 0,166, 0,5 și 1 se referă la coloanele I, II și respectiv III.

Ecuația obținută prin regresie multiliniară este dată de relația 11: (Valorile concentrațiilor au fost exprimate în $N \cdot 10^3$)

$$C_{RML} = -3,194(\pm 1,617) + 0,490(\pm 0,181) \cdot t + 0,799(\pm 0,144) \cdot C_i - 5,297(\pm 2,506) \cdot V_r \quad (11)$$

Pentru un coeficient de corelare de $R = 0,520$, un coeficient de determinare $R^2 = 0,270$, o deviație standard $DS = 3,70$ și o alură a reprezentării concentrației calculate prin regresie multiliniară C_{RML} funcție de concentrația experimentală C_{exp} de genul celei din figura 4, se constată că regresia multiliniară prin structura sa este total necorespunzătoare.

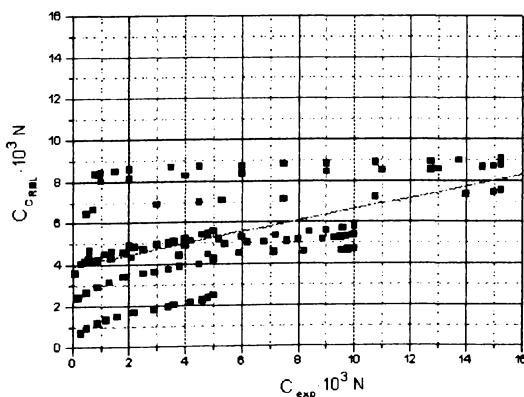


Fig. 4 Variația concentrației calculate prin regresie multiliniară C_{RML} funcție de concentrația experimentală C_{exp}

Pentru testarea modelului neural, s-au utilizat aceleași seturi de date experimentale din tabelele 2,3 și 4 ca seturi de tipare de antrenare.

S-a utilizat o rețea neurală artificială cu propagare înapoi (Backpropagation Neural Network) în scopul validării datelor experimentale¹¹² și a formării ponderilor ce corelează variabilele de intrare cu cele de ieșire ale procesului studiat⁹⁸. Sursa programului a fost aceeași ca în capitolele II și IV, fiind proiectată pentru cele 110 de tipare experimentale și având următoarele caracteristici :

- vector de intrare tridimensional: C_i , timp, V_r
- vector de ieșire unidimensional C_{cN}
- vectorul de pe stratul ascuns constituit din 24 noduri
- nu s-au utilizat unități BIAS¹¹⁰
- funcție de activare sigmoidă cu $\beta = 3$
- rata de învățare $\eta = (0,5 \dots 0,01)$
- termenul *momentum* = 0,2
- număr de epoci rulate = 5000

În figura 5 este redată variația erorii totale a rețelei cu numărul de epoci parcurse.

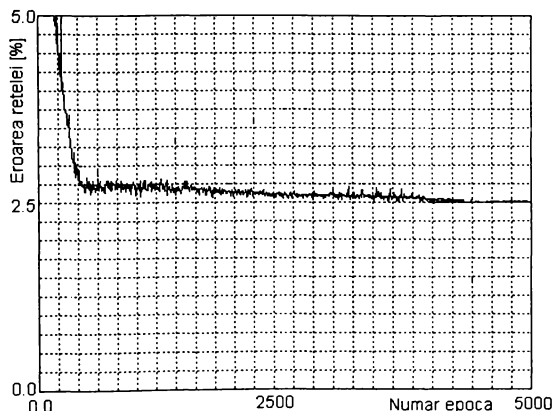


Fig. 5. Variația erorii totale a rețelei cu numărul de epoci parcurse

În urma rulării programului, valorile obținute ale concentrațiilor sunt trecute în tabelele 2,3 și 4 coloana 5.

Valoarea finală a erorii totale a rețelei de 2,49% se datorează faptului că deși de exemplu pentru primele două valori $C_{exp} = 0,31$ și $C_{cN} = 0,13$ eroarea este de 280%, semnificația experimentală a acestor diferențe nu este atât de elocventă având în vedere că ele reprezintă 6,2% și respectiv 2,6% din valoarea celei mai mici concentrații de alimentare.

Această diferență între valoarea experimentală și cea de predicție a modelului neural apare în mod frecvent la începutul unui ciclu experimental, fiind favorizată de doi factori:

1. concentrația mare a influentului
2. ordinea de dispunere a coloanelor cu schimbători de ioni - diferența este mai mare pentru prima colană, mai mică la cea de-a doua și neglijabilă la cea de-a treia.

Explicația rezultă din selecția condițiilor experimentale în privința volumului de schimbători de ioni, a concentrației de alimentare și a debitului de influent.

Pentru prima coloană ce conține 16,67% din volumul total de schimbători de ioni, echilibrul termodinamic de schimb ionic se instalează cu întârziere în comparație cu evoluția hidrodinamică a fluidului ce are implicații asupra dinamicii concentrației Sr^{2+} în influentul coloanei. Pe lângă evoluția mai rapidă a factorilor cinetici în comparație cu cei termodinamici, un alt factor perturbator este apariția unor posibile canale preferențiale în masa de schimbători de ioni.

Acest fapt este sesizat mai pregnant în valoarea concentrației efluentului în cazul în care lungimea coloanei este mai mică iar concentrația influentului este mai mare. Aceste două condiții sunt întrunite în special de prima coloană ($Vr=0,167$), cu predilecție în cazul alimentării cu soluția de concentrație maximă ($15,25 \cdot 10^{-3} N$). La celelalte coloane, concentrația influentului fiind determinată de concentrația efluentului coloanei din amonte, echilibrul termodinamic se stabilește concomitent cu apariția în influent a Sr^{2+} , inițial în concentrație redusă, aceasta crescând treptat pe măsura epuizării capacității de schimb a schimbătorilor de ioni din coloana precedentă.

Pentru o argumentare mai elocventă a capacității de corelare a modelului neural, s-a efectuat o regresie liniară între valorile obținute cu rețeaua neurală C_{CN} și cele experimentale C_{exp} . În figura 6 este redată variația liniară a valorilor concentrației calculate cu modelul neural C_{CN} funcție de valorile experimentale C_{exp} .

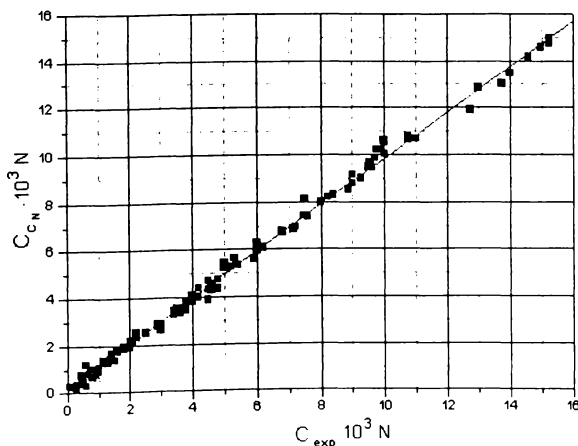


Fig. 6 Variația liniară a valorilor concentrației calculate cu modelul neural C_{CN} funcție de valorile experimentale C_{exp} .

Ecuția de regresie liniară pentru cele două seturi de determinare este dată de relația: (Valorile concentrațiilor au fost exprimate în $N \cdot 10^3$)

$$C_{cN}=0,083(\pm 0,039)+0,976(\pm 0,006) \cdot C_{exp} \quad (12)$$

iar parametrii de corelare sunt:

Volumul eșantionului :	N =110
Numarul gradelor de libertate :	NL=108
Indicele de corelare	R = 0.998
Coefficientul de determinare	R ² = 0.996
Deviatia standard	DS = 0.257
Testul Fischer	F = 28698

Valorile indicilor statistici ai corelării indică o capacitate ridicată de corelare a modelului neural.

În figurile 7 -15 sunt reprezentate grafic - punctat - valorile experimentale ale concentrațiilor la ieșirea din cele trei coloane comparativ cu valorile estimate - linie continuă - de modelul neural.

Pentru a obține o variație continuă a concentrației în timp pentru modelul neural, în urma antrenării rețelei cu seturile de date experimentale s-au generat tipare care să acopere dens în domeniul timp întreg domeniul de variație. Pe baza ponderilor create în etapa de învățare a rețelei s-au calculat valorile variabilelor de ieșire - valoarea concentrației Sr^{2+} la ieșirea din coloană.

Se observă o suprapunere fidelă a valorilor experimentale cu cele calculate pe baza modelului neural.

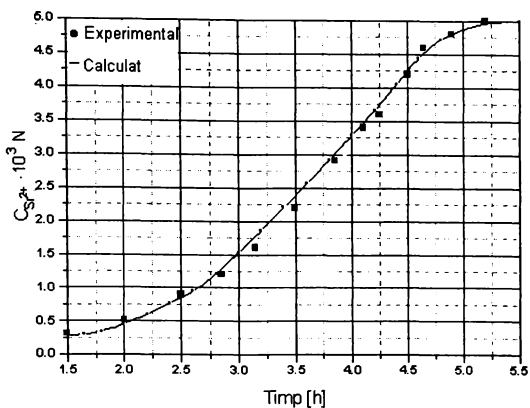


Fig. 7 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 1 pentru o concentrație a influentului coloanei 1 de $5 \cdot 10^{-3}N$

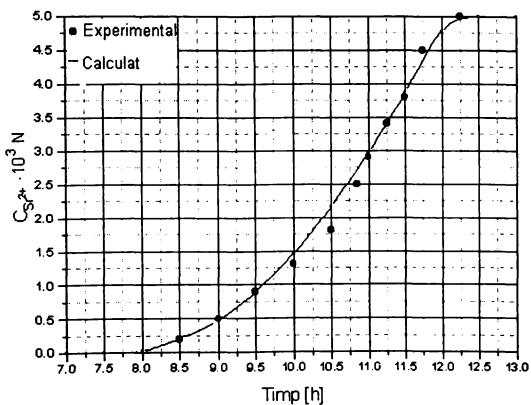


Fig. 8 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 2 pentru o concentrație a influentului coloanei 1 de $5 \cdot 10^{-3}N$

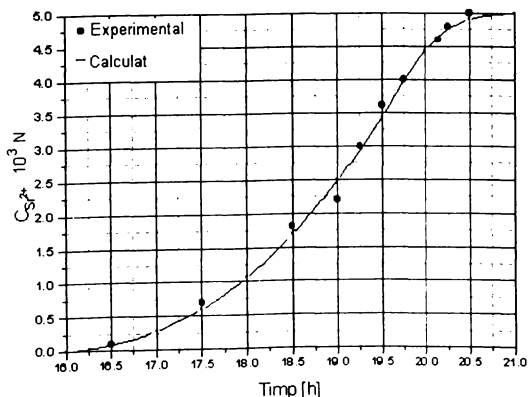


Fig. 9 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 3 pentru o concentrație a influentului coloanei 1 de $5 \cdot 10^{-3}N$

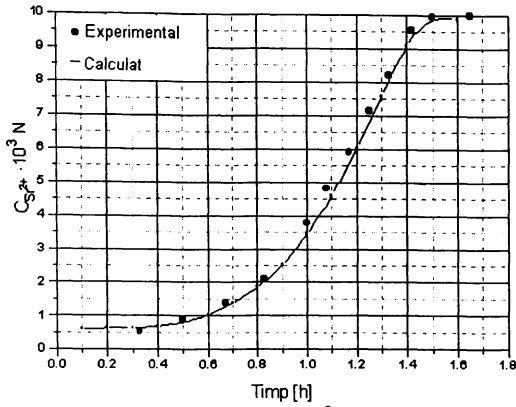


Fig. 10 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 1 pentru o concentrație a influentului coloanei 1 de $10 \cdot 10^{-3}N$

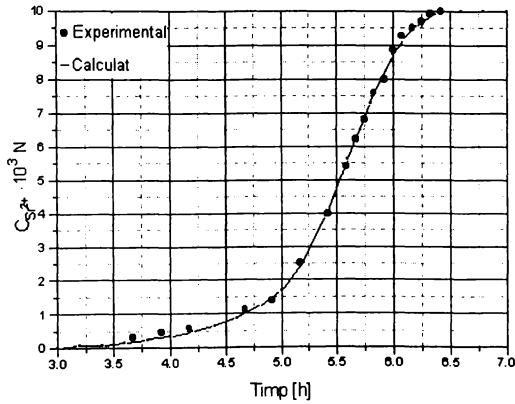


Fig. 11 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 2 pentru o concentrație a influentului coloanei 1 de $10 \cdot 10^{-3}N$

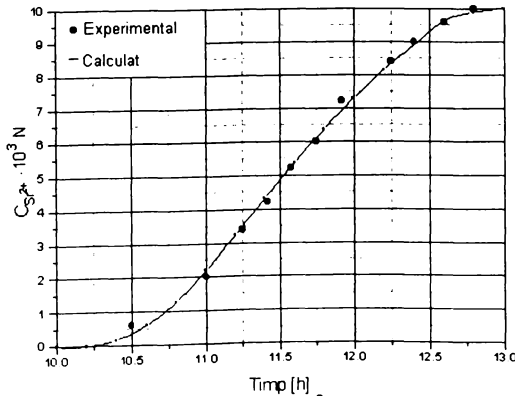


Fig. 12 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 3 pentru o concentrație a influentului coloanei 1 de $10 \cdot 10^{-3}N$

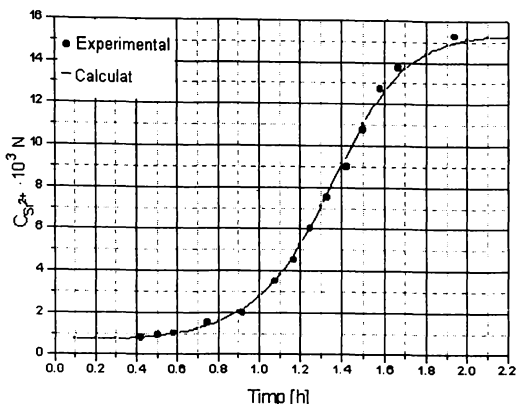


Fig. 13 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 1 pentru o concentrație a influentului coloanei 1 de $15,25 \cdot 10^{-3} N$

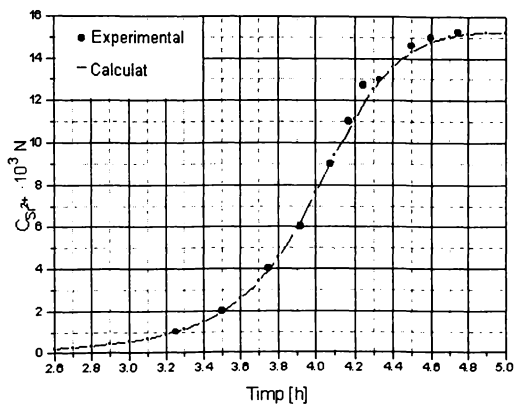


Fig. 14 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 2 pentru o concentrație a influentului coloanei 1 de $15,25 \cdot 10^{-3} N$

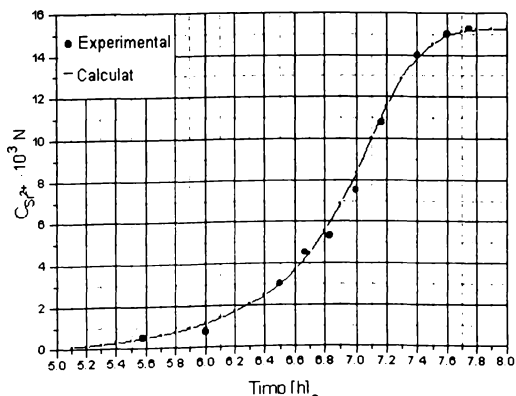


Fig. 15 Variația în timp a concentrației Sr^{2+} în efluentul coloanei 3 pentru o concentrație a influentului coloanei 1 de $15,25 \cdot 10^{-3} N$

Diferența ce apare între concentrația calculată și experimentală la coloana 1 la începutul experimentului și apoi, mai atenuat, la coloanele 2 și 3 se datorează factorilor hidrodinamici, care evoluează în mod evident mai rapid decât se stabilesc echilibrele termodinamice, la nivelul concentrației de alimentare.

Se observă că variația în timp a concentrației este în bună concordanță atât cu grosimea stratului cât și cu valorile caracteristicilor fizice ale straturilor.

V.4 Validarea modelului

Pentru validarea modelului s-a utilizat tehnica validării încrucișate descrisă anterior în capitolele II, III și IV. Pentru o divizare arbitrară a setului de date s-a adoptat drept criteriu de selecție cel al parității numărului curent pentru cele 110 tipare. Pentru validare s-a utilizat o rețea neurală care a fost antrenată succesiv cu cele două subseturi de date în vederea utilizării ulterioare pentru predicția valorilor de ieșire ale contrasetului corespunzător.

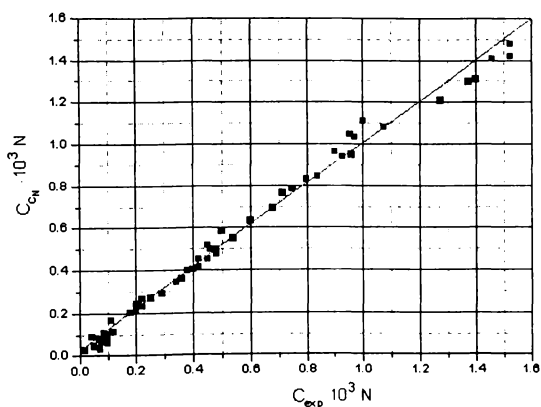


Fig. 16 Variația valorilor calculate C_{CN} funcție de cele experimentale pentru subsetul de date impar C_{exp}

Relația 13 reprezintă ecuația de corelare liniară între valorile experimentale ale concentrației C_{exp} și cele calculate pe baza modelului neural C_{CN} pentru subsetul de date impar. (Valorile concentrațiilor au fost exprimate în $N \cdot 10^3$)

$$C_{CN} = 0,028(\pm 0,012) + 0,981(\pm 0,017) \cdot C_{exp} \quad (13)$$

Parametrii de corelare sunt:

Volumul eșantionului	$N = 55$
Numarul de grade de libertate :	$NL = 53$
Coefficientul de corelare	$R = 0,992$
Coefficientul de determinare	$R^2 = 0,984$
Deviatia standard	$SD = 0,055$
Testul Fisher	$F = 3166$

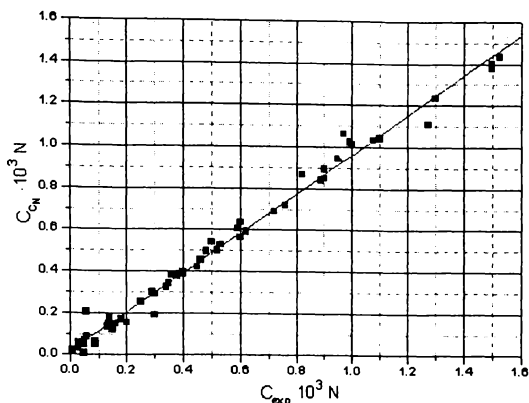


Fig. 17 Variația valorilor calculate funcție de cele experimentale pentru setul de date par

Relația 14 reprezintă ecuația de corelare liniară între valorile concentrațiilor calculate pe baza modelului neural C_{CN} valorile experimentale C_{exp} pentru subsetul de date par. (Valorile concentrațiilor au fost exprimate în $N \cdot 10^3$)

$$C_{CN} = 0,015(\pm 0,010) + 0,948(\pm 0,015) \cdot C_{exp} \quad (14)$$

Parametrii de corelare sunt:

Volumul eșantionului	:	$N = 55$
Numărul de grade de libertate	:	$NL = 53$
Coefficientul de corelare		$R = 0,993$
Coefficientul de determinare		$R^2 = 0,987$
Deviația standard		$SD = 0,047$
Testul Fisher		$F = 3886$

Din valorile parametrilor de corelare liniară a valorilor calculate față de cele experimentale în cazul ambelor subseturi de date, se apreciază o foarte bună stabilitate a modelului neural utilizat.

V.5 Testarea capacității de predicție a modelului

Până în momentul de față nu se cunosc algoritmi de modelare a procesului de schimb ionic cu proprietăți predictive.

Pentru testarea capacității de predicție a modelului, s-a utilizat o rețea neurală antrenată cu o bază de date derivată din datele experimentale. Având în vedere existența a 9 seturi de date experimentale (3 concentrații ale influentului \times 3 coloane), pentru fiecare testare a capacității de predicție, s-au utilizat pentru antrenarea unei rețele neurale 8 dintre acestea. Valorile ponderilor create în etapa de antrenare au fost utilizate pentru predicția setului de date omis în etapa de învățare.

S-a testat astfel, capacitatea de predicție internă pentru seturile de date descrise în tabelul 5. Predicția s-a efectuat pentru o variație mediană a valorii

concentrației influentului și pentru variația pe întreg domeniul a masei de schimbători de ioni parcurse.

Tabelul 5. Caracteristici ale seturilor de date utilizate la verificarea predicției modelului neural

Nr. set	C_i [$N \cdot 10^3$]	Coloana
1	5,00	2
2	10,00	2
3	15,25	2

În figurile 18, 19 și 20 este reprezentată dependența liniară dintre valorile concentrațiilor de predicție ale modelului neural C_{CN} funcție de valorile experimentale C_{exp} .

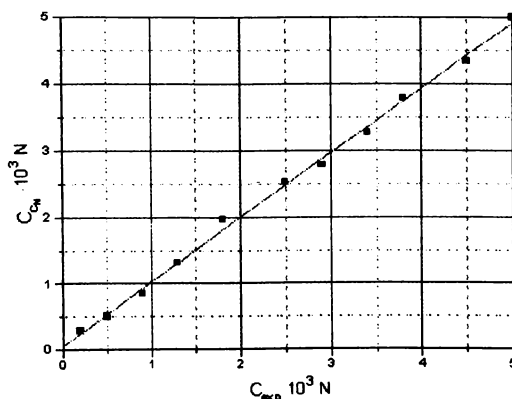


Fig 18 Variația valorilor de predicție ale modelului neural C_{CN} funcție de valorile experimentale C_{exp} pentru setul 1 de date

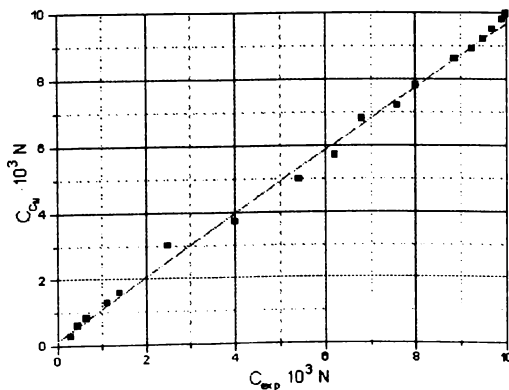


Fig 19 Variația valorilor de predicție ale modelului neural C_{CN} funcție de valorile experimentale C_{exp} pentru setul 2 de date

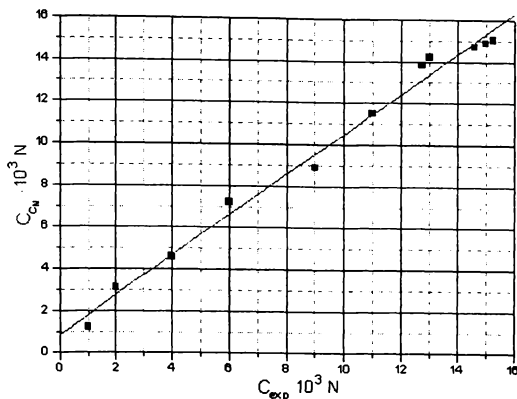


Fig 20 Variația valorilor prezise de modelul neural C_{cN} funcție de valorile experimentale C_{exp} pentru setul 3 de date

Relațiile 15, 16 și 17 redau ecuațiile de corelare liniară între valorile calculate pe baza modelului neural C_{cN} și valorile experimentale C_{exp} pentru seturile de date 1, 2 și respectiv 3. (Valorile concentrațiilor au fost exprimate în $N \cdot 10^3$)

$$C_{cN} = 0,057(\pm 0,049) + 0,969(\pm 0,017) \cdot C_{exp} \quad (15)$$

$$C_{cN} = 0,151(\pm 0,118) + 0,952(\pm 0,018) \cdot C_{exp} \quad (16)$$

$$C_{cN} = 0,825(\pm 0,358) + 0,966(\pm 0,033) \cdot C_{exp} \quad (17)$$

În tabelul 6 sunt dați parametrii de corelare pentru regresii liniare dintre valorile calculate pe baza modelului neural C_{cN} și valorile experimentale C_{exp} pentru seturile de date 1, 2 și respectiv 3.

Tabel 6 Parametrii de corelare liniară dintre valorile de predicție și cele experimentale pentru cele trei seturi de date

Parametrii de corelare	Set 1	Set 2	Set 3
Volumul eșantionului N	11	18	11
Numărul de grade de libertate NL	9	16	9
Coeficientul de corelare R	0,999	0,997	0,995
Coeficientul de determinare R^2	0,997	0,995	0,989
Deviația standard SD	0,088	0,270	0,565
Testul Fisher F	3271	2914	835.2

Din valorile parametrilor de corelare liniară dintre valorile de predicție și cele experimentale pentru cele trei seuri de date, rezultă că modelul neural prezintă capacitatea de predicție apreciabile.

V.5 Concluzii

Rezultatele obținute prin modelarea cu ajutorul rețelelor neurale a schimbului ionic sunt într-o foarte bună concordanță cu fenomenul real și prezintă în comparație cu rezultatele obținute prin modelarea fenomenului utilizând alte metode, o acuratețe mai ridicată.

Deși modelarea neurală face parte din categoria modelelor implicite, prin modalitatea de abordare globală a fenomenului de schimb ionic nu neglijează nici aspectul hidrodinamic și nici partea de chimism ce caracterizează acest proces.

De reținut faptul că este pentru prima dată când o modelare a unui proces de schimb ionic este utilizată și testată cu rezultate bune în scop predictiv.

Simularea și modelarea neurală a procesului de schimb ionic prezintă facilitățile legate de simularea computerizată ce presupune variația simplă a oricărei variabile a procesului simulat, investigațiile influenței acestora asupra procesului realizându-se fără eforturile materiale ce caracterizează celelalte posibilități de modelare.

Tematica abordată în acest capitol rămâne deschisă, soluțiile pure care cuprind un singur cation urmând a fi înlocuite inițial de soluții pure care conțin amestecuri de cationi în vederea studiilor de selectivitate și ulterior, de soluții luate din proces pentru studii practice de optimizare a procesului industrial.

Modelarea în general și în domeniul chimiei și al ingineriei chimice în particular este o noțiune vehiculată frecvent datorită importanței cunoașterii dependențelor între variabilele de intrare și ieșire ale unui sistem cât și datorită aplicabilității practice în sensul determinării unor condiții optime de procesare, de abordare optimizată a unor direcții de cercetare sau de estimare a evoluției sistemului în condiții limită.

O definiție a noțiunii de *model* trebuie să fie pe de o parte suficient de generală pentru a cuprinde tipologia actuală de modele și suficient de concretă pentru a cuprinde esența semanticii acestei noțiuni. Astfel²⁰³, un model este un sistem material sau abstract care, pe baza analogiei reale sau formale, poate fi pus în corespondență cu un sistem de referință, în scopul utilizării indirecte la studiul proprietăților acestuia.

Pentru a încadra cazul particular al modelării bazate pe rețele neurale în contextul general al modelării, se va face o tentativă de clasificare succintă a acestora pe baza unor criterii discriminatorii.

Criteriul de generalitate a modelului

- modele globale - descriu sistemele chimice urmărind evoluția unor mărimi fundamentale - energie, impuls, masa, sarcina - în timp. Posibilitățile de generalizare și predicție a acestor modele sunt remarcabile, însă aplicabilitatea lor este limitată de dificultățile matematice care intervin în cazul sistemelor complexe.

- modele locale - descriu semiempiric prin relații simple dependente între variabilele de intrare și ieșire ale sistemului. Sunt de fapt transpuneri matematice ale unor observații experimentale directe în scopul determinării unor condiții optime de procesare sau de abordare optimizată a unor direcții de cercetare.

Criteriul de formă a modelului

- modele iconice - sunt bazate pe similaritatea conceptuală sau formală a modelului cu cea a sistemului modelat. Deși între model și sistemul modelat nu există practic asemănări structurale, între formele de manifestare ale acestora există paralelisme virtuale care pot fi speculate descriptiv.

- modele analogice - se bazează pe similarități funcționale între sisteme care au expresii identice ale legității de evoluție chiar dacă parametrii celor două sisteme sunt de o natură complet diferită²⁰⁴⁻²⁰⁸. Justificarea utilizării acestor tipuri de modele este dată sau de costul ridicat al studiului sau de dificultatea de realizare a condițiilor experimentale pentru procesul de modelat.

- modelele simbolice - neglijează orice analogie cu sistemul modelat, excelează prin abstractism și au ca scop descrierea sistemului prin relații alcătuite din termeni cu semnificații acordate prin convenție.

Criteriul modalității de realizare a modelului

- modele analitice - în care interdependența dintre variabilele sistemului studiat este descrisă riguros prin relații matematice^{209,210}.

- modele de simulare - în care alura dependenței dintre variabilele de intrare și ieșire este descrisă prin relații matematice a căror corespondență cu sistemul real se realizează prin coeficienți experimentali.

Criteriul caracterului aleator al variabilelor

- modele deterministe - în care variabilele care descriu sistemul au un caracter nealeator

- modele stohastice - în care variabilele care descriu sistemul au un caracter aleator

Criteriul dependenței de timp al variabilelor modelului -

- modele statice - în care variabila **timp** nu este inclusă în model

- modele dinamice - în care variabila **timp** este inclusă în model

Rețelele neurale, deși intersectează parțial câteva din aceste categorii de modele enumerate, prin particularitățile sale, se constituie într-o categorie aparte. Astfel, în modelarea bazată pe rețele neurale nu se cunosc apriori nici variabilele de sistem cu influență asupra evoluției acestuia și nici forma matematică sau alura dependenței dintre variabilele de intrare și ieșire.

Mai mult, modelarea neurală este o modelare implicită, adică, în urma efectuării ei nu se obțin relații matematice ce caracterizează influența variabilelor de intrare asupra evoluției sistemului.

O caracteristică esențială a modelelor neurale este autodeterminarea, adică realizarea ponderilor de conexiune între variabilele de ieșire și cele de intrare printr-un mecanism de autocorecție controlată pe baza parametrilor de rețea și algoritmilor de corecție a ponderilor.

Avantajul oferit de modelarea neurală este acela al evaluării fidele a influenței variabilelor de intrare asupra sistemului, indiferent de complexitatea acesteia și de posibilele interferențe care pot apare între variabilele sistemului.

Mai mult decât atât, în măsura în care mecanismul de dependență a sistemului de variabilele de intrare rămâne același și pentru domenii neacoperite experimental, (din cauza costului ridicat al experimentelor sau din cauza unor estimări eronate a domeniului de experimentare) așa cum s-a demonstrat în capitolele IV și V, modelul neural poate fi utilizat cu bune rezultate în scop predictiv.

Din acest motiv, modelul neural se pretează mai mult ca oricare alt tip de model la studiul proceselor complexe în vederea decelării influenței variabilelor de intrare asupra sistemului și a stabilirii parametrilor optimi de funcționare ale acestuia. În continuare se vor discuta succint aspectele legate de programul de rețele neurale utilizat și modalitatea de utilizare a acestuia în aplicațiile abordate.

□ Având în vedere caracteristica de bază a produselor soft comerciale care nu urmăresc profesionalismul abordării ci accesibilitatea utilizatorului la funcțiile de bază ale programului, în studiile efectuate în lucrarea de față s-a utilizat un program de concepție proprie. Sursa a fost creată în Turbo Pascal 7, și a fost înzestrată cu următoarele facilități:

□ posibilitatea selecției a doi algoritmi de bază la corecția ponderilor în etapa de învățare, diferența dintre aceștia fiind în modalitatea de variație a ratei de învățare în funcție de evoluția erorii rețelei.

1. primul algoritm este cel clasic ce prevede o scădere valorică a ratei de învățare cu consecința scăderii vitezei de înaintare pe măsura evoluției rețelei spre minimul depistat ca global . Inițial se pornește de la o valoare

relativ mare a ratei de învățare η (de exemplu 0.5) pentru ca pe măsura involuției valorice a erorii rețelei, valoarea acesteia să scadă de 10 până la 1000 de ori.

2. în elaborarea acestui algoritm, denumit **algoritm adaptiv de convergență**, spre deosebire de cel precedent care prevede micșorarea vitezei de convergență la situarea punctului de lucru pe un gradient descendent al erorii, s-a considerat necesară o accelerare a convergenței prin creșterea valorii ratei de învățare. Astfel, η_t corespunzătoare epocii t este corectată în funcție de gradientul erorii rețelei $\frac{\partial \mathcal{E}}{\partial \hat{a}}$ conform relației:

$$\eta_t = \eta_{t-1} - m_g \frac{\partial \mathcal{E}}{\partial \hat{a}} \quad (1)$$

Valoarea coeficientului m_g este aleasă astfel încât variația ratei de învățare să prezinte semnificația necesară pentru a exploata avantajele algoritmului.

La utilizarea **algoritmului de convergență adaptivă**, programul este prevăzut cu o procedură de vizualizare grafică a variației valorii ratei de învățare funcție de numărul de epoci parcurse;

- posibilitatea învățării rețelei cu funcții diferite de activare și testarea diferitor valori ale principalelor variabile ale rețelei în vederea unei învățări optime
- vizualizarea grafică și numerică atât a erorii de rețea cât și a deviației standard a valorilor calculate ale variabilei de ieșire față de cele experimentale;
- intervenția directă în cazul paraliziei rețelei în procesul de învățare prin modificarea impuls a unor ponderi sau prin prezentarea pe parcursul unui număr redus de epoci doar a acelor tipare pentru care eroarea rețelei este maximă;
- stocarea ponderilor calibrate la sfârșitul etapei de învățare a rețelei
- utilizarea ponderilor stocate la evaluarea variabilelor de ieșire pentru vectori de intrare impuși printr-o semiepocă de propagare înainte. Această facilitate este utilă pentru studii de predicție.

- În capitolul II este exploatată periodicitatea proprietăților în sistemul periodic al elementelor. Au fost selectate două proprietăți - electronegativitatea și căldura specifică - care s-au considerat importante prima din punctul de vedere al semnificației acesteia în caracterizarea compușilor heteroatomici iar ultima din punctul de vedere al posibilității aplicării metalelor tranziționale în tehnologii speciale care presupun procese de transfer termic.

Pentru a descrie cât mai fidel schimbările configurației electronice în sistemul periodic al elementelor, **pentru variabilele independente** s-a utilizat următoarea mapare originală:

-**prima variabilă independentă** a fost dată de perioada din care face parte elementul, făcându-se astfel diferențierea între orbitalii de completare și ceilalți orbitali de același gen;

-**următoarele patru variabile independente** au fost date de numărul electronilor s, p, d, f de pe orbitalul în completare;

-**variabila dependentă**, s-a ales valoarea proprie a căldurii specifice a elementului.

La ambele aplicații s-a justificat utilizarea modelului neural în corelarea structurii electronice a elementelor cu proprietățile fizice ale acestora prin relevarea ineficienței regresiei multiliniare în același scop.

S-a constatat experimental că un domeniu larg de împrăștiere al variabilelor de ieșire conduce la rezultate slabe ale procesului de învățare.

Pentru a caracteriza domeniul larg de împrăștiere al variabilelor de ieșire, s-a definit un **factor de împrăștiere SF** (**spreading factor**) descris de relația :

$$SF = \frac{MAXout}{MINout} \quad (2)$$

unde: - MAXout este valoarea maximă a variabilei de ieșire

- MINout este valoarea minimă a variabilei de ieșire

Având în vedere scopul inițial de a efectua o predicție a proprietăților pentru elementele de la sfârșitul tabelului periodic al elementelor, în cazul căldurii specifice, pentru îmbunătățirea factorului de împrăștiere a valorilor proprietății, s-a renunțat la tiparele corespunzătoare primelor 10 elemente din tabelul periodic al elementelor.

Pentru fiecare din cele două aplicații, în scopul determinării numărului de neuroni de pe stratul ascuns, s-au testat rețele cu un număr variabil de neuroni pe acest strat. S-a observat că o creștere a numărului de neuroni de pe stratul ascuns este benefică până la atingerea unei valori de la care eroarea rețelei rămâne constantă. Valoarea respectivă a fost adoptată ca valoare optimă a numărului de neuroni de pe stratul ascuns.

Odată definiți parametrii de rețea, aceasta a fost antrenată pe parcursul unui număr de epoci relativ redus pentru a evita fenomenul de suprainvătăre care are un efect negativ asupra capacității de predicție în afara domeniului acoperit de tiparele învățate.

Din corelarea prin regresie liniară a valorilor calculate cu rețeaua neurală funcție de cele experimentale - caracterizate prin parametrii de corelare de o valoare foarte bună - se observă o bună capacitate de generalizare a rețelei.

În vederea comparării efortului de învățare cu cel prezentat în aplicații similare s-a introdus noțiunea de factor de antrenare specific definit de relația:

$$STF = (N_i \cdot N_a \cdot N_{ep}) / (N_{tip} \cdot N_o) \quad (3)$$

unde N_{ep} - numărul epocilor de antrenare a rețelei

N_i - numărul elementelor din vectorul de intrare

N_a - numărul unităților din stratul ascuns

N_{tip} - numărul de tipare din setul de date utilizat în etapa de învățare

N_o - numărul elementelor din stratul de ieșire

Având în vedere modalitatea de definire, acest factor caracterizează efortul depus de rețea pentru învățarea unei singure variabile de ieșire. Din valoarea factorului specific de antrenare pentru cele două aplicații, rezultă că efortul la învățare a fost considerabil mai mic decât cel întâlnit în aplicații similare.

Pentru validarea modelului neural, s-a utilizat metoda validării încrucișate aplicată în modul cel mai restrictiv. Setul de date s-a divizat în două subseturi de dimensiune egală, selecția făcându-se într-o modalitate cât mai arbitrară - după paritatea numărului curent al tiparelor. Fiecare din cele două subseturi a fost utilizat la antrenarea unei rețele neurale în vederea obținerii ponderilor calibrate pe dependența vector de ieșire - vector de intrare. Ulterior etapei de învățare, rețelei i s-au prezentat doar vectorii de intrare ai celui alt subset, și pe baza ponderilor calibrate, în cadrul unei semiepoci de propagare înainte s-a efectuat o predicție a valorilor vectorilor de ieșire.

Din corelarea prin regresie liniară a valorilor de predicție funcție de cele experimentale - caracterizate prin parametri de corelare de o valoare foarte bună - se observă o bună capacitate predictivă a modelului neural și totodată se validează modelul și se pune în evidență capacitatea de generalizare a acestuia în domeniul de variație al tiparelor utilizate la antrenare.

Pentru efectuarea de predicții ale proprietăților fizice ale elementelor pentru care acestea nu sunt cunoscute, inițial a fost antrenată o rețea neurală cu toate tiparele disponibile. Ulterior etapei de învățare în care ponderile au fost calibrate pe dependența vector de ieșire - vector de intrare, rețelei i-au fost prezentați doar vectorii de intrare ai elementelor pentru care s-au efectuat predicții. Pe baza ponderilor calibrate, pe parcursul unei semiepoci de propagare înainte s-a efectuat o predicție a valorilor vectorilor de ieșire, deci a valorii corespunzătoare proprietăților.

Pentru elementele cuprinse în domeniul de variație al tiparelor utilizate la antrenare, valorile obținute se încadrează foarte bine în legitatea de variație a proprietății în cadrul sistemului periodic al elementelor.

Pentru elementele care nu se încadrează în domeniul de variație al tiparelor utilizate la antrenare, valorile prezise sunt pertinente, având o bună probabilitate de a fi confirmate experimental.

□ Având în vedere că proprietățile moleculelor sunt determinate de structura lor, un grad avansat al cunoașterii structurii reprezintă practic un pas important în explicarea comportamentului fizico-chimic al moleculelor.

Astfel, în **capitolul III** s-a urmărit posibilitatea corelării punctelor de fierbere pentru un set de 100 de alcooli alifatici cu câțiva indici topologici adecvați scopului și cu indici derivați dintr-o mapare logică a conformației moleculare.

Având în vedere datele din literatură privitor la încercări anterioare, s-au calculat indicii topologici cei mai utilizați pentru întregul set de alcooli.

Pentru alegerea variabilelor de corelare, s-a testat regresia liniară a fiecăruia din indicii topologici pentru a selecta variantele optime de corelare multiplă.

S-au selectat astfel, pe baza valorii coeficientului de corelare, pentru realizarea scopului propus, următoarele perechi de indici:

$$1 - \chi + \chi d_1$$

$$2 - \chi + J$$

$$3 - J+N$$

Pentru corelarea indicilor de structură cu punctele de fierbere ale alcoolilor s-au utilizat atât rețele neurale cât și regresii multiliniare, efectuându-se comparații ale parametrilor de corelare pentru cele două alternative.

Spre deosebire de algoritmul clasic utilizat în capitoul precedent în etapa de învățare a rețelei neurale, în acest capitol s-a utilizat algoritmul *adaptiv de convergență* care s-a dovedit a fi în cazul concret al acestei aplicații superior celui clasic.

Din valorile parametrilor de corelare pentru cele două abordări - regresie multiliniară și rețea neurală - se evidențiază superioritatea modelului neural în detrimentul regresiei multilinare.

Testele de corelare a punctelor de fierbere cu indicii topologici au dat rezultate notabile, mai ales când s-au utilizat pentru corelare rețelele neurale. Caracterizarea structurală prin indici topologici s-a dovedit a fi suficient de completă, însă valorile indicilor topologici sunt lipsite de o semnificație structurală fizică.

Având în vedere acest dezavantaj al indicilor topologici, s-a realizat o hartă originală a structurii alcoolilor prin descriptorii derivați din considerente structurale ce presupun o dependență directă între valoarea acestora și valoarea temperaturilor de fierbere.

În acest sens, s-au luat în considerare principalii factori care induc o variație semnificativă a temperaturii de fierbere. Mai jos sunt redată expresiile de calcul ale discriminanților, astfel exprimați încât valoarea acestora să fie direct proporțională cu temperatura de fierbere a alcoolilor.

1. **N_c** = numărul atomilor de carbon din molecula
2. **I_a** - indice de acentricitate

$$I_a = \frac{Lc_{max}}{C_{OH}} \quad (4)$$

unde: - Lc_{max} este numărul de atomi de carbon din catena de lungime maximă a moleculei

- C_{OH} - numărul atomului de carbon din catena de lungime maximă la care este legată gruparea -OH

3. **I_i** - indice incremental de substituție care ține seama de tipul, numărul și depărtarea substituenților față de gruparea OH, calculat conform relației:

$$I_i = \frac{I}{I + \sum \frac{A_i}{N_c - c}} \quad (5)$$

în care:

- $A_i = \sqrt{N_{c_i}}$ este un coeficient al efectului inductiv al substituentului, N_c , fiind numărul atomilor de carbon din care este format substituentul

- $N_c - c$ = numărul legăturilor C-C care despart gruparea OH de substituent

Corelarea acestor discriminanți structurali cu punctele de fierbere s-a realizat atât prin regresie multiliniară cât și utilizând un model neural. În cazul modelului neural, parametrii de corelare obținuți au fost superiori celor obținuți prin regresie multiliniară.

Ceea ce trebuie subliniat, este că ambele modalități s-au dovedit a fi superioare prin valoarea parametrilor de corelare experimentului lui Hall¹⁷⁸ care a corelat punctele de fierbere ale alcoolilor cu o combinație de indici electronici, topologici și de stare de valență.

Pentru validarea modelului neural și verificarea capacității sale de predicție s-a utilizat o procedură de validare încrucișată, bazată pe divizarea setului de date în două subseturi. Fiecare dintre acestea a fost utilizat pentru antrenarea unei rețele neurale, iar pe baza ponderilor obținute s-au calculat valorile de predicție ale punctelor de fierbere pentru celălalt subset.

Din parametrii de corelare liniară a valorilor de predicție funcție de cele experimentale, rezultă o foarte bună capacitate de generalizare a modelului neural utilizat. În ceea ce privește corelarea punctelor de fierbere la alcoolii cu valorile discriminanților structurali pe lângă faptul că aceștia prezintă o semnificație structurală perceptibilă, prin valoarea parametrilor de corelare s-a dovedit a fi cea mai eficace.

□ În **capitolul IV** este considerată o aplicație a Rețelelor Neurale la modelarea procesului electrochimic de reducere a nitrobenzenului la anilină și la decelarea influenței parametrilor de proces asupra randamentului de curent.

Spre deosebire de procedeele clasice, în care spațiile electrodice sunt separate prin membrane sau diafragme în scopul evitării degradării prin oxidare a produsului obținut la catod, procesul studiat demonstrează posibilitatea realizării cu randamente de curent mari a procesului într-un reactor electrochimic necompartimentat

Pentru obținerea unor randamente de curent mari în reactoarele electrochimice necompartimentate, este necesar a stimula selectiv reducerea nitrobenzenului la anilină la catod și degajarea oxigenului la anod. În ce privește electrolitul suport utilizat, atât reducerea nitrobenzenului la catod cât și reacția de degradare prin oxidare a anilinei depind de etapa de transport a moleculelor organice din electrolit la electrod. Deci, o stimulare selectivă a reacției de electrod utilă devine posibilă dacă procesul este condus astfel încât:

- densitatea de curent catodică să fie scăzută la o valoare care să tindă spre valoarea curentului limită de reducere a nitrobenzenului fără a o depăși;
- densitatea de curent anodică să fie cât mai mare posibilă la o valoare care să depășească curentul limită de oxidare a anilinei.

Menținând constanți pe parcursul desfășurării procesului, raportul suprafață catodică/suprafață anodică, concentrația electrolitului suport și regimul hidrodinamic, principalii parametri care influențează eficiența procesului desfășurat într-un reactor necompartimentat sunt considerați:

- densitatea de curent
- temperatura
- cantitatea de electricitate trecută prin reactorul electrochimic.

Pentru justificarea utilizării rețelelor neurale la modelarea procesului, s-a încercat inițial o corelare printr-o regresie liniară multivariabilă având ca variabilă dependentă randamentul de curent experimental iar ca variabile independente densitatea de curent - i , temperatura - T și cantitatea de electricitate trecută prin reactorul electrochimic - Q .

Deși valoare coeficientului de corelare obținut este rezonabilă, valorile coeficientului de determinare și mai ales valoarea deviației standard fac regresia multiliniară improprie la modelarea procesului.

În comparație cu parametrii de corelare obținuți prin regresie multiliniară, în cazul dependenței liniare dintre valorile calculate cu rețeaua neurală și cele

experimentale deviația standard a scăzut de $\cong 5$ ori iar valoarea testului Fisher a crescut de $\cong 33$ ori, fapt ce relevă superioritatea modelului neural în detrimentul regresiei multiliniare.

Pentru validarea modelului neural și verificarea capacității sale de predicție s-a utilizat o procedură de validare încrucișată, bazată pe divizarea setului de date în două subseturi egal populate. Fiecare dintre acestea a fost utilizat pentru antrenarea unei rețele neurale, iar pe baza ponderilor obținute s-au calculat valorile de predicție ale randamentelor de curent pentru celălalt subset.

Din parametrii de corelare liniară a valorilor de predicție funcție de cele experimentale, rezultă o foarte bună capacitate de generalizare a modelului neural utilizat.

Pentru a sesiza importanța fiecărui parametru în cadrul procesului, s-au utilizat trei seturi de date în care a fost variat doar un singur parametru de intrare, ceilalți menținându-se constanți la valorile constatate. Utilizând rețeaua antrenată cu cele 55 de tipare experimentale, pe baza ponderilor calibrate la antrenare, într-o semiepocă de propagare înainte a rețelei s-au calculat randamentele de curent corespunzătoare seturilor de date de intrare.

Astfel, s-au stabilit condițiile optime de funcționare pentru procesul de reducere a nitrobenzenului la anilină într-un reactor electrochimic necompartimentat și influența fiecărui parametru de intrare asupra randamentului de curent al procesului.

□ In **capitolul V**, s-a efectuat o modelare dinamică a procesului de reținere a Sr^{2+} prin schimb ionic utilizând rețele neurale.

Partea experimentală a fost realizată pe o instalație de laborator formată din trei coloane identice ca dimensiuni, în care straturile de schimbător de ioni au fost dispuse prin multiplicarea lungimii primei coloane de 1,2 respectiv 3 ori. Masa ionică utilizată a fost un cationit puternic acid de tip PUROLITE C 100, destinată preferențial aplicațiilor speciale. Din considerente experimentale s-a lucrat la debit de alimentare constant și la 3 concentrații diferite, astfel încât în final s-au avut la dispoziție 9 seturi de date experimentale (3 coloane \times 3 concentrații ale influentului).

Inițial s-a testat o regresie multiliniară pentru a găsi o ecuația de regresie care să exprime variația concentrației C_{exp} la ieșirea din coloană în timp, funcție de cele trei variabile independente - t , concentrația inițială a soluției C_i , și volumul de schimbător de ioni parcurs, exprimat prin volumul redus V_r . Prin valorile parametrilor de corelare, se poate observa că regresia multiliniară prin structura sa este total inadecvată.

Pentru testarea modelului neural, s-au utilizat ca seturi de antrenare aceleași seturi de date experimentale testate anterior prin regresie multiliniară.

În scopul validării datelor experimentale și a formării ponderilor ce corelează variabilele de intrare cu cele de ieșire ale procesului studiat, s-a utilizat o rețea neurală artificială cu propagare înapoi. Sursa programului a fost aceeași ca în capitolele II, III și IV, fiind proiectată pentru cele 110 de tipare experimentale corespunzătoare celor 9 seturi de date experimentale. Tiparele au fost construite corespunzătoare celor 9 seturi de date experimentale. Tiparele au fost construite dintr-un vector de intrare tridimensional compus din valorile variabilelor C_i , t , V_r și

un vector de ieșire unidimensional C_{cX} - valoarea concentrației la ieșirea din coloană.

Parametrii de corelare ai regresiei liniare între valorile calculate și cele experimentale, indică o foarte bună capacitate de corelare a modelului neural și denotă descriptibilitatea tuturor tiparelor de către același mecanism de corelare a variabilelor de ieșire cu cele de intrare.

Pentru a obține o variație continuă a concentrației în timp pentru modelul neural, în urma antrenării rețelei cu seturile de date experimentale s-au generat vectori de intrare care să acopere dens în domeniul timp întreg domeniul de variație. Pe baza ponderilor create în etapa de învățare a rețelei s-au calculat valorile variabilelor de ieșire - valoarea concentrației Sr^{2+} la ieșirea din coloană, rezultând o suprapunere fidelă a valorilor experimentale peste cele calculate pe baza modelului neural.

Pentru validarea modelului neural, s-a utilizat tehnica validării încrucișate descrisă anterior în capitolele II, III, și IV. Pentru o divizare arbitrară a setului de date s-a adoptat drept criteriu de selecție cel al parității numărului curent pentru cele 110 tipare. Pentru validare s-a utilizat o rețea neurală care a fost antrenată succesiv cu cele două subseturi de date în vederea utilizării ulterioare pentru predicția valorilor de ieșire ale contrasetului corespunzător. Valorile parametrilor de corelare pentru regresia multiliniară a valorilor calculate funcție de cele experimentale conduc la validarea modelului și la relevarea capacității de generalizare a acestuia.

Ceea ce este însă notabil pentru un proces de schimb ionic, s-a testat cu bune rezultate capacitatea de predicție a modelului pentru trei seturi de date experimentale. Astfel, de fiecare dată, au fost utilizate pentru antrenarea unei rețele neurale 8 din cele 9 seturi de date experimentale, ponderile astfel calibrate fiind utilizate la predicția setului de date omis din etapa de învățare. Selecția seturilor de date s-a efectuat astfel încât variabila de bază a vectorului de intrare - C_i - concentrația influentului - să acopere întreg domeniul experimental de variație. Valorile parametrilor de corelare ai regresiei liniare dintre valorile de predicție și cele experimentale relevă o foarte bună capacitate predictivă a modelului neural utilizat.

Din considerentele explicate pe larg mai sus, utilizarea rețelelor neurale la corelarea proprietăților fizice cu variabile ce caracterizează configurația electronică elementară respectiv structura moleculară s-a dovedit net superioară atât de mult vehiculatei în acest domeniu - regresii liniare multivariabile. Mai mult decât atât, predicțiile realizate pe baza modelului neural în afara domeniului experimental s-au dovedit a fi pertinente, această particularitate găsindu-și aplicabilitatea în estimarea proprietăților elementelor pentru care acestea sunt încă nedeterminate experimental.

În ce privește utilizarea rețelelor neurale la modelarea celor două procese, prin utilizarea programului propriu, s-au efectuat atât decelarea influenței variabilelor de intrare asupra procesului electrochimic (capitolul IV) cât și elemente de predicție pentru procesul de schimb ionic (capitolul V), rezultate greu sau aproape imposibil de obținut cu ajutorul altor tipuri de modele.

Domeniul aplicării rețelelor neurale în chimie și inginerie chimică rămâne în continuare deschis, atât ca posibilitate de abordare cât și ca modalitate de exploatare intensivă a particularităților acestui tip de modelare în caracterizarea intensivă a sistemelor investigate.

Bibliografie

1. Hykin, S. *Neural Networks*, Macmillan College Publishing Company, New York, **1994**.
2. Faggin, F. *International Joint Conference on Neural Networks*, Seattle, **1991**.
3. Petrovanu, I.; Antohe, D.D.; Varlam, A. *Neuroanatomie Clinică-Sistem nervos central*, vol 1, Ed. Edit-Dan, Iași, **1996**.
4. Shepherd, G.M.; Koch, C. *Introduction to synaptic circuits*, Oxford University Press, **1990**.
5. Alecsander, I.; Morton, H. *An introduction to Neural computing*, Chapman & All, London, **1990**.
6. Hodgkin, A.L.; Huxley, A.F. *A J. Phys.*, **1952**, 117, 500.
7. McCulloch, W.S.; Pitts, W. *Bull. Math. Bioph.*, **1943**, 5, 115.
8. Ramon y Cajal *Histologie du systeme nerveux de l'homme et des vertebrates*, (Ed. Revue), Maloine, Paris, **1955**.
9. Hebb, D.O. *The Organization of Behaviour: A Neuropsychological Theory*, Willey, New York, **1949**.
10. Rochester, N.; Holland, J.H.; Haibt, L.H.; Duda, W.L. *IRE Transactions on Information Theory*, **1956**, IT-2, 80.
11. Ashby W.R. *Design for a Brain*, Wiley, New York, **1952**.
12. Minsky, M.L. *Ph.D Thesis*, Princeton University, Princeton, N.J., **1954**.
13. Minsky, M.L. *Steps towards artificial intelligence*, Proceedings of the Institute of Radio Engineers, **1961**, 49, 8.
14. Steinbuch, K. *Kybernetics*, **1961**, 1, 36.
15. Taylor, W.K. *Information Theory*, **1956**, London: Butterworth, 314.
16. Anderson, J.A. *Math. Bioscience*, **1972**, 14, 197.
17. Kohonen, T. *IEEE Transactions on Computers*, **1972**, C-21, 353.
18. Nakano, K. *IEEE Transactions on Systems, Man and Cybernetics*, **1972**, SMC-2, 380.
19. Winograd, S.; Cowan, J.D. *Reliable Computation in the Presence of Noise*, Cambridge, MA: MIT Press, **1963**.
20. Rosenblatt, F. *Psych. Rev.*, **1958**, 65, 386.
21. Rosenblatt, F. *Proceedings of the Institute of Radio Engineers*, **1960**, 48, 301.
22. Rosenblatt, F. *Report VG-1196-G-4*, Aeronautical Laboratory, Buffalo, New York, **1960**.
23. Rosenblatt, F. *Principles of Neurodynamics*, Washington, DC: Spartan Books, **1962**.
24. Van der Smagt, P.; Krose, B. "An Introduction to Neural Networks", Amsterdam, **1994**.
25. Widrow, B.; Hoff, M.E. *IRE WESCON, Convention Record*, 96, **1960**.
26. Widrow, B. *Generalisation and information storage in networks of adaline neurons*, Washington, DC: Spartan Books, **1962**.
27. Minsky, M.L.; Papert, S.A. *Perceptrons*, Cambridge, MA: MIT Press, **1969**.
28. Minsky, M.L. *Proceedings of the Institute of Radio Engineers*, **1961**, 49, 8.
29. Hopfield, J.J. *Proceedings of the National Academy of Sciences of U.S.A.*, **1982**, 79, 2554.

30. Hopfield, J.J *Proceedings of the National Academy of Sciences of U.S.A.*, **1984**, 81,3088.
31. Hopfield, J.J *Modeling and Analysis in Biomedicine*, Singapore:World Scientific, **1984**.
32. Cohen, M.A.; Grossberg, S. *IEEE Transactions on Systems, Man and Cybernetics*, **1972**, SMC-13, 815.
33. Kirkpatrick, S.;Gelatt, C.D.Jr.; Vecchi, M.P. *Science*, **1983**, 220, 671.
34. Meropolis,N.;Rosenbluth,M.;Rosenbluth,A.;Teller,A.;Teller,E. *J.Chem.Ph.*, **1953**, 21, 1087.
35. Ackley, D.H.; Hinton,G.E.; Sejnowski,T.J. *Cognitive Science*, **1985**, 9,147.
36. Breitenberg, V. *Vehicles; Experiments in Sintetic Psihology*, Cambridge, MIT, Cambridge, MA, **1984**.
37. Werbos, P.J. *Beyond regression. New tools for prediction and analysis in the behavioral sciences*, Ph.D. Thesis, Harvard University, Cambridge, MA,**1974**.
38. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. *Nature*, **1986**, 323, 533.
39. Rumelhart, D.E.; McClelland, J.L. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. I, MA:MIT Press, **1986**.
40. Parker, D.B. "Learning-logic:Casting the cortex of human brain in silicon."-*Tehncial Report TR-47. Center for Computational Research in Economics and Management Science*, MIT, Cambridge, MA, **1985**.
41. Parker, D.B. *IEE 1st International Conference on Neural Networks*, **1987**, vol.2, 593.
42. LeCun, Y. *Cognitiva*, **1985**, 85, 599.
43. Hornik, K.; Stinchcombe, M.; White, H. *Neural Networks*, **1989**, 2, 359.
44. Hornik, K.; Stinchcombe, M.; White, H. *Neural Networks*, **1990**, 3, 551.
45. Hartman, E.J.;Keeler, J.D.;Kowalski, J.M. *Neural Computation*, **1990**, 2(2),210.
46. Prados, D.L. *Electron. Lett.*, **1992**, 28, 1560.
47. Klawun, C.; Wilkins, C.L. *J.Chem.Inf.Comp.Sci.*, **1994**, 34,984.
48. Stoer, J.; Burlisch, R. *Introduction to Numerical Analysis*, Springer-Verlag, New York-Heidelberg-Berlin, **1980**.
49. Powell, M.J.D..*Math. Programing*, **1977**, 12,241.
50. van den Boomgaard, R.;Smeulders, A. *Proceedings of the I.A.S. Conference*, **1989**, 305.
51. Silva,F.M.; Almeida, L.B. *Adv. Neural Comp.*,**1990**,151.
52. Szyrbik, N. *Modelul computațional conexiunist*, TEMPUS SJEP07091, București, **1996**.
53. Hertz,A.;Krogh,A.;Palmer,R.G. "Introduction to the Theory of Neural Computation", Addison-Wesley Publishing Company, Redwood City, **1991**
54. Zurada, J.M. "Artificial Neural Systems", West Publishing Company, St. Paul, 1992
55. Jacobs, R.A. *Neural Networks*, **1988**, 1, 295.
56. Zupan, J.; Gasteiger, J. *Neural Networks for Chemists: An Introduction*, New York, **1993**, 305.
57. Egolf,M.;Jurs.P.C. *J. Chem.Inf.Comput.Sci.* **1993**, 33, 616.
58. Aoyama,T.; Ichikawa,H. *J.Chem.Inf.Comput.Sci.* **1992**, 32, 492.
59. Bodor, N.; Hargert,A.; Huang,M. *J.Am.Chem.Soc.* **1991**, 113, 9480.
60. Andrea,T.A.; Kalayeh,H. *J.Med.Chem.* **1991**,34, 2824.
61. Aoyama,T.; Suzuki,Y.; Ichikawa,H. *J.Med.Chem.* **1990**,33, 2583.
62. Hatrik, S.; Zahradnik, P. *J.Chem.Inf.Comput.Sci.* **1996**, 36, 992.

63. Schultz ,H.P. *J.Chem.Inf.Comput.Sci.* **1996**, 36, 996.
64. Hall, L.H.; Story, C.T. *J.Chem.Inf.Comput.Sci.* **1996**,36, 1004.
65. Chow, H.; Chen, H.; Myrdal, P.; Yalkowsky, S.H. *J.Chem.Inf.Comput.Sci.* **1995**, 35, 723.
66. Jordan, S.N.; Leach A.R. *J.Chem.Inf.Comput.Sci.* **1995**, 35, 640.
67. Gakh,A.; Gakh,E.G.; Sumpster,B.G.; Noid,D.W.*J.Chem.Inf.Comput.Sci.*,**1994**,34, 832.
68. Selwood,D.L.;Livingstone,D.J.;Comley,J.C.W.;Hudson,A.T.;Jackson,P.,*J.Med Chem.* **1990**, 33, 136.
69. Kireev, D.B. *J.Chem.Inf.Comput.Sci.* **1995**, 35, 175.
70. West ,G.M. *J.Chem.Inf.Comput.Sci.* **1995**, 35, 21.
71. Takahashi, Y.; Abe, H.; Sasaki, S. *J.Chem.Inf.Comput.Sci.* **1994**, 34, 109.
72. Livingstone, D.J.; Salt, D.W. *Bioorganic & Med.Chem. Letters*, vol.2, **1992**, 3.
73. Salt, D.W.; Livingstone, D.J.; Tinsley, C. *Pestic.Sci.* **1992**, 36, 161.
74. Pedersen, P.V.; Modi, N.B. *J.Pharm. Sci.*,**1991**, vol 82, 9, 918.
75. Manallack, D.T.; Livingstone, D.*J.Med. Chem. Res.*,**1992**, 2, 181.
76. Balaban, A.T.; Basak, S.C.; Colburn T.; Grunwald, G.D., *J. Chem. Inf. Comput.Sci.* **1994**, 34, 1118.
77. Burden, F.R. *J.Chem. Inf. Comput. Sci.* **1994**, 34, 1229.
78. Carrieri, A.H. *App.Optics*, **1995**, 34,(15), 2633.
79. Xu, J.; Borer, P.N. *J.Chem.Inf.Comput.Sci.* **1994**, 34, 349.
80. Ventura, S.; Silva, M.; Perez-Bendito,D. *J.Chem.Inf.Comput.Sci.* **1997**, 37, 287.
81. Svozil, D.;Sevcik, G.K.; Kvasnicka, V. *J.Chem.Inf.Comput.Sci.* **1997**, 34, 338.
82. Anker,L.S.; Jurs,P.C. *Anal.Chem.***1992**,64, 1157.
83. West,G.M.J. *J.Chem.Inf.Comput.Sci.***1993**,33, 577.
84. Meyer,B.;Hansen,T.; Nute,B.; Albersheim,P.;Darvill,A.; York,W.; Sellers, J. *Science.*, **1991**, 251, 542.
85. Doucet,J.P.;Panaye,A.;Feuilleaubois,E.;Ladd,P. *J.Chem.Inf.Comput.Sci.* **1993**, 33, 320.
86. Kvanicka,V.; Sklenak,S.; Pospichal,J. *J.Chem.Inf.Comput.Sci.* **1992**,32, 742.
87. Boger, Z.; Karpas, Z. *J.Chem.Inf.Comput.Sci.* **1994**, 34, 576.
88. Bell,S.E.;Mead,W.C.;Jones,R.D.;Eiceman,G.A.;Ewing,R.G.*J.Chem.Inf.Comput.Sci.*, **1993**, 33, 609.
89. Tanabe,K.; Tamura,T.;Uesaka,H. *Appl.Spectrosc.* **1992**,46, 807.
90. Ricard,D.;Cachet,C.;CabrolBass,D.;Forrest,T.P. *J.Chem.Inf.Comput.Sci.* **1993**,33, 202.
91. Weigel,U.M.; Herges,R. *J.Chem.Inf.Comput.Sci.***1992**,32, 723.
92. Fessenden,R.J.; Gyorgyi,L. *J.Chem.Soc.Perkin Trans.2* **1991**,1755.
93. Ricard,D.;Cachet,C. *J.Chem.Inf.Comput.Sci.* **1993**,33, 202.
94. Lohninger,H.;Stancl,F. *J.Anal.Chem.* **1992**,344, 186.
95. Gasteiger,J.;Li,X.;Simon,V.;Zupan,J. *J.Mol.Struct.* **1993**,292, 141.
96. Long,J.R.;Mayfield,H.T.;Henley,M.V.;Kromann,P.R. *Anal.Chem.* **1991**,63,1256.
97. Davalo, E.; Naim, P. *Neural networks*, MacMilan Education, **1991**
98. Woinaroschy, A.; Mihai, Mariana *Rev. Chim.*, **1994**,45,482.
99. Fu, C.S.; Poch, M. *Int.J.Systems Sci.* **1993**, 8, vol.24, 1601.
100. Aizawa, T.;Yokoshita, T. *ZKG International.* **1995**, 10, 17.
101. Aizawa, T. *IEEE Cement Industry Technical Conference.* **1992**, 254.
102. Nakamura.T.; Nakagawa,N; Aizawa,T. *Journ.Soc.Instr.Contr.Engg.* **1988**,27,84.
103. Iwassa,K.;Omatu,S. *System Control and Inf.*. **1993**, 6,(1), 11.
104. Psaltis,D.;Sideris,A.;Yamamura,A. *Control Systems Magazine.* **1988**, 17,48.
105. Chovan, T; Catfolis, T.;Meert, K. *AIChE J.*, **1996**, (42),2, 493.
106. Ishida, M.; Zhan, J. *AIChE J.*, **1995**, (41),10, 2333.

107. Pauling, L. *Chimie Generală*, Ediția a III-a, Ed. St., București, 1972, 335.
108. Sigman, N.; Rives, S. S. *J. Chem. Inf. Comput. Sci.*, **1994**, *34*, 617-620.
109. Geantă, M. *Chem. Bull. Politeh. Univ. Timisoara*, **1996**, *41*(55), 112-120.
110. Geantă, M.; Perju, Delia *Chem. Bull. Politeh. Univ. Timisoara*, **1997**, *42*(56), (sub tipar).
111. *CRC Handbook of Chemistry and Physics*, 75th Edition, Lide, D.R., CRC Press: Boca Raton, FL, 1994.
112. B. Schenkert ; M. Agarwal *Comput. Chem. Eng.*, **1996**, *2*, 175.
113. Turo, N.J. *Angev. Chem. Int. Edit. Engl.* **1986**, *25*, 882.
114. Nenițescu, C.D. *Chimie Organică*, Ed. a VII-a, Editura Didactică și Pedagogică, București, 1974, vol. I, 54.
115. Goodisman, J. *Contemporary Quantum Chemistry - An Introduction*, Plenum Press, New York, 1977.
116. Minkine, V.; Simkine, B.; Miniaev, R. *Theorie de la Structure Moleculaire (Couches Electroniques)*, Ed. Mir. Moscou, 1982.
117. Rogers, D. *Computational Chemistry using the PC*, VCH, New York 1990, p. 137.
118. Cox, Ph.J. *J. Chem. Educ.*, **1982**, *59*(4), 275.
119. Balaban, A.T. *Rev. Chim.*, **1988**, *39*(12), 1027.
120. Kier, L.B.; Hall, L.H. *Molecular Connectivity in Structure Analysis*, Research Studies Press, Wiley, New York 1986, 2.
121. Randić, M. *J. Chem. Educ.* **1986**, *26*(1), 134.
122. Randić, M. *J. Chem. Educ.* **1992**, *69*(9), 713.
123. Galvez, J.; Garcia, R.; de Julia-Ortiz, V.J.; Soler, R. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1198.
124. Galvez, J.; Garcia, R.; de Julia-Ortiz, V.J.; Soler, R. *J. Chem. Inf. Comput. Sci.* **1994**, *35*, 272.
125. Read R.C. *The Enumeration of Acyclic Chemical Compounds in Chemical Applications of Graph Theory*; Balaban, A.T.; Ed. Acad. Pry.z. London 1976, 25.
126. Randić, M. *Int. J. Quantum Chem. : Quantum Biol. Symp.* **1980**, *7*, 187.
127. Bonchev, D.; Trinajstić, N. *Int. J. Quantum Chem. : Quantum Chem. Symp.* **1978**, *12*, 293.
128. Randić, M.; Plavšić, D.; Trinajstić, N. *Gazzetta Chim. Ital.* **1988**, *18*, 441.
129. Tosić, R.; Masulović, D.; Stojmenović, I.; Brunvoll, J.; Cyvin, B.N.; Cyvin, S.J. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 181.
130. Read R.C. Preprint 1969, Univ. West Indies, Jamaica.
131. Balaban, A.T.; Moțoc, I.; Bonchev, D.; Mekenyan, O. *Topological Indices for Structure-Activity Correlation in Steric Effects in Drug Design* Charton, Springer Verlag, Berlin, 1983, 24.
132. Dubois, J.E.; Carrier, G.; Panaye, A. *J. Chem. Inf. Comput. Sci.*, **1991**, *31*, 574.
133. Fujita, S. *Bull. Chem. Soc. Japan*, **1990**, *63*, 203.
134. Fujita, S. *Bull. Chem. Soc. Japan* **1990**, *63*, 315.
135. Merrifield, R.E.; Simmons, H.E. *Topological Methods in Chemistry*, John Wiley & Sons, New York 1979, 179.
136. Dugundji, J.; Gillespie, P.; Ugi, I.; Ramirez, F. "Chemical Applications of Graph Theory", ed. Balaban, A.T., Acad. Press, London 1976, 107.
137. Mihalic, Z.; Trinajstić, N. *J. Chem. Educ.*, **1992**, *69*(9), 701
138. Balaban, A.T.; Chiriac, A.; Moțoc, I.; Simon, Z. *Steric Fit in Quantitative Structure - Activity Relations in Lectures Notes in Chemistry* 1980, 15, Springer Verlag, Berlin. 23.
139. Ruedenberg, K. *J. Chem. Phys.*, **1954**, *22*, 1878.
140. Sablic, A.; Trinajstić, N. *Acta Pharm. Jugosl.*, **1981**, *31*, 189.

141. Balaban, A. T. *Pure Appl. Chem.*, **1983**, 55, 199.
142. Harary, F. An Exposition of Graph Theory in Chemical Application of Graph Theory, Balaban, A.T.; Ed. Acad. Press, London **1976**, 5.
143. Ciubotariu, D.; Medeleanu, M.; Gogonea, V. *Chem. Bull.T.U. Timișoara* **1995**, 40 (54), 21.
144. Ciubotariu, D.; Medeleanu, M.; Gogonea, V. *Chem. Bull.T.U. Timișoara* **1996**, 41(55), 19.
145. Rouvray, D.H. *The Topological Matrix in Quantum Chemistry in Chemical Applications of Graph Theory*, Balaban, A.T.; Ed. Acad. Press, London **1976**, 175.
146. Randić, M. *J.Math.Chem.* **1992**, 9(2), 97.
147. Ivanciuc, O.; Balaban, A.T. *MATCH,Commun.Math.Chem.*,**1994**,30, 141.
148. Diudea, M.V.; Horvath, D.; Graovac, A. *J.Chem.Inf.Comput.Sci.*,**1995**, 35,129.
149. Balaban, A.T.; Ciubotaru, D. *J.Chem.Inf.Comput.Sci.* **1991**, 31(4),517.
150. M.Medeleanu;Geantă,M.;Ciubotaru,D. *The Romanian-Hungarian Phys. Joint Meeting*, Timișoara,3-5 May, **1996**.
151. Balaban, A.T. *J.Chem.Inf.Comput.Sci.* **1992**, 32, 23.
152. Wiener, H. *J.Amer.Chem.Soc.* **1947**, 69,17.
153. Wiener, H. *J.Phys.Chem.*,**1948**, 52,425
154. Wiener H. *J.Amer.Chem.Soc.*, **1947**, 69, 2636.
155. Wiener H. *J.Phys.Chem.*, **1948**, 52, 1082.
156. Boncev,D.;Trinajstic, N. *J. Chem.Phys.*, **1977**, 67,(10), 4517.
157. Rouvray, D.H.; *Math. Chem.* **1975**, 1, 125.
158. Platt, J.R. *J.Chem.Phys.* **1947**, 15, 419.
159. Altenburg, K. *Kolloid.Z.*, **1961**,178,112.
160. Altenburg, K. *Z.Phys.Chem. (Leipzig)* **1980**, 261,289.
161. Hosoya, H.; *Bull.Chem.Soc.Japan*, **1971**, 44, 255.
162. Hosoya, H.; Kawasaki, K.; Mitzutani, K. *Bull.Bhem.Soc. Japan*,**1972**, 42, 3415.
163. Narumi, H.; Hosoya, H. *Bull.Chem.Soc. Japan* **1980**, 53, 1228.
164. Gutman, I.; Riscic, M.; Trinajstic, N.; Wilcox, C.F. *J.Chem.Phys.*,**1975**, 62, 3399.
165. Gutman, I.; Randić, M. *Chem.Phys.Lett.* **1977**, 47, 15.
166. Randić, M.; Wilkins, C.L. *Chem.Phys.Lett.* **1979**, 63, 332.
167. Randić, M.; Wilkins, C.L. *J.Phys.Chem.* **1979**, 83, 1525.
168. Gutman, I.; Trinajstic, N. *J.Chem.Phys.Lett.* **1972**, 17, 535.
169. Randić, M. *J. Amer.Chem.Soc.* **1975**, 97, 6609.
170. Bonchev, D.; Mekenyan, O.;Trinajstic, N. *J.Chem.Inf.Comput.Sci.* **1981**, 2, 127.
171. Balaban, A.T. *Chem.Phys.Lett.* **1982**, 89, 399.
172. Balaban, A.T. *Commun.Math. Chem.*,**1983**, 14, 213..
173. Medeleanu,M.; Geantă,M.;Ciubotaru,D. *Quant.Struct.-Act.Relat.*, **1996**. (sub tipar).
174. Kier, L.B.;Hall,L.H.;Murray,W.J.;Randić,M. *J. Pharm.Sci.*, **1975**, 64,1971.
175. Kier, L. B.; Hall,L. *J. Pharm. Sci.*, **1981**, 70, 583.
176. Chiriac, A.;Ciubotaru,D.; Simon,Z. *Relații Cantitative Structură Chimică - Activitate Biologică(QSAR). Metoda MTD.*, Ed. Mirton, Timișoara, **1996**.
177. Kier, L.B.; Hall, L.H. *Molecular Connectivity in Chemistry and Drug Research*, Acad. Press, New York,**1976**.
178. Hall,L.H. *J.Chem.Comput.Sci.*, **1995**, 35, 1039.
179. Vassiliev,Y.B.;Bagotzky,Y.B.;Krasnova,N.;Sergeeva,A.;Bogdanovich.B. *Electrochem. Acta*, **1981**, 26, 545.
180. Vassiliev,Y.B.; Bagotzky,V.S.; Khazova.O.A.; Yastrebova.T.N.; Sergeeva.T.A. *Electrochim. Acta*, **1981**, 26, 563.
181. Nolen,T.R.; Fedkiw,P.S. *J.Appl.Electrochem.* **1990**, 20, 370.

182. Marquez, J.; Pletcher, D. *J. Appl. Electrochem.*, **1980**, *10*, 567.
183. Smeltzer, J.C.; Fedkiw, P.S. *J. Electrochem. Soc.*, **1992**, *139*, 1358.
184. Ravichandaran, C.; Noel, M.; Anantharaman, P.N. *J. Appl. Electrochem.*, **1994**, *24*, 1256.
185. Subbiah, P.; Subramanian, G.S.; Udupa, H.V.K. *Electrochim. Acta*, **1966**, *7*, 863.
186. Vaszilcsin, N.; Nemes, N.; Olah, F. *Chem. Bull. Politehn. Univ. Timisoara*, **1996**, *41(55)*, 88.
187. Vaszilcsin, N.; Nemeş, Maria; Geantă, M.; Birzescu, M. *Chem. Bull. Politehn. Univ. Timișoara*, **1996**, *41(55)*, 97.
188. Geantă, M. *Sem. Nat. Ing. Chim.*, Timișoara, **1996**, 108.
189. Coheci, V.; Strâmbeanu, N.; Geantă, M.; Păunescu, D. *Proceedings of the Scientific Communications Meeting of "Aurel Vlaicu" University, Third edition, Arad, 1996*, Vol 11, 117.
190. Coheci, V.; Strâmbeanu, N. *Proceedings of the Scientific Communications Meeting of "Aurel Vlaicu" University, Third edition, Arad, 1996*, Vol 11, 111.
191. Coheci, V.; Strâmbeanu, N.; Lungu, E. *Proceedings of the Scientific Communications Meeting of "Aurel Vlaicu" University, Third edition, Arad, 1996*, Vol 11, 129.
192. Gomella, C. *TSM L'eau*, **1975**, *9*, 182.
193. Rosen, G. *J. Chem. Phys.*, **1974**, *1*, 3676.
194. Benedict, M.; Pigford, T. *Nuclear Chemistry Engineering*, McGraw Hill, New York, **1987**.
195. Brăteanu, C. and all, *Strategii și filiere nucleare*, Ed. Tehnică, București, **1990**, 110.
196. Cecal, A. *Aspecte chimice ale energiei nucleare*, Ed. Academiei, București, **1987**, 187.
197. *** Degremont, *Memento Technique de l'eau*, ed. 50th, Reuil-Malmaison Cedex, **1989**.
198. Ungureanu, D. and all, *Hidrodinamica mediilor poroase neomogene*, Ed. Tehnică, București, **1989**, 229.
199. Roger, G.; Franks, E. "Modelarea și simularea în ingineria chimică", Editura Tehnică, București, **1979**.
200. Thompson, L.M. *Theoretical Hydrodynamics*, Mcmilland, London, **1960**.
201. *** *The PUROLYTE Co*, Division of Bro Tech Corp., London, **1991**.
202. Liteanu, C. *Chimie Analitică Cantitativă*, Ed. Did. Ped., București, **1973**, 591.
203. Ruscescu, C. Tudose, C. *Modele și modelare în fizică*, Ed. Șt. și Enc., București, **1987**, 12.
204. Perju, Delia *Chem. Bull. Techn. Univ. Timișoara*, **1993**, *38(52)*, 52.
205. Perju, Delia *Chem. Bull. Politeh Univ. Timișoara*, **1995**, *40(54)*, 1.
206. Perju, Delia; Kohn, D. *Chem. Bull. Pliteh. Univ. Timișoara*, **1996**, *41(55)*, 105.
207. Kohn, D.; Perju, Delia; Popa, Simona *Anal. West. Univ. Timișoara*, **1996**, *5*, 55.
208. Perju, Delia; Geantă, M.; Magda, A. *Sem. Nat. Ing. Chim.*, Timișoara, **1996**, 45.
209. Perju, Delia; Cofan, N. *Bull. Appl. Math.*, PAMM. Budapest, **1992**, *3*, 225.
210. Perju, Delia; Klepp, F. *Bull. Appl. Math.*, PAMM. Budapest, **1992**, *3*, 234.

CUPRINS

	pagina
INTRODUCERE	1
I. REȚELE NEURALE	3
I.1 Inspirația din Neurologie	3
I.1.1 Celula nervoasă	3
I.1.2 Sinapsa	4
I.1.3 Neuronul - unitate funcțională	4
I.1.4. Concluzii	5
I.2 Modele neurale artificiale	6
I.2.1. Modelul McCulloch & Pitts	6
I.2.2. Modelul Hebb	8
I.2.3. Modelul Rosenblatt - Perceptronul	10
I.2.3.1 Regula Delta	12
I.2.3.2 Problema XOR	15
I.2.3.3 Perceptronul multistrat	16
I.2.4. Modelul Hopfield	17
I.2.4.1 Modelul fizic Ising	18
I.2.4.2 Modelul Hopfield propriu-zis	18
I.2.5 Rețele cu propagare înapoi (<i>Backpropagation Networks</i>)	22
I.2.5.1 Generalizarea regulii delta	23
I.2.5.2 Modul de lucru cu <i>Back-propagation</i>	26
I.2.5.3 Alte funcții de activare neliniare	27
I.2.5.4 Rata de învățare și momentul	29
I.2.5.5 Invățarea unui tipar	29
I.2.5.6 Deficiențe ale algoritmului Back-propagation	30
I.2.5.7 Imbunătățiri ale algoritmului Backpropagation	31
I.2.5.8 Parametrii de antrenare a rețelei	34
I.2.5.9 Accelerarea convergenței prin utilizarea unei rate de învățare adaptive	35
I.2.6. Aplicații ale rețelelor neurale cu propagare înapoi în chimie și inginerie chimică	36
II. CORELĂRI STRUCTURĂ ELECTRONICĂ - PROPRIETĂȚI FIZICE	38
II.1 Noțiuni generale despre proprietățile fizice utilizate	38
II.1.1 Căldura specifică	38
II.1.2 Electronegativitatea elementelor	38
II.1.3 Aplicații existente în domeniu	39
II.2 Corelare structură electronică - căldură specifică	41
II.2.1 Alegerea și justificarea variabilelor de corelare	41
II.2.2 Justificarea utilizării modelului neural	41
II.2.3 Modelul neural - Caracterizare	43
II.2.3.1 Numărul straturilor ascunse și al unităților ascunse	43
II.2.3.2 Selectia tiparelor	44
II.2.3.3 Descrierea rețelei neurale utilizate	46
II.2.4 Rezultate experimentale	47
II.2.5 Validarea modelului neural	51
II.2.6 Predicții realizate pe baza modelului neural	53
II.2.7 Concluzii	53

II.3 Corelare structură electronică - electronegativitate	55
II.3.1 Alegerea și justificarea variabilelor de corelare	55
II.3.2 Justificarea utilizării modelului neural	55
II.3.3 Modelul neural - Caracterizări	56
II.3.3.1 Numărul straturilor ascunse și al unităților ascunse	56
II.3.3.2 Selecția tiparelor	57
II.3.3.3 Descrierea rețelei neurale utilizate	59
II.3.4 Rezultate experimentale	60
II.3.5 Validarea modelului	63
II.3.6 Predicții realizate pe baza modelului neural	65
II.3.7 Concluzii	66
III CORELĂRI STRUCTURĂ CHIMICĂ- PROPRIETĂȚI FIZICE	67
III.1 Introducere	67
III.2 Aplicații existente în domeniu	68
III.3 Relații numerice în teoria grafurilor	68
III.4. Indici topologici	71
III.4.1. Indicele de adiacență totală	72
III.4.2. Indicii Wiener, Rouvray-Crawford și de polaritate	72
III.4.3. Indicele Platt	73
III.4.4. Indicele polinomial Altenburg	73
III.4.5. Indicele Hosoya	74
III.4.6. Indicii M_1 și M_2	74
III.4.7. Indici de tip Randic	75
III.4.8. Indicele J	75
III.4.9. Indici derivați din indicii J și χ	76
III.4.10. Indicii Kier și Hall	77
III.5. Concluzii	77
III.6. Aplicarea indicilor topologici pentru corelarea punctelor de fierbere ale alcoolilor alifatici cu structura lor chimică	78
III.6.1. Introducere	78
III.6.2. Considerații privind alegerea variabilelor utilizate în corelare	83
III.6.3. Descrierea modelului neural	83
III.6.4. Corelarea temperaturilor de fierbere a alcoolilor alifatici cu indici topologici	87
III.6.5. Studiu corelativ între discriminanți structurali și punctele de fierbere ale alcoolilor alifatici	91
III.6.6. Validarea modelului	96
III.7. Concluzii	97
IV UTILIZAREA REȚELOR NEURALE ARTIFICIALE LA MODELAREA UNUI PROCES ELECTROCHIMIC	99
IV.1 Chimismul procesului	99
IV.2 Descrierea instalației experimentale și selecția parametrilor de proces	100
IV.3 Rezultate experimentale	102
IV.4 Modelarea procesului	103
IV.5 Validarea modelului	106
IV.6 Decelarea influenței parametrilor de proces	108
IV.7 Concluzii	114

V MODELAREA DINAMICĂ A PROCESULUI DE REȚINERE A Sr^{2+} PRIN SCHIMB IONIC	115
V.1 Noțiuni generale despre schimbul ionic	115
V.1.1 Echilibrul în schimbul ionic	115
V.1.2 Curbe izoterme	118
V.1.3 Schimbul ionic și cinetica de adsorbție	119
V.2 Determinări experimentale	120
V.3 Modelarea procesului	125
V.4 Validarea modelului	132
V.5 Testarea capacității de predicție a modelului	133
V.5 Concluzii	136
CONCLUZII FINALE	137
BIBLIOGRAFIE	147
CUPRINS	153