APLIKACE UMĚLÝCH NEURONOVÝCH SÍTÍ V CHEMICKÉ TECHNOLOGII

TOMÁŠ HERINK a ZDENĚK BĚLOHLAV

Ústav organické technologie, Vysoká škola chemicko-technologická, Technická 5, 166 28 Praha 6
e-mail: Herinkt@earthling.net, Zdenek.Belohlav@vscht.cz

Došlo dne: 14.IV.1999

Klíčová slova: umělé neuronové sítě, modelování, optimalizace, řízení, regulace

Obsah

1. Úvod
2. Umělá neuronová síť
3. Přehled aplikací
   3.1. Dopředně neuronové sítě
   3.2. Rekurentní neuronové sítě
4. Závěr

1. Úvod

Umělé neuronové sítě se díky svému širokému rozsahu uplatnění a efektivity stávají středem pozornosti v oblasti chemické technologie. Pro své schopnosti transformace vstupních signálů na výstupní bez znalosti vzájemných relací je možné využít neuronových sítí v oblastech predikce číselných a časových řad, klasifikace a filtrace datových souborů, modelování, optimalizace, řízení a regulace. Průmyslový provoz poskytuje nesčetné množství dat, které mohou být dostatečným zdrojem informací pro učení a trénování neuronových sítí. Neuronová síť naučená na správná a přesně měřená provozní data se tak může stát účinným nástrojem procesních inženýrů[1-4].

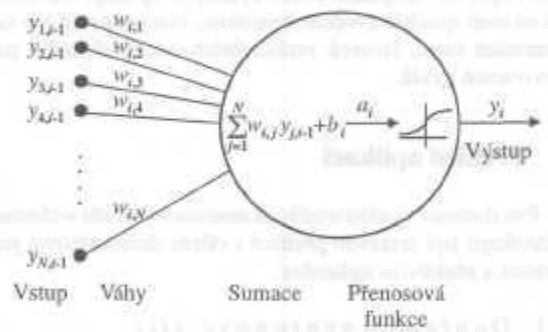
2. Umělá neuronová síť

Umělé neuronové sítě vycházejí z podobnosti s biologickým nervovým systémem. Základní stavební prvky nervové soustavy a především mozku tvoří neurony. Jsou to živé buňky, které slouží ke sběru, uchovávání, zpracování a přenosu informací. Každý neuron se skládá z těla (somy), do kterého přicházejí informace po vstupních větvích (dendritech), a z kterého informace vystupují po jediném výstupu (axonu). Výstupní signály axonů odpovídají vstupům dalších neuronů. Neurony jsou navzájem propojeny do neuronových sítí. Propojení je uskutečňováno axony, které se pomocí speciálních výběžků (synapsí) připojují na dendrity jiných neuronů. Na jeden neuron připadá od desítek do několika tisíc spojů s jinými neurony. Přenos signálu je zprostředkováván elektrickými pulsy, které však aktivují neuron až po překročení určité hodnoty elektrického potenciálu (prahu).

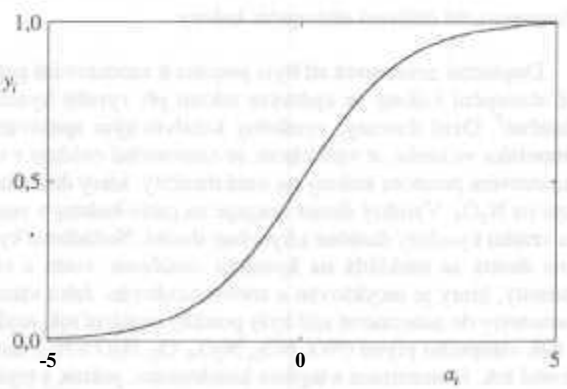
Základním prvkem umělých neuronových sítí je formální neuron (obr. 1). Skládá se z těla, dendritů přenášející signál směrem k tělu a axonu přenášející signál od těla neuronu. Každá umělá neuronová síť obsahuje několik vrstev formálních neuronů. Napojení na dendrit jiného neuronu tvoří informační rozhraní a nazývá se synapse. Přicházející signál je v synapsi zvětšen, zmenšen popř. negován vynásobením vahami. Produkt násobení je dendritem veden do těla neuronu, kde se signály z jednotlivých dendritů sečtou. Vzniklý signál, který lze považovat za celkové podráždění neuronu (aktivaci neuronu), je porovnáván s prahem a překročí-li jeho hodnotu, dochází k sumaci podle rovnice

$$a_i = \sum_{j=1}^{N} w_{i,j}\, y_{j,i-1} + b_i \qquad (1)$$

kde $y_{j,i-1}$ je signál z neuronu v předcházející vrstvě, $w_{i,j}$ a b_i jsou váhy a prahy i-té vrstvy neuronové sítě. Hodnota a, je pak transformována přenosovou funkcí na výstupní signál neuronu (y_i), který je axonem přenášen do jiného neuronu popř. přímo představuje požadovanou výstupní veličinu. Přenosová funkce, označovaná též jako aktivační, převádí hodnoty sumací do definovaného oboru výstupních hodnot. Nejběžněji používanou přenosovou funkcí neuronových sítí je sigmoida (obr. 2), která vymezuje signálu hodnoty z otevřeného intervalu (0, 1).



Obr. 1. **Schéma formálního neuronu**



Obr. 2. Sigmoida — přenosová funkce neuronu

Strukturu neuronové sítě charakterizuje její typ a geometrie (obr. 3). Typy udávají zásadní způsob uspořádám a propojení neuronů. Signál je rozváděn k vrstvám neuronové sítě prostřednictvím vrstvy neuronů označované jako vstupní, kde každý neuron představuje pouze jeden vstup. Vstupní vrstva má jen distribuční funkci a není zahrnuta do celkového počtu vrstev. Poslední vrstva se nazývá výstupní a vrstvy mezi vstupní a výstupní vrstvou jsou označovány jako skryté.

Podle způsobu přenosu signálu mezi vstupy a výstupy se rozlišují dopředně sítě (obr. 3a), u kterých se signál šíří pouze jedním směrem od vstupní vrstvy přes skryté k výstupní, a sítě zpětnovazebně (rekurentní) (obr. 3b) s šířením signálu směrem od vstupní vrstvy a současně z výstupu n-té vrstvy na vstupy předchozích vrstev. Nejjednodušší strukturu vykazuje dopředná neuronová síť bez skrytých vrstev, která se nazývá perceptron[5].

Neuronové sítě aplikované v oblastech dynamického modelování jsou označovány jako dynamické neuronové sítě. Jedná se především o zpětnovazebně neuronové sítě.

Geometrie vymezuje detailní uspořádání neuronové sítě a bývá zpravidla vyjadřována např. schématem 8-12-6, které označuje dvouvrstvou síť s osmi vstupy, dvanácti neurony v první (skryté) vrstvě a šesti neurony ve druhé (výstupní) vrstvě neuronové sítě.

Základní vlastností neuronových sítí je schopnost učit se ze vzorů. Proces učení (trénování) spočívá v předkládání vzorů (trénovacích dat) neuronové síti a vhodném nastavování vah a prahů vedoucí k minimalizaci sumy čtverců reziduí daných a vypočtených hodnot. Pro tento účel se používají různé učící algoritmy, z nichž nejběžnějším je algoritmus zpětného šíření[6]. Algoritmus zpětného šíření po vypočtení sumy čtverců chyb upravuje jednotlivé váhy a prahy směrem od výstupních vrstev zpět ke vstupním. Tento cyklus se opakuje tak dlouho, dokud není splněno zvolené kritérium, kterým může být např. minimální suma čtverců reziduálních odchylek nebo počet trénovacích cyklů.

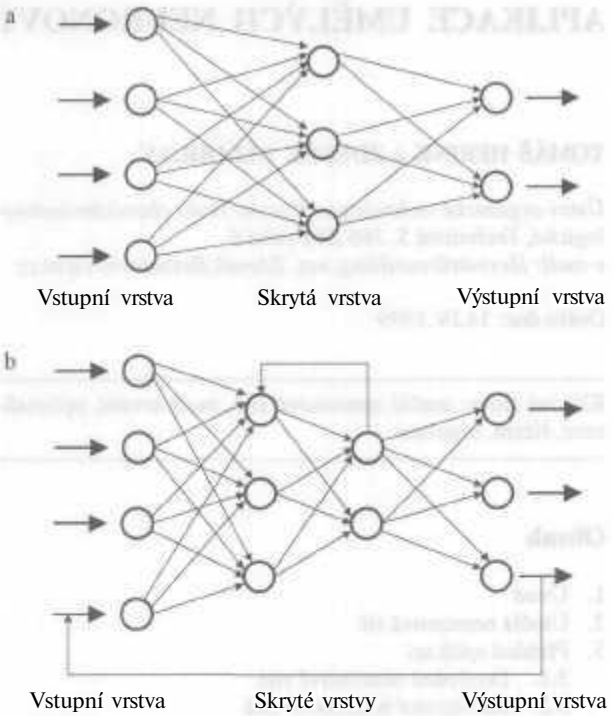
3. Přehled aplikací

Pro ilustraci využití umělých neuronových sítí v chemické technologii byl sestaven přehled s cílem demonstrovat jejich pestrost a efektivitu uplatnění.

3.1. Dopředně neuronové sítě

Parametrická citlivost absorpční kolony

Dopředná neuronová síť byla použita k modelování patrové absorpční kolony se zpětným tokem při výrobě kyseliny dusičné[7]. Oxid dusnatý, vyráběný katalytickým spalováním amoniaku ve směsi se vzduchem, se samovolně oxiduje v mezipatrovém prostoru kolony na oxid dusičitý, který dále dimeruje na N_2O_4. Vzniklý dimer reaguje na patře kolony s vodou za vzniku kyseliny dusičné a kyseliny dusité. Nežádoucí kyselina dusitá se rozkládá na kyselinu dusičnou, vodu a oxid dusnatý, který je recyklován a znovu oxidován. Jako vstupní parametry do neuronové sítě byly použity molární tok, teplota a tlak vstupního plynu (NO, NO_2, N_2O_4, O_2, H_2O a N_2), hmotnostní tok, koncentrace a teplota kondenzátu, průtok a teplota procesní vody a teplota chladící vody. Za výstupy ze sítě byly



Obr. 3. Neuronová síť; a) dopředná, b) rekurentní

zvoleny molární tok výstupního plynu obsahující NO, NO_2, O_2, H_2O a N_2, hmotnostní tok, koncentrace a teplota kyseliny dusičné, teplota procesní vody, výstupní teplota a tlak plynu. Navržená struktura sítě měla ve vstupní i skryté vrstvě 14 neuronů a ve výstupní 10 neuronů. Jako přenosovou funkci zvolili autoři standardní sigmoidu. Pro trénování sítě byl použit algoritmus zpětného šíření. Síť byla trénována na sadě padesáti dat simulovaných analytickým matematickým modelem, která pokrývala 70–110 % výrobní kapacity a tlaky od 1 do 1,5 MPa. Výsledkem simulací byla závislost koncentrace oxidů dusíku na molárním toku vstupního plynu a tlaku plynu, závislost produkce 100 %-ní kyseliny dusičné na molárním toku vstupního plynu a hmotnostním toku procesní vody a závislost koncentrace kyseliny dusičné na molárním toku vstupního plynu a hmotnostním toku procesní vody.

Modelování krakovacích procesů

McGreavy se spolupracovníky[8] navrhli neuronový model jednotky fluidního katalytického krakování. Krakovací proces štěpení těžkých ropných destilátů na cennější benzin a další lehčí produkty je velice obtížné popsat analytickým matematickým modelem. Byla proto sestavena jedna neuronová síť na predikci hustoty lože reaktoru a regenerátoru deaktivovaného katalyzátoru a druhá síť předpovídající složení produktu. Hustota lože byla predikována sítí o struktuře 6-3-1, s rychlostí a viskozitou fluidizačního média, hustotou katalyzátoru, efektivním průměrem katalyzátoru, sypnou hustotou katalyzátoru a tlakovou ztrátou jako vstupy. Simulace ukázaly, že hustotu lože lze predikovat s relativní chybou do 11 %. Vstupy do sítě na predikci složení produktu zahrnovaly vlastnosti nastřikované suroviny (hustota, obsah uhlíku a obsah síry v surovině,

destilační teplota na počátku, při 10 % a při 50 % oddestilované suroviny), vlastnosti katalyzátoru (poměr povrchu k hmotnosti čerstvého katalyzátoru, poměr povrchu k hmotnosti bilancovaného katalyzátoru) a operační podmínky (reakční teplota, teplota předehřívané suroviny, reakční tlak, poměr hmotnosti katalyzátoru k hmotnosti nástřiku a poměr hmotnosti katalyzátoru k hmotnosti zpětného toku). Výstupní data ze sítě tvořily koncentrace plynných uhlovodíků, kapalných uhlovodíků, benzinů, lehké nafty a krakovacího zbytku. Optimální struktura neuronové sítě byla nalezena ve formě 14-8-5. Pro učení a trénování sítě byl zvolen algoritmus zpětného šíření. K trénování byla použita v obou případech data z mnoha průmyslových závodů. Simulace s náhodně vybranou sadou 29 dat vykázaly v porovnání s měřenými daty průměrnou chybu u plynných uhlovodíků 8,80 %, kapalných uhlovodíků 4,07 %, benzínů 0,77 %, lehké nafty 0,86 % a u krakovacího zbytku 0,32 %.

V jiné práci[9] byla vyvinuta strategie pro diagnózu chyb a poruch procesu katalytického hydrokrakování těžkých olejů. V první fázi neuronová síť naučená na spolehlivá provozní data detegovala prostřednictvím výstupů šest možných zdrojů chyb nebo poruch - vysoký nebo nízký tok suroviny do reaktoru, vysoká nebo nízká teplota na vstupu reaktoru a lehká nebo těžká vstupní surovina. Vstupy neuronové sítě tvořilo osm výstupních parametrů z reaktoru, zejména výstupní teplota a průtok vodíku mezi dvěma loži reaktoru. Síť s dvaceti neurony ve skryté vrstvě využívala jako přenosovou funkci spojení sigmoidy s vlnovou funkcí. Na základě informací poskytnutých neuronovou sítí byla vyhodnocena kvalita produktu a režim reaktoru interpretací osmnácti pravidel o kvalitě a kvantitě produktů sestavených na základě zkušeností, výzkumu a matematického modelování. Poté následovala zpětná úprava řídících parametrů až do stabilizace systému. Ukázalo se, že uvedená strategie je schopná téměř s naprostou přesností korigovat chyby a poruchy procesu.

K detekci a diagnóze chyb při řízení teploty regenerátoru deaktivovaného katalyzátoru v procesu krakování těžkých ropných destilátů[10] byly použity vícevrstevná neuronová síť a jednovrstevný perceptron. Trénovácí data byla vypočítána vyvinutým dynamickým simulátorem, který byl schopný přesně a spolehlivě odhadnout nouzové zacházení, optimální chod, najíždění i ukončení procesu regenerace katalyzátoru. Vstupy do sítě tvořily aktuální hodnoty regenerátoru (teplota, tlak), trendy těchto veličin od neustáleného do naprosto stabilního stavu a maximální a minimální hodnoty teploty, tlaku, obsahu kyslíku, oxidu uhelnatého a množství katalyzátoru, které jsou limitní pro ustálený režim regenerátoru. Výstupy tvořily potenciální zdroje nežádoucího chování procesu, tj. maximální a minimální hodnoty regulovatelných veličin pro dané vstupní podmínky (např. nástřik, průtok vzduchu, velikost zpětného toku, teplota předehřáté suroviny, otevření ventilů, činnost kompresoru). Pro srovnání přesnosti detekce jednotlivých metod byla vygenerována sada 67 dat, v nichž některá výstupní hodnota (popř. i více) byla chybná. Takto vygenerovaná sada dat byla simulována testovanými sítěmi a metodou využívající fuzzy množinu[11]. Výsledky ukázaly, že vícevrstevná neuronová síť odhalila chybu v 65 případech, kdežto perceptron v 55 a metoda využívající fuzzy množinu v 53 případech.

Lu a Hua[12] se zabývali optimalizací ethylenové pyrolýzy uhlovodíkových směsí s vodní parou prostřednictvím dopředných neuronových sítí. V modelu použili strukturu sítě se 4 neurony ve vstupní, 2 neurony ve skryté a 1 neuronem ve výstupní vrstvě. Vstupní data pro učení a trénování sítě zpětným šířením tvořily teplota na výstupu z vlásenky, tlak, ředící poměr pára-surovina a hustota suroviny, výstupní data ze sítě představovaly hmotnostní koncentrace ethylenu v produktu. Simulace prokázaly, že provozní parametry ethylenové pyrolýzy mohou být rychle a efektivně optimalizovány umělou neuronovou sítí.

Optimalizace dřevozpracujících procesů

Aguiar a Filho[13] se zabývali modelováním a optimalizací chemických procesů na výrobu papírové drtě ze dřeva. Makroskopická struktura dřeva je tvořena vláknitými buňkami obsahujícími kromě celulosy a hemicelulosy lignin. K odstranění ligninu se používá delignifikační činidlo (směs hydroxidu sodného a sulfidu sodného), které štěpí vazby molekul ligninu a rozpouští ho, zároveň však způsobuje destrukci celulosy a hemicelulosy. Úkolem pro optimalizaci procesu bylo nalézt nejvhodnější složení varného roztoku a reakční teplotu tak, aby rychlost delignifikace byla co nejvyšší a destrukce celulosy a hemicelulosy minimální. Výsledky výpočtů neuronovými sítěmi byly porovnány s výsledky simulací s analytickým matematickým modelem navrženým Miramsem a Nguyenem[14], zahrnující popis kinetických a difuzních jevů. Vstupy do neuronové sítě tvořily koncentrace hydroxidu sodného a hydrogensulfidových aniontů, tloušťka dřevěného štěpu, poměr objemu varného roztoku ku hmotnosti dřeva, počáteční teplota, teplota varu, hustota a porozita dřeva a obsah ligninu, celulosy a hemicelulosy. Dopředná síť obsahovala jednu skrytou vrstvou a sigmoidy jako přenosové funkce v neuronech. Výpočty s umělou neuronovou sítí prokázaly dobrou shodu s chováním analytického modelu.

Složitý analytický model jednotky na zpracování dřevěných štěpů nahradili umělou neuronovou sítí Qian a Tessier[15]. Na základě operačních podmínek (rychlost produkce, měrná energie a hutnost) predikovali vlastnosti papírové drtě (délka vlákna, odvodňovací schopnost, tržná délka a součinitel dotržení). K dispozici měli sadu 110 spolehlivých průmyslových dat, kterou rozdělili na část trénovací a část testovací. Při studiu vlivu počtu trénovacích dat na přesnost predikce si ověřili, že čím více je k dispozici trénovacích dat, tím přesnější jsou výsledky prezentované neuronovou sítí. Dále zjistili, že příliš složitá struktura sítě vykazuje velkou senzitivitu a není schopná dobře aproximovat oblast, která není zastoupená v trénovací sérii. Jako optimální sítě pro jejich sadu dat navrhli struktury sítí 3-2-5 nebo 3-3-5. Při porovnání kvality predikcí poskytované pěti jednotlivými sítěmi struktury 3-2-1 ajednou sítí struktury 3-2-5 se ukázaly jako nepatrně spolehlivější jednotlivé sítě. S ohledem na výpočetní čas však bylo vhodnější použít jednu síť s více výstupními neurony.

Analýza a optimalizace rektifikace

Dopředně neuronové sítě byly použity k analýze a optimalizaci dělení heterogenního azeotropu (butylacetát-butylalkohol–voda) dvoustupňovou rektifikací se zařazeným separátorem dvou kapalných fází[16]. Byly navrženy dvě neuronové sítě. První síť s pěti vstupy (podíly butylacetátu, butylalkoholu a vody v odpadním proudu, tlaková ztráta v koloně a refluxní poměr) a třemi výstupy (složení butylacetátu v horní a dolní vrstvě separátoru a složení butylalkoholu) popisovala rektifi-

kační kolonu a separátor, jehož produktem byl butylalkohol. Druhá neuronová síť s osmi vstupy a dvěma výstupy aproximovala dvoukolonový systém rektifikace daného azeotropu. Vstupy reprezentovaly podíly butylacetátu, butylalkoholu a vody v odpadním proudu, tlaková ztráta v první a ve druhé koloně, refluxní poměr první a druhé kolony a vnější refluxní poměr. Výstupy tvořilo složení konečných produktů butylacetátu a butylalkoholu. Prostřednictvím neuronových sítí byl testován vliv počtu neuronů ve skryté vrstvě na přesnost aproximace a porovnáván výkon sítě v podobě sumy čtverců chyb dvou různých učících algoritmů - zobecněného delta pravidla[6] a Powellova algoritmu[17]. Získané výsledky jednoznačně upřednostnily delta pravidlo a ukázaly použitelnost neuronových struktur se skrytou vrstvou i bez skryté vrstvy, avšak na úkor přesnosti odhadu procesních parametrů.

Aoyama a Walsh[18] se věnovali modelování obsahu C_5 frakce ve výstupu z debutanizéru katalytického reformátoru. K dispozici měli sadu provozních dat z debutanizéru odebíraných v pětiminutových intervalech po dobu jednoho měsíce. Data rozdělili na část trénovací a testovací. Vstupy do neuronové sítě tvořily refluxní poměr, výtěžek LPG, teplota ve vařáku, průtok suroviny, průtok chladící kryogenní směsi, teplota chladící kryogenní směsi a okolní teplota. Výstup představoval obsah C_5 v LPG. Ukázalo se, že použití všech sedmi veličin jako vstupů do neuronové sítě nepřináší uspokojivé výsledky a suma čtverců chyb po natrénování sítě je příliš vysoká. Bylo proto nutné identifikovat proměnné, které největší měrou ovlivňují chování debutanizéru. Analýzou vztahů mezi vstupy a výstupy byly jako vstupní veličiny zvoleny refluxní poměr, výtěžek LPG a okolní teplota. Redukce vstupů výrazně snížila sumu čtverců chyb po natrénování a zpřesnila výsledky predikované neuronovou sítí.

Modelování a optimalizace chemických reaktorů

Krothapally a Palanki[19] testovali použití umělých neuronových sítí při modelování průmyslových vsádkových procesů polymerace styrenu a methylmethakrylátu. Analytické modely těchto procesů, založené na numerickém řešení soustav diferenciálních rovnic, byly z důvodu náročnosti výpočtu nevhodné pro přímé řízení optimálního teplotního profilu polymeračních reakcí. Autoři proto navrhli dvě dvouvrstvé neuronové sítě s 24 neurony ve skrytých vrstvách, které trénovali daty simulovanými matematickými modely v širokém rozmezí operačních podmínek. Ukázalo se, že neuronové sítě byly schopné predikovat teplotní profily s přesností srovnatelnou s analytickými modely, čas potřebný pro výpočet byl však výrazně nižší.

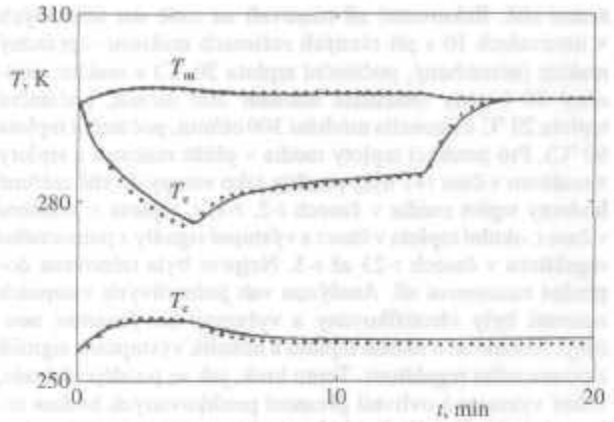
Umělé neuronové sítě využili k odhadu stupně polykondenzace ve vsádkovém reaktoru Draeger a Gesthuisen[20]. V izotermním míchaném reaktoru probíhala v přítomnosti katalyzátoru a stabilizátoru poly kondenzace oligomeru bis-(2,2)-hydroxyethyltereftalátu na polyethylentereftalát. Experimentální stanovení stupně polymerace je z hlediska řízení reaktoru příliš zdlouhavé. Empirický nebo semiempirický odhad stupně polymerace proto vychází z běžně dostupného měření počtu otáček a otáčivého momentu míchadla a teploty reakční směsi. K modelování reakčního systému byla použita dopředná neuronová síť s dvěma skrytými vrstvami 2-3-3-1. Vstupy tvořily podíl otáčivého momentu a počtu otáček míchadla a teplota reakční směsi, jediný výstup odhadovaný stupeň polymerace. Použití podílu na vstupu bylo motivováno stabilizací vstupu do neuronové sítě s ohledem na kolísám počtu otáček během polykondenzace. K trénování sítě bylo k dispozici 60 měření, k testování bylo použito 25 dat. Výsledky simulací byly kromě nízkých stupňů polymerace plně srovnatelné s experimentálními daty i s výpočty semiempirického regresního modelu.

Cílem práce Milaniče a kol.[21] bylo predikovat hodnoty rychlosti srážení TiO_2 při výrobě titanové běloby. Z důvodu nedostatku provozních dat bylo nutné použít semiempirického modelu, s jehož pomocí byla dopočítána podstatná část trénovacích dat. Vstupy neuronové sítě tvořily počáteční koncentrace TiO_2, účinná a volná kyselost, teplota varu, přídavek krystalových zárodků a průtok vody. Výstup představovala rychlost srážení TiO_2. Pokusně určili 23 neuronů ve skryté vrstvě. Testování proběhlo na sérii naměřených provozních dat (6 experimentů po 8 měřeních). Výsledky simulací ukázaly, že semiempirický model předpovídá v důsledku nedostatečné přesnosti vstupních trénovacích dat požadované hodnoty nepatrně přesněji než umělá neuronová síť.

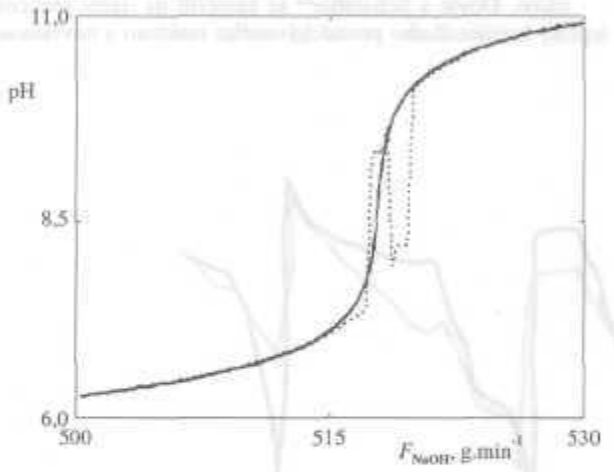
Řízení chemických reaktorů

Zaldívar a Hernandes[22] použili neuronových sítí jako regulátoru teploty reakčního kalorimetru. Řízení tepoty v izotermických, adiabatických a teplotně programovatelných experimentech zajišťoval termostat s ohřívací a chladící smyčkou a se silikonovým olejem jako teplosměnným médiem. Silikonový olej ohřívacího okruhu procházel přes elektrický ohřívač a byl čerpán do pláště reaktoru. V chladícím okruhu olej procházel výměníkem s chladící spirálou a byl napojen na ohřívací okruh řídícím ventilem. Řízení teploty zajišťoval původně regulátor PI, který měnil výkon elektrického ohřívačem a průtok silikonového oleje z chladícího okruhu regulačním ventilem. Regulátor PI, jehož nevýhodou bylo časově náročné vyhodnocování složitého dynamického systému, byl nahrazen umělou neuronovou sítí. Neuronová síť byla trénována sadou různě simulovaných časových profilů normálních i abnormálních operačních stavů. Jako nejvhodnější se ukázala síť s dvěma skrytými vrstvami o struktuře 5-10-10-1. Vstup do neuronové sítě tvořily rozdíly požadované teploty a teploty v reaktoru, teplota v plášti reaktoru, teplota v chladícím okruhu, první časová derivace rozdílu teploty v reaktoru a požadované teploty a druhá derivace teploty v reaktoru podle času. Výstup představovala tepelná energie termostatu. Simulace prokázaly značnou přesnost regulace neuronovou sítí. Porovnání mezi experimentálními a simulovanými teplotními profily během izotermně řízené neutralizace kyseliny chlorovodíkové přidáváním roztoku hydroxidu sodného je zřejmé z obr. 4.

Dirion a kol.[23] se zabývali řízením teploty polokontinuálního reaktoru s exotermní reakcí. Kapalný reaktant byl do reaktoru s duplikátorem čerpán ze zásobníku. Teplosměnné médium cirkulovalo z duplikátoru přes elektrický ohřívač a protiproudý tepelný výměník. Vzájemnou koordinací ohřívače a výměníku byl zajišťován izotermní teplotní režim reaktoru. K modelování byly použity dvě dopředně neuronové sítě se šesti neurony ve skryté vrstvě (5-6-1 a 7-6-1), výstup představovala řídící proměnná s hodnotami od -1 do 1. Záporná hodnota aktivovala chlazení tepelným výměníkem a kladná hodnota ohřev elektrickým ohřívačem. V první fázi vývoje



Obr. 4. Izotermně řízená neutralizace; T_m, T_e, T_c teploty v reaktoru, teplosměrného média v plášti reaktoru a teplosměnného média v chladící smyčce, ——— experimentální profily, ... simulované profily



Obr. 5. **Predikce pH neuronovou sítí se 4 neurony ve skryté vrstvě;** ——— experimentální hodnoty, ... predikované hodnoty

modelu zajistili autoři trénovací data pomocí již vyvinutého lineárního modelu, kterým simulovali režim reaktoru nejprve pouze s rozpouštědlem a následně s vysoce exotermickou acidobazickou reakcí s periodou vzorkování 10 až 5000 s a obdrželi tak sérii 1000 dat. Vstupy do sítě představovala hodnota řídící proměnné v čase *t*-1, teplota v reaktoru v čase *t*-1 a *t*, teplota média v čase ř+1 a f+2 (v případě sítě 7-6-1 ještě teplota média v čase ř+3 a ř+4). V závěrečné fázi byla neuronová síť trénována na sadě dat, získané monitorováním manuálně řízeného reaktoru zkušenými operátory. Neuronová síť byla schopna s požadovanou přesností řídit izotermní režim polokontinuálního reaktoru.

Wang se spolupracovníky[24] se zabýval řízením pH směsi hydroxidu sodného a kyseliny octové v průtočném promíchávaném reaktoru. Průtok kyseliny o konstantní koncentraci byl zvolen neměnný a průtok hydroxidu jako regulovatelná veličina. V časových intervalech byly odebírány vzorky a výsledné pH roztoku bylo predikováno na základě znalosti předchozích dvou hodnot pH a průtoku hydroxidu. Vstupy do sítě tvořily pH roztoku v čase ř a ř-1 a průtok hydroxidu v čase ř a ř-1. Výstup ze sítě představovala hodnota pH v čase ř+1. Časový interval snímání těchto veličin zvolili 0,2 minuty a experimentálně určily 10 neuronů ve skryté vrstvě. Takto navržená struktura neuronová síť předpovídala velmi přesně výsledky v celém rozmezí pH 6,3–10,9. Při použití jednodušší struktury sítě, pouze se 4 neurony ve skryté vrstvě, nastaly v oblasti prudkého vzrůstu pH nežádoucí oscilace hodnot pH (obr. 5).

Identifikace nestacionárních procesů

Umělé neuronové sítě ve spojení s metodou odhadu parametrů mohou být využity pro identifikaci nelineárních a nestacionárních procesů[25]. K demonstraci bylo využito řízení koncentrace produktu exotermické reakce regulací průtoku teplosměnného média v kontinuálním míchaném reaktoru pomocí neuronové sítě se třemi vstupy (aktuální reakční teplota v čase *t,* reakční teplota v čase *t*-1 a koncentrace klíčové složky v čase ř-1), třemi neurony ve skryté vrstvě a jedním výstupním neuronem (koncentrace klíčové složky v čase ř). Parametry naučené neuronové sítě byly v průběhu provozu procesu aktualizovány novým trénováním on-line snímanými provozními daty. Váhy a prahy neuronové sítě byly adaptovány v celé síti nebo jen po jednotlivých vrstvách. Autoři porovnávali rychlost odezvy a přesnost predikce dynamické neuronové sítě bez zpětné aktualizace parametrů, sítě se zpětnou adaptací všech parametrů a sítě se zpětnou adaptací parametrů po jednotlivých vrstvách. Výsledky simulací ukázaly, že nejpřesnější predikce sledovaných veličin vykazovaly sítě se zpětnou aktualizací vah a prahů po jednotlivých vrstvách, které měly též rychlejší odezvu než sítě se zpětnou adaptací všech parametrů.

Regulace tlaku v rektifikační koloně

Turner a kol.[26] se zabývali řízením tlaku rektifikační kolony dělící směs uhlovodíků převážně se třemi atomy uhlíku a porovnávali výsledek s lineárním matematickým modelem a regulátorem PI. Z třiceti monitorovatelných proměnných vybrali pět, které použili jako vstupy do neuronové sítě - spodní odtah, průtok produktu, nástřik suroviny, teplotu chladící vody a tepelnou energii vařáku. Navržená síť měla čtyři neurony ve skryté vrstvě a jeden výstupní neuron popisující tlak v koloně. Pro trénování sítě byla použita sada spolehlivých průmyslových dat, k učení Levenbergova-Marquardtova optimalizační metoda[27]. V porovnání s regulátorem PI a lineárním matematickým modelem vykazovala neuronová síť nejrychlejší a nejpřesnější dosažení požadovaného tlaku.

Predikce koncentrace znečišiujících látek v ovzduší

Pro predikci koncentrací znečišťujících látek v ovzduší použili dopřednou neuronovou síť Mudrová a kol.[28] K dispozici měli data z 22 měřících stanic monitorujících oblast severních Čech. Data obsahující informace o koncentracích škodlivých látek v ovzduší snímaných po 30 minutách byla zatížena šumem. Analýzou měření vybraných z dané monitorovací stanice pomocí „wavelet" transformace[29] byly získány dvě série trénovacích dat (u jedné série byla odstraněna šumová složka) a zároveň byl proveden výběr vzorů sloužících jako vstupy do modelů. K předpovědi koncentrací znečišťujících

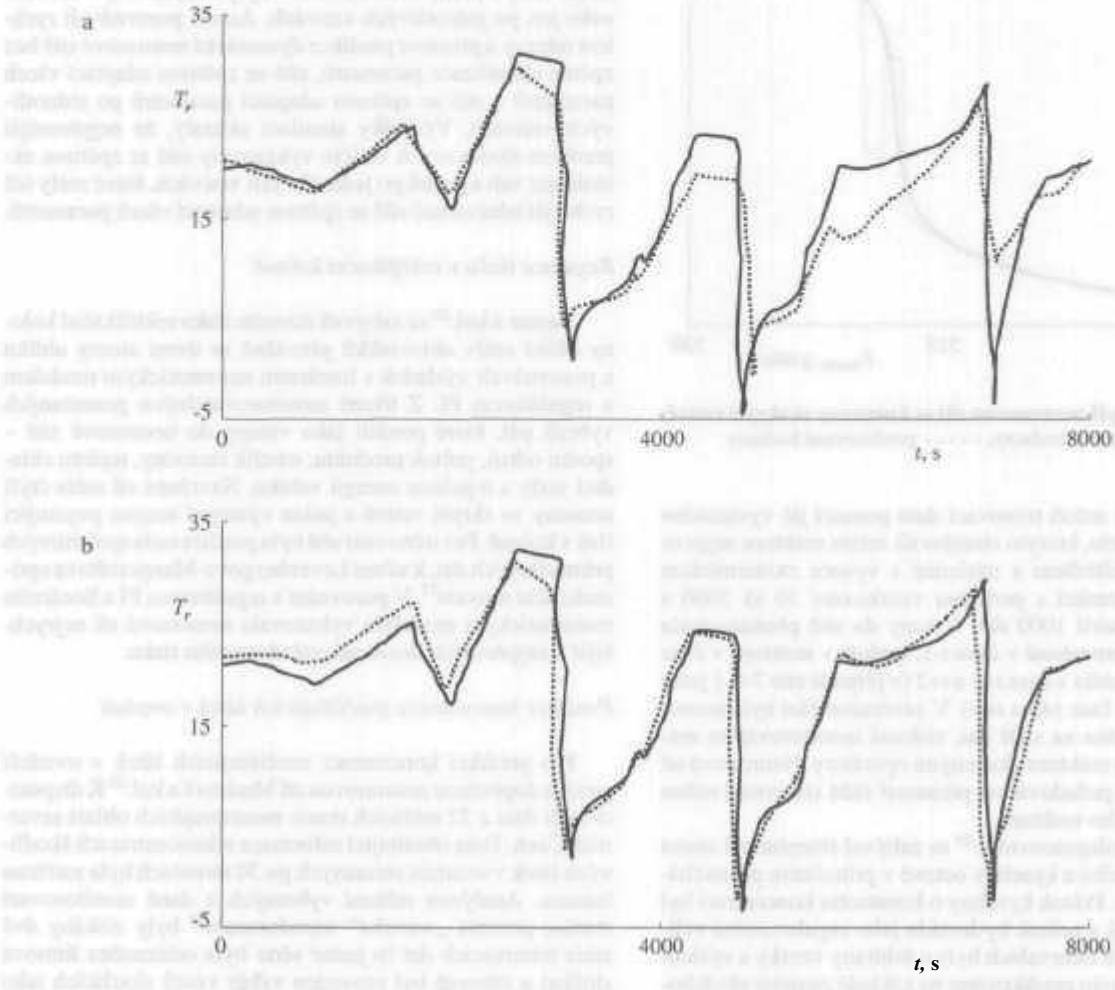
látek byla navržena dvouvrstvá neuronová síť se 3-4 neurony ve skryté vrstvě, která na základě znalosti několika hodnot koncentrací z historie predikovala hodnotu koncentrace znečišťujících látek na daný časový interval. Byly porovnávány přesnosti predikcí neuronové sítě s odpovídajícími autoregresivními modely pro původní data i data s potlačenou šumovou složkou. Kvalita predikce se u obou modelů výrazně zlepšila při odstranění šumové složky, avšak přesnější předpověď poskytovala neuronová síť.

3.2. Rekurentní neuronové sítě

Řízení chemických reaktorů

Galván a Zaldívar[30] se zabývali modelováním chování teploty teplosměnného média v plášti míchaného vsádkového reaktoru prostřednictvím rekurentních neuronových sítí a následným využitím modelu v řízení teplotního režimu reaktoru. Odhadovali časové profily teploty média v plášti reaktoru a v reaktoru. Při zahájení optimalizace rekurentní sítí dosadili prahy a váhy dopředně sítě o shodné struktuře naučené na stejná data, a tím výrazně urychlili dobu optimalizace rekurentní sítě. Rekurentní síť trénovali na sadě dat snímaných v intervalech 10 s při různých režimech reaktoru - prázdný reaktor (nemíchaný, počáteční teplota 20 °C) a reaktor naplněný 80 1 vody (intenzita míchání 200 ot/min, počáteční teplota 20 °C a intenzita míchání 300 ot/min, počáteční teplota 60 °C). Pro predikci teploty média v plášti reaktoru a teploty v reaktoru v čase t+1 byly použity jako vstupy do sítě měřené hodnoty teplot média v časech f-2, t-1, t, teplota v reaktoru v čase t, okolní teplota v čase t a výstupní signály z pomocného regulátoru v časech t-23 až t-3. Nejprve byla trénována dopředná neuronová síť. Analýzou vah jednotlivých vstupních neuronů byly identifikovány a vyřazeny nevýznamné neurony. Jednalo se o okolní teplotu a několik výstupních signálů z pomocného regulátoru. Tento krok, jak se později ukázalo, velmi významně ovlivnil přesnost predikovaných hodnot rekurentní sítě. V další fázi byly použity stejné vstupy do rekurentní sítě, místo teploty média v čase t byla dosazována zpětně hodnota predikovaná sítí. Ukázalo se, že rekurentní síť o struktuře 14-20-1 je schopna po naučení velice přesně a rychle predikovat požadované hodnoty.

Shaw, Doyle a Schwaber[31] se zaměřili na řízení reakční teploty kontinuálního promíchávaného reaktoru s nevratnou



Obr. **6. Neuronová síť;** a) dopředná b) rekurentní; —experimentální profil, ... predikovaný profil

exotermickou reakcí regulací teploty chladícího média a na simulaci produkce cyklopentenolu z cyklopentadienu kysele katalyzovanou adicí vody v kontinuálním promíchávaném reaktoru. V prvním případě tvořily vstupy do neuronové sítě reakční teplota v čase *t*-2 a *t*-1, teplota média v čase ř-2 a *t*-1. Výstup představovala reakční teplota v čase *t*. Po natrénování na známý režim reaktoru byla porovnána přesnost predikce reakční teploty lineárního regresního modelu, dopředně neuronové sítě a rekurentní dynamické neuronové sítě na náhodné vstupy a na periodické výkyvy v daném rozmezí. Z obr. 6a a 6b je patrné, že rekurentní neuronová síť předpovídala požadovanou reakční teplotu s větší přesností než dopředná neuronová síť, která není vhodná pro predikce časových řad o více než jeden časový krok vpřed. Rekurentní síť byla rovněž přesnější i ve srovnání s lineárním modelem, zatímco dopředná síť byla méně přesná. Obdobným způsobem postupovali autoři i ve druhém modelovém případě a dospěli k stejným závěrům.

Modelování membránové separace

Pirion a kol.[32] aplikovali umělé neuronové sítě v modelování procesu tlakem řízené membránové separace mikročástic ze suspenzních roztoků, tzv. mikrofiltrace. Prezentovali dvě možnosti aproximace procesu - empirický model neuronové sítě a semiempirický model obsahující diferenciální rovnici, jejíž parametry byly stanoveny pomocí neuronové sítě. Trénovací data byla získána z experimentů se suspenzemi pekařského droždí při různých filtračních tlacích a různých rychlostech tangenciálního toku. V případě empirického modelu byla použita rekurentní dynamická neuronová síť se čtyřmi vstupy (rychlost tangenciálního toku, filtrační tlak, koncentrace roztoku a hydraulický odpor v čase f-1), jedním výstupem (hydraulický odpor v čase *t)* a čtyřmi neurony ve skryté vrstvě. Parametry diferenciální rovnice semiempirického modelu byly odhadnuty pomocí dvou dopředných neuronových sítí struktur 3-5-2 a 3-4-1. Vstupy obou sítí tvořily - rychlost tangenciálního toku, filtrační tlak a počáteční koncentrace roztoku. Přesnější odhad hydraulického odporu vykázal v porovnání s rekurentní neuronovou sítí semiempirický model.

4. Závěr

Uvedený výběr prací dokládá převahu využití neuronových sítí v oblasti řízení a regulace chemických procesů. V této oblasti je možno s výhodou využít snadné a rychlé vyhodnocení stavu sledovaného systému neuronovou sítí a implementovat ji do řídících systémů. Neuronová síť dosáhne mnohem rychleji a přesněji ustálení požadovaných hodnot řízených veličin než klasické regulátory.

Velká pozornost je také věnována dělícím procesům (zejména rektifikacím), které lze jen obtížně popsat analytickými metodami a naopak snadno nahradit empirickým modelem. Neuronové sítě vykazují i v těchto případech značnou přesnost a použitelnost v průmyslovém měřítku.

Ve většině modelů jsou použity dopředně neuronové sítě, u nichž lze snadno optimalizovat parametry algoritmem zpětného šíření. Rekurentní neuronové sítě byly aplikovány pouze v případech, kdy bylo nutné předpovídat hodnoty časových řad o více než jeden časový krok vpřed, neboť dopředně neuronové sítě nemají srovnatelnou schopnost spolehlivé predikce.

Klíčovým krokem úspěšné aplikace neuronové sítě je určení optimální struktury sítě. Ta byla stanovována výhradně metodou pokusu a omylu s využitím zkušeností jednotlivých řešitelů a s ohledem na povahu vstupních trénovacích dat. Obecně platí, že pro aplikace neuronových sítí v chemické technologii jsou plně postačující dvouvrstvé (maximálně třívrstvé) sítě s méně než dvaceti neurony ve skryté vrstvě. Ukazuje se také, že vhodná volba vstupních veličin může významně ovlivnit přesnost aproximace neuronové sítě a vyhledání globálního optima, a je proto důležité důkladně analyzovat vliv vstupní veličiny na požadovanou výstupní veličinu a v případě zanedbatelného nebo minimálního vlivu tento vstup eliminovat.

LITERATURA

1. Chitra S. P.: Chem. Eng. Process. *89, AA* (1993).
2. Fogelman-Soulié F.: Rev. Inst. Fr. Pet. *47,* 383 (1992).
3. Bhat N., McAvoy T. J.: Comput. Chem. Eng. *14,* 573 (1990).
4. Mutihac R., Mutihac L.: Roum. Chem. Qart. Rev. 5, 67 (1997).
5. Šnorek M., Jiřina M.: *Neuronové sítě a neuropočítače.* Vydavatelství ČVUT, Praha 1998.
6. Bulsari A. B.: *Neural Networks for Chemical Engineers.* Elsevier Science, Lappeenranta 1995.
7. Rój E., Wilk M.: Comput. Chem. Eng. *22,* 909 (1998).
8. McGreavy C, Lu M.L., Wang X. Z., Kam E. K. T.: Chem. Eng. Sci. *49,* 24A, 4717 (1994).
9. Zhao J., Chen B., Shen J.: Comput. Chem. Eng. 27, 929 (1997).
10. Wang X. Z., Chen B. H., Yang S. H., McGreavy C: Chem. Eng. Sci. *51,* 2169 (1996).
11. Wang X. Z., Lu M. L., McGreavy C: *Proceedings of PSE'94*(Kyongju, Korea) 2, 137 (1994).
12. Lu G., Hua B.: Shiyou Lianzhi Yu Huagong *28,* 48 (1997).
13. Aguiar H. C, Filho R. M.: Comput. Chem. Eng. 22,981 (1998).
14. Mirams S., Nguyen K. L.: *International Pan Pacific Conference, Tappi Proceedings,* str. 73 (1994).
15. Qian Y., Tessier P. J. C: Chem. Eng. Technol. *18,* 337 (1995).
16. Savkovič-Stevanovič J.: Comput. Chem. Eng. *18,* 1149 (1994).
17. Powell M. J. D.: Computer J. 7, 303 (1965).
18. Aoyama A., Walsh S. P. K.: Comput. Chem. Eng. *21,* 1155 (1998).
19. Krothapally M., Palanki S.: Comput. Chem. Eng. *21,* 463 (1997).
20. Draeger A., Gesthuisen R.: Chem. Ing. Tech. *68,* 1320 (1996).
21. Milanič S., Šel D., Hvala N., Strmčnik S., Karba R.: Comput. Chem. Eng. *21,* 637 (1997).
22. Zaldívar J. M., Hernández H.: Chem. Eng. Process. *31,* 173 (1992).
23. Dirion J. L., Ettedqui B., Cabassud M., Le Lann M.V., Gasamatta G.: Chem. Eng. Process. *35,* 225 (1996).

24. Wang H., Oh Y., Yoon E. S.: Comput. Chem. Eng. *22,* 823 (1998).
25. Nikravesh M., Farell A. E., Stanford T. G.: Comput. Chem. Eng. *20,* 1277(1996).
26. Turner P., Montague G., Morris J.: Comput. Chem. Eng. *20,* 937 (1996).
27. Marquardt D.: SIAM, J. Appl. Math. *11,* 431 (1963).
28. Mudrová M., Procházka A., Fiala J.: *Signal Analy sis and Prediction I. Proc. ESCAP-97, Praha, 24-27. červen 1997* (Procházka A., Uhlíř J., Sovka P., ed.), str. 247. EURASIP, ICT Press, Praha 1997.
29. Daubechies I.: IEEE Trans. Theory *36,* 961 (1990).
30. Galván I. M., Zaldívar J. M.: Chem. Eng. Process. *36,* 505 (1997).
31. Shaw A. M., Doyle F. J., Schwaber J. S.: Comput. Chem. Eng. 27, 371 (1997).
32. Pirion E., Latrille E., René F.: Comput. Chem. Eng. *21,* 1021 (1997).

T. Herink and Z. Bělohlav *{Department of Organic Technology, Institute of Chemical Technology, Prague):* **Applications of Artificial Neural Networks in Chemical Technology**

The review on applications of artificial neural networks is aimed at modelling, optimisation, control and regulation of chemical processes. Attention has been predominantly paid to the types of neural networks and to the choice of geometries. A substantial proportion of applications is focused on modelling and optimisation of refining and separation processes and on the control of chemical reactors. Feed-forward neural networks are used in most models. Recurrent neural networks are usually implemented in dynamic models depending on their capability of prediction of time series. A common feature of the reviewed methods is their efficiency in solving problems by neural networks.