

Факултет инжењерских наука Универзитета у Крагујевцу



Назив студијског програма: Машинско инжењерство

Ниво студија: Мастер академске студије

Модул: Информатика у инжењерству

Предмет: Вештачка интелигенција

Број индекса: 317/2013

Марина Ј. Радовић

Имплементација и примена рекурентних неуронских мрежа

мастер рад

Комисија за преглед и одбрану:

1. Др Весна Ранковић - ментор

2. _____

3. _____

Датум одбране: _____

Оцена: _____

У оквиру овог мастер рада кандидат треба да

1. изврши систематизацију литературе која се односи на примену рекурентних неуронских мрежа,
2. детаљно опише структуру рекурентних неуронских мрежа и алгоритама за њихово обучавање,
3. објасни начин имплементације рекурентних неуронских мрежа,
4. примени различите врсте рекурентних неуронских мрежа у предикцији потрошње природног гаса.

Препоручена литература:

- [1] Gorucu, F.B., Geris, P.U., Gumrah, F.: "**Artificial neural network modeling for forecasting gas consumption**", *Energy Sources* 26 (2004) 299–307.
- [2] Azadeh, A., Asadzadeh, S.M., Saberi, M., Nadimi, V., Tajvidi, A.: "**A neuro-fuzzy-stochastic frontier analysis approach for long-term natural gas consumption forecasting and behavior analysis: the cases of Bahrain, Saudi Arabia, Syria, and UAE**", *Applied Energy* 88 (2011) 3850–3859.
- [3] Azadeh, A., Saberi, M., Asadzadeh, S.M., Hussain, O.K., Saberi, Z.: "**A neuro-fuzzy-multivariate algorithm for accurate gas consumption estimation in South America with noisy inputs**", *Electrical Power and Energy Systems* 46 (2013) 315–325.
- [4] Hagan, M., Jesus, O. D. and Schultz, R.: "**Training Recurrent Networks for Filtering and Control**", *Chapter 12 of Recurrent Neural Networks: Design and Applications*, L.R. Medsker and L.C. Jain, Eds., CRC Press, pp. 327-389, 2001.

Крагујевац, 19.06.2015.

ментор:

Др Весна Ранковић, редовни професор

Резиме

Рад описује тренирање, тестирање и примену модела слојевите дигиталне рекурентне мреже (Layered Digital Recurrent Network - LDRN) и NARX (Нелинеарни ауто-регресиони модел са спољашњим улазом) модела за обрачунавање потрошње природног гаса у земљама Европске Уније. Динамички алгоритам учења са пропагацијом уназад и сигмоидална функција преноса су коришћени у мрежама. Да би показали употребљивост и ефикасност разматраног приступа, подаци су прикупљени за 32 године (1980 - 2011) у земљама Европске Уније. Модел рекурентне мреже укључују бруто домаћи производ (Gross domestic product-GDP) и број становника (Population-POP) као улазне променљиве. Добијени резултати потврђују да коришћење модела рекурентне мреже за предвиђање природног гаса је прихватљиво.

Кључне речи: Дугорочно, Потрошња гаса, LDR мрежа, предвиђање, NARX

Abstract

The paper describes the training, testing and application of the Layered Digital Recurrent Network (LDRN) and NARX models for computing the natural gas (NG) consumption in the European Union (EU) countries. The dynamic back propagation learning algorithm and logistic sigmoid transfer function were used in the networks. To show the applicability and effectiveness of the considered approach, data were collected for 32 years (1980–2011) in the EU countries. The recurrent network models included Gross Domestic Product (GDP) and population as the input variables. The reported results confirmed that the use of recurrent network models for NG prediction is acceptable.

Keywords: Long-term, Gas consumption, LDR Network, prediction, NARX

САДРЖАЈ

1. УВОД	1
2. НЕУРОНСКЕ МРЕЖЕ	6
2.1. Биолошки неурон	7
2.2. Архитектура и учење неуронских мрежа без повратних веза	9
2.2.1. Неурон и модел статичког неурона. Архитектура једнослојног перцептрона.	10
3. РЕКУРЕНТНЕ НЕУРОНСКЕ МРЕЖЕ	12
3.1. Динамички неурон	14
3.2. Елманова мрежа	14
3.3. Хопфилдова мрежа.....	17
3.4. Дигитална рекурентна неуронска мрежа	18
3.4.1. Принципи динамичког учења.....	20
3.5. Динамичка пропација уназад за LDRN (Layered Digital Recurrent Network).....	24
3.5.1. Експлицитни изводи	25
3.5.2. Комплетан FP алгоритам за LDRN	26
4. ПРЕДВИЂАЊЕ ДУГОРОЧНЕ ПОТРОШЊЕ ПРИРОДНОГ ГАСА У ЗЕМЉАМА ЕВРОПСКЕ УНИЈЕ ПРИМЕНОМ РЕКУРЕНТНИХ НЕУРОНСКИХ МРЕЖА	30
4.1. Формулација проблема	30
4.1.1. Специфичности природног гаса.....	31
4.1.2. Тражња природног гаса у Европи	32
4.1.3. Улога природног гаса у укупној примарној енергији	33
4.2 Резултати симулација.....	34
5. ЗАКЉУЧАК	60
ЛИТЕРАТУРА.....	62

1. УВОД

Мозак је грађен од преко 100 милијарди неурона, при чему је сваки од њих у вези са хиљадама других неурона у мрежи која омогућава сложену контролу и комуникацију. Начин рада биолошке неуронске мреже није у потпуности познат, али се зна да све неуронске функције, укључујући памћење, обављају неурони и њихове међусобне везе. Учење се може дефинисати као процес успостављања нових веза или промена постојећих.

Мреже које се разматрају у овом раду нису биолошке, већ крајње поједностављене, вештачке неуронске мреже које користе принципе људског мозга и његове структуре како би развиле стратегију процесирања података. Код вештачке неуронске мреже заступљена је масовна паралелна обрада података.

Између класичних рачунара и неуронских мрежа постоје одређене разлике. Конвенционални рачунари раде на логичкој основи, детерминистички, секвенцијално или са веома ниским степеном паралелизма, при чему би софтвер написан за такве рачунаре морао бити готово савршен како би исправно радио. Са друге стране, неуронске мреже спадају у категорију паралелног асинхроног дистрибуираног процесирања, при чему су овакве мреже толерантне на оштећење или испадање из рада релативно малог броја неурона, као и на присуство шума у улазном сигналу. Иако су конвенционални рачунари бржи у рачунарским алгоритмима и прецизним аритметичким операцијама, постоје бројне области примене неуронских мрежа, због чега их многи истраживачи називају последњом људском технологијом.

У инжењерству, физици и биомедицини се са развојем нових технологија стварају многобројни проблеми у којима нестационарност, нелинеарност, неизвесност и сложеност играју главну улогу. Да би се овакви проблеми решили неопходно је коришћење нелинеарних процесора, при чему неуронске мреже представљају један од најмоћнијих. Оне су привлачне јер уче на примерима, снажно су подржане од стране статистичких и оптимизационих теорија и стога су погодна алтернатива за проширење могућности обраде сигнала.

Према смеру простирања информација, неуронске мреже се могу сврстати у две класе: без повратних веза (нерекурентне) код којих се сигнали простиру само у једном смеру, од улаза ка излазу и неуронске мреже које се овде разматрају, мреже са повратним везама (рекурентне).

Током 1990. рекурентне мреже су биле тема од значајног интереса. Оне су дизајниране да нуче секвенцијалне или временски-промељиве обрасце, тако да су данас применљиве на широком спектру проблема. Посебно је њихова употреба као идентификатора и предиктора у нелинеарним динамичким системима значајно порасла. У последњих неколико година, рекурентне мреже су показале одличне перформансе у задацима за обраду природног језика [1,2,3,4,5,6]. У [7] аутори су користили

рекурентну неуронску мрежу за конструисање фази неуронске мреже којом су извршили реализацију фази закључивања на проблему контроле роботског праћења. Посебна врста рекурентних неуронских мрежа, дубоке неуронске мреже, показале су се изванредним за задатке као што су препознавање говора [8] и препознавање руком написане цифре [9]. Хопфилдова мрежа, један је од представника рекурентних неуронских мрежа и користи се за чување једног или више стабилних циљних вектора који се могу посматрати као меморије које опозивају мрежу када је то предвиђено помоћу сличних вектора који делују као знак на меморију мреже. У [10] Хопфилдове неуронске мреже су прилагођене тако да се на њима заснива дизајн хардверског модула намењеног обављању параметарске идентификације динамичких система. Предности имплементације овог модула огледају се у томе што приликом рачунски интензивних задатака, као што је препознавање облика, ова неуронска кола су у стању да надиграју софтвер, а поред тога они могу бити успешно уграђени у аутономне роботе због својих карактеристика као што су смањене димензије, тежина и утрошак снаге. Танк-Хопфилдове мреже у [11] имају битну улогу у представљеном алгоритму избегавања препрека за редундантне роботе, као и за извршену контролу роботске руке. У [12] енергетска функција, која представља ограничење решења је пресликана на Хопфилдову неуронску мрежу у циљу минимизације. Још један пример примене Хопфилдових мрежа описан је у [13]. У овом раду природна формулација концепата у неуронском смислу, као што су инверзна кинематика редундантних робота, избегавање препрека и представљање флексибилних објеката, се добија помоћу генерализоване Хопфилдове мреже чије јединице одговарају физичким ентитетима, као што су делови манипулатора или окружења, а везе међу њима пропадају интеракције снаге и међусобних померања. У [14] описан је неуро-фази модел који може бити коришћен за идентификацију и (реално-временску) симулацију вискоеластичних модела, при чему је такав модел мотивисан кооперативним неуро-фази приступом заснованом на архитектури векторизоване рекурентне неуронске мреже. Употребљивост овог модела у пракси је демонстрирана његовом применом у области хируршке симулације. Рекурентне неуронске мреже су у [15] примењене на проблему регулације нелинеарног излаза тако што су предложене за реално-временску синтезу и аутоматско-подешавање повратних контролера за нелинеарни излаз регулационих система. У [16] описан је модел когнитивне роботике у коме секвенце лингвистичких улаза доводе до развоја концепата вишег реда заснованих на основним концептима и акцијама. У овом процесу, коришћење рекурентне неуронске мреже омогућава учење концепата вишег реда заснованих на временским секвенцама примитивне акције. На овај начин је извршено истраживање односа између развоја могућности манипулације симболима и сензоримоторног знања код хуманоидног робота iCub. Елманова мрежа је врста рекурентних неуронских мрежа. Захваљујући повратној вези коју садржи, ова мрежа је у стању да научи да препозна и генерише временске и просторне обрасце. Тако је у [17] изабрана и дизајнирана Елманова мрежа према предложеној методи за решавање инверзне кинематике робота. При томе је приступ решавању овог проблема заснован на поузданости и примењен на робот манипулатор са шест степени слободе. Хибридна квантизирана Елманова неуронска мрежа је у [18] примењена да научи комплексну динамику временских серија оптерећења мотора по сату и предвиђање оптерећења у

блиској будућности. Модел Елманове неуронске мреже (Elman neural network-ENN) примењен је у [19] за реализацију предиктивног контролног система притиска ваздушне коморе у тунелском штиту од муља испод реке Јангцекјанг у Вухану (Кина). Посредством ENN предиктора и ENN контролера извршено је подешавање оптималних контролних параметара аутоматског праћења жељеног ваздушног притиска у комори. У [20] предложен је предиктивни контролер за шесто-степени Stanford роботски манипулатор, при чему је неуронска мрежа контролера била дизајнирана коришћењем генерализоване предиктивне контроле и Елманове мреже. С обзиром на то да предиктивни алгоритам, који је класа дигиталне методе контроле, захтева дуго време израчунавања, Елманова мрежа контролера је била дизајнирана да смањи време процесирања избегавањем високе математичке и рачунарске комплексности поменутог алгоритма. Високе перформансе Елманових неуронских мрежа су у [21] примењене за праћење стања и дијагностику у системима нуклеарне електране и ротационој машинерији. У првој примени, студија се бави коришћењем рекурентних неуронских мрежа за детекцију аномалија уведеном са симулираном снагом функционисања високо-температурног гасно охлађеног нуклеарног реактора, док се у другој обавља детекција оштећења лежаја мотора коришћењем приступа кохерентности функција. У [22] нов предиктивни приступ заснован на упитницима коришћењем Елманове неуронске мреже је предложен за предикцију особина коже учесника. Овај систем предикције је много бржи и много исплатљивији од традиционалних клиничких тестова особина коже, а резултати предвиђања су показали добру прецизност. У [23] су представљени експерименти на учењу понашања руковањем објектом од стране малог хуманоидног робота коришћењем модела динамичке неуронске мреже, рекурентне неуронске мреже са параметарским праговима активација. Како су динамичке мреже генерално моћније од статичких мрежа и имају способност да науче секвенцијалне или временски-променљиве обрасце, њихова примена је вишеструка. Поред претходно наведене примене, такође се примењују и у областима као што су: предвиђање у финансијским тржиштима [24], уједанченост канала у комуникационим системима [25], детекција фаза у електроенергетским системима [26], сортирање [27], детекција грешака [28], препознавање говора [29], па чак и предвиђање структуре протеина у генетици [30].

Овај рад предлаже поступак за примену техника рачунарске интелигенције за предвиђање дугорочне потрошње природног гаса (natural gas-NG). Рад описује примену модела слојевите дигиталне рекурентне мреже и NARX (Нелинеарни ауто-регресиони модел са спољашњим улазом) модела за обрачунавање потрошње природног гаса у земљама ЕУ (Италија, Аустрија, Бугарска, Француска, Уједињено краљевство, Ирска, Белгија и Шпанија).

Анализа понашања потрошње гаса је важна за тачно креирање политике. Потрошња NG је изузетно порасла у последњој деценији због све већег економског развоја и будућа процена NG је важна за утврђивање будућег нивоа равнотеже понуде/потражње. Потражња гаса у Европи је изгубила око 40 билиона кубних метара (Bcm) сама у једној години или око 6% од укупне потрошње гаса. У последњих

неколико година фокус истраживачких пројеката је на великим променама које су се догодиле у индустрији природног гаса и евалуацији која се догодила у електроенергетском сектору. Неколико студија имају изграђена сценарија за истицање изазова и прилика до 2030. године. Да би се упоредила статистика на различитим тржиштима, IEA (International Energy Agency) прикупља статистичке податке од званичних органа земаља чланица коришћених од стране Honoré [31].

Различити модели су примењени да предвиде еволуцију тражње NG. Khotanzad et al. [32] користе вишеслојни перцептрон у предвиђању дневне потрошње природног гаса. Gogucu et al. [33] укључују приступ да процене и прогнозирају потрошњу гаса са ANN моделирањем за краткорочне примене. Siemek et al. [34] су описали хипотетички модел за процену потражње природног гаса, на основу просечног тренда развоја привреде током последњих деценија; модел разматра природну производњу/максималну потражњу енергената. У [35] извршена је процена функције потражње домаћинства за природним гасом од стране Yoo et al. коришћењем модела селекције узорака применом података из прегледа домаћинства у Сеулу. Прилагодљив приступ анализи мрежно-заснованог фази закључивања системски-стохастичке границе за предвиђање дугорочне потрошње природног гаса и анализа су предложени од стране Azadeh et al. [36]. Предложени модели су засновани на улазним променљивама бруто домаћем производу (Gross domestic product-GDP) и броју становника (Population-POP). Azadeh et al. [37] су презентовали интегрисани адаптивно-мрежно-засновани систем фази закључивања (ANFIS) – алгоритам анализе обавијања фази података за побољшање дугорочне прогнозе потрошње гаса и анализе.

Истраживања се у [38,39,40,41] фокусирају на побољшање процене и прецизности предвиђања. У [42] Huang и Lin су представили значај предиктивних метода у циљу решавања задатака предвиђања усталасане тражње. Gonzales Chavez et al. [43] су дали формулацију модела прогнозирања за предвиђање производње и потрошње енергије у Астурији (Северна Шпанија) применом модела анализа временских-серија Box-Jonkis (ARIMA). Kadoshin et al. [44] су узимајући у обзир два фактора – повећање становништва и економски развој, разматрали тренд потрошње енергије – тренутни, као и у блиској будућности. Jai Persavd и Kumar [45] представили су пројекције понуде и потражње нафте и природног гаса до 2020. године за Канаду. У [46] Gogucu и Gumrah су извршили предикцију краткорочне потрошње гаса применом анализе мултиваријабилне регресије за главни град Турске, Анкару. Gutierrez et al. [47] уочили су процес раста потрошње природног гаса у Шпанији испитивањем примене Gompertz – типе иновације процеса дифузије за стохастичко моделирање и снимање истраженог проблема. Sanchez-Udeda и Berzosa [48] су извршили предикцију краја индустријске потрошње природног гаса у средњорочном периоду (1-3 године) на разложеном приступу, тако да је извршена предикција са веома великом резолуцијом (дани). Реч је о прогнози која је добијена комбинацијом три различите компоненте: једна која снима тренд временских серија, сезонска компонента и пролазна компонента. Parikh et al. [49] процењују пројекције потражње нафтних деривата и природног гаса у Индији са разматрањем GDP-а и броја становника као улазима у модел.

Предвиђање потражње /потрошње природног гаса због своје изузетне важности у будућности зелених енергетских тржишта, све више добија посебну пажњу у литератури.

Рекурентне неуронске мреже могу решити неке неодређене проблеме увођењем извесног степена интелигентности, што је у овом раду потврђено њиховом имплементацијом на значајним, изазовним проблемима 21. века.

2. НЕУРОНСКЕ МРЕЖЕ

Прихватљив ниво разумевања о грађи и функционисању људског мозга представља полазну тачку за истраживање и развој вештачких неуронских мрежа које се могу применити за решавање многих сложених проблема. Вештачке неуронске мреже су у стању да на одређен начин функционишу по принципима асоцијативног мишљења и препознавања које постоји код човека. Другим речима, функционални процеси који се одвијају у мрежи подсећају на нервни систем код човека.

Неуронска мрежа која се овде разматра, представља систем састављен од одређеног броја међусобно повезаних процесора или чворова, или процесних елемената чији је заједнички назив вештачки неурони. Сваки од тих неурона има локалну меморију која му омогућава памћење података које обрађује.

У једној неуронској мрежи постоји много више веза него процесорских јединица, а њихов број представља снагу неуронске мреже. Коефицијенти који су тим везама у сваком тренутку додељени називају се тежински коефицијенти веза или тежине веза. Како се неуронске мреже не програмирају, већ се “обучавају”, ови се коефицијенти током обучавања ажурирају. Поред ажурирања ових коефицијената које се најчешће примењује, понекад се могу ажурирати само вредности улазних процесорских јединица.

У наставку је дат кратак историјски преглед развоја вештачких неуронских мрежа:

1943. године неуро-психолози McCulloch и Pitts су увели први једноставни систем који функционише користећи механизме људског нервног система. О томе сведочи њихов често цитирани чланак “Логички рачун идеја својствених нервној активности”.

1949. је Donald Hebb у својој књизи “The Organization of behavior” (Организација понашања) предложио одређени закон учења за синапсе, помоћу кога је дао објашњење неких експерименталних резултата из психологије.

1951. Marvin Minsky је конструисао неурорачунар под називом Snark, са синапсама које имају могућност прорачунавања тежина веза у зависности од успешности обављања одређеног задатка. Тај процес је заснован на Hebb-овом начину учења.

1956. година је позната по томе што је тада представљена прва симулација Hebb-овог модела.

1958. је година када је Frenk Rosenblatt изумео прву (једнослојну) неуронску мрежу звану перцептрон, која није била у стању да решава проблеме класификације који нису линеарно сепарабилни (као што је случај са XOR проблемом).

1960. године Widrow i Hoff су развили нови алгоритам учења који се и данас користи. Реч је о правилу које се примењује за адаптацију параметара перцептрона са линеарним активационим функцијама.

Даља истраживања у области вештачких неуронских мрежа су се одвијала у два смера: развој нових архитектура и побољшање алгоритама обучавања. После привременог застоја током 70-их, поновним почетком развоја ове области сматра се 1986. година када Rumelhart, Hinton и Williams објављују свој рад на процедури учења вишеслојних неуронских мрежа са пропагацијом грешке уназад (тзв. backpropagation algorithm).

Иако се вештачке неуронске мреже последњих година успешно примењују у многим областима (електроници, медицини, економији, телекомуникацијама, војним апликацијама итд.) и даље се истражују нове могућности њихове примене.

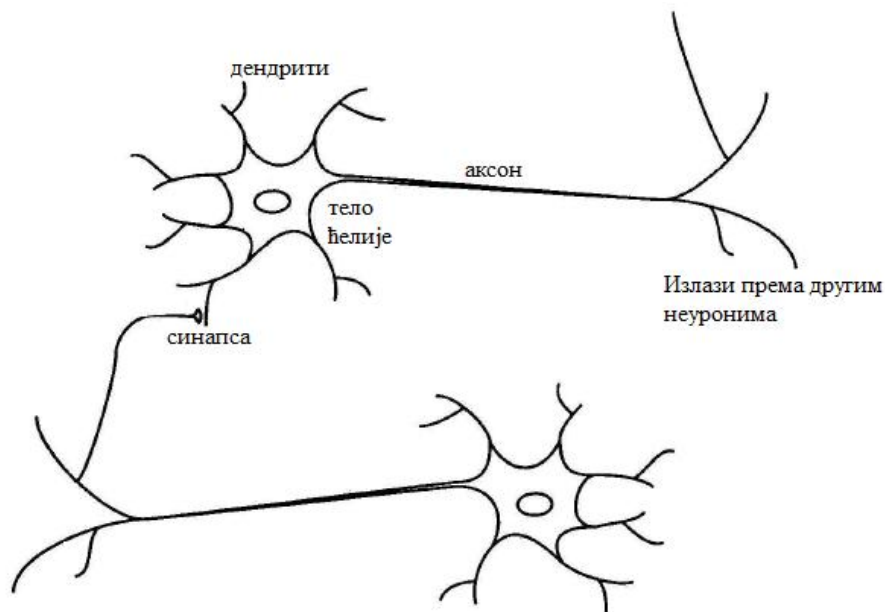
2.1. Биолошки неурон

Појам вештачка неуронска мрежа је настао по угледу на природне неуронске мреже, односно нервне системе живих бића. Иако су вештачке неуронске мреже по одређеним својствима сличне биолошким, биолошке неуронске мреже су изразито компликованије од својих математичких модела.

Мозак се састоји од неурона (нервних ћелија), којих има приближно око 10^{11} са 10^{15} међусобних веза које они остварују. Основни делови типичног биолошког неурона су:

- тело ћелије (сома)
- улазни део ћелије (дендрити)
- излазни део ћелије (аксон)

На слици 2.1 дат је шематски приказ два биолошка неурона. Дендрити представљају улазни део ћелије. То су кратке, разгранате нити које имају улогу да примљене сигнале од других неурона доводе до тела ћелије. У телу ћелије се одвија процес сабирања сигнала са више улаза. Излазни део ћелије (аксон) представља једно дуго влакно (обично око 1cm, а понекад и до 1 m) које има функцију да преноси сигнал од тела ћелије до других неурона. Овај пренос није директан, већ се сигнали достављају до синапси које га даље прослеђују дендритима друге нервне ћелије. Од претходних синаптичких преноса зависиће и ефикасност текућег синаптичког преноса кроз неку синапсу. На основу тога може се рећи да су синапсе меморијски чланови биолошке неуронске мреже. Сигнали који се од синапси, преко дендрита доводе до тела неурона могу за тело неурона бити побуђујући или смирујући, зависно од тога да ли је синапса инхибиторна или екситаторна.



Слика 2.1 Шематски приказ биолошких неурона [50]

Да ли ће тело неурона генерисати импулсе зависи од тога да ли је кумулативна вредност ових сигнала током кратког временског интервала мања или већа од прага осетљивости неурона. У случају да је мања, неурон неће генерисати импулсе.

Колико је вештачка неуронска мрежа једноставнија од природне, показује број неурона који она садржи (реда 10^3 неурона, док је код биолошких реда 10^{11} неурона). У погледу брзине обраде информација, вештачке неуронске мреже су у предности у односу на биолошке. Међутим, постоје и неке сличности између вештачких и биолошких неуронских мрежа:

- процес обучавања ради усвајања знања
- тежине међунеронских веза, односно јачина синаптичких веза
- функција мреже је одређена везама између неурона.

Захваљујући проучавањима и сазнањима о биолошким неуронским мрежама, данас вештачке неуронске мреже показују следеће способности:

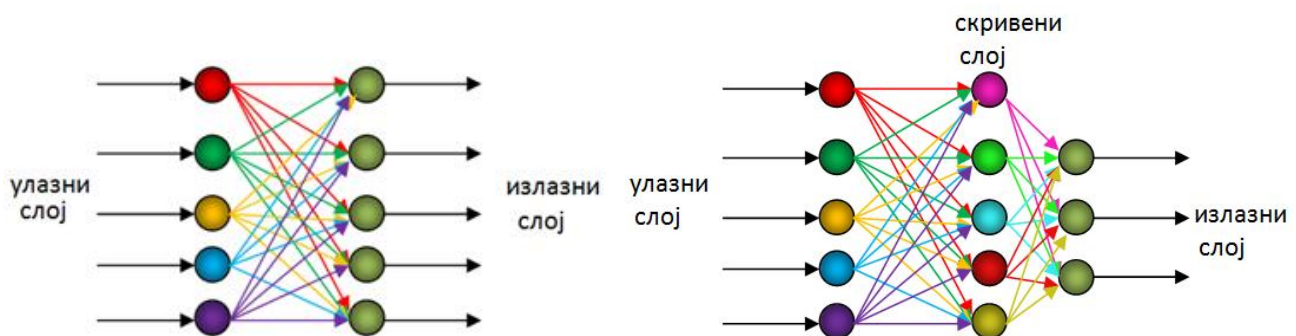
- способност да уче на основу искуства,
- имају способност генерализације,
- способност апстракције,
- примена у многим областима (класификација, предвиђање, идентификација, оптимизација и сл.).

2.2. Архитектура и учење неуронских мрежа без повратних веза

Појам архитектура (топологија) вештачке неуронске мреже означава специфично уређење и повезивање неурона у облику мреже. Неурони у мрежи су распоређени по слојевима. Први слој се назива улазни и излази из неурона који припадају овом слоју истовремено представљају улаз у један или више других неурона. Последњи слој је излазни, а он је специфичан по томе што његови неурони генеришу излазне сигнале. Остали слојеви садрже неуроне који се називају скривени (hidden).

Постоје два основна типа неуронских мрежа према начину на који су њихови неурони повезани. То су нерекурентна (feedforward) неуронска мрежа код које се сигнали простиру само у једном смеру и рекурентна (feedback) неуронска мрежа о којој ће више речи бити у следећем одељку.

Неуронске мреже без повратних веза могу бити једнослојне (слика 2.2 а) и вишеслојне (слика 2.2 б).



Слика 2.2 Приказ (а) једнослојне и (б) вишеслојне неуронске мреже без повратних веза

Најпознатија једнослојна неуронска мрежа је једнослојни перцептрон, мада се данас највише изучавају и примењују неуронске мреже са више слојева (мреже које поред улазног и излазног слоја садрже један или више скривених слојева). Пре него што се мрежа може употребити у корисне сврхе потребно је извршити њено обучавање (тренирање).

Обучавање вештачких неуронских мрежа обавља се итеративним поступком предочавања улазних примера (узорака) и евентуално очекиваних излаза. Током овог процеса тежински коефицијенти се ажурирају како би се у наредној итерацији добио излаз ближи задатој вредности. Разлог због кога се ажурирају тежине веза је то што се знање о излазу, као функцији улаза налази имплицитно у њима. Овај итеративни процес се наставља све дотле док се за све облике из скупа за обучавање не добију задовољавајући резултати. Тада, када је обучавање мреже завршено тежине веза остају фиксне (непромењене). Поред обучавања са адаптивним тежинским коефицијентима веза (које је најчешће), постоје и мреже чије су тежине непроменљиве и током

обучавања када се мењају само нивои активација појединих неурона. Дакле, разликују се два начина учења неуронских мрежа. Први је контролисано (“supervised”) обучавање које је претходно описано. Њега карактерише учење примерима у облику пара (улаз, излаз). Са друге стране, постоји неконтролисано обучавање (“unsupervised”) код кога је мрежи излаз непознат.

У зависности од средине у којој се обучавање обавља, оно може бити on-line и off-line. У првом случају мрежа се обучава у току експлоатације достављањем података који се добијају из радног окружења. Насупрот томе, off-line (лабораторијско) учење је оно које се одвија ван средине у којој ће мрежа радити.

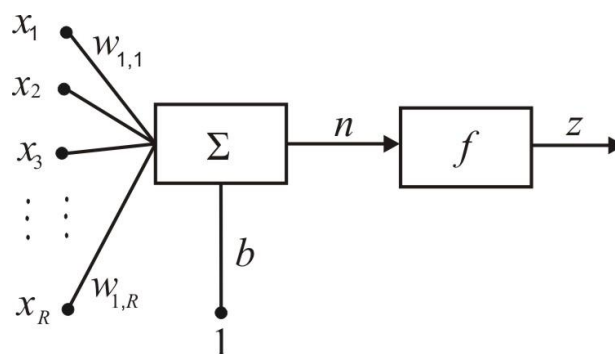
У претходном делу текста описана је процедура учења код које се мењају тежински коефицијенти мреже, тзв. статичка процедура учења. За разлику од овог учења, примењује се и динамичко учење код кога се мења и архитектура. Приликом формирања вештачке неуронске мреже потребно је решити неке проблеме као што су:

- одредити број скривених слојева (који је у општем случају произвољан)
- одредити број неурона сваког слоја
- извршити избор активационе функције неурона и
- изабрати скуп података за учење мреже.

Понашање неуронске мреже зависи од укупног броја неурона, њихове међусобне повезаности, тежинских коефицијената и облика активационих функција.

2.2.1. Неурон и модел статичког неурона. Архитектура једнослојног перцептрона.

На слици 2.3 приказан је модел статичког неурона.



Слика 2.3 Модел статичког неурона

Излаз неурона је:

$$z = f(\mathbf{W}\mathbf{x} + b) \quad (2.1)$$

где су:

$x_1, x_2, \dots, x_R$ — улазни сигнали

$w_{1,1}, w_{1,2}, \dots, w_{1,R}$ — појачања по синапсама

b — праг активације

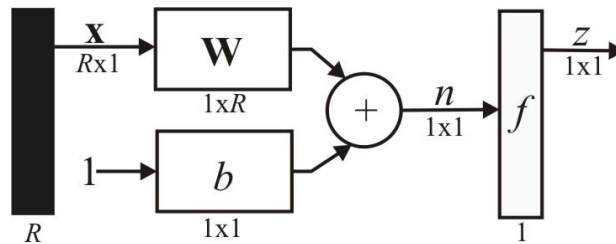
f — активациона функција

R — број елемената улазног вектора

Збир n који је аргумент активационе функције f добија се на следећи начин:

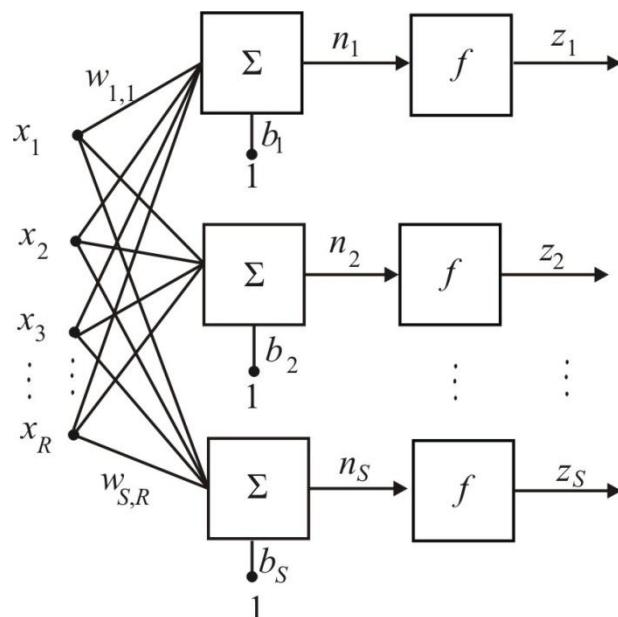
$$n = w_{1,1}x_1 + w_{1,2}x_2 + \dots + w_{1,R}x_R + b \quad (2.2)$$

Статички неурон може да се прикаже на једноставнији начин (слика 2.4).



Слика 2.4 Једноставнији приказ статичког неурона

На слици 2.5 приказан је једнослојни перцептрон.



Слика 2.5 Једнослојни перцептрон

Излаз перцептрона је:

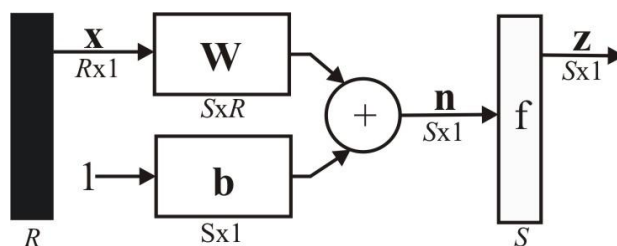
$$\mathbf{z} = \mathbf{f}(\mathbf{W}\mathbf{x} + \mathbf{b}) \quad (2.3)$$

На слици 2.5 са R је означен број елемената улазног вектора, а са S број неурона у слоју.

Матрицу која садржи све тежинске коефицијенте једнослојног перцептрона са слике 2.5 означићемо са:

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,R} \\ w_{2,1} & w_{2,2} & \dots & w_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S,1} & w_{S,2} & \dots & w_{S,R} \end{bmatrix} \quad (2.4)$$

Једнослојни перцептрон може да се прикаже на једноставнији начин (слика 2.6).



Слика 2.6 Једноставнији приказ једнослојног перцептрона

Вишеслојни перцептрон се састоји из више невидљивих слојева.

3. РЕКУРЕНТНЕ НЕУРОНСКЕ МРЕЖЕ

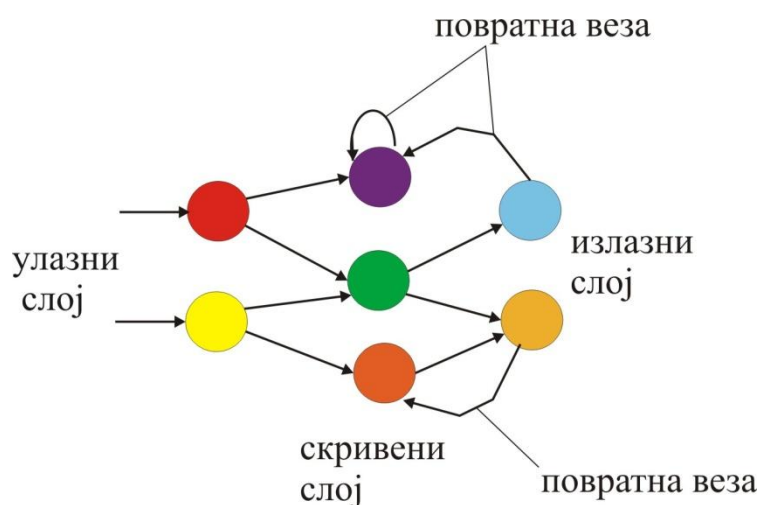
Рекурентне неуронске мреже датирају из краја 80-их, а већ током 1990. су биле важан фокус истраживања и развоја.

Оне су сачињене од више нелинеарних рачунарских елемената специфичних по томе што функционишу паралелно и распоређени су на начин који подсећа на повезаност биолошких неурона.

За разлику од мрежа без повратних веза које су статичке, тј. дати улаз може произвести само један сет излаза, код рекурентних (feedback) неуронских мрежа информације се са виших слојева враћају нижим слојевима. Осим тога, могућа је и повратна веза чвора са самим собом.

Може се рећи да код ове мреже излази неурона формирају динамички систем с обзиром на то да неурони могу да прослеђују излазе на своје улазе. Дакле, излази се мењају и поново враћају на улазе неурона тако да после одређеног времена мрежа може доћи у стабилно стање, осциловати или показивати хаотично понашање.

Кључна ствар је да рекурентне мреже имају краткорочну меморију, што их чини интересантнијим за моделирање мозга али и тежим за разумевање. Код неких проблема, ове мреже су се испоставиле тешким за тренирање (нпр. проблеми са компликованом дуго-дометном временском структуром). Међутим, у поређењу са мрежама без повратних веза оне имају много веће процесне способности. На слици 3.1 приказан је пример рекурентне неуронске мреже.



Слика 3.1 Пример *feedback* неуронске мреже

Како су рекурентне неуронске мреже дизајниране да науче секвенцијалне или временски-променљиве обрасце, успешно се примењују на задацима као што су: несегментно повезано препознавање рукописа, обрада природног језика, обрада говора, обрада слика итд.

Захваљујући свом супериорном временском и просторном понашању (стабилне и нестабилне фиксирани тачке и гранични циклуси; хаотична понашања) помоћу рекурентних мрежа могуће је моделирати одређене когнитивне функције, као што су асоцијативна меморија, ненадгледано учење, само-организујуће мапе и временско резоновање.

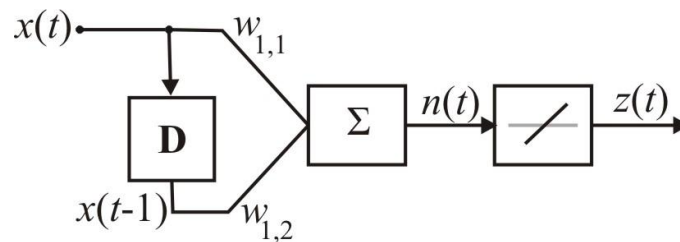
Представници ове врсте вештачких неуронских мрежа су:

- Хопфилдова неуронска мрежа
- Елманова неуронска мрежа
- Целуларна неуронска мрежа (CNN)
- Кохененова неуронска мрежа

- Двоструко асоцијативна неуронска мрежа и сл.

3.1. Динамички неурон

Да би се боље разумела архитектура рекурентних неуронских мрежа чији је механизам функционисања динамички, у наставку је дат приказ типичног динамичког неурона (Слика 3.2).



Слика 3.2 Динамички неурон

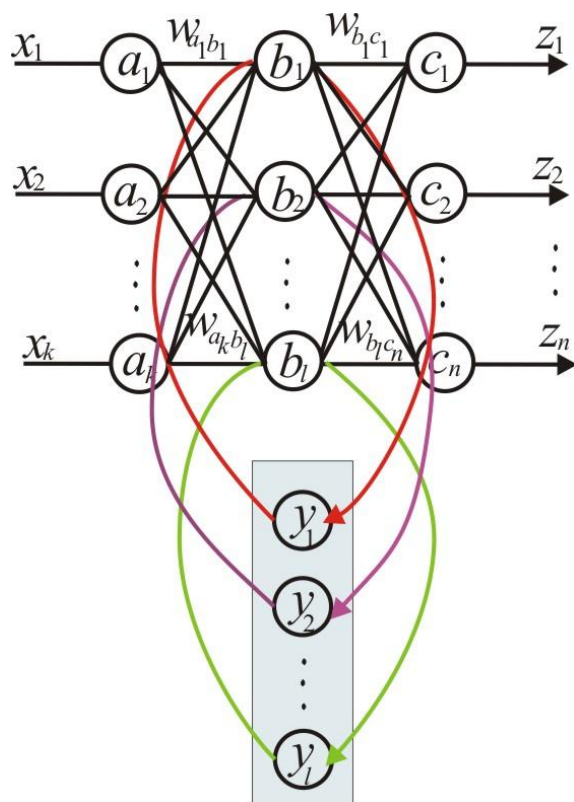
Излаз из динамичког неурона се рачуна као:

$$z(t) = w_{1,1}x(t) + w_{1,2}x(t-1) \quad (3.1)$$

3.2. Елманова мрежа

Елманова рекурентна неуронска мрежа добила је име по свом изумитељу Jeff-у Elman-у. Како би се омогућило обављање послова као што су секвентна предвиђања, Елман је додао у стандардну мрежу без повратних веза слој „контекстних јединица“.

У примеру на слици 3.3 коришћена је трослојна мрежа. Контекстне јединице су означене са у.



Слика 3.3 Пример трослојне Елманове мреже

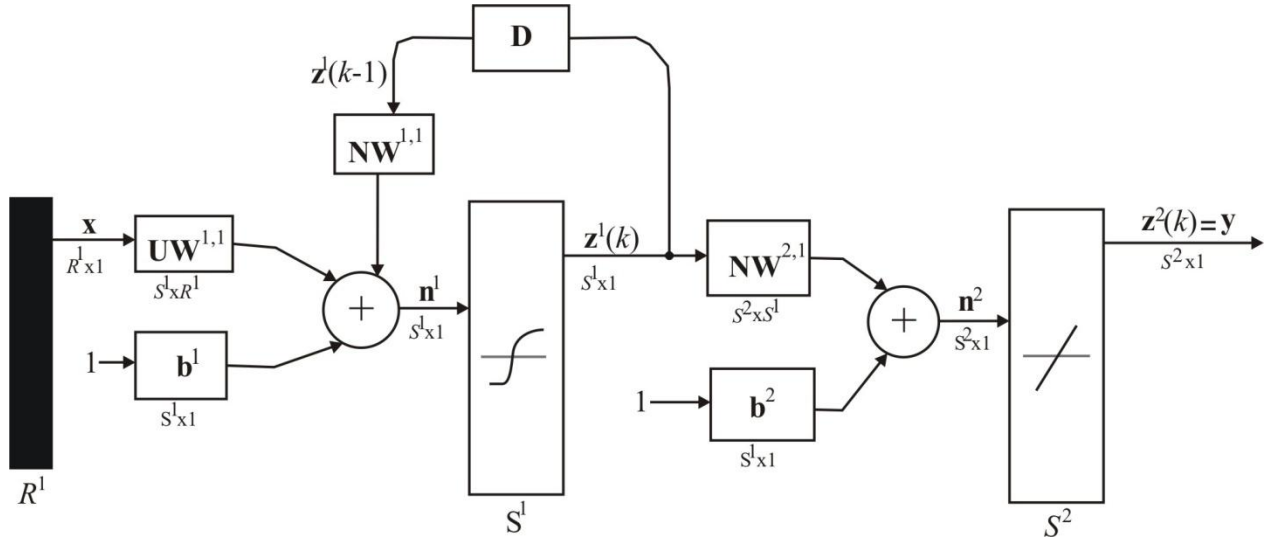
Механизам функционисања је такав да се улаз доводи на стандардан начин (као код мрежа без повратних веза), затим фиксираним везама са тежином један се из скривених јединица вредност доводи у контекстне јединице, а након тога се примењује правило учења. Понављањем овог поступка, у сваком временском кораку мрежа има своје интерно стање које мрежи даје могућност предвиђања секвенци што код стандардног вишеслојног перцептрона није могуће.

Поред претходно датог објашњења, принцип функционисања ове мреже најбоље описује Елман (1990.):

“ И улазне јединице и контекстне јединице активирају скривене јединице, а затим скривене јединице се простиру унапред да активирају излазне јединице. Скривене јединице се простиру такође и уназад да активирају контекстне јединице. Ово представља активацију унапред. Зависно од задатка, може бити или не фаза учења у овом временском циклусу. Ако је тако, излаз се пореди са улазом учитеља и пропагација грешке уназад се користи да постепено прилагоди јачине веза. Рекурентне везе су фиксне на 1.0 и не подлежу прилагођавању. У следећем временском кораку $t+1$ горња секвенца се понавља. Овај пут контекстне јединице садрже вредности које су управо вредности скривених јединица у времену t . Ове контекстне јединице тако обезбеђују мрежи меморију.“

На слици 3.4 приказана је двослојна Елманова мрежа, која је и најчешћа. Код ове мреже уочава се да се повратне информације простиру од излаза првог слоја до улаза истог тог слоја. Кашњење у овој повратној вези омогућава да се вредност у претходном

временском кораку ускладишти како би се искористила у текућем, чинећи тако Елманову мрежу способном да детектује и генерише временски-променљиве обрасце. На слици се такође уочава да мрежа има сигмоидалне неуроне у рекурентном и линеарне неуроне у излазном слоју, тако да мрежа има могућност да са произвольном прецизношћу апроксимира неку функцију (са ограниченим бројем дисконтинуитета). Како би се то остварило, потребно је да буде довољно неурона у скривеном слоју, зависно од сложености.



Слика 3.4 Пример двослојне Елманове мреже

Излаз из скривеног (рекурентног) слоја се рачуна као:

$$\mathbf{z}^1(k) = \mathbf{tansig}(\mathbf{n}^1), \text{ односно:} \quad (3.2)$$

$$\mathbf{z}^1(k) = \frac{2}{1 + e^{-2\mathbf{n}^1}}, \text{ где је:} \quad (3.3)$$

$$\mathbf{n}^1 = \mathbf{x} \cdot \mathbf{UW}^{1,1} + \mathbf{NW}^{1,1} \cdot \mathbf{z}^1(k-1) + \mathbf{b}^1. \quad (3.4)$$

Излаз из неуронске мреже се рачуна као:

$$\mathbf{z}^2(k) = \mathbf{purelin}(\mathbf{n}^2), \text{ односно:} \quad (3.5)$$

$$\mathbf{z}^2(k) = k \cdot \mathbf{n}^2, \text{ где су:} \quad (3.6)$$

k – коефицијент пропорционалности,

$$\mathbf{n}^2 = \mathbf{NW}^{2,1} \cdot \mathbf{z}^1(k) + \mathbf{b}^2. \quad (3.7)$$

$\mathbf{UW}^{1,1}$, $\mathbf{NW}^{1,1}$ и $\mathbf{NW}^{2,1}$ су матрице тежина веза.

Излаз мреже је функција не само тежина, прагова активација, и мрежног улаза, већ такође и излаза скривених неурона у претходним тачкама у времену.

3.3. Хопфилдова мрежа

Хопфилдова мрежа је облик рекурентне структуре неуронске мреже и служи као меморија адресибилна по садржају (content addressable memory).

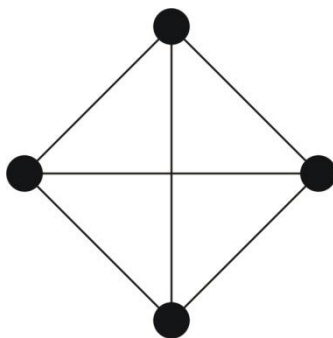
Ову мрежу је изумео John Hopfield 1982. године, по коме је добила назив. Сваки од чворова (“неурона”) у Хопфилдовој мрежи може бити у стању 1 и 0 или 1 и -1 зависно од тога да ли улаз премашује праг. Дакле, сваки појединачни чвор у мрежи је у основи перцептрон.

Везе у овој мрежи су структуриране тако да чворови нису повезани сами са собом и све везе морају бити симетричне. Пре објашњења процеса тренирања Хопфилдове мреже потребно је размотрити појам “енергија”. Овде енергија представља скаларну вредност мреже повезану са сваким стањем, при чему ће мрежа током израчунавања вредности (update стања) случајним редоследом конвергирати ка стању које представља локални минимум у овој функцији енергије (које је уједно и стабилно стање у мрежи).

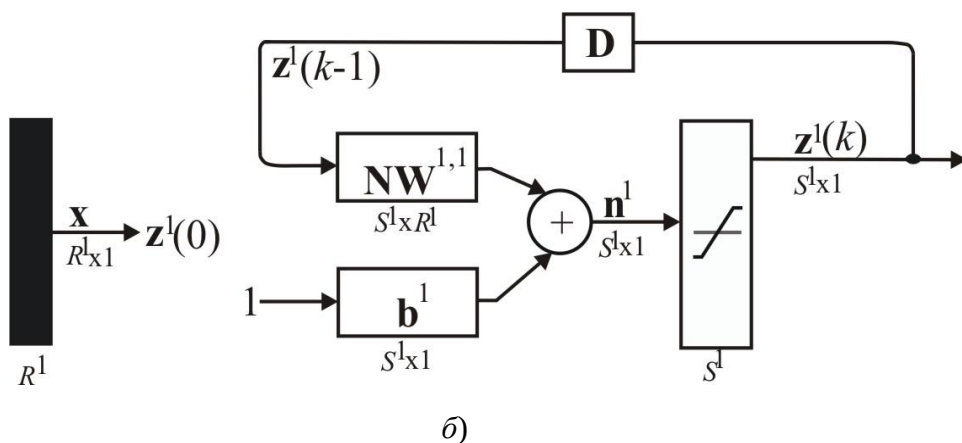
Дакле, тренирање Хопфилдове мреже се састоји у томе да стања која мрежа треба да запамти буду уједно и локални минимуми функције енергије.

Чињеница да мрежа служи као меморија адресибилна по садржају објашњава се тиме да ако на улаз доведемо стање чији је један део уједно и део стабилног стања, мрежа ће конвергирати ка том стабилном стању.

На слици 3.5 а) приказана је структура, док је на 3.5 б) приказан пример Хопфилдове мреже са елементима који дефинишу њено функционисање.



а)

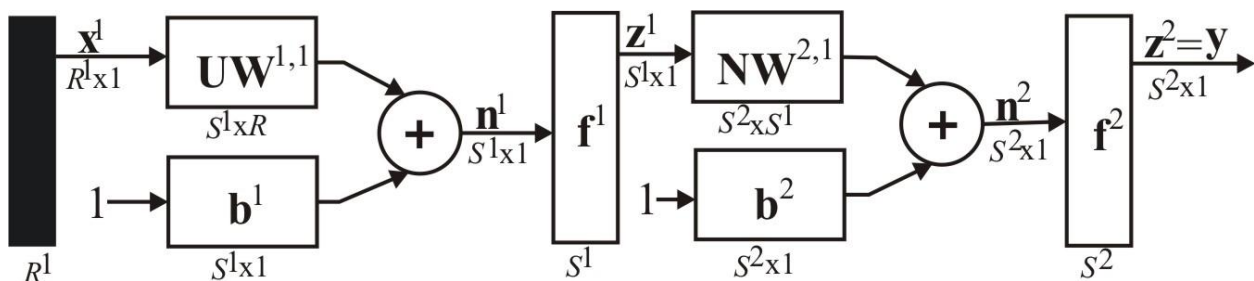


Слика 3.5 (а) Структура и (б) пример Хопфилдове неуронске мреже

Хопфилдова мрежа није дизајнирана за обраду секвенци узорака и захтева стационарне улазе.

3.4. Дигитална рекурентна неуронска мрежа

На слици 3.6 приказана је двослојна неуронска мрежа без повратних веза.



Слика 3.6 Пример двослојне неуронске мреже без повратних веза

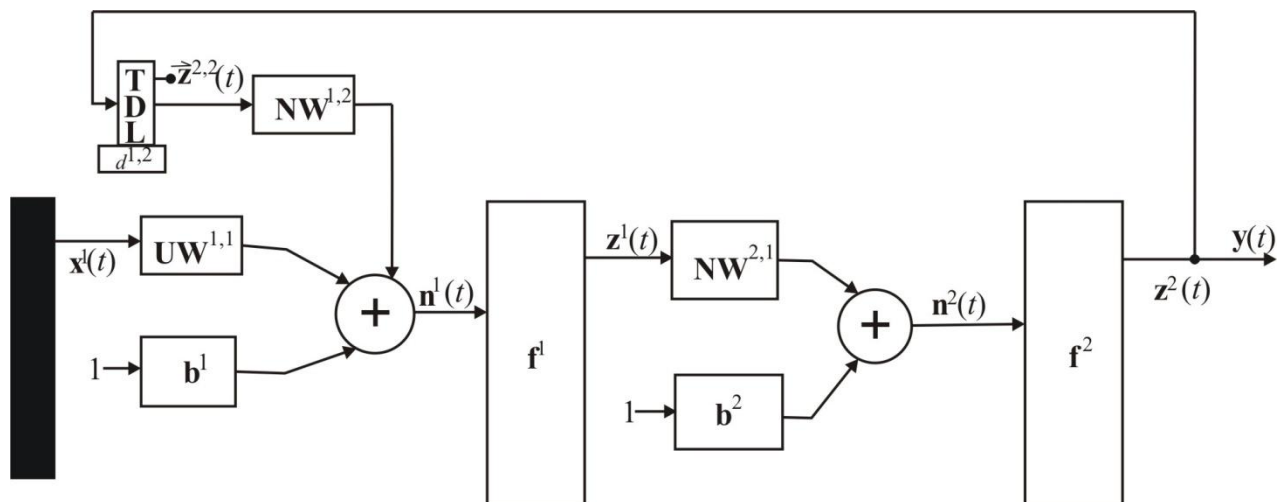
Излаз из првог слоја неуронске мреже се рачуна као:

$$\mathbf{z}^1 = \mathbf{f}^1(\mathbf{UW}^{1,1} \cdot \mathbf{x}^1 + \mathbf{b}^1) \quad (3.8)$$

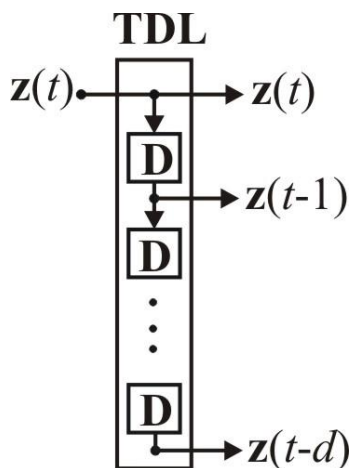
Излаз из неуронске мреже се рачуна као:

$$\mathbf{z}^2 = \mathbf{f}^2(\mathbf{NW}^{2,1} \cdot \mathbf{z}^1 + \mathbf{b}^2) \quad (3.9)$$

Двослојна неуронска мрежа са повратним везама које повезују излазе из неуронске мреже и јединице у невидљивом слоју приказана је на слици 3.7. Ова мрежа је позната као NARX мрежа (нелинеарни ауто-регресиони модел са спољашњим улазом). На слици 3.8 приказана је структура блока TDL.

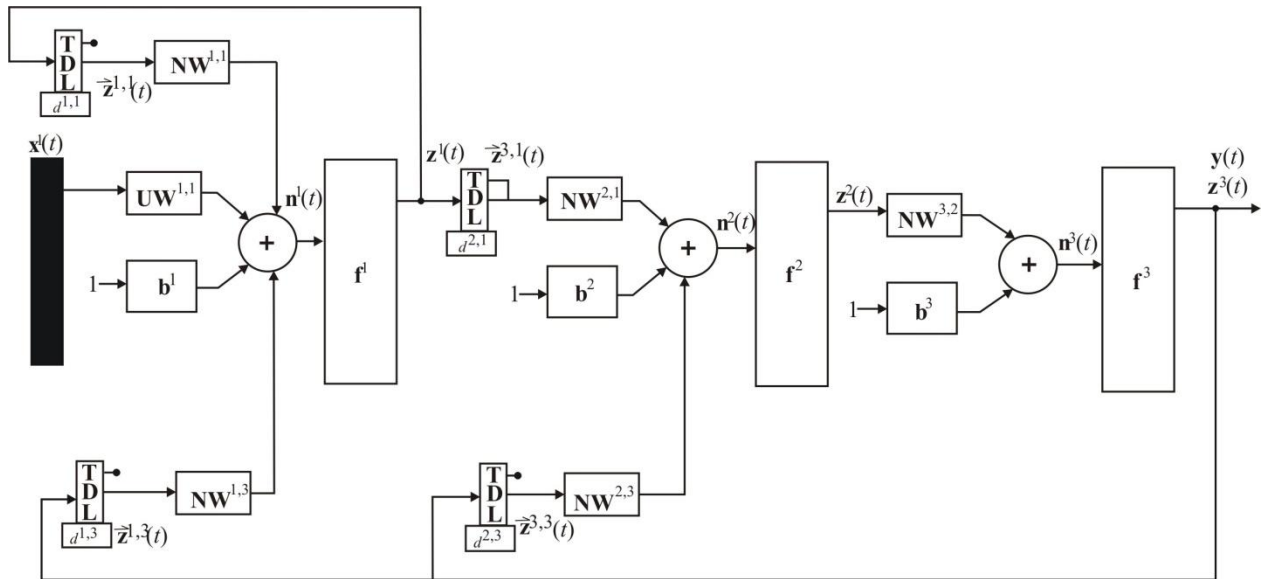


Слика 3.7 Пример двослојне дигиталне неуронске мреже



Слика 3.8 Блок којим се представља кашњење сигнала $z(t)$

Пример трослојне дигиталне рекурентне неуронске мреже дат је на слици 3.9.



Слика 3.9 Трослојна дигитална рекурентна неуронска мрежа

3.4.1. Принципи динамичког учења

У претходном делу текста, на слици 3.6, приказана је вишеслојна мрежа без повратних веза. Уколико се мрежи задају скупови парова улаз/излаз: $\{(\mathbf{x}_1, \mathbf{t}_1), (\mathbf{x}_2, \mathbf{t}_2), \dots, (\mathbf{x}_Q, \mathbf{t}_Q)\}$, њен задатак биће да научи удруживања између тих парова. То се може постићи коришћењем основне једначине за ову мрежу, која гласи:

$$\mathbf{z}^k = \mathbf{f}^k \left(\sum_i \mathbf{UW}^{k,i} \mathbf{x}^i + \sum_j \mathbf{NW}^{k,j} \mathbf{z}^j + \mathbf{b}^k \right), \quad (3.10)$$

при чему се k повећава у току процеса симулације. Индекс перформансе за мрежу је:

$$E(\mathbf{p}) = \sum_{q=1}^Q (\mathbf{t}_q - \mathbf{y}_q)^T (\mathbf{t}_q - \mathbf{y}_q) = \sum_{q=1}^Q \mathbf{e}_q^T \mathbf{e}_q, \quad (3.11)$$

при чему су:

$E(\mathbf{p})$ – критеријумска функција,

$\mathbf{y}_q$ – излаз мреже, када је дат q -ти улаз $\mathbf{x}_q$ и

$\mathbf{p}$ – вектор који садржи све тежине и прагове активација мреже.

Задатак мреже је да научи вектор $\mathbf{p}$ да минимализује E . У случају стандардног алгоритма са пропагацијом грешке уназад користи се најоштрије силазно правило учења. Индекс перформансе се може апроксимирати на следећи начин:

$$\hat{E} = \mathbf{e}_q^T \mathbf{e}_q. \quad (3.12)$$

У претходној једначини уочава се да је укупна сума квадрата замењена квадратним грешкама за један улазно/излазни пар, а приближно најстрмији (градијент) силазни алгоритам тада ће бити:

$$\Delta w_{ij}^{k,l} = -\alpha \frac{\partial \hat{E}}{\partial w_{ij}^{k,l}}, \quad \Delta b_i^k = -\alpha \frac{\partial \hat{E}}{\partial b_i^k} \quad (3.13)$$

при чему је α стопа учења.

Једначина:

$$s_i^k \equiv \frac{\partial \hat{E}}{\partial n_i^k} \quad (3.14)$$

представља осетљивост индекса перформансе на промене у мрежном улазу јединице i у слоју k .

Следеће три једначине показују једнакости добијене коришћењем правила ланца:

$$\frac{\partial \hat{E}}{\partial w_{ij}^{m,l}} = \frac{\partial \hat{E}}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial w_{ij}^{m,l}} = s_i^m x_j^l, \quad \frac{\partial \hat{E}}{\partial w_{ij}^{m,l}} = \frac{\partial \hat{E}}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial w_{ij}^{m,l}} = s_i^m z_j^l, \quad (3.15)$$

$$\frac{\partial \hat{E}}{\partial b_i^m} = \frac{\partial \hat{E}}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial b_i^m} = s_i^m.$$

Као што се у наредној једначини може видети осетљивости одговарају приказаном принципу понављања, када се m повећава у току редоследа пропагације уназад (који је инверзан редоследу симулације):

$$s^m = \dot{E}^m(\mathbf{n}^m) \sum (\mathbf{N} \mathbf{W}^{i,m})^T s^i \quad (3.16)$$

при чему је

$$\dot{E}^m(\mathbf{n}^m) = \begin{bmatrix} \dot{f}^m(n_1^m) & 0 & \dots & 0 \\ 0 & \dot{f}^m(n_2^m) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \dot{f}^m(n_{s^m}^m) \end{bmatrix} \quad (3.17)$$

и

$$\frac{\partial f^m}{\partial n_j^m}(n_j^m) = \frac{\partial f^m(n_j^m)}{\partial n_j^m} . \quad (3.18)$$

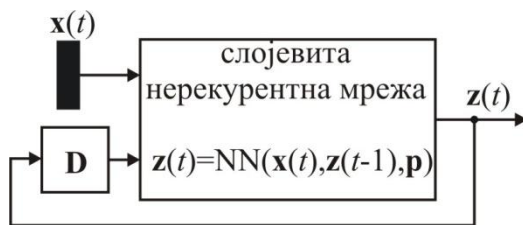
Може се показати да је на излазном слоју овај процес враћања иницијализован и дефинисан је као:

$$s^M = -2 \mathbf{E}^M(\mathbf{n}^M)(\mathbf{t}_q - \mathbf{y}_q) . \quad (3.19)$$

Кораци кроз које целокупни алгоритам учења треба да прође су следећи:

- применом једначине (3.10) улаз се простире напред
- осетљивост се пропагира уназад применом једначина (3.19) и (3.16)
- ажурирање тежина и прагова активација применом једначина (3.13) и (3.15).

У наредном делу текста обављена је анализа LDRN, слика 3.7. Уколико би се користио исти алгоритам опадања градијента (једначина 3.13), који је био коришћен у стандардном поступку пропације уназад јавио би се одређени проблем. Приликом покушаја да се пронађе еквивалент једначине (3.15) уочава се да постоје два различита дејства на излаз мреже од стране тежинских коефицијената и прагова активације. Један од њих је директан ефекат (објашњен једначином 3.15) и индиректан ефекат. Овај ефекат је назван индиректан јер постоје улази у мрежу, као нпр. $\bar{\mathbf{z}}^{1,2}(t)$, који су и сами функције тежина и прагова активација. Због бољег појашњења овог индиректног ефекта потребно је користити динамичку пропацију уназад, илустровану сликом 3.10. Проста рекурентна мрежа, приказана на слици 3.10 има просту повратну спрегу, која је путем једног кашњења доведена из излаза на улаз мреже.



Слика 3.10 Проста рекурентна мрежа

На истој слици вектор $\mathbf{p}$ означава параметре мреже (тежине и прагове активација), док је вектор $\mathbf{z}(t)$ излаз дате мреже на временском кораку t .

Да би се минимизирала једначина

$$E(\mathbf{p}) = \sum_{t=1}^Q (\mathbf{t}(t) - \mathbf{z}(t))^T (\mathbf{t}(t) - \mathbf{z}(t)) \quad (3.20)$$

приликом примене опадања градијента, потребно је претходно пронаћи градијент од E узимајући у обзир параметре мреже. Овом проблему могуће је приступити на два различита начина, којима је заједничко то да оба користе правило ланца:

$$\frac{\partial E}{\partial \mathbf{p}} = \sum_{t=1}^Q \left[\frac{\partial \mathbf{z}(t)}{\partial \mathbf{p}} \right]^T \times \frac{\partial^e E}{\partial \mathbf{z}(t)} \quad (3.21)$$

или

$$\frac{\partial E}{\partial \mathbf{p}} = \sum_{t=1}^Q \left[\frac{\partial^e \mathbf{z}(t)}{\partial \mathbf{p}} \right]^T \times \frac{\partial E}{\partial \mathbf{z}(t)} \quad (3.22)$$

при чему натпис e означава експлицитни извод, не дајући при томе објашњење индиректних ефеката кроз време. Стандардним алгоритмом пропагације уназад (једначина 3.16) могуће је добити експлицитне изводе, међутим, потребне су додатне једначине како би се пронашли комплетни изводи захтевани у једначини (3.21) и једначини (3.22):

$$\frac{\partial \mathbf{z}(t)}{\partial \mathbf{p}} = \frac{\partial^e \mathbf{z}(t)}{\partial \mathbf{p}} + \frac{\partial^e \mathbf{z}(t)}{\partial \mathbf{z}(t-1)} \times \frac{\partial \mathbf{z}(t-1)}{\partial \mathbf{p}} \quad (3.23)$$

и

$$\frac{\partial E}{\partial \mathbf{z}(t)} = \frac{\partial^e E}{\partial \mathbf{z}(t)} + \frac{\partial^e \mathbf{z}(t+1)}{\partial \mathbf{z}(t)} \times \frac{\partial E}{\partial \mathbf{z}(t+1)}. \quad (3.24)$$

Разматрањем претходних једначина, примећује се да је кључни термин у једначинама (3.21) и (3.23):

$$\frac{\partial \mathbf{z}(t)}{\partial \mathbf{p}}, \quad (3.25)$$

и он мора бити пропагиран унапред кроз време. На основу тога, може се рећи да једначине (3.21) и (3.23) чине алгоритам пропагације унапред (енг. forward perturbation – FP). Насупрот томе, једначине (3.22) и (3.24) чине алгоритам пропагације уназад кроз време (енг. backpropagation-through-time (BTT)). У њима се уочава да је кључни термин

$$\frac{\partial E}{\partial \mathbf{z}(t)}, \quad (3.26)$$

и он мора бити пропагиран уназад кроз време.

Са једне стране, FP алгоритам захтева више рачунања од ВТТ алгоритма, док са друге стране имплементација ВТТ алгоритма није могућа из разлога што излази морају бити рачунати за све временске кораке, а затим изводи морају бити враћени путем пропагације уназад до почетне временске тачке. Међутим, када је реч о имплементацији у реалном времену, FP алгоритам је сасвим погодан јер изводи могу бити рачунати у сваком временском кораку.

3.5. Динамичка пропагација уназад за LDRN (Layered Digital Recurrent Network)

Како би се FP алгоритам применио на произвољне LDRN потребно га је генерализовати. У циљу бољег појашњења овог алгоритма, у наставку су дате одређене дефиниције које се односе на LDRN.

Пре свега, битно је навести да се слој ове мреже састоји од скупа тежина, коришћених блокова за кашњење, функције сумирања и функције преноса. Улазне тежине, означене са $\mathbf{UW}^{i,j}$ представљају посебне тежине са којима су повезани улази мреже, при чему j представља број улазног вектора који доставља тежину, а i означава број слоја до кога је тежина повезана. Тежине, означене са $\mathbf{NW}^{i,j}$ представљају тежине слоја које повезују један слој са другим, при чему је j број слоја који долази у тежину, док i означава број слоја на излазу тежине. Како би се израчунао одговор мреже у фазама, слој по слој потребно је применити поступак назван редоследом симулације. Он је специфичан по томе што је потребно наставити у правилном слојном редоследу, како би неопходни улази у сваки слој били доступни.

За разлику од овог поступка, редослед пропагације уназад карактерише настављање у супротном редоследу. Овај поступак се примењује у циљу пропагације уназад извода за рачунање градијента. У циљу једноставнијег описа алгоритма обуке, LDRN је подељена на подмреже. Подмрежа као саставни део мреже, нема текуће кашњење осим на улазу подмреже.

Да би се пронашла последња подмрежа, потребно је кренути од првог слоја у редоследу пропагације уназад и кретање наставити овим редоследом до проналаска слоја са кашњењима. Тај слој постаје први слој у последњој подмрежи. На исти начин могуће је пронаћи претходну подмрежу. Потребно је кренути од следећег слоја у редоследу пропагације уназад и наставити до наилаaska следећег слоја са кашњењима. Овај процес могуће је понављати све док се не дође до последњег слоја у редоследу пропагације уназад, а до тог тренутка прва подмрежа је већ дефинисана. Редослед симулације за слојеве је дефинисан у претходном делу текста, међутим постоји и подмрежни редослед симулације који се одвија од прве до последње подмреже.

Као што се на слици 3.9 може видети, подмрежа је представљена бројем њеног излазног слоја. Поред тога, на овој слици се уочава да постоје три слоја и две подмреже, да је редослед симулације за мрежу 1-2-3, редослед пропагације уназад 3-2-1 и да је подмрежни редослед симулације 1-3.

3.5.1. Експлицитни изводи

Да би се алгоритам пропагације унапред (FP) могао применити на произвољној LDRN потребно га је генерализовати (код једначине (3.21) и (3.23)). На десној страни једначине (3.23) уочавају се два термина. Први термин се односи на експлицитни извод перформансе са поштовањем тежина. Он чини директни утицај тежина на перформансе и могуће га је израчунати стандардним алгоритмом са пропагацијом уназад (као у једначини 3.16). Други термин који постоји у једначини означава то да тежине имају секундарни ефекат кроз претходни мрежни излаз. Када је у питању општа LDRN, постоји могућност да се она састоји од много различитих повратних петљи. У том случају може бити много различитих ознака десно у једначини (3.23), док би свака од тих ознака могла имати посебну једначину за ажурирање укупног извода кроз време. Да би се омогућио развој FP алгоритма који је овде описан, постоји један термин (и додатна једначина) за свако место где је једна подмрежа улаз у другу подмрежу. Границе подмреже, како је претходно речено, одређене су локацијама коришћених блокова кашњења. Како би се извршило пропагирање експлицитних извода у оквиру подмреже може се применити стандардна пропагација уназад (као у једначини 3.16). Међутим, у границама подмреже мора бити коришћена једначина слична (3.23) да би се израчунао укупни извод, који укључује директне и индиректне ефекте. Да би се израчунали експлицитни изводи, код сваке подмреже је потребно извршити процес пропагације уназад, а саме једначине подразумевају рачунање извода излаза подмреже са поштовањем сваког излаза слоја у подмрежи.

Једначина:

$$\frac{\partial^e \mathbf{z}^{ja}(t)}{\partial \mathbf{z}^i(t)} = \sum_k \frac{\partial^e \mathbf{n}^k(t)}{\partial \mathbf{z}^i(t)} \times \frac{\partial^e \mathbf{z}^{ja}(t)}{\partial \mathbf{n}^k(t)} \quad (3.27)$$

представља основну једначину са следећим ознакама:

$\mathbf{z}^{ja}(t)$ - излаз подмреже,

$\mathbf{z}^i(t)$ - излаз слоја у подмрежи ja ,

$\mathbf{n}^k(t)$ - улаз мреже слоја k , са везом од слоја i .

У овој једначини индекс i се увећава током редоследа пропагације уназад.

С обзиром на једначину:

$$\mathbf{S}^{k,ja} \equiv \frac{\partial^e \mathbf{z}^{ja}(t)}{\partial \mathbf{n}^k(t)} \quad \text{и} \quad \frac{\partial^e \mathbf{n}^k(t)}{\partial \mathbf{z}^i(t)} = \mathbf{N}\mathbf{W}^{k,i} \quad (3.28)$$

једначина (3.27) се може писати у облику:

$$\frac{\partial^e \mathbf{z}^{ja}(t)}{\partial \mathbf{z}^i(t)} = \sum_k \mathbf{N}\mathbf{W}^{k,i} \times \mathbf{S}^{k,ja} . \quad (3.29)$$

Матрица идентитета $\mathbf{I}$ се рачуна као:

$$\mathbf{I} = \frac{\partial^e \mathbf{z}^{ja}(t)}{\partial \mathbf{z}^{ja}(t)} \quad (3.30)$$

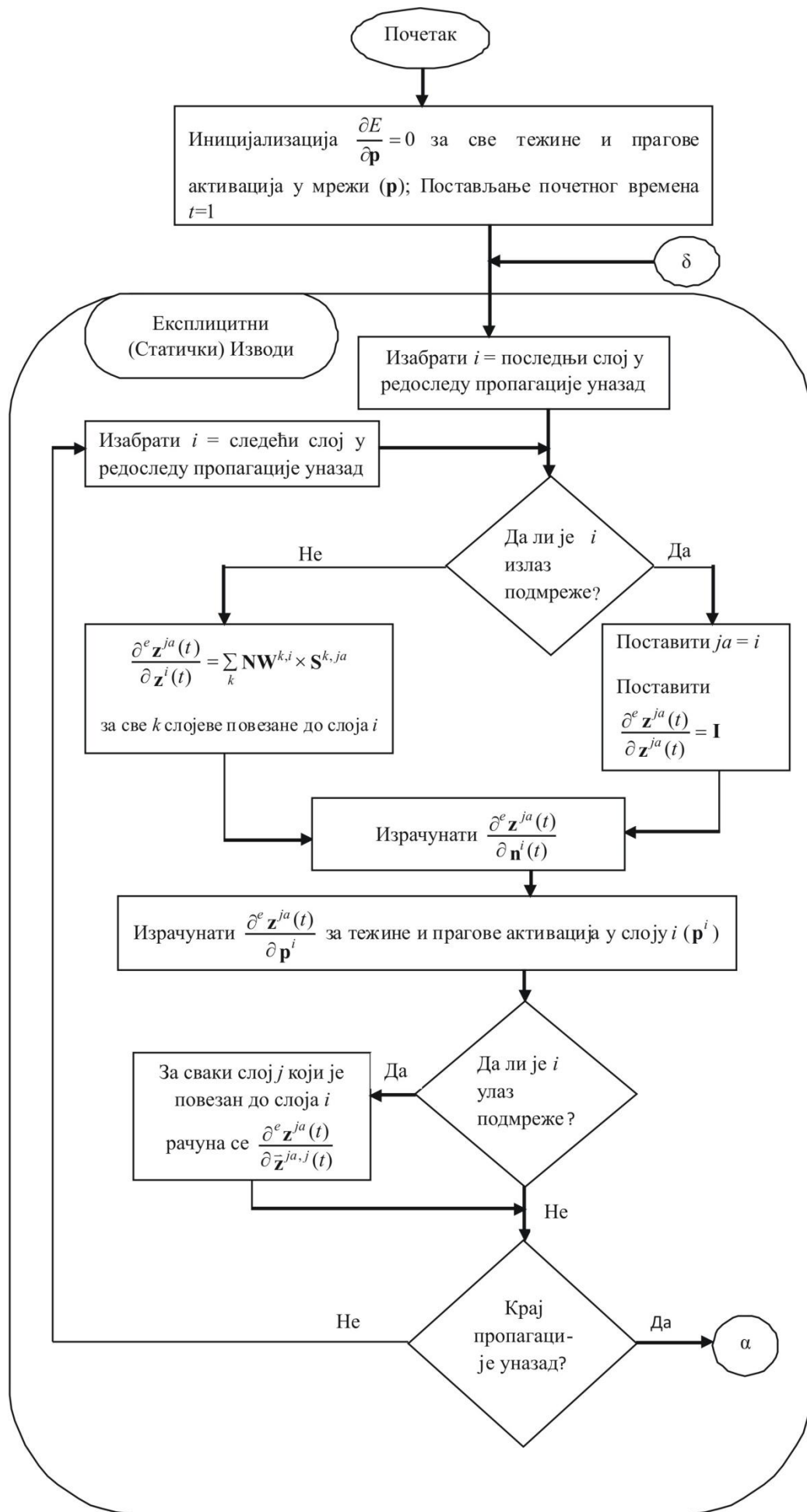
и њена димензија је величина слоја ja . Ова једначина показује да понављање током кога је i увећавано дуж редоследа пропагације уназад почиње у последњем слоју у подмрежи.

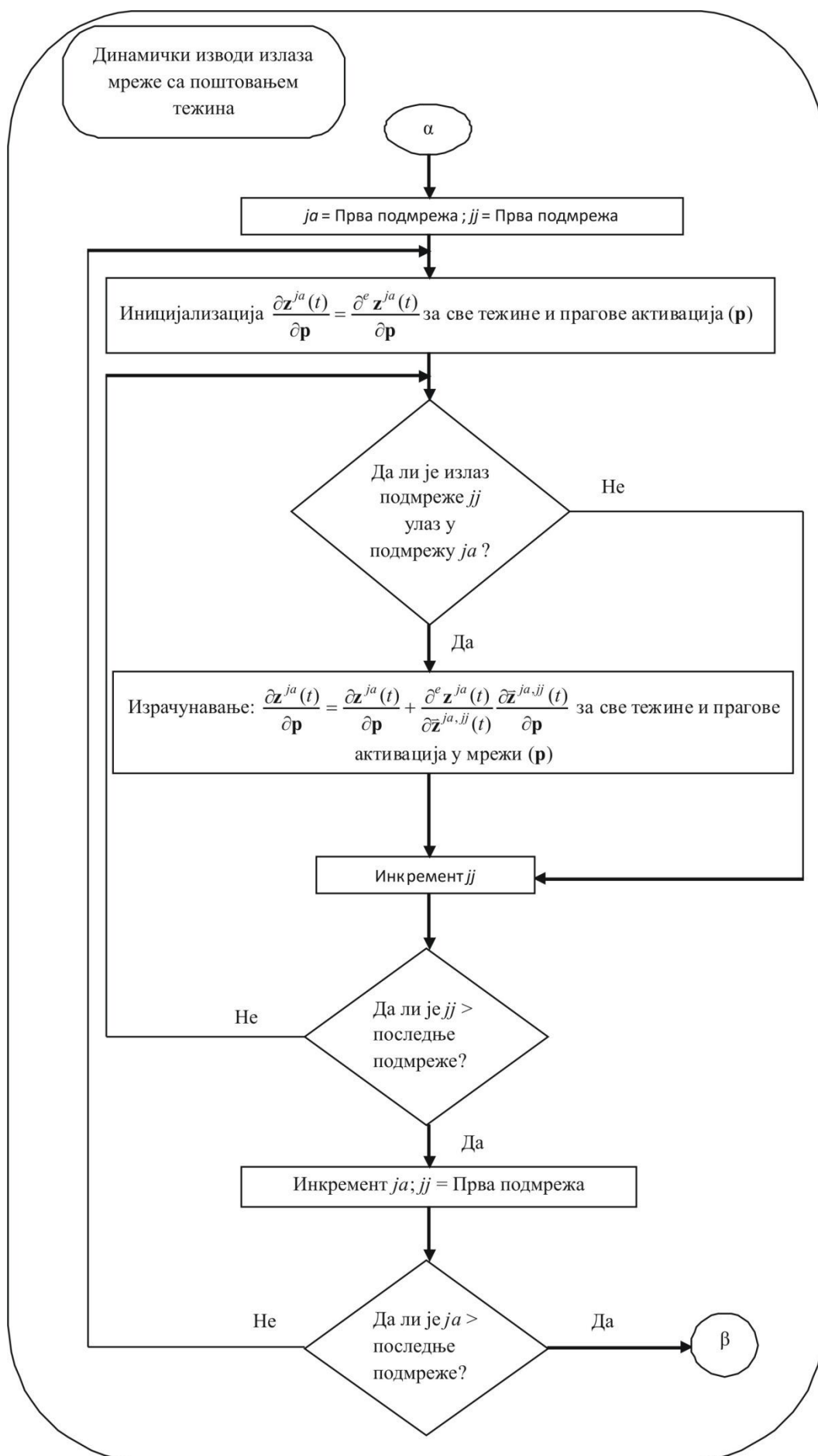
3.5.2. Комплетан FP алгоритам за LDRN

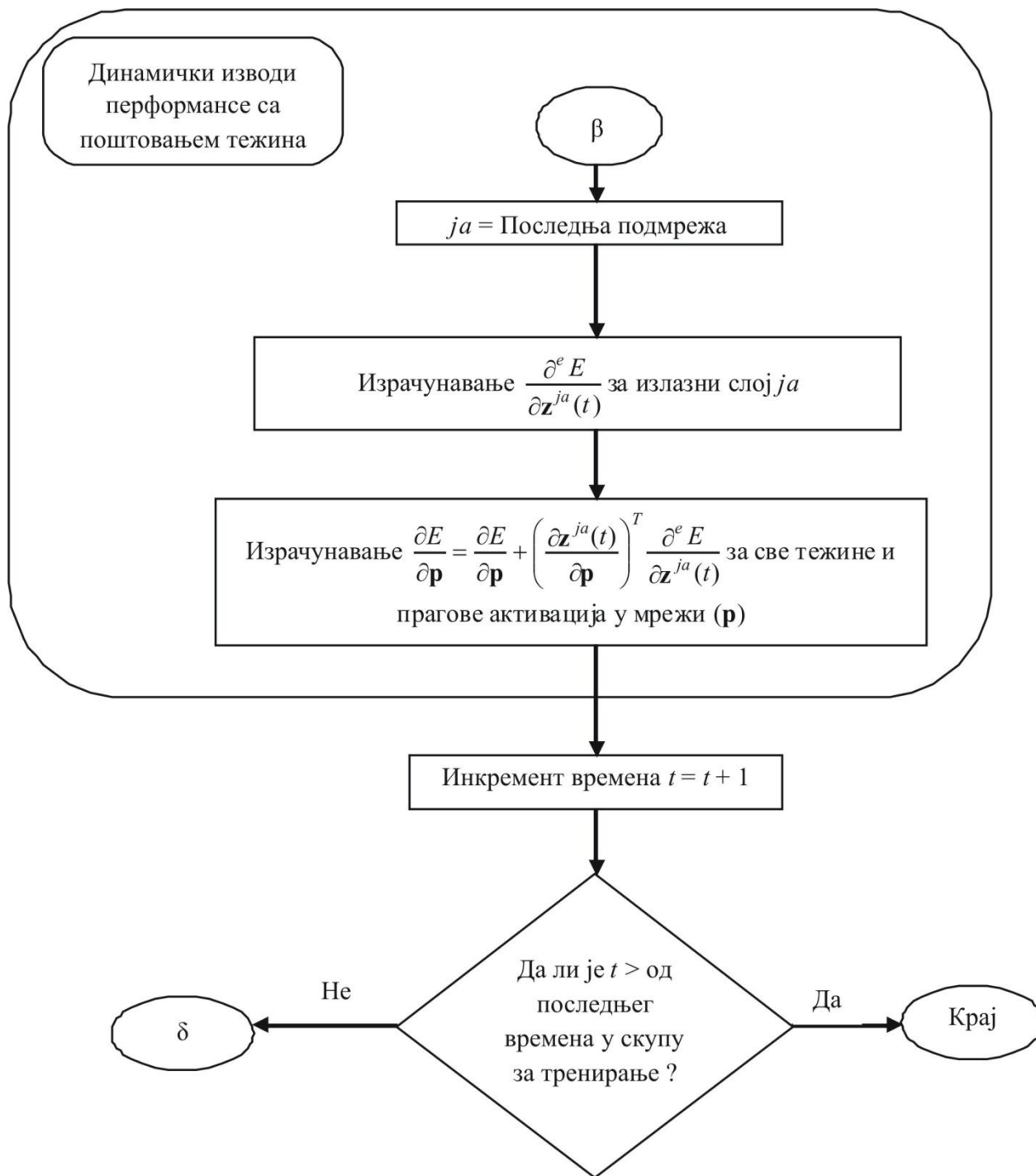
У овом одељку је дат опис генерализованог FP алгоритма за произвољну LDRN. Овај алгоритам се састоји од два дела представљена једначином (3.21) и (3.23). За разлику од једначине (3.21) која остаје иста за LDRN као за просту мрежу представљену на слици 3.10, једначина (3.23) мора бити модификована. Другим речима, за сваки подмрежни излаз имамо једну једначину као једначина (3.23), а свака од тих једначина има један термин за експлицитни извод и додатни термин за сваки подмрежни излаз.

Дијаграм тока, представљен на слици 3.11, садржи комплетни FP алгоритам за LDRN мрежу, а састоји се од три главна дела:

1. У овом делу израчунавају се експлицитни (статички) изводи, као у једначини (3.29). Они су потребни као део динамичких једначина.
2. Израчунавање динамичких извода подмрежних излаза са уважавањем тежина као у једначини (3.23).
3. Израчунавање динамичких извода перформансе са уважавањем тежина као у једначини (3.21).







Слика 3.11 Дијаграм тока комплетног FP алгоритма за LDRN [51]

4. ПРЕДВИЂАЊЕ ДУГОРОЧНЕ ПОТРОШЊЕ ПРИРОДНОГ ГАСА У ЗЕМЉАМА ЕВРОПСКЕ УНИЈЕ ПРИМЕНОМ РЕКУРЕНТНИХ НЕУРОНСКИХ МРЕЖА

У овом раду двослојна дигитална рекурентна мрежа (Слика 3.4) и двослојна рекурентна NARX мрежа (слика 3.7) са сигмоидалним неуронима у скривеном слоју и линеарним неуроном у излазном слоју је коришћена за предикцију потрошње природног гаса у земљама ЕУ (Италија, Аустрија, Бугарска, Француска, Уједињено краљевство, Ирска, Белгија и Шпанија). Обучавање и тестирање неуронских мрежа извршено је помоћу софтвера MATLAB R211a. У оквиру одељка о резултатима симулације биће нешто више речи о самом поступку имплементације поменутих рекурентних мрежа.

4.1. Формулација проблема

Енергија је битан фактор за друштвени и економски развој сваке нације, при чему је раст потрошње енергије суштински повезан са растом у привреди.

Природни гас је важан извор енергије и представља центар пажње за ствараоце енергетске политике. Такође, моделирање енергетске потражње и процена потрошње гаса су важни због свог утицаја на развој друштва. Како су конвенционалне методе у моделирању потрошње гаса праћене различитим тешкоћама, проналажење ефикасне и погодне функције апроксимације је од виталног значаја.

На потрошњу гаса у некој земљи утичу различити фактори. Економски раст и конкурентност цена су кључни, међутим утичу и националне карактеристике као што су домаћи извори, уговори о снабдевању, транспортни капацитет (и приступ), енергетски микс, политика, порези и механизми финансијске подршке. У будућности ће улога гаса у Европи све више зависити од успостављене енергетске политике и мера заштите животне средине, како на регионалном тако и на националном нивоу.

Значај тачне процене и предвиђања је двострука: поред тога што су помоћ менаџерима у дефинисању исправних стратегија, оне такође могу бити коришћене у понашању потрошње у циљу стварања ефикасне и тачне политике.

4.1.1. Специфичности природног гаса

Природни гас или земни гас је природно фосилно гориво које се у природи јавља самостално, као суви земни гас, или пак у нафтним налазиштима заједно са нафтом, као влажни земни гас.

Природни гас се примарно састоји од метана, најпростијег угљоводоника (једињење угљеника и водоника), са којим су помешане мање количине других гасовитих угљоводоника као што су етан, бутан и пропан. Састојци овог гаса су важне сировине за хемијске процесе.

У односу на конвенционална горива, природни гас поседује значајне технолошко-економске и еколошке предности. Посматрано у односу на сва остала фосилна горива, природни гас има најмањи коефицијент емисије угљен-диоксида по јединици ослобођене енергије што је битни допринос смањењу глобалне емисије угљен-диоксида.

Још једна од предности природног гаса је то што је уз угаљ и остала чврста горива, једини примарни облик енергије који се, уз елементарну припрему, може непосредно употребљавати. Како се сви остали облици енергије (сирова нафта, водена снага, нуклеарно гориво и сл.) морају трансформисати у погодније облике (деривате нафте, електричну енергију, водену пару или врелу воду), то захтева градњу постројења за енергетске трансформације. Да би се то реализовало потребне су велике инвестиције, тако да је оправдано настојање да се што више природног гаса непосредно употребљава.

Природни гас је готово идеално гориво које не садржи загађиваче, лако се меша са ваздухом и има велику брзину сагоревања без дима, чађи и чврстих остатака. Не загађује околину јер се његовим сагоревањем добија само угљен-диоксид и водена пара.

Како показује искуство земаља са дугом традицијом коришћења, природни гас је и један од најбезбеднијих енергената. Он чини једну петину светских резерви енергије и сматра се енергентом у експанзији.

Употреба природног гаса је вишеструка:

- користи се у индустрији (са уделом од 43%, индустрија је највећи потрошач природног гаса.)
- за производњу електричне енергије
- као сировина у хемијској индустрији (као извор водоника при производњи азотних ђубрива, потенцијално и за гориве ћелије).
- за комерцијалну употребу
- у широкој потрошњи

- за погон моторних возила, као компримовани природни гас и евентуално као течни природни гас. Сматра се једним од најперспективнијих алтернативних горива, с обзиром на то да се карактерише изузетно ниском емисијом моно-азотних оксида, чађи и сумпор диоксида.

Земље са највећим налазиштима природног гаса су: Русија, Сједињене Америчке Државе, Иран, Холандија, Алжир и Блиски исток.

У највеће произвођаче земног гаса убрајају се: Сједињене Америчке Државе, Русија и Канада.

За дистрибуцију природног гаса се користи, пре свега, разграната мрежа гасовода. Европа углавном користи природни гас који испоручује Русија.

4.1.2. Тражња природног гаса у Европи

Иако се откриће и развој Европског домаћег конвенционалног природног гаса догодило касније него у Северној Америци, већ од 1960. и 1970. је постало јасно да ће увоз бити у обавези да задовољи потражњу широм Европске територије.

Међу првима по развоју природног гаса на скали били су северна Европска тржишта, односно, најближа копну Холандије, Уједињеног Краљевства, Норвешке и Данске, замењујући при томе угљ и нафтне производе у подручју грејања и индустријском сектору.

Јужно и централно Европско национално тржиште почиње да се развија и расте захваљујући допремању увозним токовима гасовода из Совјетског Савеза и северне Африке, и залиха течног природног гаса од растуће мреже добављача.

У наредном периоду, тј. од 1990. на овамо, распрострањено усвајање врло ефикасног комбинованог циклуса гасне турбине, пружило је додатни подстицај за европску тражњу гаса у електроенергетском сектору, поново узимајући тржишни удео од угља и нафтних деривата.

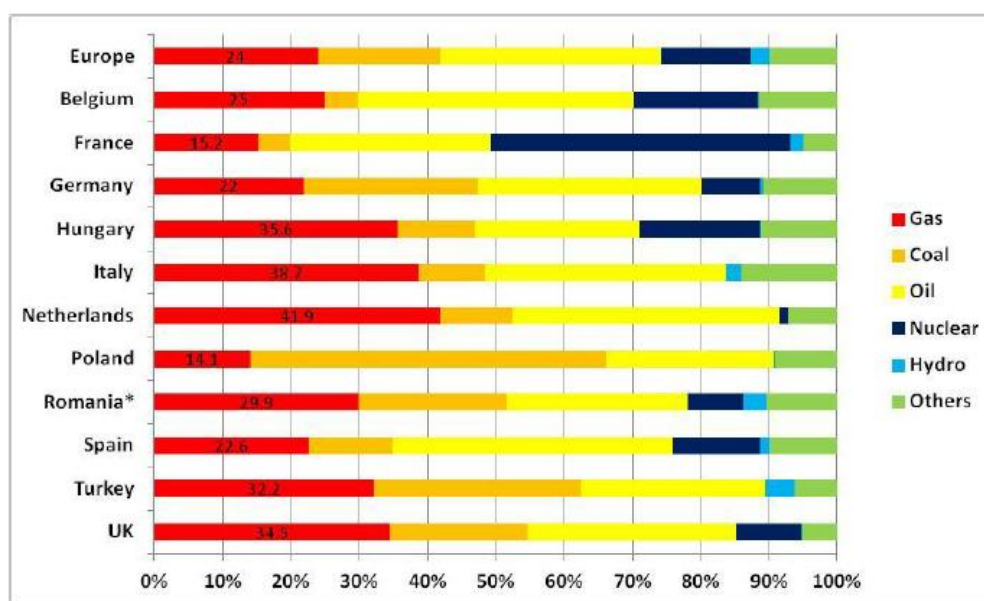
У раним 2000. изгледи за наставак здравог раста будуће Европске тражње гаса су изгледали осигураним, захваљујући сталном подстицању Русије да обезбеди оштар темпо узводних инвестиција за реализацију предвиђеног.

У наредном периоду Европски раст тражње је успорен, чак и пре финансијске кризе 2008. и накнадне финансијске рецесије. Неки од разлога за овакву ситуацију су сазревање тржишта, низак раст становништва, више цене гаса, као и миграција прерађивачке индустрије у друге делове света.

Како је енергетски сектор жариште већине неизвесности, методе за предикцију раста тражње гаса представљају важан фактор за извођење закључка о будућим трендовима.

4.1.3. Улога природног гаса у укупној примарној енергији

Током времена коришћење гаса је све више замењивало нафту и у мањој мери угаљ, с обзиром на то да је тржишни удео природног гаса нагло порастао са мање од 10% од укупне количине примарне енергије у раним 1970. на око 24% у 2012. Међутим, иза ове глобалне слике могу се уочити значајне разлике имајући у виду да продор гаса варира од земље до земље. Тако је удео коришћења гаса изнад 30% од укупне потражње за енергијом у Холандији, Италији, Мађарској и Великој Британији, али 15% или мање од укупне количине примарне енергије у Француској и Пољској, како је показано на слици 4.1. Поменуте разлике могу се објаснити националном задужбином природних ресурса и историјским развојем гаса и тржишта електричне енергије.



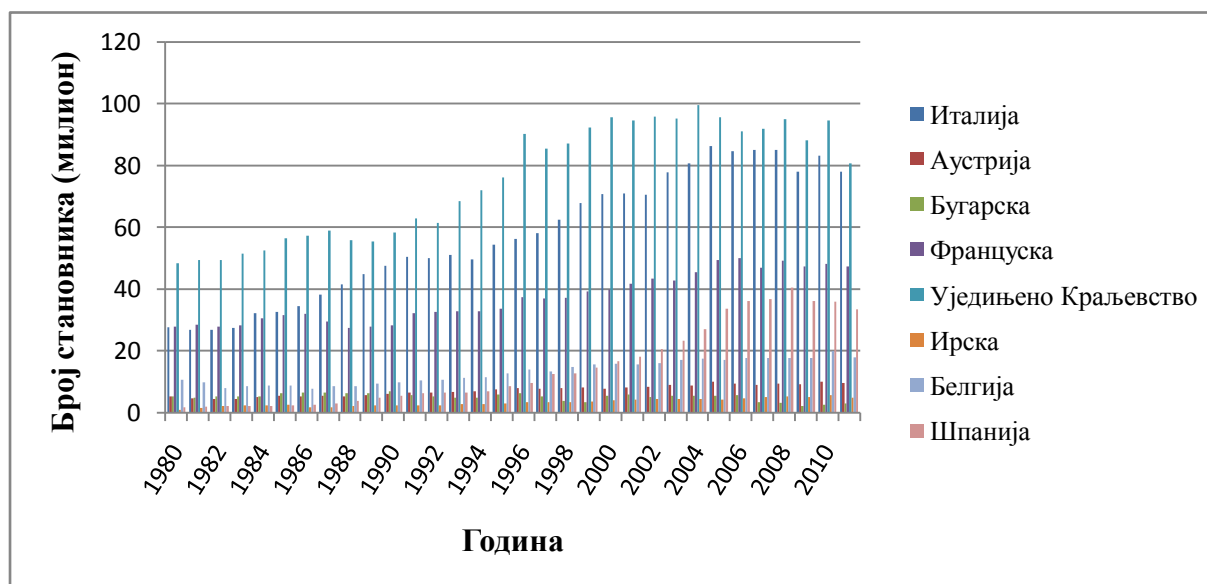
* Romania: the data is for 2011

Слика 4.1. Удео природног гаса у укупној количини примарне енергије у одабраним Европским тржиштима у 2012, у % [32]

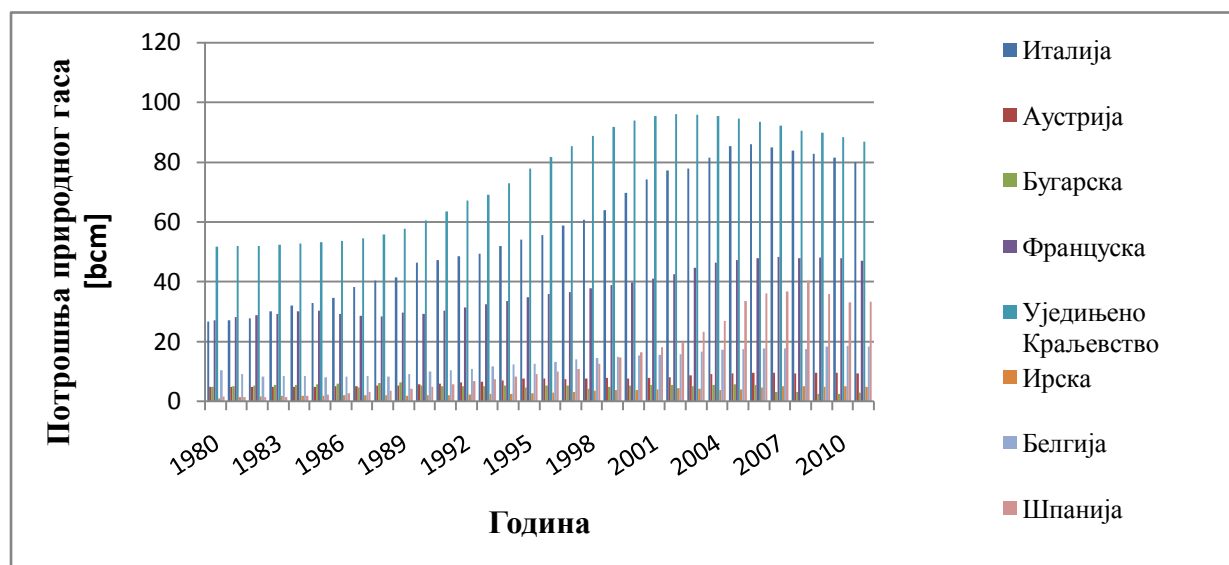
4.2 Резултати симулација

Предложени алгоритам се примењује за предвиђање потрошње природног гаса у Италији, Аустрији, Бугарској, Француској, Уједињеном краљевству, Ирској, Белгији и Шпанији од 1980. до 2011.

Слика 4.2 и слика 4.3 приказују вредности за број становника и потрошњу природног гаса током поменутих година.



Слика 4.2 Број становника у разматраним земљама од 1980. до 2011



Слика 4.3 Потрошња природног гаса у разматраним земљама

Предложени алгоритам је примењен на подацима кроз следеће кораке:

1. Одређивање улаза модела.
2. Подела података на два подскупа: обучавање и тестирање података. 32 сета података је подељено на скуп од 25 података за тренирање (1980. – 2004.) и скуп од 7 података за тестирање (2005. – 2011.).
3. Развијање модела рекурентне мреже за скуп података за обуку.
4. Процена тачности модела.

У овом раду, број становника и бруто домаћи производ се бирају као улазне променљиве модела дигиталне рекурентне мреже. Модел слојевите дигиталне рекурентне мреже коришћен у овом раду је двослојна мрежа састављена од скривеног слоја и излазног слоја. У излазном и скривеном слоју, линеарна (purelin) и тангенс хиперболичка (tansig) / сигмоидална (logsig) активациона функција су биле коришћене.

Неуронска мрежа са линеарним излазним неуроном и једним скривеним слојем неурона са тангенс хиперболичком активационом функцијом су универзални функцијски апроксиматор [52].

Неколико LDR мрежа су биле генерисане и тестиране. За процену добијених резултата, коришћена је средња квадратна грешка (MSE - Mean Squared Error). Број скривених неурона за разматрање је између 5 и 9. Оптимална величина мреже била је једна која је резултирала у минималној MSE скупа за тренирање и тестирање.

Сагласно структури података, вредности улазних променљивих од 2005. до 2011. су снабдевале изабрану LDR мрежу, а затим су добијене излазне вредности (за скуп за тестирање).

У табелама у наставку приказане су вредности средњих квадратних грешака за сваку разматрану земљу - применом LDR мреже, при чему су означене вредности које су изабране као оптималне.

Табела 4.1. Вредности средњих квадратних грешака (применом LDR мреже) - Италија

Италија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	13,9380	29,7895	4,7861	6,2672	9,1085	7,2716
Скуп за тестирање	20,5333	5,4500	4,8761	16,7165	8,3002	4,7282

Табела 4.2. Вредности средњих квадратних грешака (применом LDR мреже) - Аустрија

Аустрија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,1386	0,1267	0,0747	0,3278	0,0931	0,1336
Скуп за тестирање	0,1517	0,1301	0,1820	0,4331	0,1874	0,3637

Табела 4.3. Вредности средњих квадратних грешака (применом LDR мреже) - Бугарска

Бугарска (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,5958	0,5236	0,4596	0,5128	0,4631	0,4660
Скуп за тестирање	0,6656	0,4649	0,0515	0,1533	0,0650	0,7510

Табела 4.4. Вредности средњих квадратних грешака (применом LDR мреже) - Француска

Француска (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	1,7706	1,6791	1,3754	1,5063	1,2949	1,5503
Скуп за тестирање	2,6603	2,1881	2,5811	5,7559	1,6859	3,0568

Табела 4.5. Вредности средњих квадратних грешака (применом LDR мреже) -
Уједињено краљевство

Уједињено краљевство (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	6,3486	8,1238	9,6536	5,7854	8,5402	7,7474
Скуп за тестирање	36,4522	16,7097	15,7402	19,6433	14,9443	27,0882

Табела 4.6. Вредности средњих квадратних грешака (применом LDR мреже) - Ирска

Ирска (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,0692	0,0634	0,0716	0,0530	0,0715	0,0589
Скуп за тестирање	0,0435	0,0828	0,0434	0,0884	0,0268	0,0817

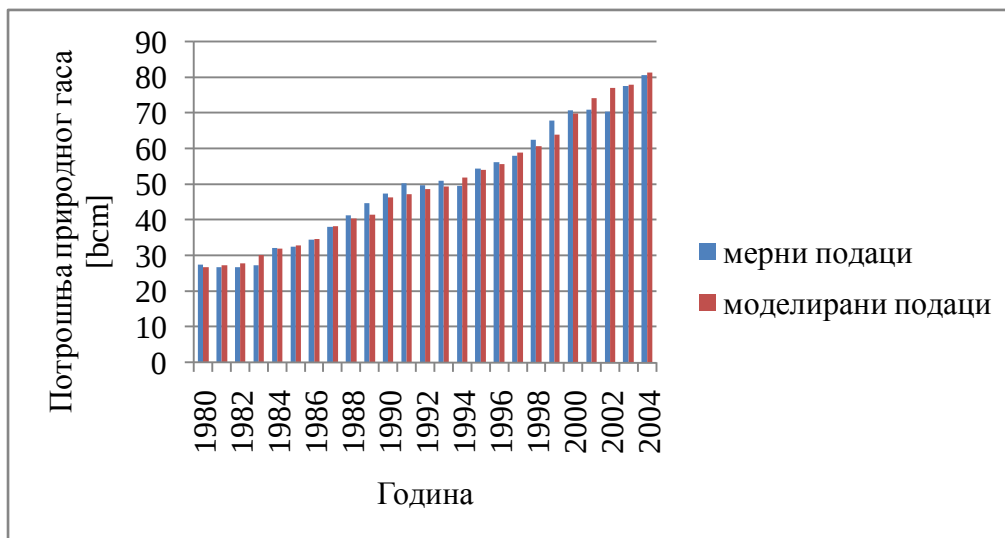
Табела 4.7. Вредности средњих квадратних грешака (применом LDR мреже) - Белгија

Белгија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,3991	0,4561	0,5343	0,4267	0,1768	0,3435
Скуп за тестирање	0,2761	0,3621	0,5000	0,5079	0,3918	0,4724

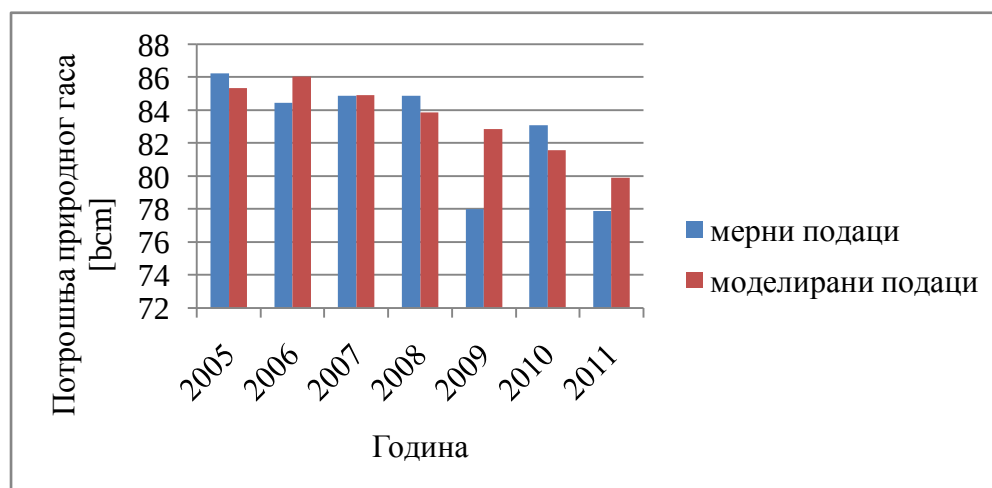
Табела 4.8. Вредности средњих квадратних грешака (применом LDR мреже) - Шпанија

Шпанија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	2,6271	0,3723	3,6129	0,5507	0,9811	2,5095
Скуп за тестирање	3,8411	1,0203	1,6022	1,3191	3,6421	2,4776

На наредним сликама је у виду хистограма приказано предвиђање потрошње природног гаса применом скупа података за тренирање и скупа података за тестирање, за сваку разматрану земљу ЕУ (применом LDR мреже).

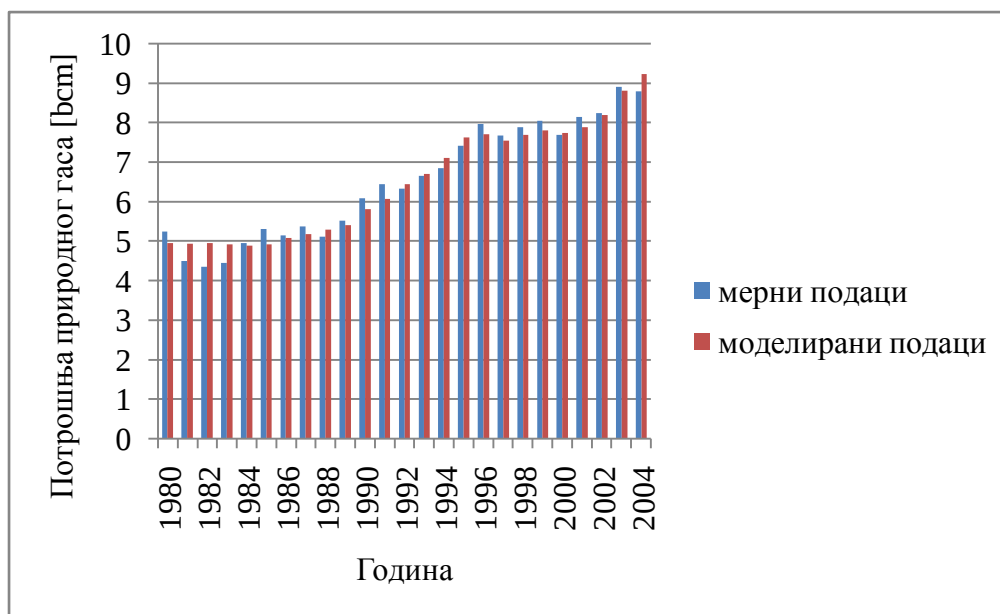


а)

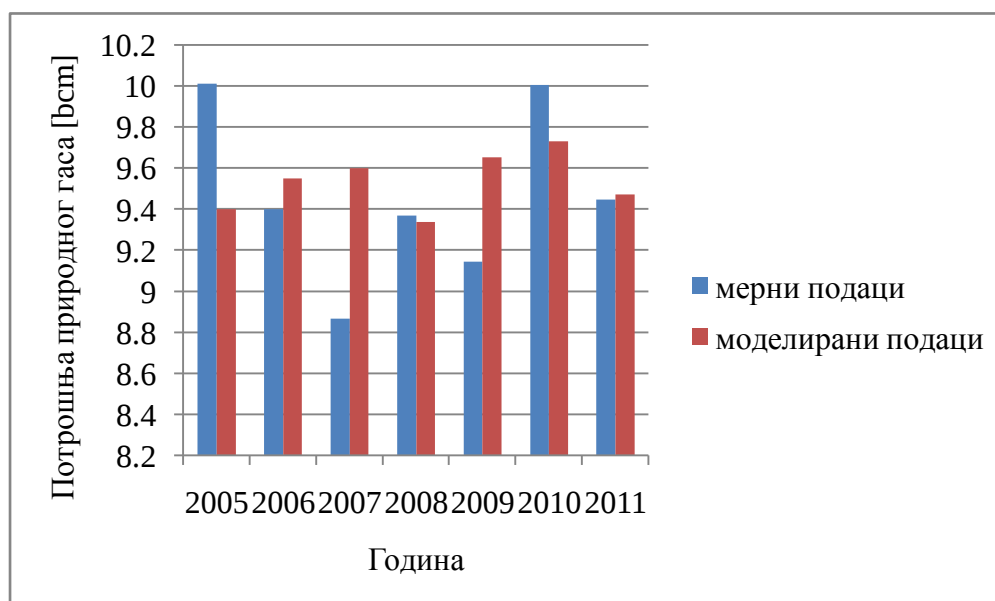


б)

Слика 4.4. Предвиђање потрошње природног гаса у Италији - применом LDR мреже (а - скуп за тренирање, б - скуп за тестирање)

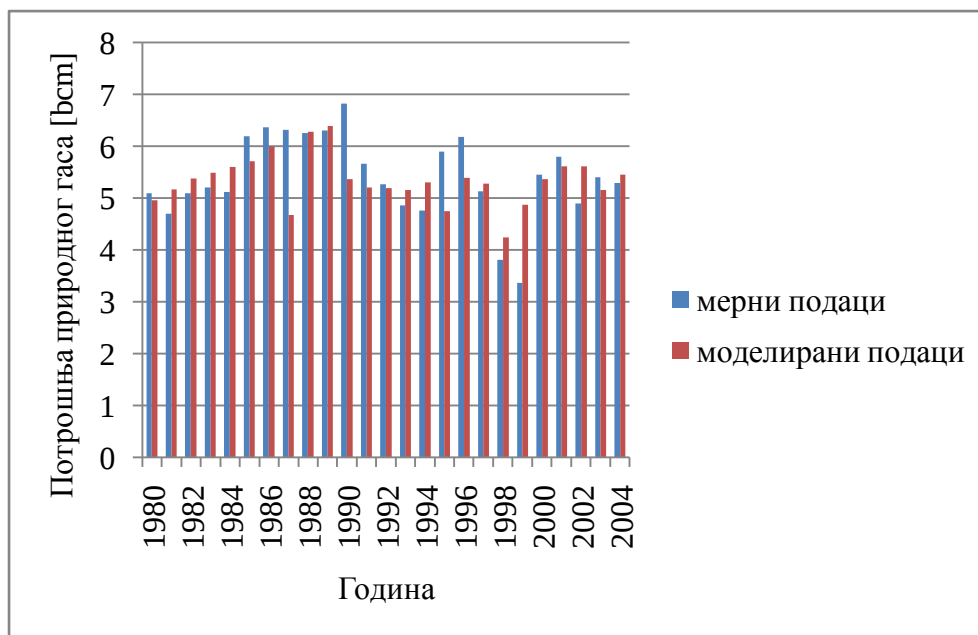


а)

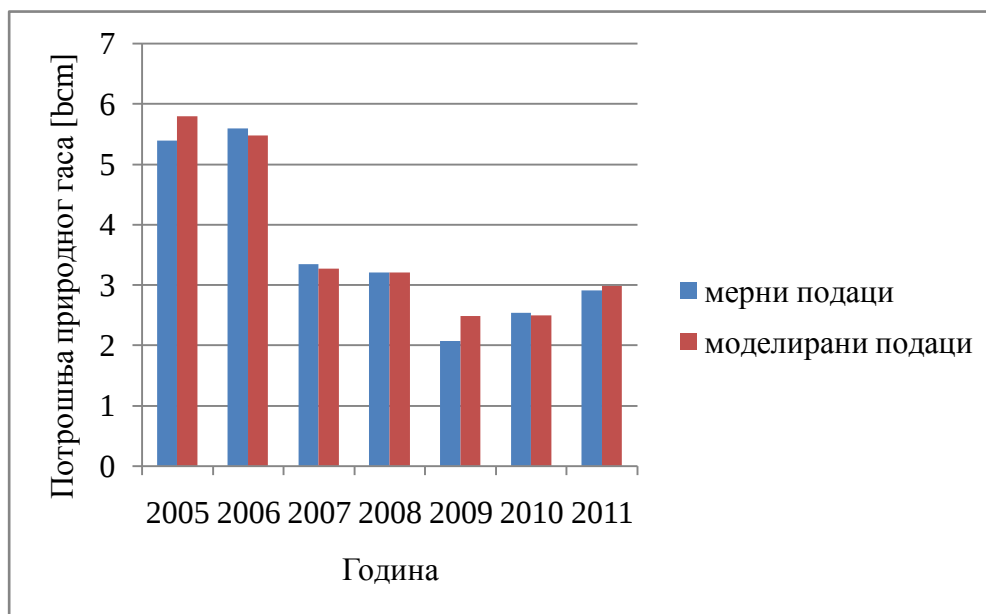


б)

Слика 4.5. Предвиђање потрошње природног гаса у Аустрији - применом LDR мреже (а - скуп за тренирање, б - скуп за тестирање)

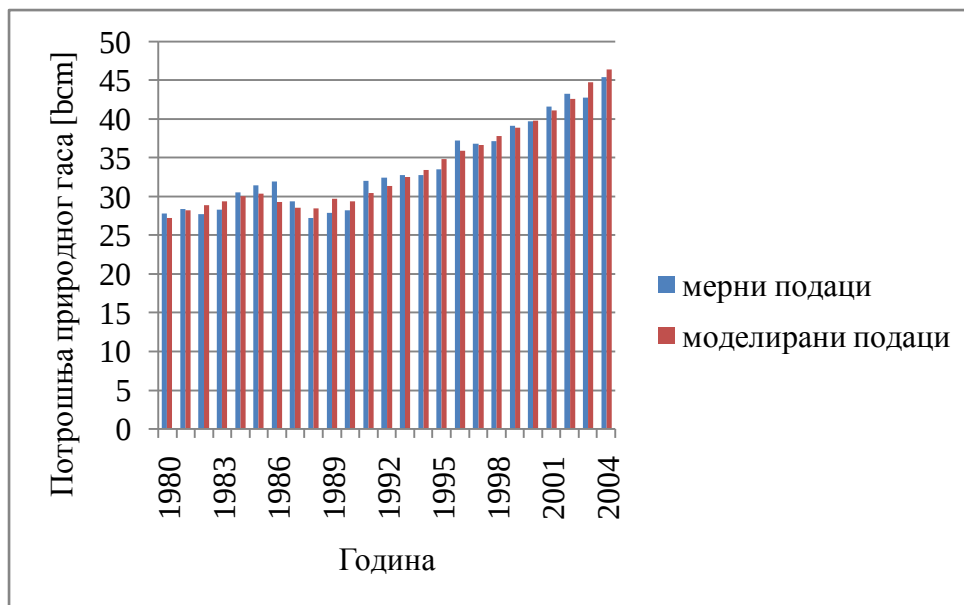


а)

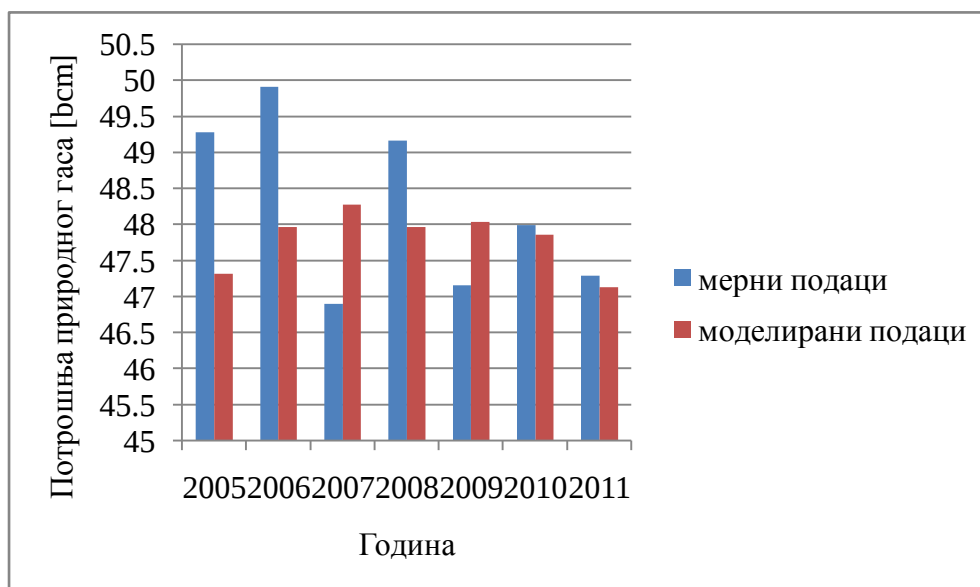


б)

Слика 4.6. Предвиђање потрошње природног гаса у Бугарској - применом LDR мреже (а - скуп за тренирање, б - скуп за тестирање)

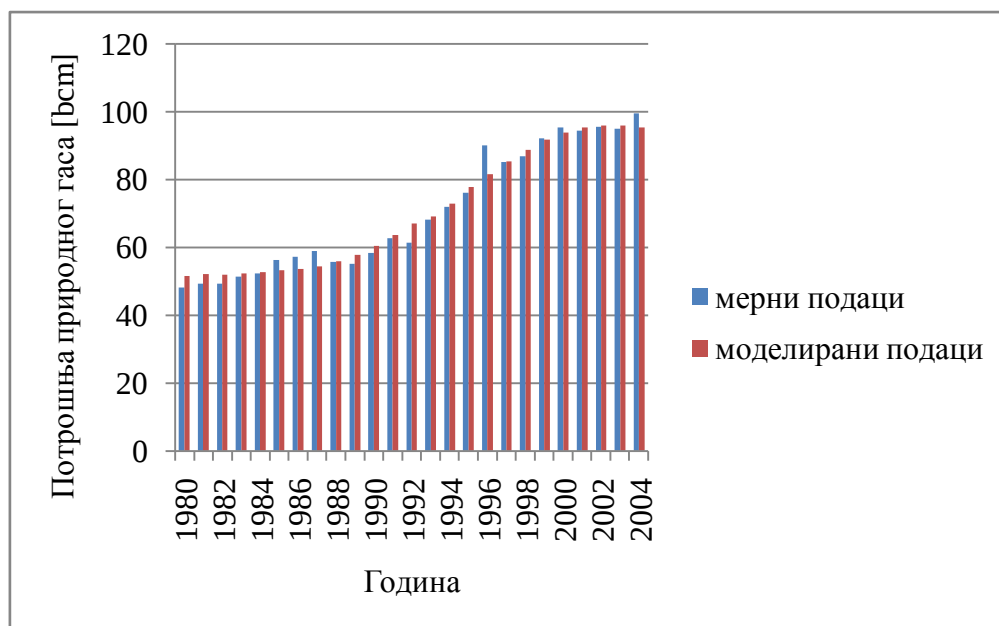


а)

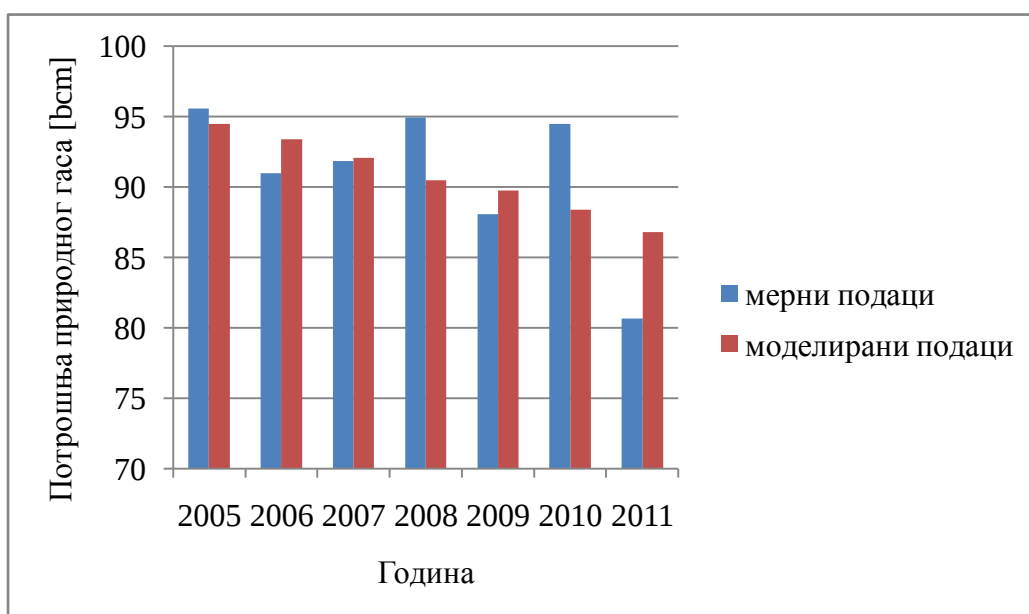


б)

Слика 4.7. Предвиђање потрошње природног гаса у Француској - применом LDR мреже
(а - скуп за тренирање, б - скуп за тестирање)

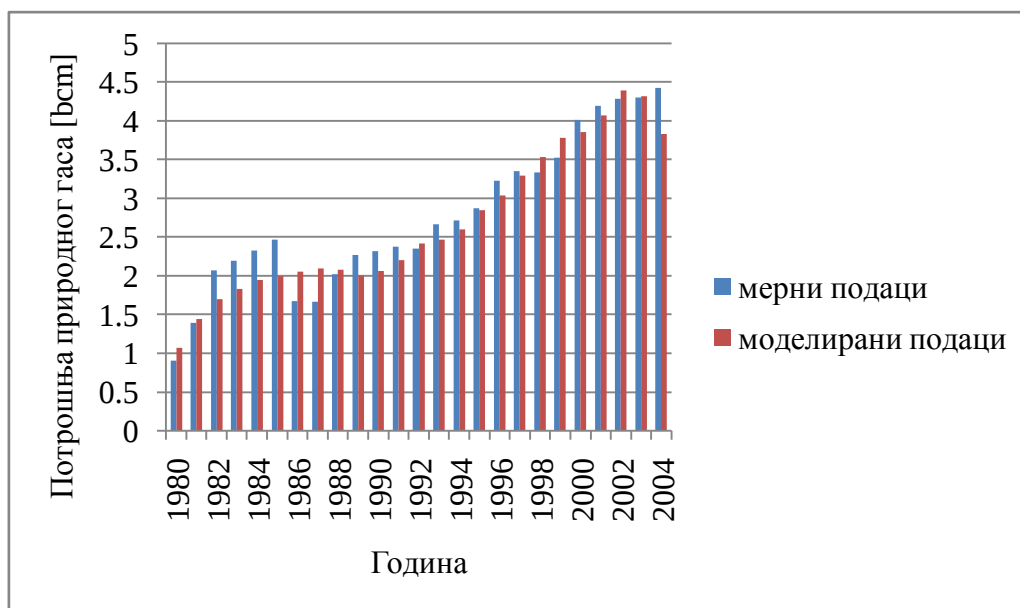


а)

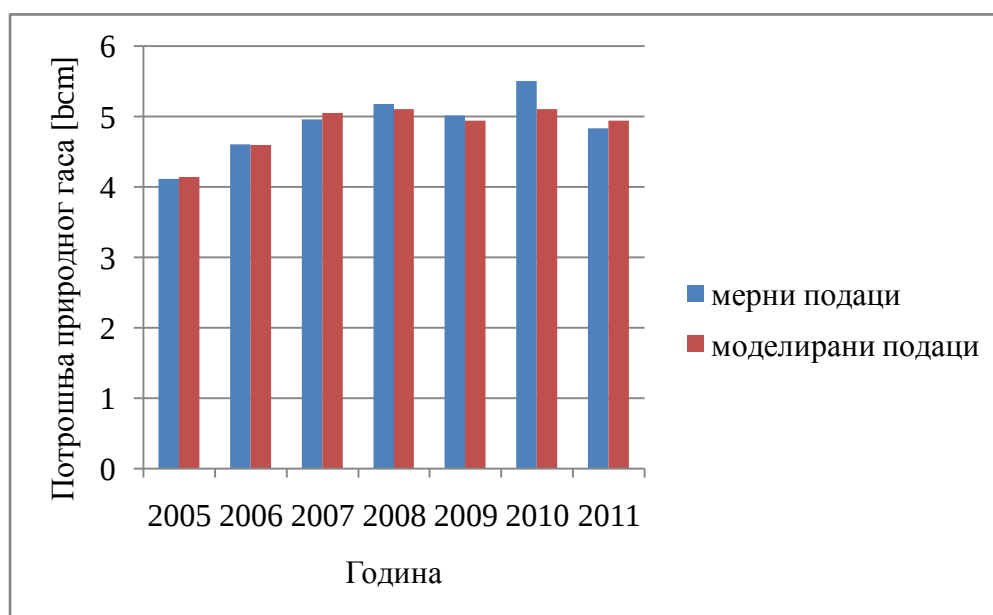


б)

Слика 4.8. Предвиђање потрошње природног гаса у Уједињеном Краљевству - применом LDR мреже (а - скуп за тренирање, б - скуп за тестирање)

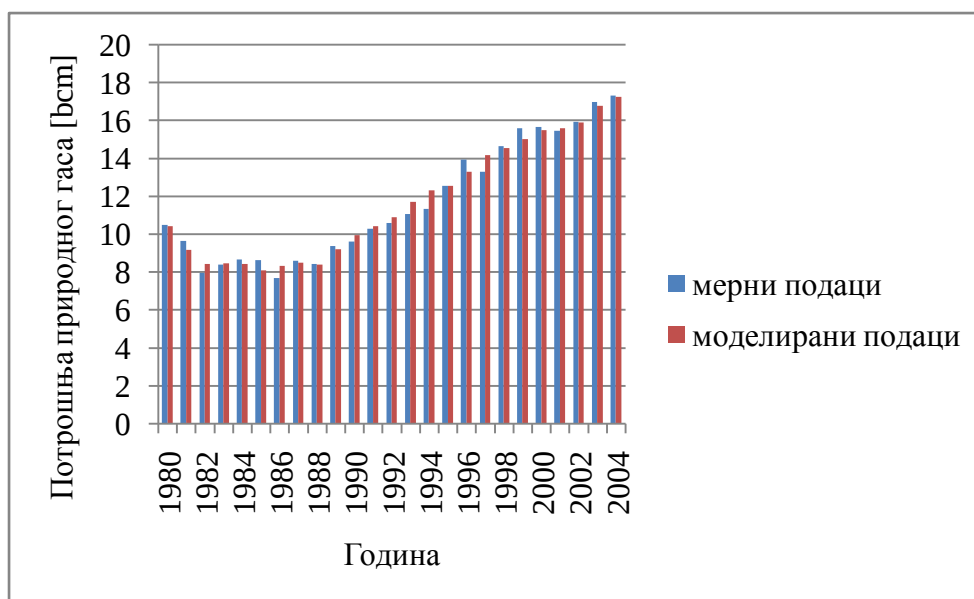


а)

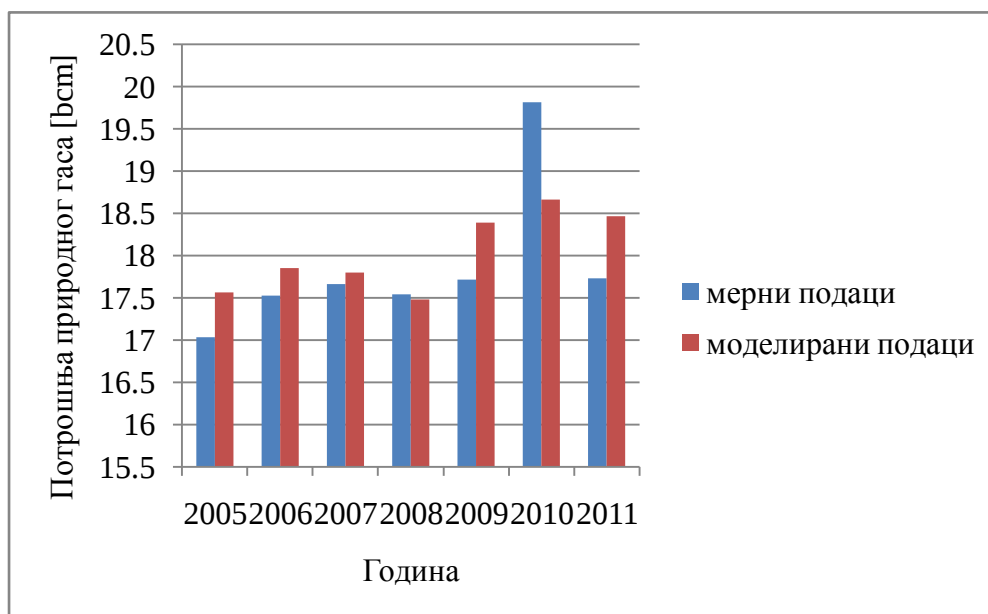


б)

Слика 4.9. Предвиђање потрошње природног гаса у Ирској - применом LDR мреже (а - скуп за тренирање, б - скуп за тестирање)

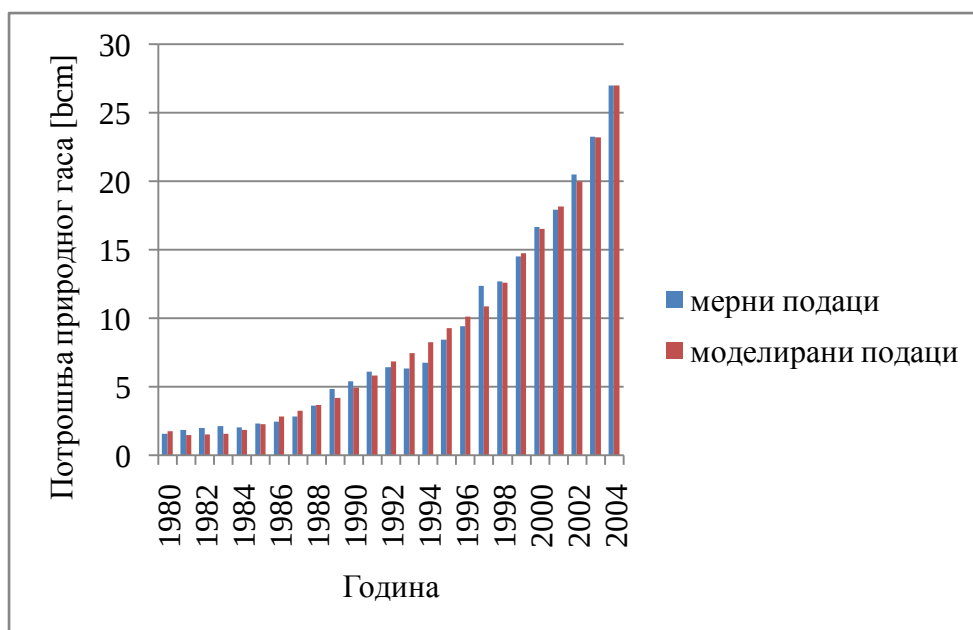


а)

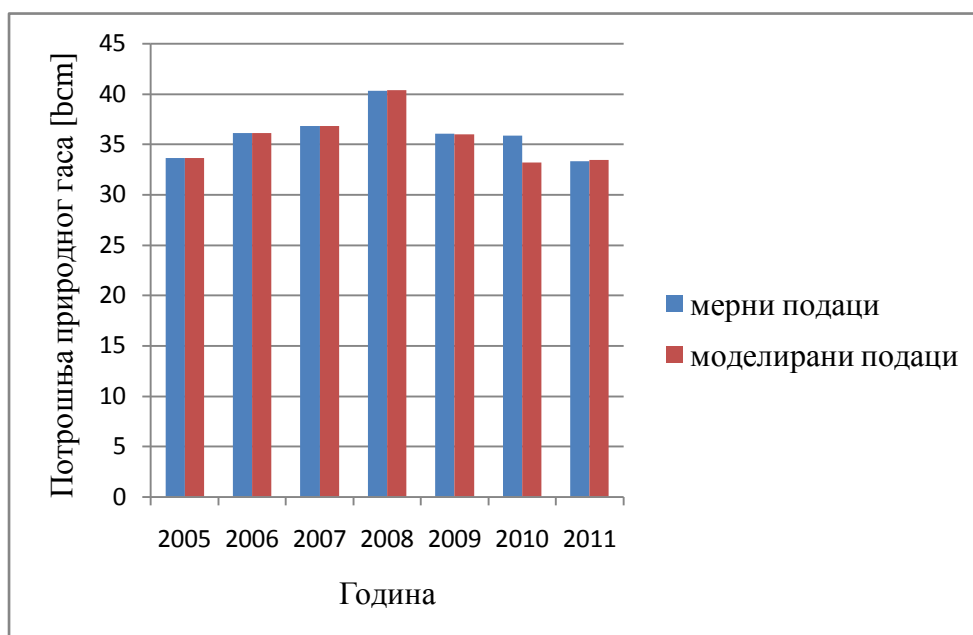


б)

Слика 4.10. Предвиђање потрошње природног гаса у Белгији - применом LDR мреже (а - скуп за тренирање, б - скуп за тестирање)



а)



б)

Слика 4.11. Предвиђање потрошње природног гаса у Шпанији - применом LDR мреже (а - скуп за тренирање, б - скуп за тестирање)

Решавање проблема предвиђања потрошње природног гаса у разматраним земљама ЕУ, извршено је применом још једног типа рекурентне неуронске мреже. У питању је NARX модел неуронске мреже (слика 3.7). Поступак имплементације који је предходно описан за случај LDR мреже, важи и код овог модела неуронске мреже.

У табелама у наставку приказане су вредности средњих квадратних грешака за сваку разматрану земљу – применом NARX модела, при чему су означене вредности које су изабране као оптималне.

Табела 4.9. Вредности средњих квадратних грешака (применом NARX модела) – Италија

Италија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	3,3282	4,9234	2,3661	3,2461	8,1400	4,1954
Скуп за тестирање	27,9492	2,7335	7,7298	4,7164	0,4252	3,1200

Табела 4.10. Вредности средњих квадратних грешака (применом NARX модела)- Аустрија

Аустрија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,0594	0,0960	0,0850	0,0785	0,4174	0,2353
Скуп за тестирање	0,1596	0,2381	0,2489	0,1544	0,2122	0,1626

Табела 4.11. Вредности средњих квадратних грешака (применом NARX модела) – Бугарска

Бугарска (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,3367	0,4521	0,4672	0,6661	0,3857	0,7794
Скуп за тестирање	0,0445	0,4028	0,5615	1,0268	0,3295	0,4985

Табела 4.12. Вредности средњих квадратних грешака (применом NARX модела) – Француска

Француска (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	2,1309	1,0789	1,6745	1,8385	1,5840	0,7025
Скуп за тестирање	2,0060	1,3911	1,0851	2,6645	2,2123	1,6671

Табела 4.13. Вредности средњих квадратних грешака (применом NARX модела) -
Уједињено краљевство

Уједињено краљевство (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	6,5523	6,3174	21,3542	15,2076	6,9532	7,8168
Скуп за тестирање	30,4316	28,4649	25,3616	17,3206	7,5781	17,5325

Табела 4.14. Вредности средњих квадратних грешака (применом NARX модела) –
Ирска

Ирска (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,0510	0,0520	0,0522	0,0613	0,0354	0,2721
Скуп за тестирање	0,1040	0,2127	0,0623	0,0105	0,0224	0,0943

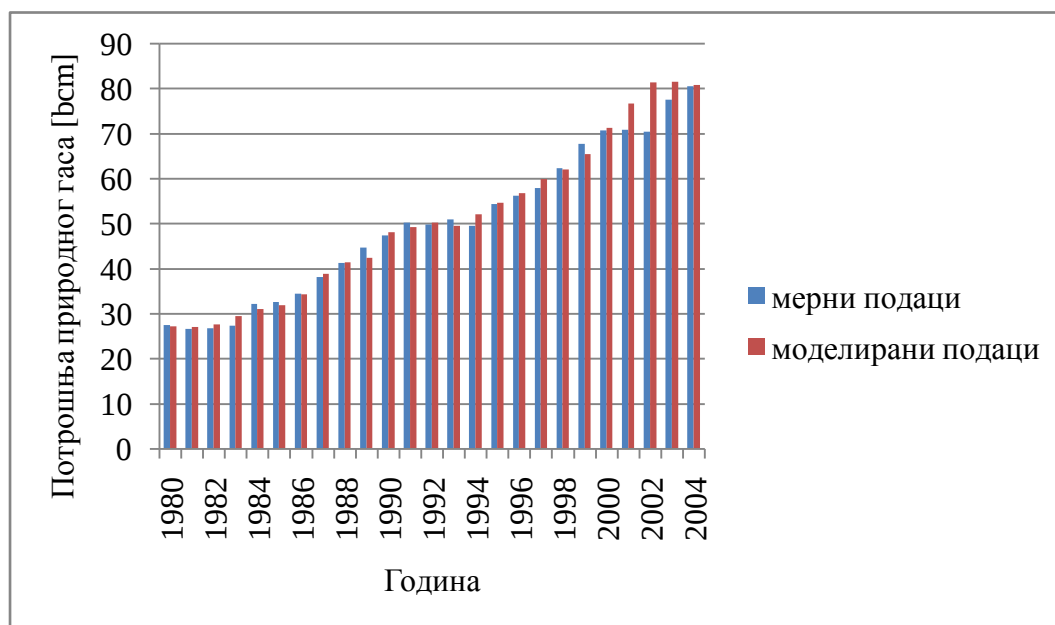
Табела 4.15. Вредности средњих квадратних грешака (применом NARX модела) –
Белгија

Белгија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,4485	0,2434	0,0877	0,4420	0,2346	0,5082
Скуп за тестирање	0,6446	0,7565	0,7651	0,5786	0,3279	0,4241

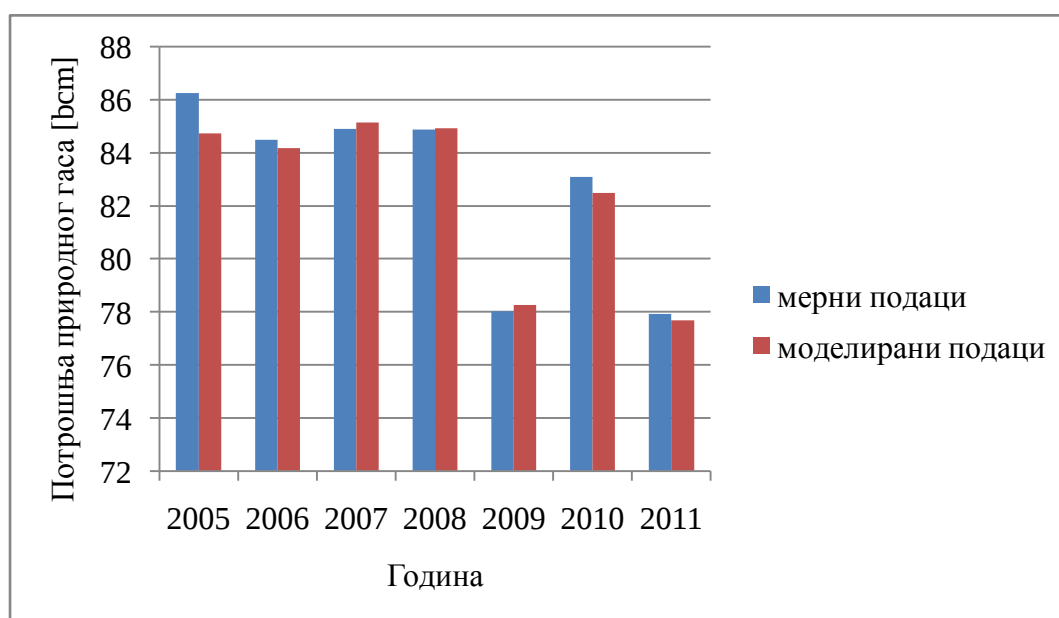
Табела 4.16. Вредности средњих квадратних грешака (применом NARX модела) –
Шпанија

Шпанија (1980. – 2011.)						
Број скривених неурона	5	7	9	5	7	9
Активациона функција у скривеном слоју	tansig	tansig	tansig	logsig	logsig	logsig
Скуп за тренирање	0,6924	1,4207	0,4202	0,4890	0,4053	0,3470
Скуп за тестирање	6,7699	4,4134	2,9058	3,1813	6,6949	4,9091

На наредним сликама је у виду хистограма приказано предвиђање потрошње природног гаса применом скупа података за тренирање и скупа података за тестирање, за сваку разматрану земљу ЕУ (применом NARX модела).

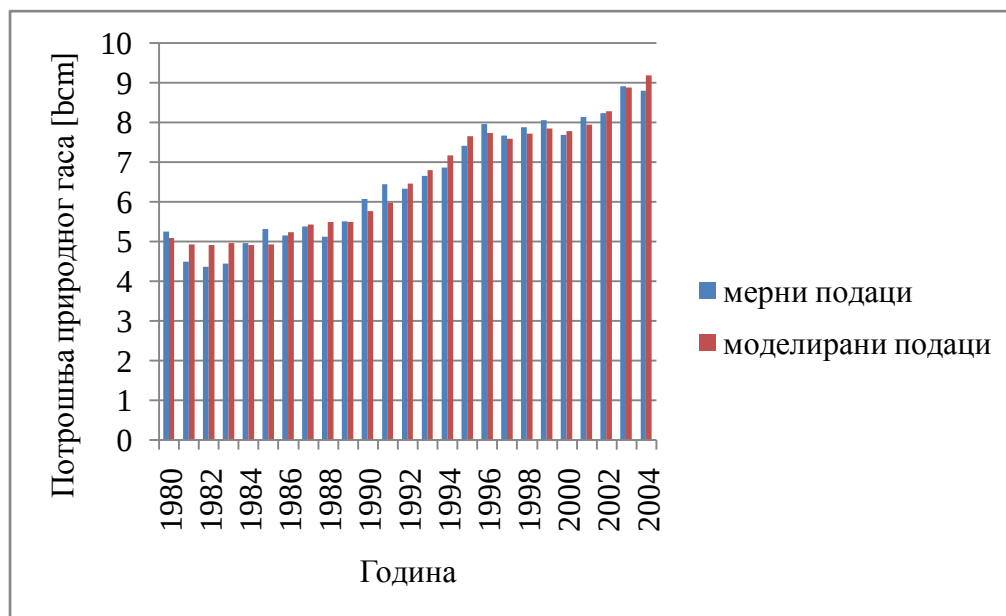


а)

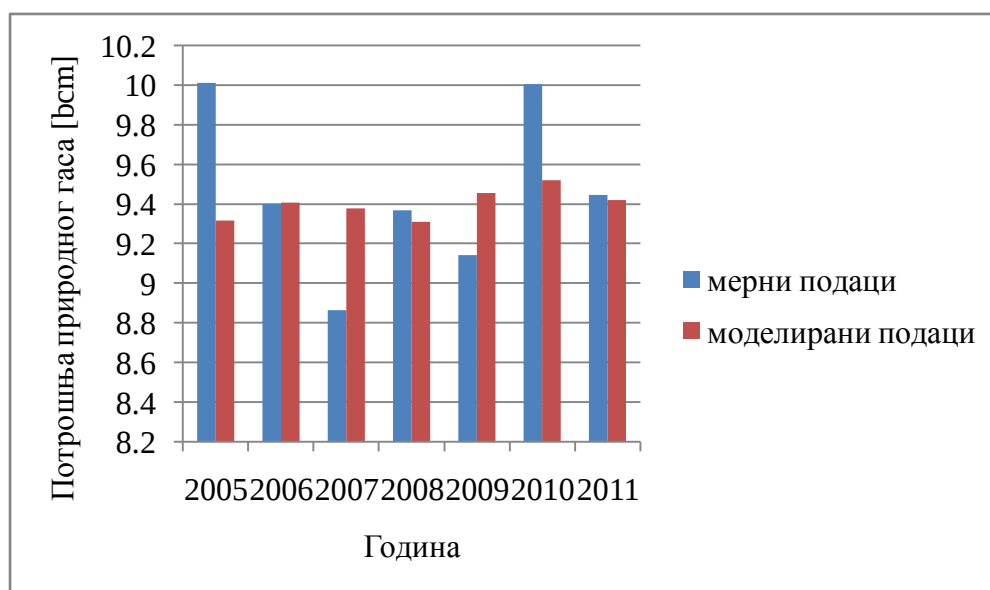


б)

Слика 4.12. Предвиђање потрошње природног гаса у Италији - применом NARX модела
(а - скуп за тренирање, б - скуп за тестирање)

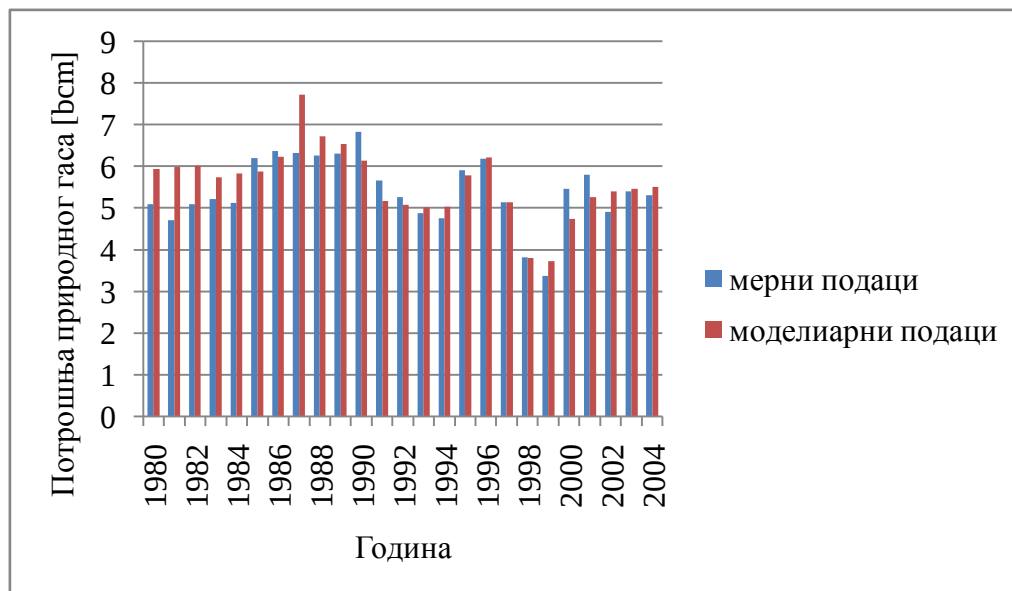


a)

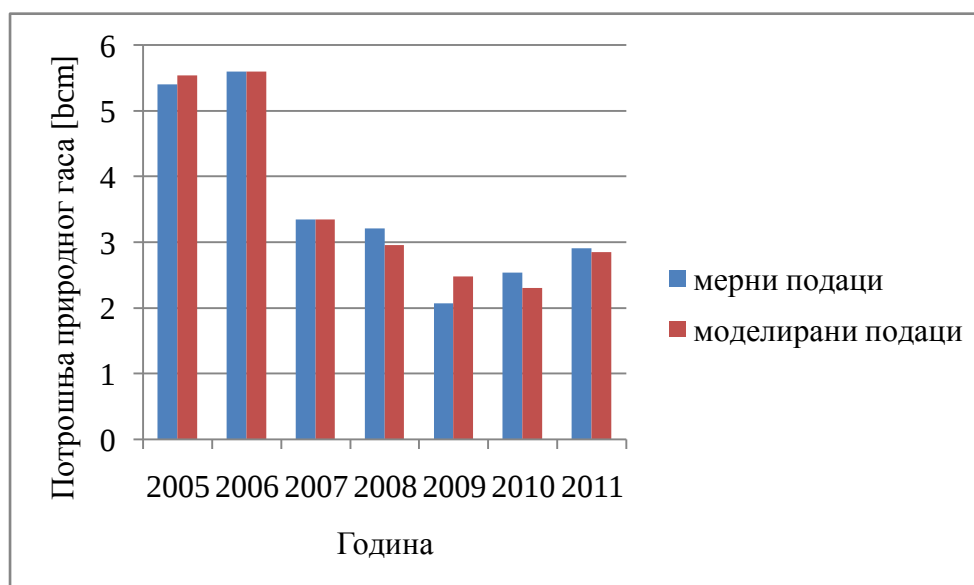


б)

Слика 4.13. Предвиђање потрошње природног гаса у Аустрији - применом NARX модела (а - скуп за тренирање, б - скуп за тестирање)

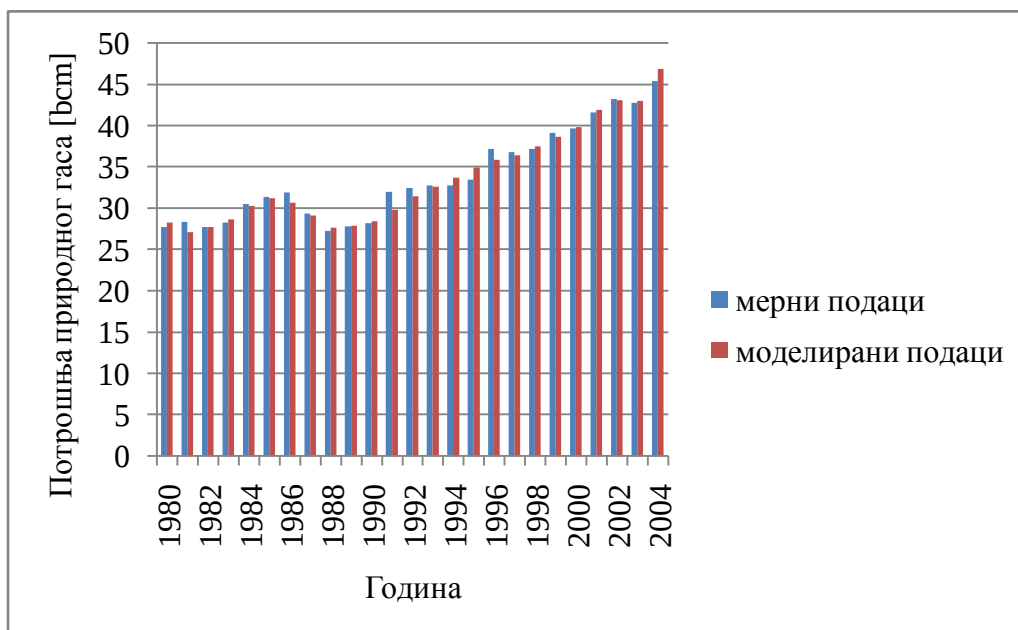


а)

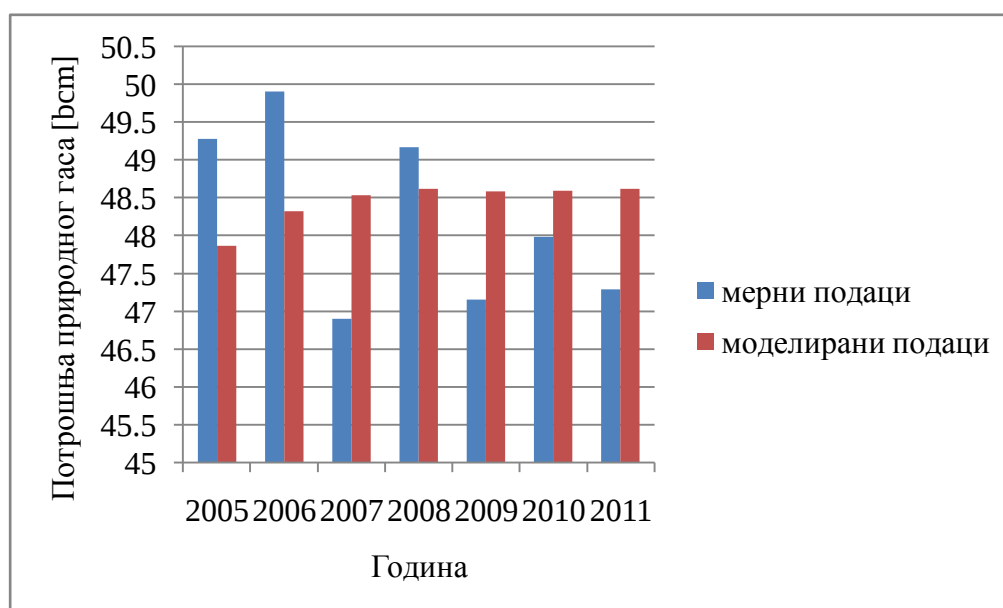


б)

Слика 4.14. Предвиђање потрошње природног гаса у Бугарској - применом NARX модела (а - скуп за тренирање, б - скуп за тестирање)

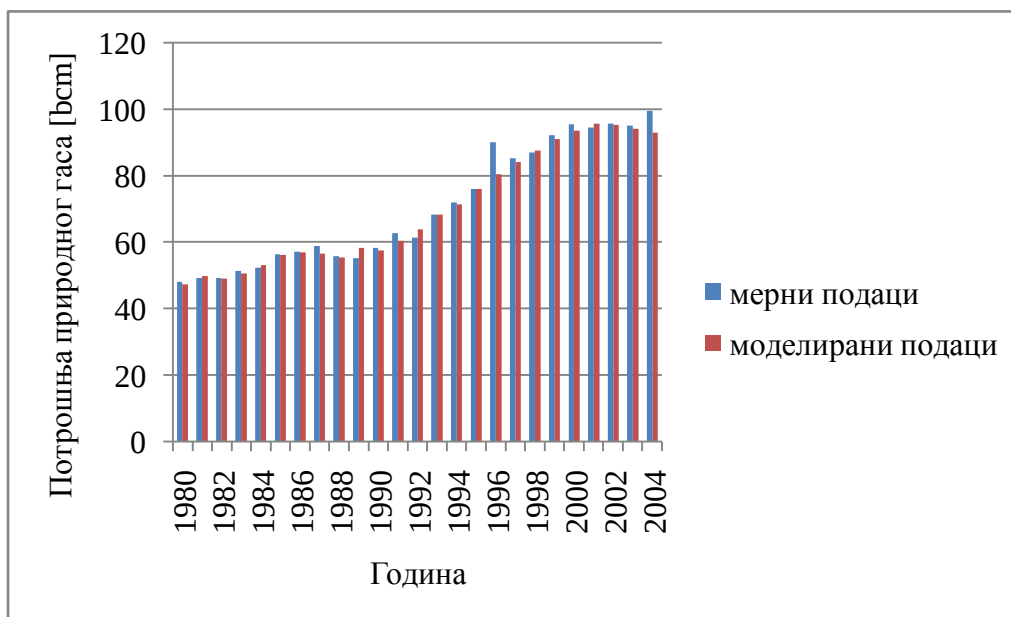


а)

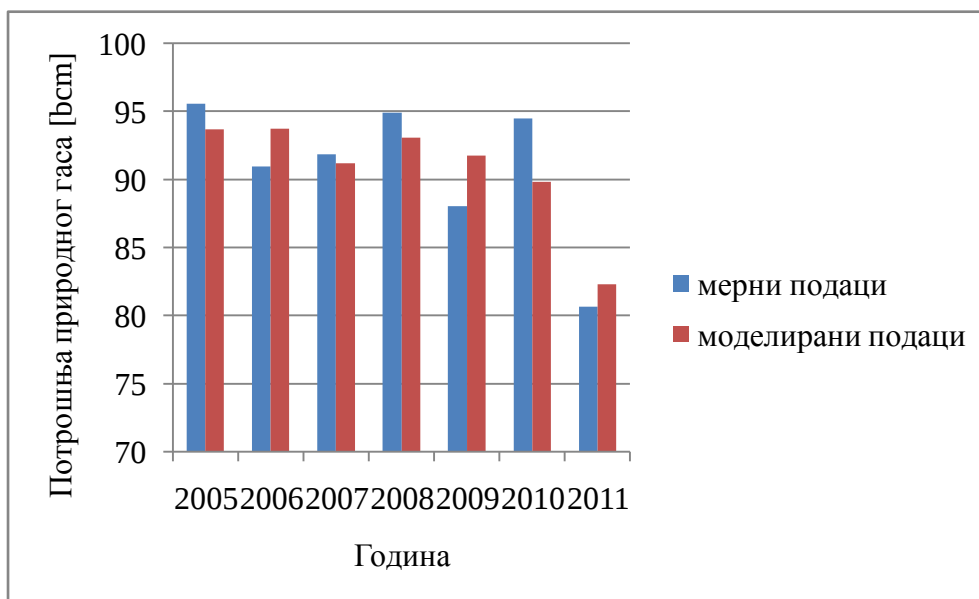


б)

Слика 4.15. Предвиђање потрошње природног гаса у Француској - применом NARX модела (а - скуп за тренирање, б - скуп за тестирање)

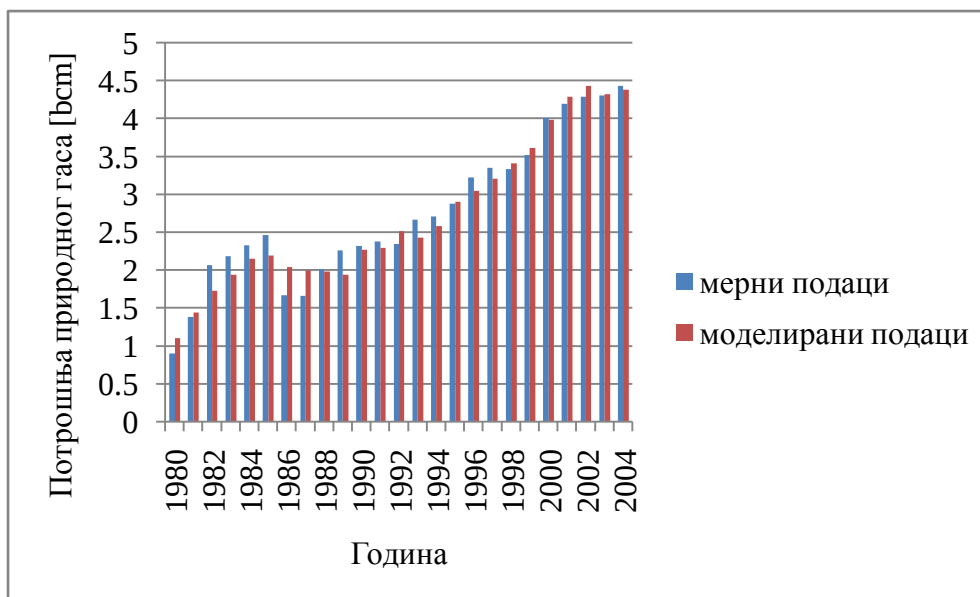


а)

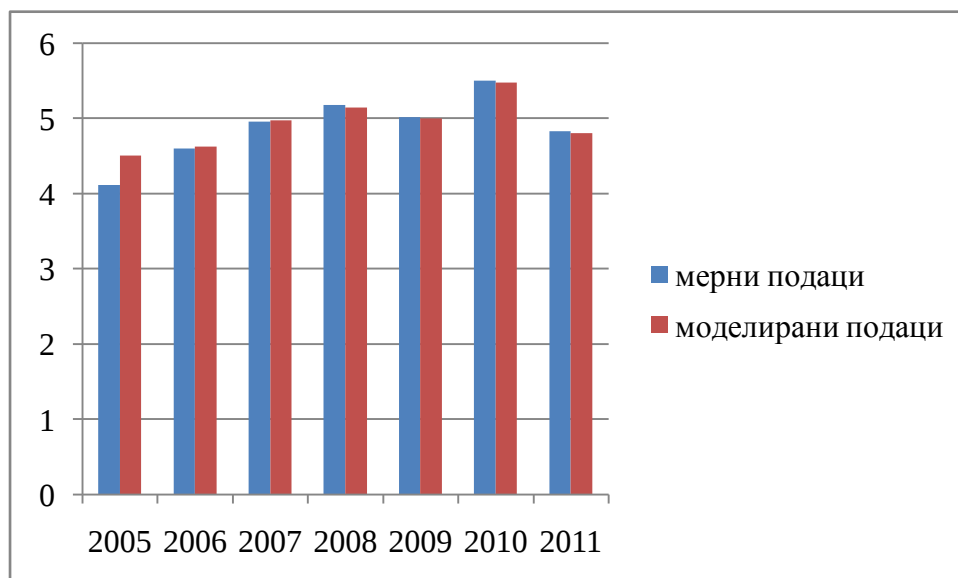


б)

Слика 4.16. Предвиђање потрошње природног гаса у Уједињеном Краљевству - применом NARX модела (а - скуп за тренирање, б - скуп за тестирање)

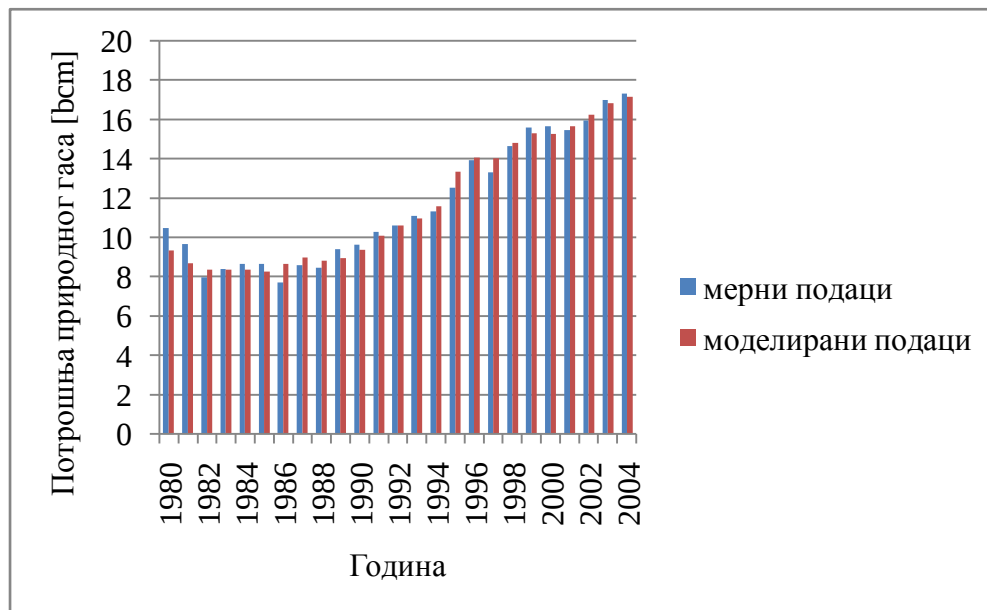


а)

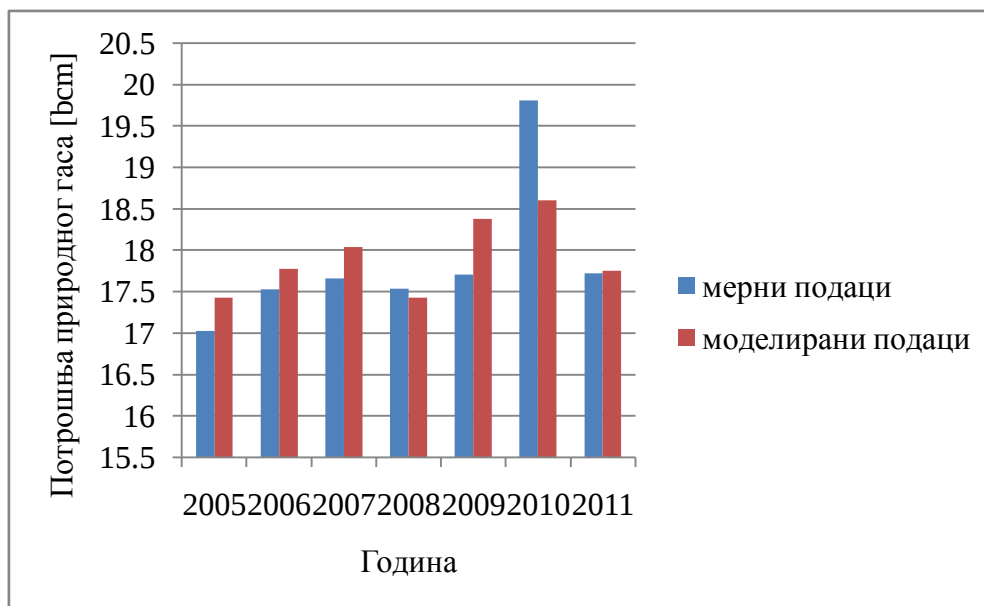


б)

Слика 4.17. Предвиђање потрошње природног гаса у Ирској - применом NARX модела (а - скуп за тренирање, б - скуп за тестирање)

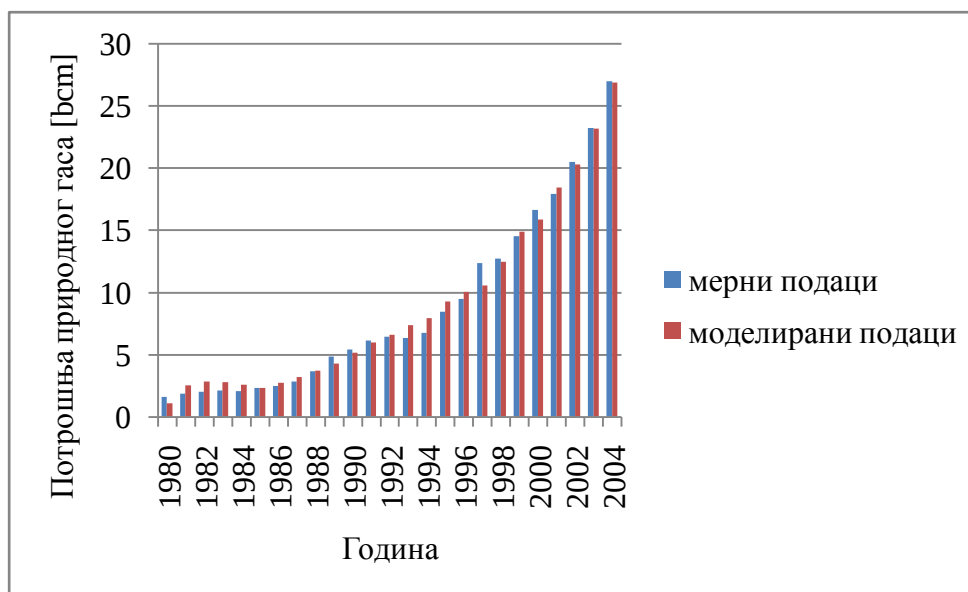


а)

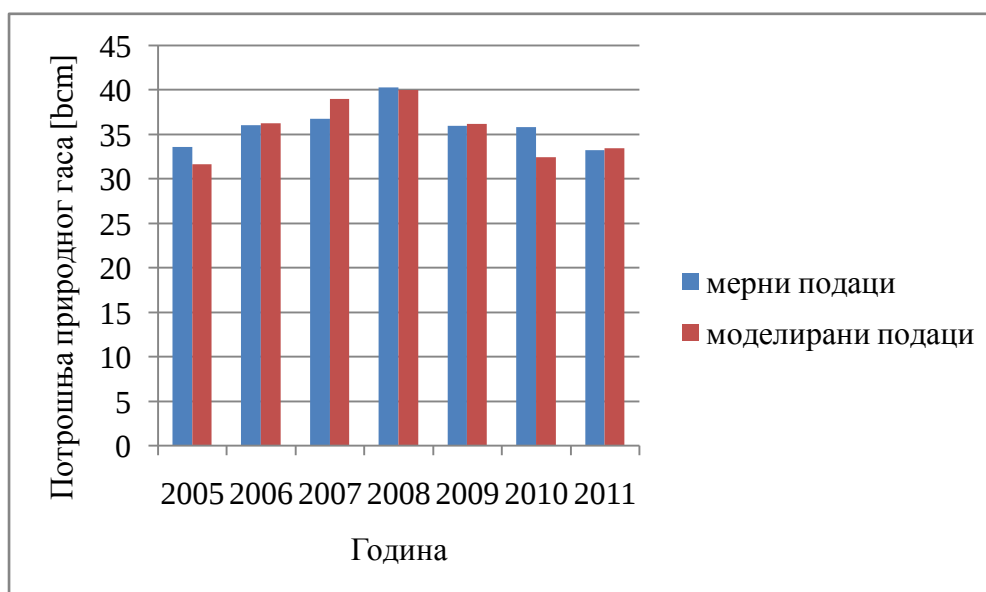


б)

Слика 4.18. Предвиђање потрошње природног гаса у Белгији - применом NARX модела (а - скуп за тренирање, б - скуп за тестирање)



а)



б)

Слика 4.19. Предвиђање потрошње природног гаса у Шпанији - применом NARX модела (а - скуп за тренирање, б - скуп за тестирање)

У односу на резултате добијене помоћу LDR мреже, NARX модел је показао бољу способност предвиђања за Италију, Аустрију, Бугарску, Француску, Уједињено Краљевство, Ирску и Белгију, док је бољи резултат за Шпанију добијен коришћењем LDR мреже.

5. ЗАКЉУЧАК

Из перспективе повезаности неурона, неуронске мреже се могу груписати у две категорије: нерекурентне, чији графови немају петље и рекурентне, са постојањем повратних веза којима се информације са виших слојева враћају нижим слојевима.

У поређењу са мрежама без повратних веза, архитектура рекурентних мрежа омогућава информацији да буде привремено меморисана у мрежама, што сугерише да рекурентне мреже имају много веће процесне способности.

Такође, модели рекурентних неуронских мрежа имају потенцијал да представе одређене рачунарске структуре на штедљивији начин, с обзиром на то да користе повратне везе.

Рекурентне неуронске мреже се карактеришу широком применом, што треба да подстакне веће интересовање за истраживање и развој, као и пораст даљих теоријских питања и питања дизајна.

У овом раду, рекурентне неуронске мреже су коришћене за предвиђање дугорочне потрошње природног гаса у земљама Европске Уније (Италији, Аустрији, Бугарској, Француској, Уједињеном краљевству, Ирској, Белгији и Шпанији) од 1980. до 2011. године.

Моделирање потрошње природног гаса укључује методолошке проблеме. Развој и имплементација разматраних модела рекурентних неуронских мрежа остварена је помоћу софтвера MATLAB R211a. Модели који су том приликом коришћени су модел двослојне дигиталне рекурентне мреже и модел двослојне NARX мреже, са једним скривеним и једним излазним слојем. Као улазне променљиве модела изабрани су број становника и бруто домаћи производ.

Подаци су подељени на два подскупа: обучавање и тестирање података. 32 сета података је подељено на скуп од 25 података за тренирање (1980. – 2004.) и скуп од 7 података за тестирање (2005. – 2011.), након чега је извршен развој и процена тачности модела рекурентне мреже за скуп података за обуку. Број скривених неурона за разматрање био је 5, 7 и 9. Приликом тренирања и тестирања модела је коришћена $\tansig$ и $\logsig$ преносна функција у скривеном слоју, док је у излазном слоју била коришћена $\purelin$ преносна функција.

За процену добијених резултата, коришћена је средња квадратна грешка.

Резултат примењених мрежних приступа указује да су предложени модели тачни за предикцију потражње природног гаса у посматраним земљама ЕУ.

У односу на резултате добијене помоћу LDR мреже, NARX модел је показао бољу способност предвиђања (мања вредност MSE) за Италију, Аустрију, Бугарску,

Француску, Уједињено Краљевство, Ирску и Белгију, док је бољи резултат за Шпанију добијен коришћењем LDR мреже.

На основу добијених резултата симулације, може се закључити да је NARX модел ефикаснији за решавање овакве врсте проблема.

ЛИТЕРАТУРА

- [1] Mikolov, T., Karafizt M, Cernocky J, and Khudanpur S: "Recurrent neural network based language model", INTERSPEECH (2010) 1045-1048.
- [2] Mikolov,T., Kombrink, S., Burget, L., Cemocky, J., Khudanpur, S.: "Extensions of recurrent neural network based language model", ICASSP (2011) 5528-5531.
- [3] Mikolov,T., Deoras, A., Kombrink, S., Burget, L., Cernocky J.: "Empirical evaluation and combination of advanced language modeling techniques", INTERSPEECH (2011) 605- 608.
- [4] Mikolov,T., Deoras, A., Povey, D., Burget, L., Cernocky, J.: "Strategies for training large scale neural network language models", ASRU (2011)
- [5] Mikolov,T., Zweig G.: "Context dependent recurrent neural network language model," Proc. SLT (2012) 234-239.
- [6] Arisoy, E., Sainath, T. N., Kingsbury, B., Ramabhadran, B.: "Deep neural network language models", Proceedings of the NAACL-HLT 2012 Workshop: Will We Ever Really Replace the N-grain Model? On the Future of Language Modeling for HLT (2012) 20-28.
- [7] Sun, W.,Wang, Y.: "A Recurrent Fuzzy Neural Network Based Adaptive Control and its Application on Robotic Tracking Control", Neural Information Processing-Letters and Reviews, (2004) 19-26.
- [8] Mohamed, A., Dahl, G., Hinton, G.: "Acoustic modeling using deep belief networks. Audio, Speech, and Language Processing", IEEE Transactions on, 20(1), (2012) 14-22.
- [9] Ciresan, D., Meier, U., Gambardella, L., Schmidhuber, J.: "Deep, big, simple neural nets for handwritten digit recognition", Neural computation, 22(12) (2010) 3207-3220.
- [10] Atencia, M., Boumeridja, H., Joya, G., Garcia-Lagos, F., Sandoval, F.,: "FPGA implementation of a systems identification module based upon Hopfield networks" Neurocomputing, (2007) 2828-2835
- [11] Han, D., Chan, S.P.: "Collision-free motion planning for redundant robots using neural-network processing", Engineering Applications of Artificial Intelligence, (1997) 179-188.
- [12] Ruichek, Y., Postaire, J.G: "A new neural real-time implementation for obstacle detection using linear stereo vision", Real-Time Imaging, (1999) 141-153.
- [13] Morasso, P., Solari, M.: A "NEURAL IMPLEMENTATION OF ANALOGIC PLANNING METHODS", Artificial Neural Networks (1991) 1281-1284.

- [14] Numberger, A., Radetzky A., Kruse R.: "Using recurrent neuro-fuzzy techniques for the identification and simulation of dynamic systems", *Neurocomputing*, (2001) 123-147
- [15] Zhang, Y., Wang, J.: "Recurrent neural networks for nonlinear output regulation", *Automatica*, (2001) 1161-1173
- [16] Marocco, D., Stramandinoli, F., Cangelosi, A. : "The grounding of higher order concepts in action and language: A cognitive robotics model Neural Networks", (2012) 165-173.
- [17] Koker, R.: "Reliability-based approach to the inverse kinematics solution of robots using Elman's networks", *Engineering Applications of Artificial Intelligence*, (2005) 685-693
- [18] Li,P., Li, Y., Xiong, Q., Chai, Y., Zhang, Y.: "Application of a hybrid quantized Elman neural network in short-term load forecasting" , *International Journal of Electrical Power & Energy Systems*, (2014) 749-759
- [19] Zhou, C., Ding, L.Y., He, R. : "PSO-based Elman neural network model for predictive control of air chamber pressure in slurry, shield tunneling under Yangtze River", *Original Research Article Automation in Construction*, (2013) 208-217
- [20] Koker, R.: "Design and performance of an intelligent predictive controller for a six-degree-of-freedom robot using the Elman network", *Original Research Article Information Sciences* (2006) 1781-1799
- [21] Seker, S., Ayaz, E., Turkcan, E.: "Elman's recurrent neural network applications to condition monitoring in nuclear power plant and rotating machinery", *Research Article Engineering Applications of Artificial Intelligence*, (2003) 647-656
- [22] Wan, W., Xu, H., Zhang, W., Hu, X., Deng, G.: "Questionnaires-based skin attribute prediction using Elman neural network", *Neurocomputing*, (2011) 2834-2841
- [23] Ito, M., Noda, K., Hoshino, Y., Tani, J.: "Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model" *Original Research Article Neural Networks*, (2006) 323-337
- [24] Roman, J., Jameel, A.: "Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns", *Proceedings of the Twenty-ninth Hawall, International Conference on System Science*, (1996) 454-460.
- [25] Feng J. ,Tse, C.K., Lau, F.C.M.: "A neural-network-based channel-equalization strategy for chaos-based communication systems", *IEEE Transactons on Circuits and Systems : Fundamental Theory and Applications*, (2003) 954-957.
- [26] Kamwa, I., Grondin, R., Sood, V.K., Gagnon, C., Van Thich Nguyen, Mereb, J.: "Recurrent neural networks for phasor detection and adaptive identification in power system control and protection", *IEEE Transactions on Instrumentation and Measurement* (1996) 657-664.

- [27] Jayadeva, Rahman, S.A.: "A neural network with $O(N)$ neurons for ranking N numbers in $O(1/N)$ time", *IEEE Transactions on Circuits and Systems I: Regular Papers*, (2004) 2044-2051.
- [28] Chengyu, G., Danai, K.: "Fault diagnosis of the IFAC Benchmark Problem with a model-based recurrent neural network", *Proceedings of the 1999 IEEE International Conference on Control Applications*, (1999) 1755-1760.
- [29] Robinson, A.J., "An application of recurrent nets to phone probability estimation ", *IEEE Transactions on Neural Networks*, (1994)
- [30] Gianluca, P., Przybylski, D., Rost, B., Baldi, P.: "Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles", *Proteins: Structure, Function, and Genetics*, (2002) 228-235.
- [31] Honoré, A.: "The Outlook for Natural Gas Demand in Europe", Oxford Institute for Energy Studies, 2014.
- [32] Khotanzad, A., Elragal, H., Lu, T.: "Combination of artificial neural-network forecasters for prediction of natural gas consumption", *IEEE Transactions on Neural Networks* 11 (2000) 464–73.
- [33] Gorucu, F.B., Geris, P.U., Gumrah, F.: "Artificial neural network modeling for forecasting gas consumption", *Energy Sources* 26 (2004) 299–307.
- [34] Siemek, J., Nagy, S., Rychlicki, S.: "Estimation of natural-gas consumption in Poland based on the logistic-curve interpretation", *Applied Energy* 75 (2003) 1–7.
- [35] Yoo, S., Lim, H.J., Kwak, S.J.: "Estimating the residential demand function for natural gas in Seoul with correction for sample selection bias", *Applied Energy* 86 (2009) 460–465.
- [36] Azadeh, A., Asadzadeh, S.M., Saberi, M., Nadimi, V., Tajvidi, A.: "A neuro-fuzzy-stochastic frontier analysis approach for long-term natural gas consumption forecasting and behavior analysis: the cases of Bahrain, Saudi Arabia, Syria, and UAE", *Applied Energy* 88 (2011) 3850–3859.
- [37] Azadeh, A., Saberi, M., Asadzadeh, S.M., Hussain, O.K., Saberi, Z.: "A neuro-fuzzy-multivariate algorithm for accurate gas consumption estimation in South America with noisy inputs", *Electrical Power and Energy Systems* 46 (2013) 315–325.
- [38] Vilar, J.M., Cao, R., Aneiros, G.: "Forecasting next-day electricity demand and price using nonparametric functional methods" *International Journal Electrical Power & Energy Systems* (2012) 48-55.

- [39] Wang, J., Ma, X., Wu, J., Dong, Y., "Optimization models based on GM (1,1) and seasonal fluctuation for electricity demand forecasting", *International Journal Electrical Power & Energy Systems* (2012) 109-17.
- [40] Amina, M., Kodogiannis, V.S., Petrounias, I., Tomtsis, D.: "A hybrid intelligent approach for the prediction of electricity consumption", *International Journal Electrical Power & Energy Systems* (2012) 99-108.
- [41] Lei, M., Feng, Z.: "A proposed grey model for short-term electricity price forecasting in competitive power markets", *International Journal Electrical Power & Energy Systems* (2012) 531-8.
- [42] Huang, Y.L., Lin, C.T.: "Developing an interval forecasting method to predict undulated demand", *Quality & Quantity* (2011) 513-524
- [43] Gonzales Chavez, S., Xiberta Bernat, J., Llaneza Coalla, H.: "Forecasting of energy production and consumption in Asturias (Northern Spain) ", *Energy* (1999) 183-98.
- [44] Kadoshin, S., Nishiyama, T., Toshihide, I.: "The trend in current and near future of energy consumption from a statistical perspective". *Applied Energy* (2000) 407-17,
- [45] Jai Persaud, A., Kumar, U.: "An eclectic approach in energy forecasting: a case of natural resources Canada's (NRCan's) oil and gas outlook", *Energy Policy* (2001) 303-13.
- [46] Gorucu, F.B., Gumrah, F.: "Evaluation and forecasting of gas consumption by statistical analysis", *Energy Sources* (2004) 267-76.
- [47] Gutierrez, R., Nafidi, A., Gutierrez Sanchez, R.: "Forecasting total natural-gas consumption in Spain by using the stochastic Compertz innovation diffusion model", *Applied Energy* (2005) 115-24,
- [48] Sanchez-Ubeda, E.F., Berzosa, A.: "Modeling and forecasting industrial end-use natural gas consumption", *Energy Economics* (2007) 710-42.
- [49] Parikh, J., Purohit, P., Maitra, P.: "Demand projections of petroleum products and natural gas in India", *Energy* (2007) 1825-37.
- [50] Ranković, V.: "Inteligentno upravljanje", Mašinski fakultet Kragujevac, 2008.
- [51] Hagan, M., Jesus, O. D. and Schultz, R.: Training Recurrent Networks for Filtering and Control, *Chapter 12 of Recurrent Neural Networks: Design and Applications*, L.R. Medsker and L.C. Jain, Eds., CRC Press, pp. 327-389, 2001.
- [52] Radović, M., Ranković, V., Bojić, M., Radulović, J.: "LAYER DIGITAL RECURRENT NETWORK FOR FORECASTING OF LONG-TERM NATURAL GAS CONSUMPTION IN EU COUNTRIES", 9th International Quality Conference, Kragujevac, Serbia (2015) 543-546