

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

**Metode i alati za predviđanje događaja na
veleprodajnom tržištu električne energije**

Marija Mijić

Voditelj: *Vedran Podobnik*

Zagreb, svibanj 2013.

Sadržaj

1.	Uvod	1
2.	Power Trading Agent Competition.....	2
3.	Metode predviđanja	4
3.1.	Kvalitativne metode predviđanja.....	4
3.2.	Kvantitativne metode predviđanja.....	4
3.2.1.	Vremenski nizovi.....	5
3.2.2.	Ekonometrijske metode	19
4.	Zaključak	20
5.	Literatura	21
6.	Sažetak.....	22

1. Uvod

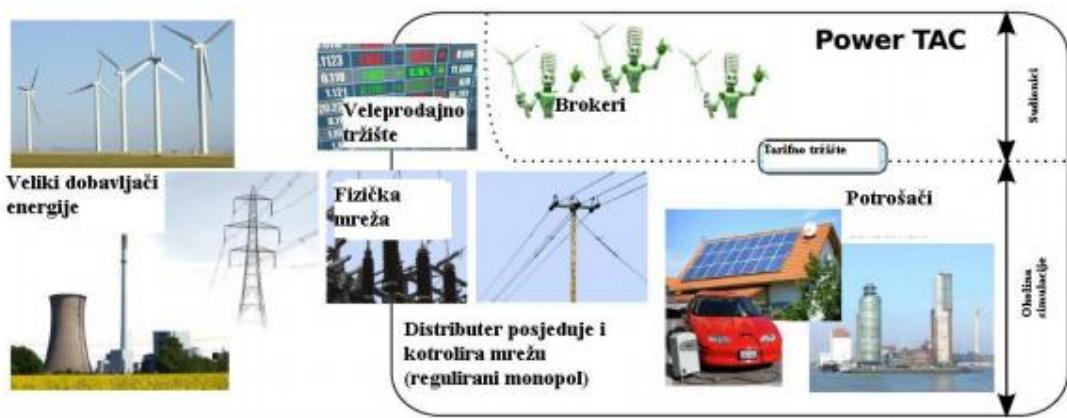
Analiza podataka sve češće podrazumijeva i stvaranje modela koji, ne samo dobro opisuju već postojeće podatke, nego mogu dati i dobru procjenu za buduće događaje. Posebno je zanimljivo tržište električne energije zbog svoje nestalnosti i promjenjivosti cijena. Poznavajući činjenicu da se na takvom tržištu trguje "dan unaprijed", jasno je da modeli moraju imati sposobnost brzog "učenja" na temelju novih podataka koji stižu svaki sat. I dok se na ostalim tržištima, poput tržišta dionica, najčešće govori samo o predviđanju cijena, na tržištu električne energije važno je i predvidjeti potrebnu količinu energije.

Ovaj rad će se baviti analizom podataka dobivenih iz igara natjecanja Power Trading Agent Competition, s ciljem pronalaska odgovarajućih metoda koje dobro predviđaju događaje na Veleprodajnom tržištu. Podatci će biti analizirani alatom R.

2. Power Trading Agent Competition

Power Trading Agent Competition je natjecanje u kojem se agenti, brokeri, bore za maksimizaciju profita i prevlast na tržištu električne energije.

Glavna zamisao Power TAC-a je simulacija liberaliziranog tržišta električne energije koje više odgovara modelu slobodnog tržišta, nego monopolja kako je to dosad bilo u većini država. Više ne postoji samo jedna tvrtka od koje svi kupuju električnu energiju bez obzira na cijenu, već mnoštvo samostalnih brokera.



Slika 1- Glavni elementi Power TAC natjecanja¹

Slika 1 predstavlja glavne elemente natjecanja. Tri su osnovna dijela natjecanja:

- veleprodajno tržište
- tarifno tržište i
- distributer.

Broker kupuje energiju na veleprodajnom tržištu (engl. *wholesale market*) od proizvođača (engl. *Generation Companies - GenCos*), a kasnije je prodaje potrošačima (engl. *customers*) koji su dio tarifnog tržišta (engl. *tarrif market*).

Svaki od brokera nastoji pridobiti potrošače nudeći im pretplate na različite tarife koje mogu biti vrlo povoljne u odnosu na cijene na klasičnom monopoliziranom tržištu. Brokeri dobavljaju energiju na veleprodajnom tržištu te, osim zarade od tarifa, mogu i na tom

tržištu ostvariti značajan prihod samo strateškim kupovanjem i prodajom električne energije.

Vrijeme u simulaciji je podijeljeno na vremenske odsječke (engl. *timeslots*). Svaki vremenski odsječak predstavlja 1 sat, a u stvarnom vremenu traje 5 sekundi.

3. Metode predviđanja

3.1. Kvalitativne metode predviđanja

Kvalitativne metode predviđanja ne zahtijevaju da podaci budu numerički. Moguće je da neki od podataka budu numerički, ali to nije nužan uvjet. Podaci su često produkt mišljenja ili procjene, a to čak ne mora biti mišljenje stručnjaka (npr. mišljenje kupaca dobiveno na temelju ankete). No, to ne znači da su ove metode išta jednostavnije od kvantitativnih metoda.

Ponekad problem predstavlja procjena korisnosti kvalitativnih metoda; ova vrsta metoda se najčešće koristi kao pomoć pri planiranju i dodatak kvantitativnim metodama, a ne daje specifičan brojčani proračun. Međutim, unatoč sumnji u njihovu vrijednost, ove metode nerijetko pružaju korisne informacije menadžerima pri donošenju odluka, posebno uzimajući u obzir da su dostupni podaci često podložni samo jednoj od ovih metoda. U kvalitativne metode ubrajaju se: metoda „Delphi”, istraživanje tržišta, statističke ankete itd. Ovaj rad se dalje neće baviti ovim metodama jer su ulazni podaci koji će se koristiti isključivo numerički i podložni velikoj većini kvantitativnih metoda predviđanja.

3.2. Kvantitativne metode predviđanja

Kvantitativne metode predviđanja se mogu primijeniti ako su sljedeća tri uvjeta ispunjena:

1. informacije o prošlosti su poznate,
2. te iste informacije su u numeričkom obliku ili se mogu transformirati u taj oblik,
3. može se očekivati da će se neki aspekti prošlih uzoraka ponoviti u budućnosti.²

Zadnji uvjet je poznat pod nazivom „prepostavka kontinuiteta” (engl. *assumption of continuity*) te je najvažnija premlisa svih kvantitativnih i mnogih kvalitativnih metoda.

Kvantitativne metode mogu se podijeliti na dvije veće skupine:

- vremenski nizovi i
- ekonometrijske metode.

3.2.1. Vremenski nizovi

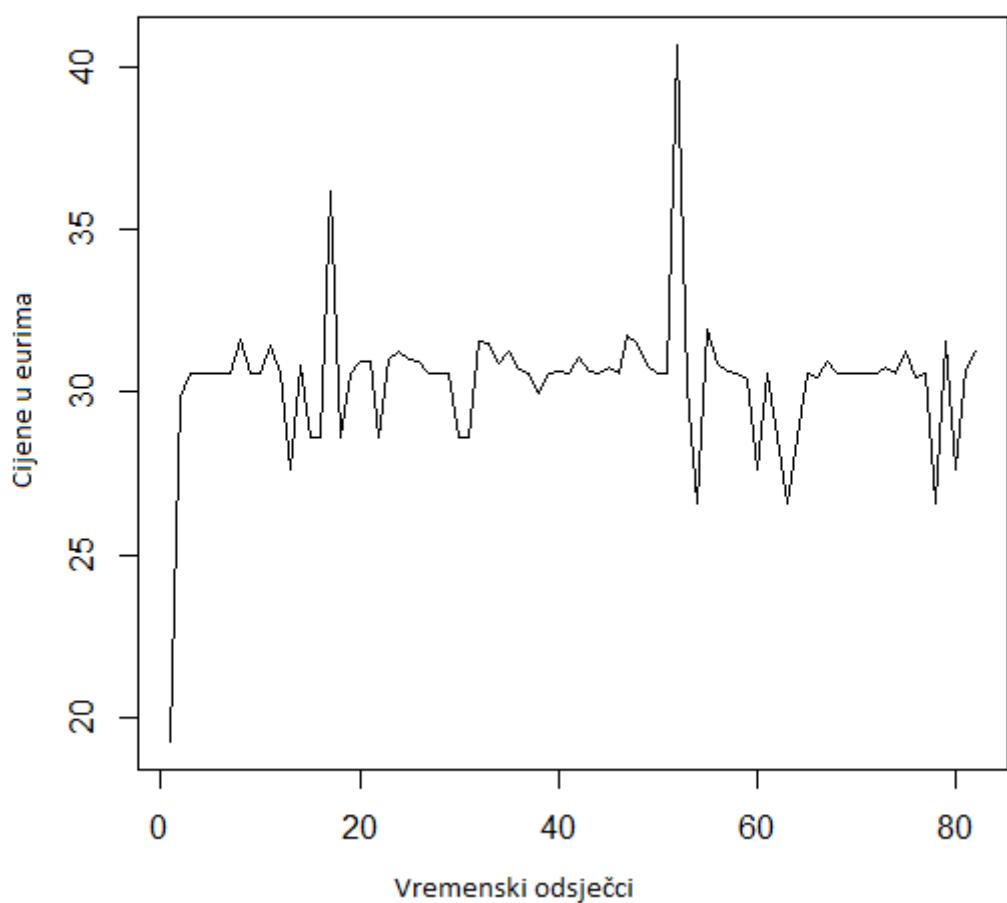
Kod prve skupine metoda predviđanje se temelji isključivo na podacima iz prošlosti, a ne uključuje nikakve druge varijable koje bi mogle biti korelirane s varijabljom koja se analizira. Cilj je takve analize pronaći uzorke koji se ponavljaju u prošlosti i koji bi se mogli ponoviti u budućnosti. Dva su glavna razloga zašto je takvo tretiranje podataka korisno. Kao prvo, sustav koji se proučava nije uvijek u potpunosti shvatljiv i mjerjenje korelacije između varijabli koje ga čine može biti iznimno zahtjevno. Drugi razlog je da ponekad predmet razmatranja uopće nije zašto se nešto događa, već samo predviđanje događaja u budućnosti.

U nastavku će biti objašnjene neke od metoda koje pripadaju skupini vremenskih nizova.

- **Pomični prosjek**

Pomični prosjek je jedna od jednostavnijih metoda za određivanje trendova kod vremenskih nizova. To je ujedno i jedna od metoda kojima se postiže „izglađivanje” krivulja da bi se uklonila manja slučajna kretanja koja ne predstavljaju niti karakterističan uzorak, niti imaju neki značajan utjecaj na niz. U nastavku će na primjeru biti opisana ova metoda.

Slika 2 prikazuje graf kretanja cijena na veleprodajnom tržištu simulacije PowerTAC u prvih 80 vremenskih odsječaka. Budući da se za svaki vremenski odsječak može trgovati 24 puta, trebalo je odrediti na koji način pristupiti podacima. Podaci prikazani na slici 2 su cijene za različite odsječke i dobiveni su iz prošlih igara.



Slika 2- Graf kretanja cijena na veleprodajnom tržištu

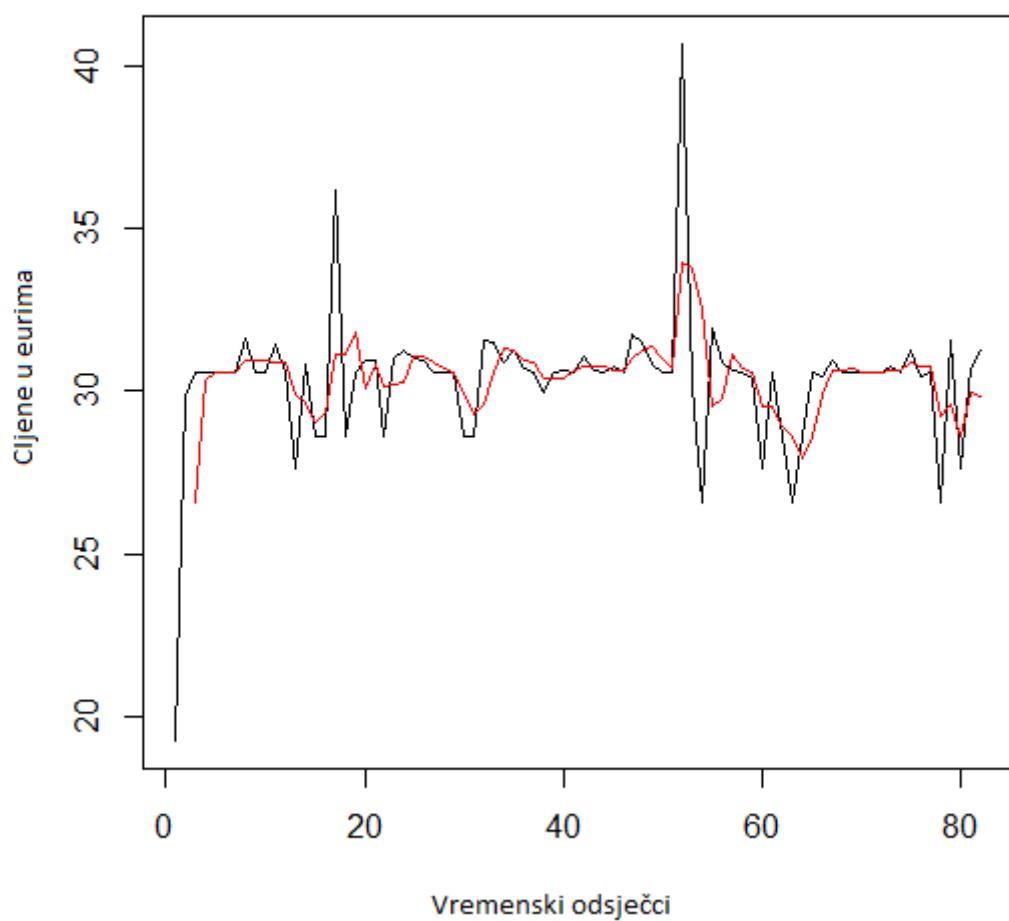
Na ovakvim „sirovim” podacima teško je uočiti trend kretanja zbog raznoraznih manjih skokova koji su u ovom slučaju samo šum. Zbog toga se u ovom primjeru primjenjuje metoda pomičnog prosjeka koja će na primjeru i biti opisana.

Vremenski odsječak	Promatrane vrijednosti	Pomični prosjek s okvirom 3	Pomični prosjek s okvirom 3
1	19.29	-	-
2	29.87	-	-
3	30.6	26.59	-
4	30.6	30.36	-
5	30.6	30.6	28.19
6	30.6	30.6	30.45
7	30.6	30.6	30.6
8	31.62	30.94	30.8
9	30.6	30.94	30.8
10	30.6	30.94	30.8
11	31.45	30.88	30.97
12	30.6	30.88	30.97

Slika 3- pomični prosjek

Slika 3 prikazuje princip centriranog pomičnog prosjeka. Za odsječke od 3 ili 5 uzastopnih vrijednosti računa se prosjek. Nakon toga, „prozor” od skupine vrijednosti pomiče se jedno mjesto naprijed te se postupak ponavlja. Na slikama 4 i 5 se vidi rezultat ovog postupka u grafičkom obliku. Originalni podaci su prikazani crnom, a izglađeni crvenom bojom.

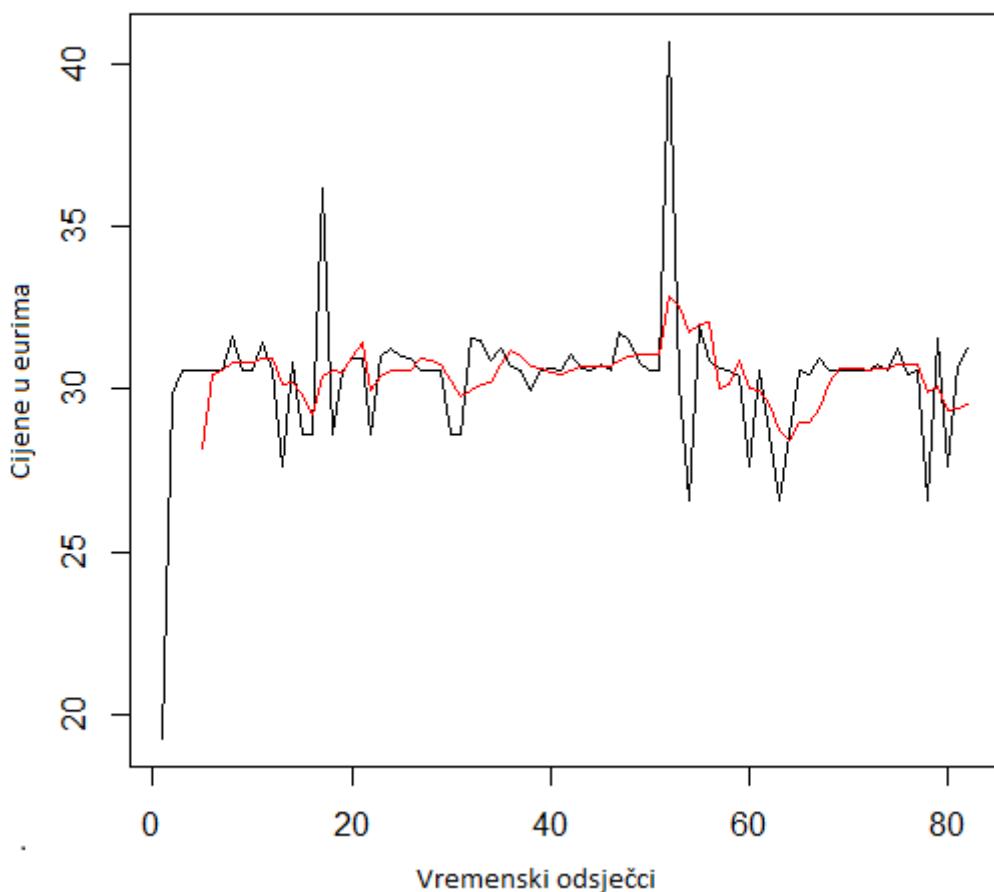
Pomični prosjek s okvirom 3



Slika 4- pomični prosjek s okvirom veličine 3

Na slici 4 vidljivo je kako je kretanje cijena nakon primjene izglađeno, vrhovi su manji, a lakše je primijetiti i trendove kretanja. Ipak, vidljiva su i odstupanja iz kojih se naziru nedostatci ove metode.

Pomični prosjek s okvirom 5



Slika 5-pomični prosjek s okvirom veličine 5

Gledajući sliku 5, vidljivo je da povećanje okvira na veličinu od 5 vremenskih odsječaka nije promijenilo stvari nabolje. Nova krivulja ne prati dobro originalne podatke i nije pogodna za korištenje u igri. U nastavku slijedi pojašnjenje nedostataka ove metode.

- **Eksponencijalno izglađivanje**

Prethodno opisana metoda pomičnog pomaka ima jednu značajnu manu- uzima u obzir samo posljednjih k vrijednosti i tretira ih kao jednakov vrijedne, a vrijednosti koje su ostvarene prije toga u potpunosti zanemaruje. Taj problem rješava metoda eksponencijalnog izglađivanja na način da se uvodi parametar α . To je broj između 0 i 1 koji se množi sa prethodnim vrijednostima na sljedeći način³:

$$\hat{x}_{t+1} = S_t = \alpha x_t + (1 - \alpha)S_{t-1}$$

Ako se jednadžba proširi raspisivanjem elementa S_{t-1} , postaje jasno zašto metoda ima u nazivu riječ "eksponencijalno":

$$\hat{x}_{t+1} = S_t = \alpha x_t + (1 - \alpha)[\alpha x_{t-1} + (1 - \alpha)S_{t-2}] = \alpha x_t + \alpha(1 - \alpha)x_{t-1} + (1 - \alpha)^2 S_{t-2}$$

Nadalje, nastavimo li raspisivanjem članova dobivamo sljedeći oblik jednadžbe:

$$\hat{x}_{t+1} = S_t = \alpha x_t + \alpha(1 - \alpha)x_{t-1} + \alpha(1 - \alpha)^2 x_{t-2} + \alpha(1 - \alpha)^3 x_{t-3} + \dots + \alpha(1 - \alpha)^{t-1} x_1$$

Ili

$$\hat{x}_{t+1} = S_t = \alpha \sum_{j=0}^t (1 - \alpha)^j x_{t-j}$$

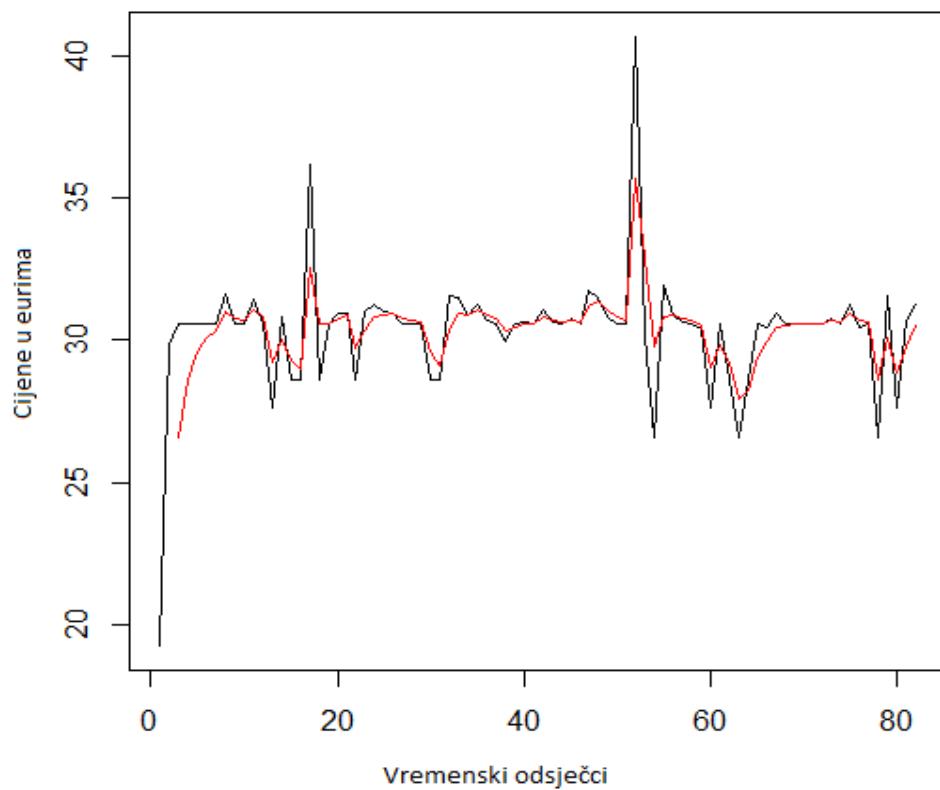
Tako vrijednosti koje su recentnije imaju veći značaj. Kako bi eksponencijalno izglađivanje dalo što bolje rezultate, valja pravilno izabrati vrijednost parametra α . Najbolja vrijednost parametra α je ona koja minimizira sumu kvadriranih odstupanja (engl. *Sum of Squared Errors*) :

$$SSE = \sum (x_i - \hat{x}_i)^2$$

Nažalost, ta vrijednost se pronalazi metodom pokušaja i pogrešaka.

U sljedećem primjeru korišten je parametar $\alpha=2/(n+1)$, gdje je n broj vremenskih odsječaka u okviru. Za testiranje metode ponovno se koriste podaci sa slike 2. Na njih se primjenjuje eksponencijalno izglađivanje s okvirom veličine 3 vremenska odsječka.

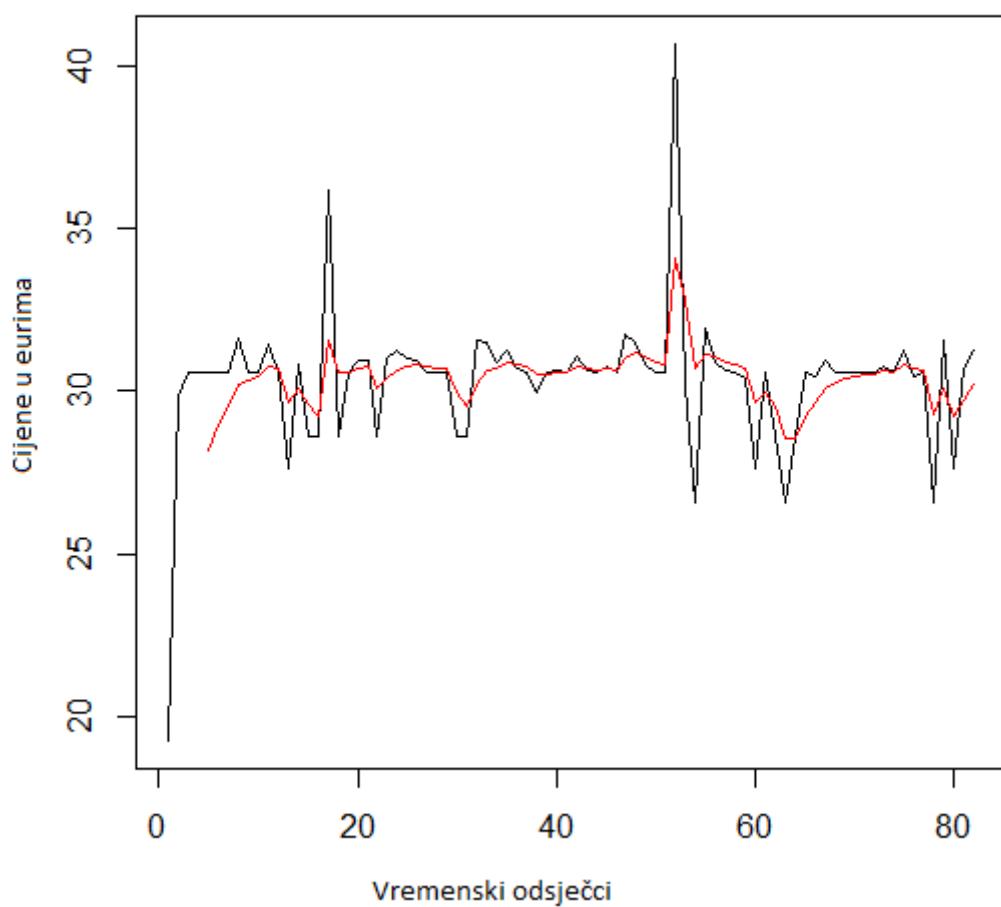
Eksponencijalno izglađivanje, n=3



Slika 6- eksponencijalno izglađivanje s okvirom 3

Na slici 6 vidljivo je da izglađeni graf jako dobro prati originalne podatke te eliminira smetnje i omogućuje preglednost kretanja cijena. Zatim je na podatke primijenjena ista metoda, ali s okvirom veličine 5 vremenskih odsječaka.

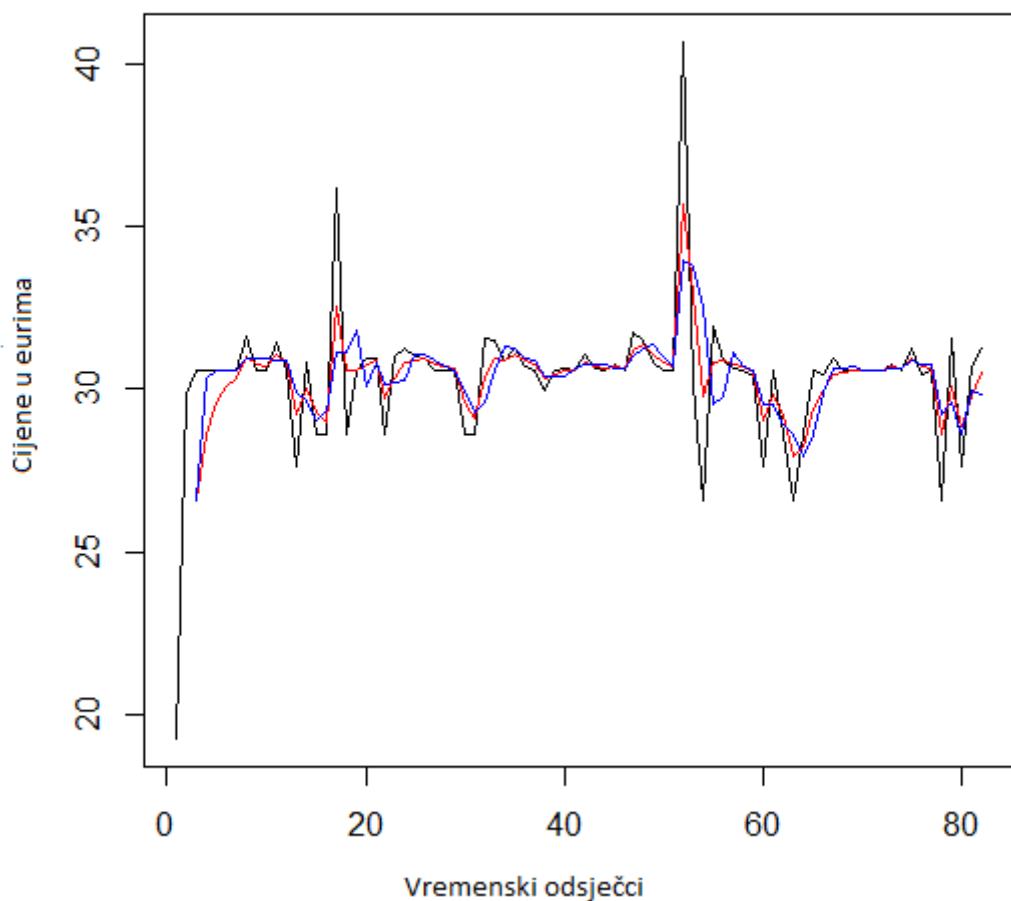
Eksponencijalno izglađivanje, n= 5



Slika 7- eksponencijalno izglađivanje s okvirom 5

Kao i kod pomičnog prosjeka, povećanje okvira nije donijelo poboljšanje, ali metoda i dalje dobro opisuje podatke. Na sljedećem je grafu, radi lakše usporedbe, prikazan graf pomičnog prosjeka, zajedno s grafom eksponencijalnog izglađivanja. Obje metoda u ovom slučaju primjenjuju okvir veličine 3 vremenska odsječka.

Usporedba pomičnog prosjeka i eksponencijalnog izglađivanja



Slika 8- usporedba eksponencijalnog izglađivanja i pomičnog prosjeka

Na slici 8 je plavom bojom prikazana metoda pomičnog prosjeka, a crvenom eksponencijalno izglađivanje. Jasno je vidljivo da metoda eksponencijalnog izglađivanja ima bolji učinak.

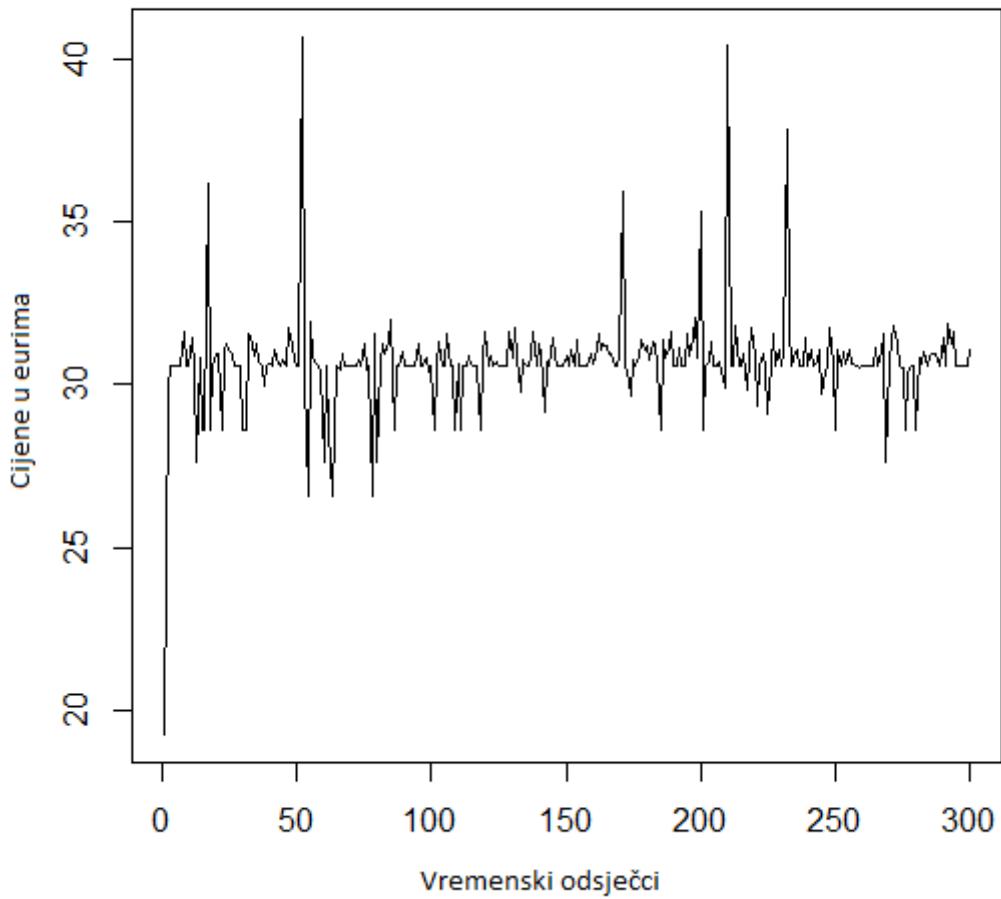
● Autoregresivni pomični prosjek

Model autoregresivnog pomičnog prosjeka (engl. *Autoregressive moving average, ARMA*) nastao je sjedinjenjem dva modela- autoregresije i procesa pomičnog prosjeka. Regresija u ekonometriji označava model kojim je opisana veza između neke zavisne varijable i jedne ili više nezavisnih. Autoregresija je model u kojem se pretpostavlja da trenutna vrijednost zavisne varijable ovisi o prethodnim vrijednostima te iste varijable. Autoregresivni proces

u kojem x_t zavisi o p prethodnih vrijednosti naziva se AR(p) proces i može se zapisati u sljedećem obliku⁴:

$$x_t = \alpha + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + w_t$$

gdje su α i φ konstante, a w_t bijeli šum. Na slici 10 je prikazan rezultat primjene ovog modela na podatke iz igre PowerTAC. Uzeti su podaci vidljivi na slici 9, prvih 300 cijena različitih vremenskih odsječaka.



Slika 9

```

R C:\Users\marijamijic\Documents\Telecom\Seminar\skripta.R - R Editor
(mydata=c(19.29,29.87,30.6,30.6,30.6,30.6,30.6,31.62,30.6,30.6,31
(mydata=as.ts (mydata))

plot (mydata)

predict(arima (mydata,order=c(10,0,0)), n.ahead=1)|

<      >

$pred
Time Series:
Start = 301
End = 301
Frequency = 1
[1] 30.71262

```

Slika 10- rezultati primjene autoregresije na podatke

Predviđena vrijednost, kao što je vidljivo na slici 10, iznosi 30,71262 eura. To je dosta dobar rezultat, s obzirom da je očitana vrijednost za taj vremenski odsječak iznosila 30,75 eura. Ovdje je za varijablu p uzeta vrijednost 10, jer je ta vrijednost dala najbolje predviđanje, ali su testirani i sljedeći slučajevi autoregresije reda 5, 20, 30 i 40, kao što je vidljivo u tablici 1.

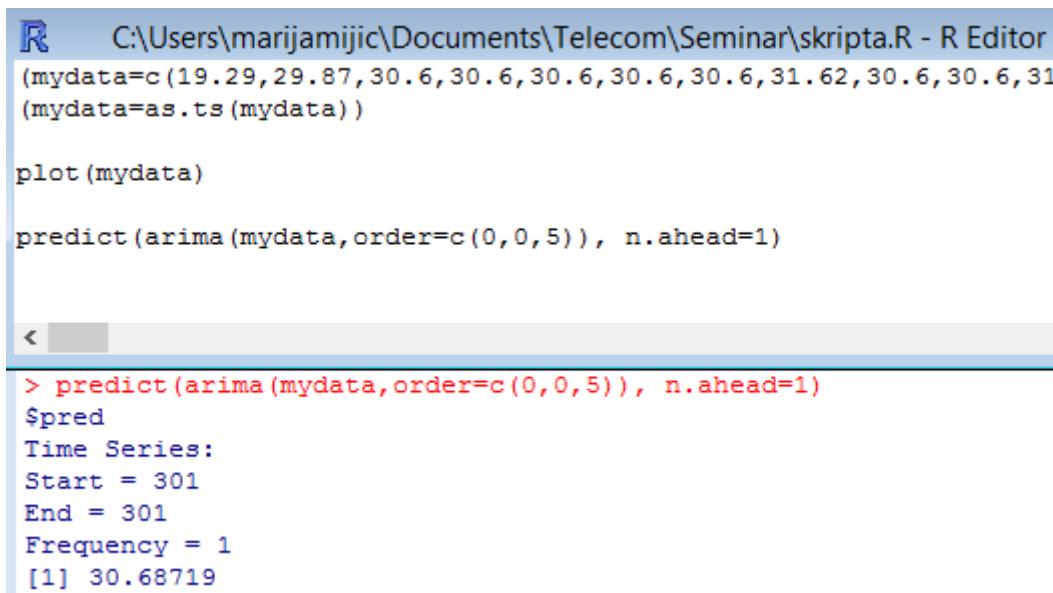
AR(5)	30,68818
AR(20)	30,61269
AR(30)	30,79693
AR(40)	30,48271

Tablica 1

Drugi dio, proces pomičnog prosjeka (engl. *moving average process*), prepostavlja da trenutna vrijednost zavisne varijable ovisi o prethodnim vrijednostima bijelog šuma⁴:

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}$$

Proces pomičnog prosjeka koji ovisi o q prethodnih vrijednosti bijelog šuma naziva se proces pomičnog prosjeka reda q i označava s $MA(q)$. Ovdje treba primijetiti da je ovaj proces, unatoč sličnosti imena, drukčiji od izglađivanja pomičnim prosjekom.



```
R C:\Users\marijamijic\Documents\Telecom\Seminar\skripta.R - R Editor
(mydata=c(19.29,29.87,30.6,30.6,30.6,30.6,30.6,31.62,30.6,30.6,31
(mydata=as.ts(mydata))

plot(mydata)

predict(arima(mydata,order=c(0,0,5)), n.ahead=1)

< 
> predict(arima(mydata,order=c(0,0,5)), n.ahead=1)
$pred
Time Series:
Start = 301
End = 301
Frequency = 1
[1] 30.68719
```

Slikica 11- rezultati primjene procesa pomičnog prosjeka na podatke

Na slici 11 se vide rezultati testiranja procesa pomičnog prosjeka. Varijabla q u ovom slučaju iznosi 5. Iako je predviđanje još uvijek blizu stvarne vrijednosti, lošije je od predviđanja autoregresijom. Rezultati testiranja iste metode s različitim redovima prikazani su tablici 2:

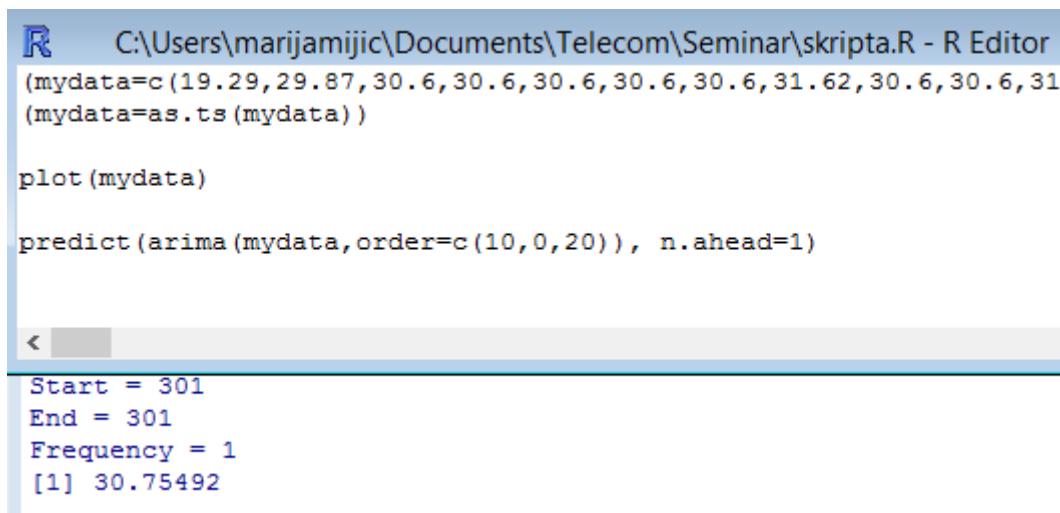
MA(10)	30,67602
MA(20)	30,67418
MA(30)	30,68413
MA(40)	30,92308

Tablica 2

Napokon, dolazimo do modela $ARMA(p,q)$. P i q označavaju red autoregresivnog procesa, ondnosno procesa pomičnog prosjeka⁴:

$$X_t = c + w_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i w_{t-i}$$

ARMA se najčešće koristi za stacionarne vremenske nizove. Ukoliko se niz pokaže kao nestacionaran (npr. slučajna šetnja), za predviđanje se koristi Autoregresivni integrirani pomični prosjek (engl. Autoregressive integrated moving average, ARIMA).



```
R      C:\Users\marijamijic\Documents\Telecom\Seminar\skripta.R - R Editor
(mydata=c(19.29,29.87,30.6,30.6,30.6,30.6,30.6,31.62,30.6,30.6,31
(mydata=as.ts(mydata))

plot(mydata)

predict(arima(mydata,order=c(10,0,20)), n.ahead=1)

< [REDACTED]
Start = 301
End = 301
Frequency = 1
[1] 30.75492
```

Slika 12- testiranje ARMA-e

Kod testiranja ARMA-e, kombinacija $(p,q)=(10,20)$ se pokazala kao pun pogodak, budući da je očitana vrijednost 30,75. Ostale kombinacije varijabli p i q prikazane su u tablici 3:

ARMA(5,5)	30,54811
ARMA(5,10)	30,57419
ARMA(5,20)	30,65757
ARMA(10,5)	30,88921
ARMA(10,10)	30,58927
ARMA(20,5)	30,92297
ARMA(20,10)	30,88216

Tablica 3

- Autoregresivni integrirani pomični prosjek

Uz modele autoregresije i pomičnog prosjeka ARIMA, za razliku ARMA-e, obuhvaća još i red integracije vremenskog niza. Red integracije označava koliko puta vremenski niz treba derivirati da bi postao stacionaran. Dakle, ARIMA je zapravo generalizirani oblik ARMA-e za nizove koji nisu stacionarni. Za proces kažemo da je ARIMA(p,d,q) ako je

$$\nabla^d x_t = (1 - B)^d x_t$$

ARMA(p,q)⁴. Ovdje varijabla d označava broj deriviranja niza. Općenito ARIMA(p,d,q) se zapisuje u sljedećem obliku:

$$\varphi(B)(1 - B)^d x_t = \theta(B)w_t$$

gdje je operator B definiran kao:

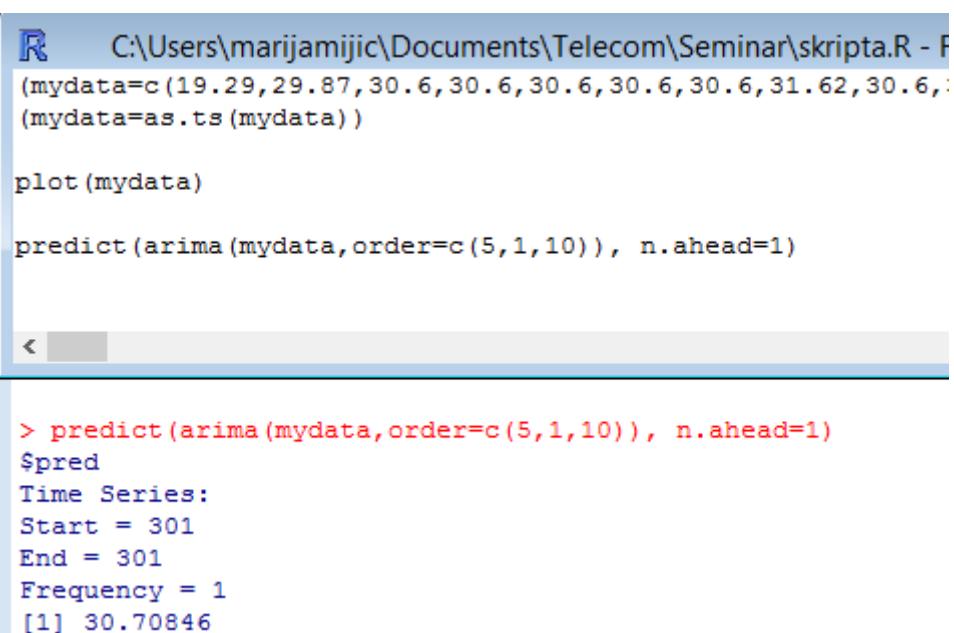
$$Bx_t = x_{t-1}$$

te su njegove potencije:

$$B^2 x_t = B(Bx_t) = Bx_{t-1} = x_{t-2}$$

Općenito se može zapisati u obliku:

$$B^k x_t = x_{t-k}$$



```
R C:\Users\marijamijic\Documents\Telecom\Seminar\skripta.R - F
(mydata=c(19.29,29.87,30.6,30.6,30.6,30.6,30.6,31.62,30.6,
(mydata=as.ts(mydata))

plot(mydata)

predict(arima(mydata,order=c(5,1,10)), n.ahead=1)

< [REDACTED]

> predict(arima(mydata,order=c(5,1,10)), n.ahead=1)
$pred
Time Series:
Start = 301
End = 301
Frequency = 1
[1] 30.70846
```

Slika 13-testiranje ARIMA-e

Od svih testiranih kombinacija, kombinacija $(p,i,q)=(5,1,10)$ se pokazala najboljom kod testiranja ARIMA-e:

ARIMA(5,1,5)	30,61983
ARIMA(10,1,20)	30,67317
ARIMA(5,1,20)	30,56114
ARIMA(10,1,5)	30,53246
ARIMA(10,1,10)	30,88471
ARIMA(20,1,5)	30,57708
ARIMA(20,1,10)	30,61793

Tablica 4

Općenito, bolji rezultati su dobivani korištenjem ARMA-e, što bi značilo da podatke vjerojatno ne treba integrirati jer su stacionarni.

3.2.2. Ekonometrijske metode

Kod ekonometrijskih metoda, za razliku od vremenskih nizova, predviđanje se temelji na korelaciji varijabli. Prepostavlja se da jedna ili više varijabli ulaznih varijabli utječe na predviđanu varijablu (izlaz) te da se svaka promjena u ulaznim vrijednostima odražava na izlaznu varijablu na način koji je moguće predvidjeti. Od ekonometrijskih metoda najpoznatija je regresija.

4. Zaključak

Područje predviđanja događaja zahtjeva dobro poznavanje matematičkog aparata koji stoji u pozadini. U ovom seminarskom radu korištene su neke od najpopularnijih metoda predviđanja, a testiranje se odvijalo po principu pokušaja i pogrešaka. Tu je za budući rad ostavljeno prostora za "pametniji" odabir modela. Podaci bi se prije samog testiranja trebali detaljno analizirati da bi se utvrdilo postoje li neke važne karakteristike poput stacionarnosti ili korelacije među podacima. Također, budući da ovdje nisu korištene ekonometrijske metode, valjalo bi istražiti postoje li varijable koje utječu na cijenu električne energije i u kojem razmjeru. Jedna od varijabli koja bi mogla imati značajan utjecaj su vremenske prilike.

Ipak, i bez detaljne analize, iz testiranja se nazire da korištene metode dobro prate podatke te ponekad čak i izvrsno predviđaju događaje. Stoga, vrijedi ovoj temi posvetiti još vremena, posebno uzimajući u obzir činjenicu da profit u igri PowerTAC izravno ovisi o mogućnostima brokera da pametno trguje.

5. Literatura

- [1] Ketter, W., Collins, J., Reddy, P., Flath, C., Weerdt, M. The Power Trading Agent Competition, Erasmus Research Institute of Management Report, ERS-2011-011-LIS, SSRN 1975237, 2011.
- [2] Makridakis, Spyros G. Forecasting: Methods and Applications, Wiley; 3rd edition (December 1997)
- [3] Cooray, T.M.J.A., Applied Time Series Analysis and Forecasting, Series Oxford U.K, 2007, ISBN 978-81-7319-901-1
- [4] Shumway, R., Stoffer, D., Time Series Analysis and Its Applications , Springer; 3rd edition (November 24, 2010)

6. Sažetak

Predviđanje događaja danas se vrši uz pomoć složenih alata koji mogu vrlo brzo obraditi veliku količinu podataka. Neke od metoda predviđanja su opisane u ovom radu i testirane na podacima iz igre PowerTAC koristeći alat R. PowerTAC je simulacija liberaliziranog tržišta električne energije u kojem brokeri kupuju energiju na veleprodajnom tržištu te je kasnije prodaju potrošačima po maloprodajnim cijenama. Metode koje su testirane su redom: pomični prosjek, eksponencijalno izglađivanje, autoregresija, proces pomičnog prosjeka, autoregresivni pomični prosjek te autoregresivni integrirani prosjek. Metode su testirane na podacima kretanja cijena s veleprodajnog tržišta. Najbolji rezultati predviđanja su dobiveni primjenom autoregresivnog pomičnog prosjeka reda.