

# ДОКУМЕНТАЦИЈА ТЕХНИЧКОГ РЕШЕЊА

## „SeLaR – Информациони систем – база и анализа података“

### Аутори техничког решења

- *др Ивана Радојевић, научни сарадник,*
- *др Душан Стефановић, ванредни професор,*
- *др Љиљана Чомић, редовни професор,*
- *др Александар Остојић, ванредни професор,*
- *др Марина Топузовић, ванредни професор,*
- *др Ненад Стефановић, доцент.*

### Наручилац техничког решења

- Институт за биологију и екологију, Природно-математички факултет,  
Универзитет у Крагујевцу

### Корисник техничког решења

- Институт за биологију и екологију, Природно-математички факултет,  
Универзитет у Крагујевцу

### Година када је техничко решење урађено

- 2008 - 2012

### Област технике на коју се техничко решење односи

- Хидробиологија
- Мониторинг вода
- Екологија водених екосистема
- Микробиологија вода
- Квалитет вода - еколошки аспект и санитарно-еколошки аспект

## **1. Опис проблема који се решава техничким решењем**

Увидом у стање акумулација и језера у Србији, утврђено је постојање бројних и различитих проблема (неадекватан мониторинг, еутрофизација, недостатак еколошке заштите и сл.) који угрожавају рационално управљање и њихову одрживу експлоатацију. Помоћ у решавању наведених проблема огледа се у адекватнијој и напреднијој примени резултата истраживања из претходних периода. Поставка обједињених података стандардних и научних хидробиолошких истраживања омогућила би да се на репрезентативним сетовима података, који садрже микробиолошке и остале параметре, испитају њихови односи и зависности. Претпоставка је да се овим истраживањем могу добити нова научна сазнања о заједницама микроорганизама у акватичним екосистемима, а да она могу имати и шири апликативни значај приликом праћења стања и побољшања квалитета вода.

Србија је земља сиромашна природним језерима, али је у њој изграђен већи број акумулација са различитом наменом. Тако је у нашој земљи направљено више од 150 хидроакумулација од којих 26 има запремину већу од  $10 \times 10^6 \text{ m}^3$ . Највећи број акумулација се користи за снабдевање становништва водом за пиће или за потребе хидроенергетике.

Наведени проблеми су у непосредној вези са недостатаком и/или недоступношћу валидних информација о језерима и процесима у њима. Лимнолошки подаци се налазе у оквиру различитих институција које се баве различитим аспектима проучавања вода, нису у довољној мери систематизовани, нити припремљени за савремено коришћење. Из наведених разлога је веома често тешко доћи до неопходних података и потребно је уложити доста времена да се они сакупе, јер их је немогуће наћи на једном месту. Посебан проблем је и што су постојећи подаци на различите начине ускладиштени, што додатно отежава претрагу и доступност података.

Превазилажење оваквог стања је могуће применом рачунарских информационих технологија које омогућавају да информације буду архивирани на адекватан начин, да су лако доступне, да се могу брзо претраживати и ефикасно користити. На овај начин би велики број корисника могао да брзо дође до неопходних података који могу да послуже као основа за истраживања различитих нивоа стручности. Поред научне и стручне јавности, оваква база података би била од великог значаја и за остале кориснике, јер би на најбржи могући начин могли доћи до потребних информација.

Последњих година се и у нашој земљи уводе савремене информационе технологије. Проблем је што не постоји увек и координација у раду различитих институција и служби које се баве проучавањима језера и акумулација, тако да је приступ траженим подацима парцијалан и отежан.

SeLaR информациони систем о језерима и акумулацијама Србије даје уређене податке о њиховом стању. Систем је намењен великом броју корисника, нарочито у области биологије и екологије, али и других учесника у управљању акумулацијама и омогућава њихову сарадњу. SeLaR подржава све видове управљања и одрживу експлоатацију водених ресурса у целини. Пружа на увид систематизоване информације у овој области научној и стручној светској јавности.

## **2. Стање решености проблема у свету – приказ и анализа постојећих решења**

Брз развој цивилизације, потпуно неусклађен са очувањем природних ресурса, довео је до тога да је све мање вода остало незагађено или барем у задовољавајућим

границама. Све израженији процеси загађивања навели су људе широм наше планете да се предузму неопходни кораци за обнову и заштиту вода.

Из ових разлога, у многим земљама су уложени напори да се оформи база података о акватичним екосистемима. Те базе пружају корисницима податке из различитих области: порекло и тип акватичних екосистема, физичке и хемијске карактеристике, биолошке карактеристике, степен загађености и угрожености и сл. Овакве базе на националном нивоу су послужиле и као полазна основа за формирање података на интернационалном нивоу (<http://www.livinglakes.org>; <http://www.worldlakes.org>, <http://www.ilec.or.jp>).

У свету постоји велики број информационих система који обухватају и обрађују проблематику језера и акумулација са различитих аспеката. На пример, Мое et al. (2008) су креирли јединствени информациони систем језера Европе где су подаци организовани у оквиру релационе базе података и то у пет основних табела: основне информације, хемијске информације о узорку, биолошке информације о узорку, хемијске вредности (које укључују вредности притиска, рН или фосфора) и биолошке вредности (као што су биомаса и абунданца по таксону). Специјална пажња у овој области се посвећује системима за управљање воденим ресурсима и системима за праћење квалитета вода:

- информациони систем за одређивање квалитета воде (Parinet et al., 2004),
- информациони систем који се базира на релационој бази података и садржи фази модел за одређивање квалитета воде (Mitreski et al., 2004, Menshutkin et al., 2009a, 2009b),
- информациони систем за управљање воденим ресурсима интеграцијом хидро-хемијских, хидролошких, метеоролошких, микробиолошких података, података о фито- и зоопланктону, макрофитама итд. (Kondrat'ev et al., 2006),
- информациони систем за управљање који се базира на подацима добијеним у реалном времену (Imberger, 2004),
- информациони систем за управљање воденим ресурсима базиран на релационој бази података која садржи хидролошке податке, податке о квалитету воде, расположиви потенцијал водених ресурса, наводњавање, привредне активности у околини и становништво (Fulazzaky & Akil, 2009),
- информациони системи за менаџмент водама комбиновани са географским информационим системима (Park et al., 2006).

Последњих година различити алати и технике информационих технологија, као и методе истраживања података (data mining), постају важан део у праћењу стања квалитета вода, омогућавају предвиђање промена и значајни су у процесима одрживог управљања воденим ресурсима (Kumar et al., 2006). Data mining, позната и као "knowledge-discovery in databases", подразумева аутоматско или полуаутоматско истраживање и анализу велике количине података у циљу откривања шаблона и односа сакривених међу подацима (Han et al., 2010, Gorunescu, 2011). Анализу ризика од загађења различитих вода помоћу анализе података и математичких модела истраживао је Ganoulis (2009). Методе DM успешно су примењиване у праћењу стања квалитета вода (Conrads & Roehl, 2006a), али и неких микробиолошких индикатора квалитета воде (Ogwueleka & Ogwueleka, 2010). Анализу, моделирање и предвиђање бројности индикатора фекалног загађења различитим статистичким техникама, вештачким неуронским мрежама и другима алатима применио је већи број аутора (Bergstein et al., 2001, Neelakantan et al., 2001, Idakwo & Abu, 2004, Chandramouli et al., 2007, Iscen et al., 2008, Syed Ahmad et al., 2009, Kazemi Yazdi & Scholz, 2010).

Неки од аутора указују на чињеницу да параметри квалитета вода сакупљени у стандардним хидробиолошким истраживањима могу бити искоришћени за прављење

модела који ће ефикасно омогућити праћење динамике и предвиђање стања различитих бактеријских заједница у воденим екосистемима (Brion et al., 2001, Neelakantan et al., 2001). Најбржи и најефикаснији начин за истраживања ове врсте јесте примена информационог система. У управљању воденим екосистемима информационе технологије налазе све ширу примену.

### 3. Суштина техничког решења

Информациони систем SeLaR створен је у циљу интегрисања свих битних фактора за надзор и управљање акумулацијама и језерима у Србији. Овакав систем омогућава брзо и ефикасно добијање потребних информација за потребе научно-истраживачког рада, али и за потребе водопривредних организација и осталих учесника у процесу управљања воденим ресурсима.

SeLaR информациони систем садржи: морфометријске податке, физичко-хемијска својства, хидробиолошке податке о заједницама организама, податке о квалитету воде, начин коришћења, детериорацију околине и проблеме загађења, податке о методама еколошке заштите, податке о предузетим мерама за побољшање квалитета, као и друге податке од значаја за језеро/акумулацију. Посебна пажња посвећује се карактеристикама бране, а код акумулација за водоснабдевање и процесима технолошке обраде. Саставни део базе података су фотографија, мапе, шеме, библиографски подаци: име истраживача; институција; назив публикације; датум издавања; апстракт. Обезбеђене су везе до различитих датотека (књига, чланака, монографија). Важни параметри система су временска и просторна димензија. Подаци о истраживачима и публикацијама које су издали

Извори информација су: резултати истраживања ПМФ-а у Крагујевцу, литературни извори, документација водопривредних и других организација које се баве квалитетом воде и газдовањем акумулацијама. Унети подаци су конвертовани у потребан формат. Информациони систем обезбеђује потребне форме за унос ових података са елементима вишекорисничког рада и заштитом, водећи рачуна о људским аспектима коришћења информационог система. Извештаји обухватају потребе свих група корисника. Претраживања података је: по индексу или кључним речима.

Он је неопходан корак ка унапређењу постојећег стања и омогућава:

- Прикупљање и одржавање података;
- Просторно и временско праћење стања физичких, климатских, топографских карактеристика околине језера, као и хидролошких, физичко-хемијских и биолошких карактеристика;
- Откривање и анализирање веза између променљивих;
- Анализу динамике;
- Откривање образаца динамике променљивих или групе променљивих;
- Стварање различитих data mining модела за предвиђање и подршку у одлучивању, односно моделирање и прогнозу стања;

Све наведено је омогућено уз помоћ оригиналног интерфејса који повезује базу података информационог система са структурама података који су неопходни за статистичку обраду и data mining. ИС SeLaR омогућава стварање научно заснованог система за управљање воденим ресурсима и омогућава процену ефикасности постојећих система и праћење препорука у погледу њиховог побољшања. Систем је намењен за велики број корисника, нарочито у области науке и образовања, али и за остале учеснике у управљању акумулацијама и омогућава њихову сарадњу. Он подржава све облике управљања и одрживо коришћење водних ресурса у целини и омогућава пренос информација и знања и стварање мреже истраживача.

На тај начин SeLaR представља виртуелну лабораторију на располагању великом броју корисника који могу да координирају, синхронизују и размењују резултате својих истраживања.

#### **4. Детаљан опис техничког решења (укључујући и пратеће илустрације и техничке цртеже)**

Информациони систем SeLaR је целовит и сврсисходан информациони систем о језерима и акумулацијама у Србији. Омогућава складиштење, обраду, анализу и релативно лак приступ и манипулацију научним подацима, нов, савремен и актуелан увид у стање језера и хидроакумулација у Србији, као и обједињено и координирано управљање подацима и добијање информација које су потребне за мониторинг, управљање и едукацију. Део интранет апликације информационог система SeLaR који је описан у овом техничком решењу, чија је насловна страница приказана на **Слици 1**, доступан је преко интернет портала ([www.selar.pmf.kg.ac.rs](http://www.selar.pmf.kg.ac.rs)). Детаљан опис поставке Информациони систем SeLaR дат је у радовима Ćomic et al. (2006), Radojević (2007a), Radojević et al. (2007b, 2008a,b). Систем је унапређен у правцу одређивања квалитета вода и анализе података (Stefanović et al., 2012).

ИС SeLaR, осим различитих прегледа комбинованих података, представља основу за формирање структура података на којима могу директно да се примене различите методе анализе података - статистичке методе и методе DM. База података је дизајнирана тако да може да омогући и симулацију. На пример, да са аспекта микробиолошких параметара у једној акумулацији, а на основу унетих података, утврди састав и динамику заједнице микроорганизама, изврши процену степена трофичности екосистема, утврди битне утицаје на дате параметере и на основу добијених резултата предложи мере адекватне еколошке заштите ради побољшања стања екосистема, и слично.

За изградњу информационог система примењен је објектни приступ уз коришћење UML-a, Unified Modelling Language (Naiburg & Maksimchuk, 2001). UML омогућава моделирање информационог система различите намене уз помоћ више дијаграма од којих неки дају приказ система са становишта корисника, односно дају приказ шта информациони систем треба да садржи. Дијаграм класа даје приказ класа у систему од којих се оне које се односе на ентитете реалног система могу директно трансформисати у модел података, а затим и у релациону базу података. Посебна група дијаграма приказује понашање објеката у систему. Основне компоненте ИС SeLaR јесу одржавање базе података, извештаји и анализа података.

Слика 1. Насловна страна централне базе података

База података, као модел реалног система, садржи све ентитете релевантне за реализацију постављених циљева. Ентитети су приказани у **Табели 1**, на пример, језера и акумулације, физичке димензије, локација, догађаји, итд. Ентитети имају својства која су садржана у бази података. У даљем излагању ова својства ће се означавати термином параметри, на пример физичко-хемијски параметри, микробиолошки параметри, локација, дубина итд.

У оквиру алата за анализу података ова својства се називају променљивим. Под објектима се подразумевају појединачне вредности својстава (у бази података одговарају врстама у табели). Циљ креирања базе података је да се омогуће различити аспекти истраживања. Избор битних компоненти за креирање ИС SeLaR (у овом случају ентитета) базиран је на научним чињеницама које су добијене на основу досадашњих истраживања језера и акумулација, а које се односе на њихово функционисање и параметре који утичу на њих (Jørgensen, 2003, Jones et al., 2004, Bruhn & Soranno, 2005, Jørgensen et al., 2005b, Martin & Soranno, 2006, Cheruvilil & Soranno, 2008). При избору ентитета водило се рачуна да су они у складу са препорукама Оквирне директиве Европске уније о водама, WFD CIS (2000/60/EC).

**Табела 1.** Преглед ентитета ИС SeLaR

Назив ентитета	Опис ентитета
Екосистем	Тип екосистема (језеро или хидроакумулација)
Врсте ЈХА	Врста <i>језера и хидроакумулација</i> у односу на начин настанка (тектонско, еолско, хидроакумулација и тд).
Типови мешања воде	Подаци о типу мешања воде.
Време задржавања	Средње време које вода проведе у акумулацији.
Догађаји	Релевантни догађаји који имају утицај на језеро/акумулацију, као што су поплава, плављење ораница, цветање воде, нагло опадање нивоа воде итд.
Активности	Активности на језеру/акумулацији или у околини као што су туризам, спортске активности, риболов, саобраћај итд.
Климатске карактеристике	Температура ваздуха, воде, падавине итд.
Коришћење земљишта	Начин коришћења земљишта у околини - поља, шуме, пашњаци, воћњаци, спортски терени, пут, паркинг итд.
Популација у околини	Број насеља, број становника, густина популације итд.
Интервенције на акумулацији	Интервенције као што су додавање $\text{CuSO}_4$ , примена хиполимнетичке аерације итд.
Физичке димензије	Подаци о површини, запремини, максималној дубини, просечној дубини, средњој годишњој амплитуди, површини, дужини и ширини слива, батиметријска мапа језера/акумулације итд.
Пречишћавање	Санитарна заштита и пречишћавање отпадних вода.
Језеро Хидроакумулација	Основне, временски непроменљиве карактеристике језера и хидроакумулација, као што су опис, локација (географска ширина и дужина, надморска висина, географска карта), подаци о брани (висина, дужина, круна, дужина куполе, слика).
Реке и канали	Реке и канали који утичу или истичу из језера/акумулације.
Филум, Класа, Ред, Фамилија, Род, Врста, Биолошка заједница	Систематске категорије припадника забележених биолошких заједница.
Физичко-хемијски параметри	Физичко-хемијски параметри са одговарајућом јединицом мере.
Биолошки параметри	Биолошки параметри са одговарајућом јединицом мере.
Карактеристике биолошка заједница	Карактеристике које се односе на одређену биолошку заједницу.
Метода карактеристике	Методе које се примењују за појединачне карактеристике. На пример, методе за оцену квалитета воде: стандардне методе (АРНА 1985, 1995, 1998, 2005).
ЈХА КГ	Веза <i>језера/акумулације</i> и <i>карактеристика биолошких заједница</i> са следећом семантиком: једна карактеристика се односи на једно или више језера/акумулације и једно језеро/акумулација садржи више карактеристика.
ЈХА КГ вредности	Приказ <i>вредности карактеристика</i> које су измерене одређеног датума за микролокацију (локација на <i>језеру/акумулацији</i> , дубина). Меморисање је за више вредности једне карактеристике, нпр. вредности карактеристике воде, односе се на различите акумулације, локације у оквиру језера, датуме итд.
КГ ФХП	Веза категорије <i>биолошка заједница</i> и <i>физичко-хемијских параметара</i> са следећом семантиком: једна категорија биолошке заједнице може да садржи један или више физичко-хемијских параметара и један физичко-хемијски параметар може да се нађе у више категорија биолошке заједнице.
КГ МБП	Веза категорије <i>биолошка заједница</i> и <i>микробиолошких параметара</i> са следећом семантиком: једна категорија биолошке заједнице може да садржи један или више микробиолошких параметара и један микробиолошки параметар може да се нађе у више категорија биолошке заједнице.
Врста ЈХА	Веза ентитета <i>врста</i> и <i>језеро/хидроакумулација</i> са семантиком: једна врста се налази у једном или више језера/акумулација и једно језеро/акумулација садржи једну или више врста.
Јединице мере	Шифарник јединица мера којим се изражавају вредности свих параметара.

Сви потребни подаци у релационој бази података су концентрисани на једном месту. Једна група ових параметара односи се на физичке, климатске и топографске карактеристике околине акумулација/језера, друга на физичко-хемијске и биолошке карактеристике акумулација/језера (Clair et al., 2003). Важни параметри су дубина и локација у самој акумулацији/језеру (Roehl & Cook, 2010). Посматрајући важну улогу акумулација/језера са аспекта водоснабдевања у ИС SeLaR примењене су различите методе анализе и приступа квалитету воде (May et al., 2008). У односу на наведено, три групе ентитета су уграђене унутар базе података: ентитети који се односе на карактеристике акумулације/језера и њихове околине, ентитети са вредностима различитих параметара (физичко-хемијски и биолошки) и израчунате вредности које представљају системе, као што су квалитет воде и ентитети који се односе на систематске категорије (Табела 1).

База података садржи суштинске податке језера и акумулација који обезбеђују реализацију циљева информационог система. Ови подаци су структурирани и повезани тако да обезбеђују квалитетно и ефикасно одржавање (уношење нових података, измену или брисање) и једноставно добијање различитих информација у виду упита, табела или графичких приказа. То омогућава праћење релевантних параметара за управљање језерима и акумулацијама, као што су стање биљног и животињског света, стање квалитета воде кроз простор и време и слично. Примери различитих начина за упите и извештаје дати су на **Сликама 2 - 5**. Извештај се добија упитом за одређену акумулацију, за временски период, за локацију или/и дубину, као и методу по којој се одређује квалитет воде.

The screenshot shows the SeLaR web application interface. At the top, there is a navigation bar with 'File', 'Home', and 'Add-Ins' menus. The main header features the 'SeLaR' logo and the text 'SERBIAN LAKES AND RESERVOIRS'. Below the header, there are buttons for 'Unos', 'Izmena', 'Brisanje', and 'Povratak'. The main content area is titled '>>> Brana jezera ili hidroakumulacije'. On the left, there is a search form labeled 'PRETRAGA' with dropdown menus for 'Naziv jezera ili hidroakumulacije' (set to 'Gruza'), 'Izbor' (set to 'Brana'), and 'Prikaz' (set to 'Forma'), along with a 'Prikazi' button. The central part of the page displays a table of data under the heading 'PREGLED'. The table lists the following information:

Šifra jezera ili hidroakumulacije:	10
Naziv jezera ili hidroakumulacije:	Gruza
Visina:	51.5 m
Kruna:	207.89 m
Odnos između krune i visine:	4.04 m
Dužina kupole krune:	230.3 m
Dužina:	288.05 m
Opis lokacije:	Izgradnje brane započela je 1979. god. a akumulacija je definitivno formirana 1985. god. (Stepanović, 1974). Građevinski a visina brane iznosi 51.50m, a letiva u kruni je 207.89m. Odnos letive u kruni i visine brane je 4.04. Dužina kupole u kruni je 230.30m. Lucna brana se u gornjim delovima koritune na obema padinama oslanja na stenu preko oporaca. Ukupna dužina brane zajedno sa oporcama je 238.05m. Krana preliva nalazi se na 270mm. Na brani su postavljena četiri ispusta o 1300mm koji imaju osovinu na koti 245.000. Pored ova četiri ispusta postoji još jedan o 200mm. Voda iz akumulacije se zahvata sa tri nivoa: prvi, najviši je na koti 265.00, srednji na koti 257.50, a treći, najniži, na koti 250.00

On the right side of the page, there is a section titled 'Slika brane (kliknuti na sliku za uvećanje):' which contains a photograph of the Gruza dam, showing the concrete structure and the surrounding landscape.

**Слика 2.** Начин претраге/упита података централне базе података и резултат



## Kvalitet vode

Naziv jezera ili hidroakumulacije

Lokacija

Dubina

Reka

Vrsta klasifikacije

Period  
OD   
DO

### Kvalitet vode

WQI

**Jezero ili hidroakumulacija**

**Šifra:** 10  
**Naziv:** Gruza

**Od datuma:** 1/1/2001  
**Do datuma:** 1/1/2002  
**Lokacija:** Brana  
**Dubina:** 10 m  
**Reka:**

Kategorija:	84	Serbian WQI
vrednost:	II klasa	Veoma dobar
r. br.	parametar	vrednost
1	Temperatura (Celsius)	8.63
2	Saturacija	72.1
3	BOD (BPK 5) mg/l	0.99
4	Amonijak (mg/l)	0.09
5	E. coli (br/100cm <sup>3</sup> )	
6	pH	8.01
7	Total Oxidised Nitrogen (mg/l,N)	5.03
8	Phosphate (mg/l,P)	0.02
9	SS (mg/l)	
10	Conductivity (micro S/cm)	

Слика 3. Пример упита за извештај и резултат

Примери извештаја дати су на следећим сликама:

### Kvalitet vode

Limnološka klasifikacija (OECD, Vollenweider 1968)

**Jezero ili hidroakumulacija**

**Šifra:** 23  
**Naziv:** Groznica

**Od datuma:** 1/1/2007  
**Do datuma:** 1/1/2009  
**Lokacija:**  
**Dubina:**  
**Reka:**

Parametar	Ocena	Vrednost
Ukupan P, sr. [µg/l]	Oligotrofo	5.21
Hlorofil sr. [µg/l]	Oligotrofo	2.16
Hlorofil max [µg/l]	Mezotrofo	10.44

Слика 4. Лимнолошка класификација (OECD, Vollenweider 1968) за акумулацију  
Грошница

Kvalitet vode		
IST (Carlson 1977)		
Jezero ili hidroakumulacija		
Šifra:	10	
Naziv:	Gruza	
Od datuma:	1/1/2007	
Do datuma:	1/1/2009	
Lokacija:		
Dubina:		
Reka:		
Parametar	Ocena	Vrednost
Chl, sr. [mg/m <sup>3</sup> ]	Eutrofo	7.85
TP, sr. [µg/m <sup>3</sup> ]	Eutrofo	84.38
Secchi-disk, sr. [m]	Eutrofo	1.49

Слика 5. Лимнолошка класификација (IST, Carlson 1977) за акумулацију Гружа

Поред ових својстава, ИС SeLaR омогућава знатно напредније методе анализе података, као што су статистичке анализе и DM (Tong & Chen, 2002, Conrads & Roehl, 2006b, Soranno et al., 2008). Изграђен је посебан јединствени интерфејс који омогућава аутоматску трансформацију података из базе података у структуре података коју захтевају софтвери за статистичку обраду и DM. То пружа јединствено радно окружење у коме се подаци уносе у систем само једном, када настану, после чега су расположиви за реализацију упита и анализу.

#### *Истраживање података*

У бази података све релевантне промене на ентитетима су регистроване по датуму и локацији, што обезбеђује анализу података у временској и просторној димензији. Анализа података омогућава извођење закључака и предузимање одговарајућих мера у складу са постављеним циљевима. Пре него што се приступи анализи података врши се денормализација (Han & Kamber, 2000). Денормализација се спроводи преко специјализованог корисничког интерфејса који трансформише релације из релационе базе података у структуре погодне за обраду на софтверима за анализу. За анализу података се користе специјализовани софтвери: SQL Server, Analysis Services и Excel.

Анализа података у Excel-у се спроводи на бази коришћења компоненти истраживања података у SQL Serveru 2008 (Tang & MacLennan, 2005). Подаци се из релационе базе података трансформишу у одговарајуће структуре преко специјализованог корисничког интерфејса. То значи да кориснику стоје на располагању одговарајуће структуре података које су трансформисане из базе података без икаквог његовог ангажовања. Microsoft DM нуди различите могућности за анализу података (Hart, 2008). Свака од анализа садржи следеће кораке: моделирање, реализацију модела и добијање извештаја. Моделирање обухвата одређивање својстава који улазе у модел, што зависи од циља анализе. Пре извршавања алгоритма потребно је да се провере и очисте подаци и одреде параметри алгоритма који се примењују. Реализација модела представља извршавање одговарајућег алгоритма на моделу.

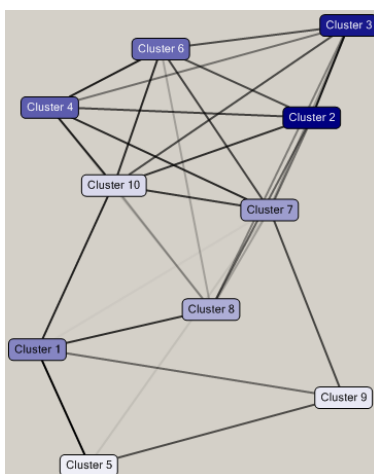
Процеси екстракције, трансформације и учитавања се реализују кроз посебан интерфејс под називом Unified Dimensional Model - UDM (Mundy et al., 2011). Изградња UDM као додатног слоја над изворима података нуди јаснији модел, изолацију од различитих платформи података и различитих формата података, као и побољшане

перформансе за комплексне упите и DM обраду. UDM омогућава креирање новог приказа извора података (data source view - DSV) који се користи од стране система. Као окружење за развој софтвера коришћен је Microsoft SQL server 2008 R2 пакет: релациона база података која покреће податке у ИС SeLaR, интеграциони сервис за трансформацију и учитавање интегрисане податке и сервис за анализу за OLAP (On-line analytical processing) и моделовање DM.

За реализацију постављених циљева користе се описне методе и методе предвиђања. Од описних метода могу се применити статистичке (средина, корелација) и вишеваријантне истраживачке технике (кластер анализа, анализа утицајних фактора). Од метода предвиђања користе се класификација, временско предвиђање и анализа сценарија. Ове методе користе различите алгоритме: за кластер се користе различити кластер алгоритми, за класификацију се користе стабло одлучивања, логистичка регресија, naïve bayes и неуронске мреже, за временско предвиђање временске серије, за кључне утицаје naïve bayes (Hart, 2008).

#### *Приказ резултата неких анализа података у информационом систему SeLaR*

Резултати кластерована могу да се прикажу на различите начине који омогућавају даљу анализу и доношења одлука. То су: приказ кластер профила, приказ кластер дијаграма (Слика 6), приказ дискриминације кластера (Табела 2, 3) и приказ карактеристика кластера (Табела 4). Прва два приказа описују све кластере заједно, а последња два приказа се односе на одређене кластере. Преглед кластер профила приказује колону за сваки кластер у моделу и ред за сваки атрибут. Ова поставка омогућава да се лако уоче занимљиве разлике кроз простор кластера, како између кластера, тако и између атрибута, као и водоравно простирање сваког интересантног атрибута кроз све кластере. Кликом на било коју ћелију у координатној мрежи добијају се детаљи о информацијама садржаним у легенди. Кластер Дијаграм даје визуелну презентацију свих кластера, где су кластери са већим бројем објеката са сивом тамнијом бојом. Такође, дебље линије између кластера представљају чвршће везе. У приказу карактеристике кластера, атрибути се приказују са вероватноћама и сортирају у опадајућем редоследу. Атрибути са највећим вероватноћама одређују карактеристике кластера и његово име. Приказ дискриминација кластер даје поређење једног кластера са осталим кластерима или заједно, или са било којим појединачним кластером (MacLennan et al., 2010).



**Слика 6.** Кластер дијаграм за укупне колиформне бактерије у акумулацији Гружа

**Табела 2.** Разлика кластера 2 и осталих кластера за укупне колиформне бактерије у акумулацији Гружа

Variables (units)	Values	Favors Cluster 2	Favors Complement of Cluster 2
TC (MPN/100ml)	< 120	100	
H (cfu/ml)	< 1540	62	
Hm (cfu/ml)	< 320	56	
FO (cfu/ml)	< 1520	49	
Chl a ( $\mu\text{g/l}$ )	< 25	46	
TC (MPN/100ml)	> 120		26
EC ( $\mu\text{S/cm}$ )	300 - 330	21	
H (cfu/ml)	> 1540		20
Hm (cfu/ml)	> 320		19
FO (cfu/ml)	> 1520		16
Fe (mg/l)	< 0.1	15	
M alkalinity (ml/l)	26- 29	10	
Chl a ( $\mu\text{g/l}$ )	> 25		7
BOD <sub>5</sub> (mg/l)	< 6.3	5	
EC ( $\mu\text{S/cm}$ )	330 - 815		3

**Табела 3.** Разлика кластера 6 и кластера 2 за укупне колиформне бактерије у акумулацији Гружа

Cluster 2		Cluster 6	
Location	Probability	Location	Probability
1 - 2	63 %	4 - 5	66 %
3	33 %	3	13 %
4 - 5	1 %	1 - 2	1 %

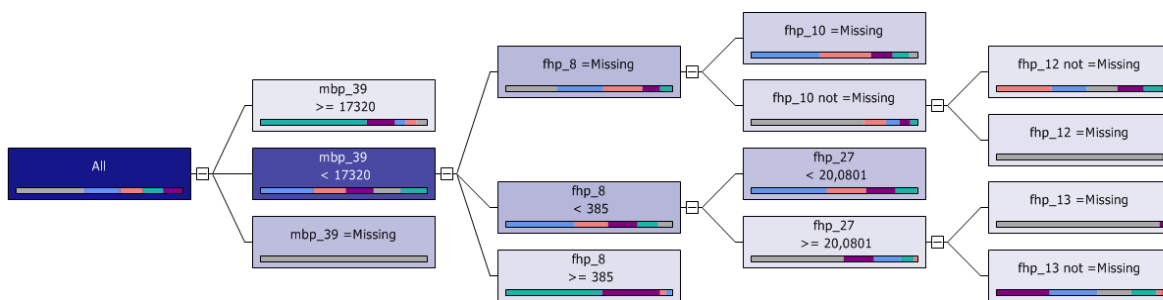
**Табела 4.** Основне карактеристике кластера 2 и кластера 6 за укупне колиформне бактерије у акумулацији

Variables (units)	Values	Favors Cluster 6	Favors Complement of Cluster 6
Depth (m)	15	100	
TC (MPN/100ml)	0 - 2340	51	
Mn (mg/l)	0.0 – 0.1	51	
Depth (m)	2 - 14		43
Chl-a ( $\mu\text{g/l}$ )	0.0 – 7.7	25	
TC (MPN/100ml)	> 2340		11
Chl-a ( $\mu\text{g/l}$ )	> 7.7		7
Mn (mg/l)	0.1 – 4.6		5
Hm (cfu/ml)	< 7795		2
Hm (cfu/ml)	> 7796	2	

Класификација и метод стабла одлучивања може бити коришћен за регресију и процену. У **Табели 5** приказан је пример за класификацију података који је основа за стабло одлучивања које је као пример приказано на **Слици 7**.

**Табела 5.** Груписање вредности укупних колиформних бактерија на основном нивоу у акумулацији Гружа употребом класификације

Value	Cases	Probability	Color
< 22	224	13.99	pink
≥ 1500	203	12.70	green
150 – 1500	195	12.20	purple
22 – 150	346	21.50	blue
missing	640	39.61	gray



**Слика 7.** Стабло одлучивања за укупне колиформне бактерије у акумулацији Гружа

*Анализа кључних утицаја* омогућава да се изабере променљива (колона, параметар) која садржи жељени исход или циљне вредности, а затим да се анализирају узорци у неком скупу података да би се утврдило који фактори имају најјачи утицај на исход. На пример, ако имамо податке укупног броја бактерија у колони са вредностима у протеклој години, може се анализирати табела за одређивање параметара који имају кључне утицаје. Постоји могућност избора пар вероватних исхода, њиховог поређења, што помаже при утврђивању могућих фактора одлучивања. Резултат анализе кључних утицаја су нове табеле података које извештавају о факторима у вези са сваким исходом и графички приказују вероватноћу односа. Табеле могу да се филтрирају од различитих фактора и исхода, па се резултати истражују на више нивоа. Алат аутоматски подешава све параметре након спроведене анализе података за одређивање оптималне поставке. Извештаји који су креирани укључују четири колоне, са следећим информацијама: колона које садрже фактор разлике, оцена вредности која је најјаче повезана са циљем, фаворизација исхода или циљна вредност коју фактор предвиђа, релативни утицај указује на снагу удружења (Han et al., 2010).

*Анализа сценарија* обједињује алате *Тражени циљ* и *Шта ако*. Ови алати омогућавају да се анализира утицај промене. Сценарио *Тражени циљ* је комплементаран сценарију *Шта ако* (где *Шта ако* каже о утицају промене, *Тражени циљ* показује основни фактор који мора да се промени да се оствари жељена промена). Пример приказа за анализу сценарија *Шта ако* дат је у **Табели 6**.

*Временско предвиђање* омогућава предвиђање будућих вредности на основу претходно посматраних вредности временске серије. Код ове анализе опционо се виде вероватноће повезане са сваком предвиђеном вредношћу. Нпр. ако колона приказује вредности параметра по датумима, могу се предвидети вредности за будуће дане. Постоји могућност да се наведе број предвиђања који треба да се направи. Креира се нови радни лист назван извештај предвиђања. Он показује да ли је предвиђање

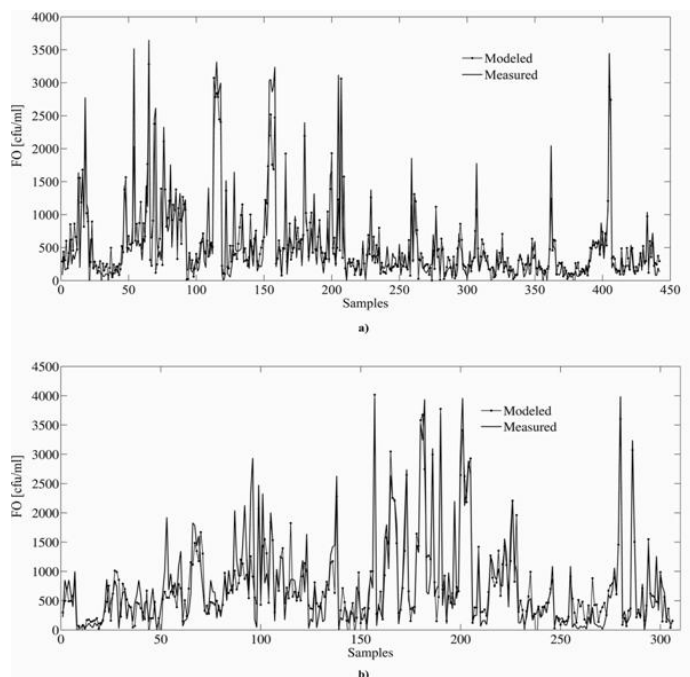
креирано успешно и садржи графикон који приказује линијске трендове. Анализе дају предвиђања која задовољавају минимални праг вероватноће (Han et al., 2010).

**Табела 6.** Пример анализе сценарија Шта ако за хетеротрофе психрофиле у акумулацији Бован

Измерене вредности за параметре*				Ако је растворен кисеоник +5% онда су хетеротрофи психрофили	Ако су нитрати -5% онда су хетеротрофи психрофили	Ако су укупни фосфати -5% онда су хетеротрофи психрофили
растворен кисеоник	нитрати	укупни фосфати	хетеротрофи психрофили			
8.7	4.3	0.071	3400	1673	6256	5990
5.1	3.8	0.034	5940	9576	5569	5326
4.5	3.8	0.053	7120	8417	5459	5221
7.2	1.3	0.01	4254	4578	9560	9476
2.3	2.9	0.045	7980	7055	10681	10501
2	2.4	0.051	8260	7764	9898	9752
6.4	1.1	0.006	8120	5905	11097	11026
0.9	1.2	0.69	14503	8644	8449	8550
0.95	1.4	0.097	12470	9301	11718	11650
7.2	0.6	0.024	8920	7573	9185	9152

\*јединице мере дате у Табели 2 и 4; Поверење 56.26

Веиштакче неуронске мреже као модели користе структуру неуронске мреже за моделирање комплексних нелинеарних односа, посебно у ситуацијама када је облик односа између укључених варијабли непознат (Smith, 1994). За предвиђање и анализу података користи се нерекурентна неуронска мрежа (feed-forward neural network, FFNN), са различитим бројем променљивих, у зависности шта се моделира. Она има интензивну употребу за предвиђање променљивих код водних ресурса (Palani et al., 2008). Из SeLaR информационог система подаци се конвертују у MATLAB Neural Network Toolbox који се користи за имплементацију неуронских мрежа. Пример резултата модела добијених на овај начин дат је на **Слици 8**.



**Слика 8.** Измерене и предвиђене вредности факултативних олиготрофа (FO) за: а- акумулацију Гружа, и б – акумулацију Грошница,

Валидност анализа и предвиђање стања предложеним алатима и моделима проверена је на примеру предвиђања концентрација раствореног кисеоника у акумулацији Гружа (Ranković et al., 2010, Ranković et al., 2012, Stefanović et al., 2012), као и за анализу различитих микробиолошких параметара (Radojević et al., 2012, Radojević et al., 2013).

Улазне варијабле за различите анализе, моделирања и предвиђања базиране су углавном на лако мерљивим параметрима који одсликавају квалитет воде. То омогућава да се подаци сакупљени у стандардним хидробиолошким истраживањима искористе за креирање модела за предвиђање. На тај начин број лабораторијских анализа из редовног праћења стања вода може бити смањен (било би омогућено смањење потрошње материјала, времена и ангажовања у лабораторијским анализама). Са оваквим моделима успешно се предвиђа стање различитих параметара, а моделе треба повремено допунити новијим мерењима.

Безбедност података решена је дефинисањем система заштите на нивоу транспорта, базе података, апликације и порука. Обезбеђење софтвера укључује оцењивање и тестирање при крају сваке фазе и подфазе развоја. Примена стандарда квалитета обезбеђује да се успостављени развој система, процеси управљања пројектом и процедуре поштују и да се могући ризици и грешке брзо идентификују и решавају.

#### **Главни ефекти које SeLaR омогућава:**

- стварање информатичке основе за архивирање расположивих лимнолошких података у електронском облику и њихову обраду у информације потребне свим корисницима. Централни део информационог система је дистрибуирана база података која садржава свеобухватне информације и омогућава нов савремен и актуелан увид у стање језера и акумулација у Србији;

-могућност бољег истраживања у области фундаменталних и примењених наука;

-омогућавање учења у овој области;

-успостављање мреже истраживача и институција;

-стварање услова за рационализацију и повећање ефикасности лимнолошких истраживања у Србији;

-омогућава савремен и актуелан увид у стање језера и акумулација у Србији чиме се ствара основа за правилно планирање и оптимизацију истраживања без непотребних понављања;

-допринос рационалном управљању и одрживој експлоатацији водених ресурса. Ово се посебно односи на успостављање адекватног мониторинга, као и на предузимање мера еколошке заштите и на примену мера за побољшање стања ових екосистема;

-карактеризацију акумулација/језера Србије у складу са Оквирном директивом Европске уније о водама (WFD) и стварање услова за повезивање са Integrated Water Quality Information System-ом.

## **5 Литература**

Bergstein B-DT, Shteinman B, Kamenir Y, Itzhak O, Hochman A (2001). Hydrodynamical Effects on Spatial Distribution of Enteric Bacteria in the Jordan River - Lake Kinneret Contact Zone. Water Research, 35(1): 311-314.

Bruhn CL, Soranno PA (2005). Long term (1974-2001) Volunteer Monitoring of Water Clarity Trends in Michigan Lakes and their Relation to Ecoregion and Land Use/Cover. *Lake and Reservoir Management*, 21(1): 10-23.

Chandramouli V, Brion G, Neelakantan TR, Lingireddy S. (2007). Backfilling missing microbial concentrations in a riverine database using artificial neural networks. *Water Research*, 41(1): 217-227.

Cheruvilil KS, Soranno PA (2008). Relationships between lake macrophyte cover and lake and landscape features. *Aquatic Botany*, 88: 219-227.

Clair T, Dennis I, Cosby BJ (2003). Probable Changes in lake chemistry in Canada's Atlantic Provinces under proposed North American emission reduction. *Hydrology and Earth System Sciences*, 7(4): 574-582 .

Conrads P, Roehl E (2006a). Transforming large databases into critical knowledge using Data mining-three case studies in south Carolina and Georgia. WEFTEC (Water environment federation, technical exhibition), 6249-6267.

Conrads P, Roehl E (2006b). Estimating Water Depths Using Artificial Neural Networks. HIC, 7th International Conference on Hydroinformatics, Nice.

Čomić Lj, Radojević I, Ćurčić S, Ostojić A (2006). Database for Monitoring, Protection and Management of Reservoirs in Serbia. *Kragujevac Journal of Sciences*, 28: 97-106.

Fulazzaky MA, Akil AH (2009). Development of Data and Information Centre System to Improve Water Resources Management in Indonesia. *Water Resources Management*, 23(6): 1055-1066.

Ganoulis J (2009). Risk Analysis of Water Pollution: Second, Revised and Expanded Edition. WILEY-VCH Verlag GmbH and Co. KGaA, Weinheim, 311 pp.

Gorunescu F (2011). Data Mining: Concepts, Models and Techniques. Intelligent Systems Reference Library, Springer - Verlag Berlin Heidelberg,.

Han J, Kamber M (2000). Data Mining: Concepts und Techniques, Morgan Kaufmann Publishers.

Han J, Kamber M, Pei J (2010). Data mining: concepts and techniques. 3rd ed., Elsevier Inc., 327-349.

Hart D (2008). Microsoft Office 2007 Business Intelligence, McGraw-Hill.

Idakwo PY, Abu GO (2004). Distribution and Statistical analysis of bacteria in Lake Alau in the arid northern Nigeria. *Journal of Applied Sciences and Environmental Management*, 8(1): 5-9.

Imberger J (2004). A lake diagnostic system for managing lakes and reservoirs. *Water resources impact*, 6(1): 7-10.

Iscen CF, Emiroglu Ö, Ilhan S, Arslan N, Yilmaz V, Ahiska S (2008). Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat lake, Turkey. *Environmental Monitoring and Assessment*, 144: 269-276.

Jones JR, Knowlton MF, Obrecht DV, Cook EA (2004). Importance of landscape variables and morphology on nutrients in Missouri reservoirs. *Canadian Journal of Fisheries and Aquatic Sciences*, 61: 1503-1512.

Jørgensen SE (2003). The application of models to find the relevance of residence time in lake and reservoir management. *Journal of Limnology*, 62(1): 16-20.

Jørgensen SE, Xu F-L, Costanza R (2005b). Handbook of Ecological Indicators for the Assessment of Ecosystem Health, Elsevier, New York, 464 pp.

Kazemi Yazdi S, Scholz M (2010). Assessing Storm Water Detention Systems Treating Road Runoff with an Artificial Neural Network Predicting Fecal Indicator Organisms. *Water Air and Soil Pollution*, 206(1-4): 35-47.



Kondrat'ev SA, Ul'yanova TYu, Moiseenkov AI, Izmailova AV, Shkrebets AE (2006). Information Support for Studying Water Resources of Lake Ladoga and Its Basin. *Water Resources*, 33(5): 494-498.

Kumar P, Alameda J, Bajcsy P, Folk M, Markus M (2006). *Hydroinformatics: Data Integrative Approaches in Computation. Analysis and Modeling*. CRC Press, 534 pp.

MacLennan J, Tang ZH, Crivat B (2010). *Data Mining with Microsoft SQL Server 2008*, Wiley Publishing, Inc, 291-318.

Martin SL, Soranno PA (2006). Lake landscape position: relationships to hydrologic connectivity and landscape features, *Limnology and Oceanography*, 51: 801-814.

May R, Dandy G, Maier H, Nixon J (2008). Application of partial mutual information variable selection to ANN forecasting of water quality in water distribution systems. *Environmental Modelling and Software*, 23 (10-11): 1289-1299.

Menshutkin VV, Filatov NN, Potakhin MS (2009a). "Karelian Lakes" Expert System: 1. Ordinal and Nominal Characteristics of Lakes, *Water Resources*, 36(2): 148-159.

Menshutkin VV, Filatov NN, Potakhin MS (2009b). "Karelian Lakes" Expert System: 2. Classification of Lakes. *Water Resources*, 36(3): 284-295.

Mitreski K, Koneski, Z, Naumov N, Davcev D, (2004). Web - Based Information System for Pollution Monitoring of Lake Ohrid. *Water, Air and Soil Pollution: Focus*, 4(4-5): 189-199.

Moe J, Dudley B, Ptacnik R (2008). REBECCA databases: experiences from compilation and analyses of monitoring data from 5,000 lakes in 20 European countries. *Aquatic Ecology*, 42(2): 183-201.

Mundy J, Thornthwaite W, Kimball R (2011). *The Microsoft Data Warehouse Toolkit: With SQL Server 2008 R2 and the Microsoft Business Intelligence Toolset*. Wiley Publishing, Inc., 29-78.

Naiburg E, Maksimchuk R (2001). *UML for Database Design*, Addison Wesley.

Neelakantan T, Brion GM, Lingireddy S (2001). Neural network modeling of *Cryptosporidium* and *Giardia* concentrations in the Delaware river. *Water Science and Technology*, 43(12): 125-132.

Ogwueleka TC, Ogwueleka FN (2010). Data mining application in predicting *Cryptosporidium* spp. oocysts and *Giardia* spp. cysts concentrations in rivers. *Journal Of Engineering Science and Technology*, 5(3): 342-349.

Palani S, Liong S, Tkalich P (2008). An ANN application for water quality forecasting. *Marine Pollution Bulletin*, 56: 1586-1597.

Parinet B, Lhote A, Legube B (2004). Principal component analysis: an appropriate tool for water quality evaluation and management - application to a tropical lake system. *Ecological Modelling*, 178: 295-311.

Park S-Y, Choi JH, Wang S, Park SS (2006). Design of a water quality monitoring network in a large river system using the genetic algorithm. *Ecological modelling*, 199: 289-297.

Radojević I (2007). Stvaranje informatičke osnove za praćenje kvaliteta akumulacionih jezera Srbije na osnovu mikrobioloških pokazatelja, Magistarska teza, Univerzitet u Kragujevcu, Prirodno matematički fakultet, 145 str.

Radojević I, Čomić Lj, Spasojević D, Ostojić A (2008b). Informatički sistem - SeLaR, monitoring, zaštita i menadžment akumulacija u Srbiji. Naučno - stručni skup o prirodnim vrednostima i zaštiti životne sredine Ekološka istina. Zbornik radova, 399-402.

Radojević I, Stefanović D, Čomić Lj, Nikolić N (2007). Database - Serbian Lakes and Reservoirs - SeLaR. *Kragujevac Journal of Sciences*, 29: 85-90.

Radojević I, Stefanović D, Čomić Lj, Ostojić A (2008a). Information system of lakes and reservoirs of the Serbia - SeLaR. *Kragujevac Journal of Sciences*, 30: 81-98.

Radojević ID, Stefanović DM, Čomić LjR, Ostojić AM, Topuzović MD, Stefanović ND (2012). Total coliforms and data mining as a tool in water quality monitoring. *African Journal of Microbiology Research*, 6(10): 2346-2356.

Radojević I, Ranković V, Comić Lj, , Ostojic A, Topuzovic M (2013). Applying Neural Networks for Predicting the Facultative Oligotrophic Bacteria in Two Reservoirs with Different Trophic State. *Journal of Environmental Protection and Ecology*, 14(1): 55-63.

Ranković V, Radulović J, Radojević I, Ostojić A, Čomić Lj (2010). Neural network modeling of dissolved oxygen in the Gruža reservoir, Srbija. *Ecological modelling*, 221: 1239-1244.

Ranković V, Radulović J, Radojević I, Ostojić A, Čomić Lj (2012). Prediction of dissolved oxygen in reservoirs using adaptive network-based fuzzy inference system. *Journal of Hydroinformatics*, 14(1): 167-179.

Roehl E, Cook J (2010). Effects of Nonpoint Source Marsh Loading on Complex Estuaries. NEIWPC, Conference on Nonpoint Source Loading, Plymouth, MA. [www.neiwpc.org/npsconference/10-presentations/Cook%20-%20NPS%20Complex%20Estuaries%20\(ppt\).pdf](http://www.neiwpc.org/npsconference/10-presentations/Cook%20-%20NPS%20Complex%20Estuaries%20(ppt).pdf)

Smith M (1994). *Neural Networks for Statistical Modelling*. Van Nostrand Reinhold, NY, 235 pp.

Soranno PA, Cheruvilil KS, Stevenson RJ, Rollins SL, Holden SW, Heaton S, Torng E (2008). A framework for developing ecosystem - specific nutrient criteria: Integrating biological thresholds with predictive modeling. *Limnology and Oceanography*, 53(2): 773-787.

Stefanović D, Radojević I, Čomić Lj, Ostojić A, Topuzović M, Kaplarević-Mališić A (2012). Management Information System of Lakes and Reservoirs. *Water Resources*, 39(4): 488-495.

Syed Ahmad SM, Turki MB, Malek S (2009). Intelligent Computational Modeling and Prediction of Coliform Growth in Tropical Lakes based on Hybrid Self Organizing Maps (SOM) and Fuzzy Logic Approaches. *eJCSIT Electron. Journal Computer Science Information Technology*, 1(1): 18-22.

Tang Z, MacLennan J (2005). *Data Mining with SQL Server 2005*. Wiley Publishing, Inc.

Tong S, Chen W (2002). Modeling the relationship between land use and surface water quality, *Journal Environmental Management*, 66(4): 377-393.

WFD CIS (2006). Common Implementation Strategy for the Water Framework Directive (2000/60/EC) Guidance Document No. 7 “Monitoring under the Water Framework”.

ИНТЕРНЕТ ИЗВОРИ:

[www.selar.pmf.kg.ac.rs](http://www.selar.pmf.kg.ac.rs)

[www.livinglakes.org](http://www.livinglakes.org)

[www.worldlakes.org](http://www.worldlakes.org)

[www.ilec.or.jp](http://www.ilec.or.jp)