Analiza zagađenja otpada i otpadnih voda

Nina, Perić

Undergraduate thesis / Završni rad

2018

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Organization and Informatics / Sveučilište u Zagrebu, Fakultet organizacije i informatike**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:211:877384

Rights / Prava: Attribution 3.0 Unported/Imenovanje 3.0

Download date / Datum preuzimanja: **2024-03-28**

SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
VARAŽDIN

Repository / Repozitorij:

Faculty of Organization and Informatics - Digital Repository

zir.nsk.hr

DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

Nina Perić

ANALIZA ZAGAĐENJA OTPADA I OTPADNIH VODA

ZAVRŠNI RAD

Varaždin, 2018.

SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

Nina Perić

Matični broj: 43989/15–R

Studij: Poslovni sustavi

ANALIZA ZAGAĐENJA OTPADA I OTPADNIH VODA

ZAVRŠNI RAD

Mentor/Mentorica:

Izv. Prof. dr. sc. Jasminka Dobša

Varaždin, kolovoz 2018.

Nina Perić

Izjava o izvornosti

Izjavljujem da je moj završni rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristila drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

Autor/Autorica potvrdio/potvrdila prihvaćanjem odredbi u sustavu FOI-radovi

_____

Sažetak

Ovaj rad bavi se statističkom analizom podataka prikupljenih analizom otpadnih voda na području primorskog grada u razdoblju od 2014. do 2017. godine. Metodama inferencijske statistike ćemo provjeriti odnose dobivenih vrijednosti prije i poslije projekta izgradnje obalnog kolektora 2015. godine i usporediti vrijednosti prije i poslije pročišćavanja na uređaju za pročišćavanje. Također ćemo usporediti prosječne vrijednosti dobivene mjerenjima s vrijednostima za otpadne vode koje propisuje Ministarstvo poljoprivrede.

Ključne riječi: otpad, otpadne vode, statistička analiza, Shapiro-Wikls test, Wicoxonov test, T-test, test predznaka

Sadržaj

1. Uvod 1
2. Metode i tehnike rada 2
   2.2. Studentov T-test 5
   2.3. Test predznaka 6
   2.4. Wilcoxonov test na osnovi ekvivalentnih parova 7
   2.5. Mann-Whitney-Wilcoxonov test za nezavisne uzorke 7
3. Opis skupa podataka 9
4. Provođenje analize i rezultati 12
   4.1. Usporedba uzoraka istih datuma 12
      4.1.1. Pročišćivač 1 12
      4.1.2. Pročišćivač 2 17
      4.1.3. Pročišćivač 3 18
   4.2. Usporedba prije i poslije obalnog kolektora 19
      4.2.1. Pročišćivač 1 19
      4.2.2. Pročišćivač 2 23
      4.2.3. Hotel 25
   4.3. Usporedba s propisanim dozvoljenim količinama 26
      4.3.1. Hotel 27
      4.3.2. Prepumpna stanica 29
      4.3.3. Stari kanal 30
5. Diskusija 31
   5.1. Analiza testova na uzorcima istih datuma 31
   5.2. Analiza testova učinkovitosti obalnog kolektora 32
   5.3. Analiza testova uspoređivanja s propisanim vrijednostima 33
6. Zaključak 35
7. Popis literature 36
8. Popis slika 38
9. Popis tablica 39
10. Prilozi 40

1. Uvod

Ovaj rad će obraditi temu zagađenja otpada i otpadnih voda na području primorskog grada. Analizom skupa podataka ćemo prepoznati uzorke i trendove koji se javljaju u periodu od 2014. do 2017. godine. Detaljno ćemo opisati skup podataka te metode analize pomoću kojih ćemo obraditi podatke.

Otpadne vode su nužan nusproizvod svake uređene urbane sredine. Život i industrija bi većini nas bili nezamislivi bez tekuće vode. Ipak, rijetko tko razmišlja što se događa s vodom nakon što ju upotrijebimo. Većina nastale otpadne vode otpušta se izravno u okoliš. To ne predstavlja problem sve dok ta voda u sebi ne sadrži tvari koje zagađuju okoliš. Kako bi se to spriječilo neki gradovi u sustav odvoda ugrađuju uređaje za pročišćavanje za otpadne vode. Na području gdje su obavljena mjerenja nema industrije koja bi značajno zagađivala otpadne vode ali živi približno 60 000 ljudi (DZS, 2011). Otpadne vode proizvode i turisti za vrijeme sezone kojih je 2017. godine bilo približno 320 000.

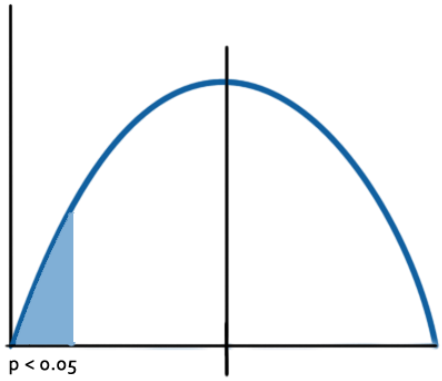
Sustav odvodnje otpadnih voda na tom području je velikim djelom bio zastario i neuređen. Neki dijelovi sustava poput starog kanala izgrađeni su u vrijeme Austro-Ugarske i na njih je samo spajana daljnja mreža koja ih je dodatno opterećivala. Na području primorskog grada je 2015. godine završen projekt izgradnje obalnog kolektora čime je saniran sustav kanalizacijske odvodnje (Hrvatske vode, 2015). S obzirom da priloženi skup podataka obuhvaća mjerenja prije i poslije puštanja u rad obalnog kolektora, postoji prilika da se provjeri učinkovitost tog višemilijunskog ulaganja.

2. Metode i tehnike rada

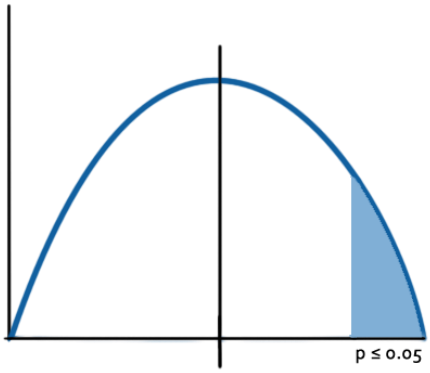
Temelj ovog rada je skup podataka i alat R pomoću kojeg je taj skup obrađen. O načinu pripreme skupa podataka će biti više riječi u sljedećem poglavlju, dok ćemo u ovom opisati metode statističke obrade koje će se koristiti. Kako bi se donijeli zaključci o podacima iz skupa koristit će se metode inferencijalne statistike. To je vrsta statistike koja se bavi provjeravanjem postavljenih hipoteza. Za nju je također karakteristično da polazi od uzorka podataka, tj. podskupa od realnog potpunog skupa podataka (Šošić, 2006, str. 3). Iz tog razloga se za nju veže pojam vjerojatnosti. S obzirom da se analiziraju samo uzorci, postoji mogućnost da ti uzorci nisu dobra reprezentacija procesa ili neke vrijednosti koju analiziramo.

Na početku svakog testiranja potrebno je postaviti nultu (H_0) i alternativnu hipotezu (H_1). One predstavljaju sudove koje ćemo ovisno o rezultatima testa prihvatiti ili odbaciti. Alternativna hipoteza je tvrdnja koju ispitivač želi potvrditi testom i suprotna je nultoj hipotezi. Kako u radu s uzorcima nepotpunog skupa ne može se biti potpuno siguran da će rezultat testa biti istinit za cijelu populaciju, postoji određena mogućnost da ćemo prihvatiti pogrešnu hipotezu. Razlikujemo pogrešku tipa I. kada se odbacuje istinita nulta hipoteza i pogrešku tipa II. kada se prihvati lažna nulta hipoteza. Vjerojatnost pogreške tipa I. označavamo sa α a vjerojatnost pogreške tipa II. sa ß (Šošić, 2006, str. 237). Iz toga izvodimo vrijednost snage testa koja predstavlja vjerojatnost odbacivanja lažne nulte hipoteze jednaka je 1 - ß. U statistici se često koristi i p-vrijednost koja predstavlja vjerojatnost da je H_0 točna. Vrijednost koja se veže uz nju je razina značajnosti i jednaka je 1 – p (Eterović, 2011).

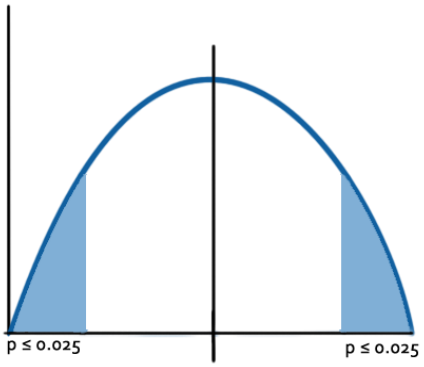
Kod postavljanja hipoteze potrebno je odrediti da li će test biti jednosmjeran da gornju granicu, jednosmjeran da donju granicu ili dvosmjeran. Jednosmjeran test može se koristiti u slučajevima kad se može sa određenom dozom sigurnosti očekivati da će rezultat biti na gornjoj ili donjoj granici, kao na primjer kad se uspoređuje koliko je temperatura na nekom području niža po zimi u odnosu na ljeti. Ako se ne može sa sigurnošću odrediti na kojoj granici se očekuje razlika u procijenjenoj aritmetičkoj sredini najsigurnije je odabrati dvosmjeran test (Bacarea & Marusteri, 2010). Jednosmjeran test može imati jednosmjernu hipotezu na „desno“ ili na „lijevo“. Jednosmjerna hipoteza na desno provjerava da li je razlika srednjih vrijednosti negativna, odnosno da li je došlo do smanjenja u odnosu na očekivanu vrijednost. Jednosmjerna hipoteza na lijevo provjerava da li je razlika srednjih vrijednosti pozitivna, odnosno da li je došlo do povećanja u odnosu na pretpostavljenu vrijednost. Dvosmjerna hipoteza se prihvaća u slučaju da je došlo do povećanja ili smanjenja u odnosu na pretpostavljenu vrijednost. Vrste hipoteza su ilustrirane na slikama na sljedećoj stranici. Područje obojano plavom bojom predstavlja područje odbacivanja H_0.



Slika 1: Jednosmjerna hipoteza na lijevu stranu (autorski rad prema Bacarea & Marusteri, 2010)



Slika 2: Jednosmjerna hipoteza na desnu stranu (autorski rad prema Bacarea & Marusteri, 2010)

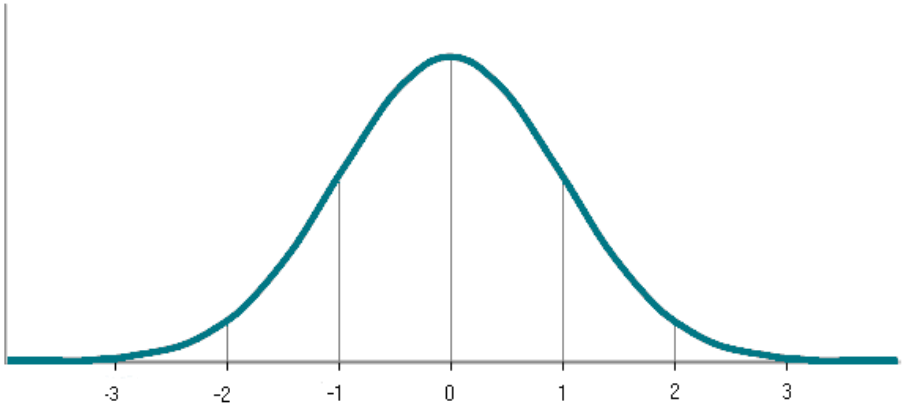


Slika 3: Dvosmjerna hipoteza (autorski rad prema Bacarea & Marusteri, 2010)

Kako bi se odabrali prikladni testovi za provođenje analize potrebno je odrediti tip podataka, broj uzoraka, zavisnost uzoraka, njihovu distribuciju i veličinu uzorka. Sve vrijednosti dobivene analizom otpadne vode su kvantitativni kontinuirani podaci osim datuma. Ovisno o vrsti hipoteze testovi na ovom skupu podataka mogu se provoditi na jednom ili dva

uzorka. Uzet ćemo da su zavisni uzorci oni koji su mjereni na istoj lokaciji istog dana prije i poslije počišćivanja, a za ostale ćemo reći da su nezavisni. Veličina uzorka određuje se tako da uzorke s manje od 30 članova smatramo manjim uzorkom, a s više od 30 članova velikim uzorkom (Šošić, 2006, str.210).

Važno je odrediti i distribuciju podataka jer o njoj ovisi da li se bira parametarski ili neparametarski test. Parametarski testovi poput t-testa i ANOVA testa podrazumijevaju normalnu ili Gaussovu distribuciju. To je distribucija podataka koja odražava činjenicu da kad se stvari mijenjaju nastoje ostati bliže prosječnoj vrijednosti a rjeđe radikalno od nje odstupaju.



Slika 4: Gaussova distribucija (autorski rad)

Kao što je vidljivo na slici 4. Gaussova distribucija je krivulja u obliku zvona na kojoj je većina podataka koncentrirana na srediti, tj. većina podataka je približne vrijednosti kao i artimetička sredina. Ako distribucija podataka ne odgovara Gaussovoj krivulji koriste se primjereni neparametarski testovi. Kako bi provjerili normalnost distribucije koriste se brojni testovi poput D'Agostino-Pearsonov, Kolmogorov-Smirnov, Jarque-Beratestov, Anderson-Darlingov, Cramér-von-Misesovt i Lillieforsov test od kojih je nama najprimjereniji Shapiro-Wilksov test (Bacarea & Marusteri, 2010). Shapiro-Wilksov test je primjereniji za testiranje manjih uzoraka s kojima ćemo se susretati a i prema simulacijama ima najjaču snagu testa (Razali & Yap, 2011).

2.1. Shapiro-Wilksov test

Shapiro-Wilksov test se koristi za provjeru normalnosti distribucije skupa podataka. Koristi ga se na skupovima od 3 do 50 elemenata (Zaionts, 2013). Prvi korak kod provođenja testa je sortirati elemente skupa od najmanjeg prema najvećem. Zatim računamo vrijednost SS prema formuli

$$SS = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Gdje je x element iz skupa a i njegov redni broj u sortiranom skupu. Vrijednost $\bar{x}$ je aritmetička sredina skupa. Zatim računamo vrijednost *b* prema formuli

$$b = \sum_{i=1}^{n} a_i (x_{n+1-i} - x_i)$$

Vrijednosti a uzimamo iz Shapiro-Wilks tablica a vrijednost n predstavlja broj elemenata u skupu (Zaiontz, 2013.). Zatim iz dobivenih vrijednosti računamo W koji dobijemo pomoću formule

$$W = \frac{b^2}{SS}$$

Krajnji rezultat testa je p-vrijednost koja nam govori s kojom vjerojatnošću možemo tvrditi da testirani skup podataka ima normalnu distribuciju. Pomoću izračunate vrijednosti W možemo dobiti p-vrijednost preko Shapiro-Wilks tablica (Zaiontz, 2013.). Ako smo odabrali razinu značajnosti od 95% i dobili vrijednost $p < 0.05$ smatramo da postoji statistički značajna vjerojatnost da skup nema normalnu distribuciju. Ovaj test ćemo provoditi prije svakog drugog testa kako bi smo mogli odabrati parametarske ili neparametarske testove. Za provođenje testa koristit ćemo alat R i Rstudio.

2.2. Studentov T-test

Studentov T-test koristi se za testiranje hipoteze o aritmetičkoj sredini jednog ili dva uzorka. Ako testiramo samo jedan uzorak njegovu aritmetičku sredinu uspoređujemo s očekivanom vrijednošću. T-test je parametarski test te se stoga primjenjuje samo na uzorcima koji imaju Gaussovu distribuciju. Test je osmišljen kako bi omogućio testiranje manjih uzoraka do 30 članova s nepoznatim varijancama. Za veće uzorke može se koristiti z-test. T-test se provodi tako da se računa t-omjer. On je slučajna veličina koja pripada Studentovoj distribuciji s određenim stupnjem slobode (Šošić, 2006, str. 241). Njega tumačimo kao omjer razlika između skupova i razlika unutar skupova (Statistics How To, 2018). T-test na jednom uzorku se računa prema formuli

$$t = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

gdje su $\bar{x}$ aritmetička sredina uzorka, μ vrijednost pretpostavljene aritmetičke sredine osnovnog skupa i $\sigma_{\bar{x}}$ standardna devijacija sampling distribucije aritmetičkih sredina (Šošić, 2006, str.238). Dobiveni t-omjer usporedimo s vrijednošću iz t-tablice koja sadrži teorijske vrijednosti Studentove distribucije. Vrijednost iz tablice odabiremo na temelju razine značajnosti i stupnja slobode koji iznosi (n-1).

Kod provođenja T-testa potrebno je uzeti u obzir zavisnost varijabli. Za zavisne uzorke pretpostavka testa je da je distribucija razlika parova normalna. T-omjer za zavisne uzorke se računa po formuli

$$t = \frac{\frac{\sum_{i=1}^{n}(x_i - y_i)}{n}}{\sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2 - \left(\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}\right)}{(n-1)\cdot n}}}$$

gdje je n broj parova a $\sum_{i=1}^{n}(x_i - y_i)$ suma razlika između parova (Statistics How To, 2018). Brojnik ove jednadžbe je procjena razlike sredina osnovnih skupova brojem, a nazivnik je standardna pogreška razlike (Šošić, 2006, str.300). Stupanj slobode se računa na isti način kao i kod testa na jednom uzorku a pomoću razine značajnosti potraži se odgovarajuća vrijednost u t-tablici.

Za testiranje nezavisnih uzoraka prvo se izračuna vrijednost sampling distribucije D koja je jednaka razlici aritmetičke sredine prvog i drugog skupa. Potrebno je uzeti u obzir i varijance skupova. Ako su varijance skupova poznate i jednake standarda pogreška razlike se računa pomoću formule

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sigma\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

Gdje je n_1 broj članova prvog skupa, n_2 broj članova drugog skupa i σ poznata varijanca koja je jednaka za oba skupa. Ako varijance skupova nisu poznate standarda pogreška razlike se računa pomoću formule

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sum_{i=1}^{n}(x_{i1} - \frac{\sum_{i=1}^{n_1} x_{i1}}{n_1})^2}{n_1} - \frac{\sum_{i=1}^{n}(x_{i2} - \frac{\sum_{i=1}^{n_1} x_{i2}}{n_2})^2}{n_2}}$$

gdje je su s x_{i1} označeni članovi prvog skupa a x_{i2} članovi drugog skupa (Šošić, 2006, st. 282). Dalje se test provodi kao i u ostalim slučajevima. U slučaju kad varijance skupova nisu jednake koristi se Welchov t-test. U tom testu se t-omjer računa prema formuli

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

gdje je $\bar{x}$ aritmetička sredina skupa i s^2 varijanca skupa (Statistics How To, 2015).

2.3. Test predznaka

Test predznaka (*eng. Sign Test*) testira hipotezu o medijanu osnovnog skupa. Nema pretpostavku o Gaussovoj distribuciji podataka pa se može koristiti kao alternativa T-testu za

jedan uzorak. Uzorak se uspoređuje s pretpostavljenom vrijednosti medijana koju označavamo s η_0. Nulta hipoteza H_0 je da je medijan skupa jednak pretpostavljenom medijanu. Kod provođenja testa uzimaju se u obzir razlike vrijednosti pojedinih članova skupa i pretpostavljenog medijana. Tim razlikama se zatim pridružuju predznaci + ako su veće od 0 ili – ako su mane od 0. Članove čija je razlika s pretpostavljenom vrijednosti medijana jednaka 0 se izostavlja iz ispitivanja. Test-veličina je broj pozitivnih predznaka c_+ ili broj negativnih predznaka c_- ovisno o hipotezi. Ako je hipoteza dvosmjerna i ispituje se razlika od pretpostavljenog medijana u pozitivnom ili negativnom smjeru, nulta hipoteza se odbacuje ako je vjerojatnost pojavljivanja c ili manje predznaka u uzorku od n članova manja ili jednaka odabranoj razini značajnosti. Kod dvosmjernog testa za c vrijednost se uzima manja od c_+ i c_-. Kod jednosmjernog testa na desnu stranu je test veličina c_- a kod jednosmjernog testa na lijevu stranu je c_+ (Šošić, 2006, str.332).

2.4. Wilcoxonov test na osnovi ekvivalentnih parova

Wilcoxonov test na osnovi ekvivalentnih parova (*eng. Wilcoxon Matched-pairs signed Rank Test*) je neparametarski test koji se provodi na zavisnim uzorcima. Koristi se za testiranje hipoteze o jednakosti medijana dva simetrično raspoređena skupa (Šošić, 2006, str. 340). Ako distribucija razlika nije simetrična test je manje pouzdan. Hipoteza H_0 je da ne postoji statistički značajna razlika između medijana osnovnih skupoova, a H_1 tvrdi da postoji takva razlika.

Test započinjemo tako da utvrdimo razlike vrijednosti koje su u paru. Ako je par (x_i, y_i) razlika vrijednosti para je (y_i - x_i). Ako dobijemo razliku para koja je jednaka nuli, taj par isključujemo iz daljnjeg testiranja. Zatim apsolutne razlike $|y_i – x_i|$ rangiramo od najmanje do najveće i svakoj pridružimo redni broj počevši od 1 do n gdje je n broj parova. Tom rangu zatim pridružujemo predznake + ili – ovisno o tome da li je razlika (y_i - x_i) pozitivna ili negativna. Nakon toga određujemo test-veličinu tako da zbrajamo rangove s pozitivnim predznakom T_+ ili rangove s negativnim predznakom T_-. Test-vrijednosti koristimo ovisno o tome da li je test jednosmjeran ili dvosmjeran. Ako je test dvosmjeran test-veličina koju uzimamo je ona T koja je najmanja. Ako je test jednosmjeran na desnu stranu test-veličina je T_-. Ako je test jednosmjeran na lijevu stranu test-veličina je T_+ (Šošić, 2006, str.340).

2.5. Mann-Whitney-Wilcoxonov test za nezavisne uzorke

Mann-Whitney-Wilcoxonov test (*eng. Wilcoxon rank-sum test*) je neparametarski test za testiranje hipoteze o jednakosti oblika dviju distribucija dvaju nezavisnih uzoraka. Također se može koristit za testiranje hipoteze da su medijani osnovnih skupova jednaki. Ako se

uspoređuju aritmetičke sredine skupova test se može primijeniti ako su distribucije skupova simetrično raspoređene (Šošić, 2006, str.344). U praksi se Mann-Whitney-Wilcoxonov test koristi kad skupovi podataka ne zadovoljavaju uvjete za parametarski T-test.

Test se provodi na dva uzorka od kojih se onom s manje članova veličina označava s n_1 a onom s više članova s n_2. Računa se test-veličina T_1 tako da se spoje svi članovi iz oba skupa u novi niz. Članovi tog niza se zatim poredaju po veličini od najmanjeg do najvećeg i pridruže im se rangovi od 1 do n_1+n_2. Test-veličinu zatim dobijemo tako da zbrojimo rangove samo članova koje smo dobili iz prvog niza (Šošić, 2006). Dobivenu test-veličinu uspoređujemo s kritičnim vrijednostima i na temelju toga donosimo odluku. U slučaju da oba skupa imaju više od 10 članova, može se koristiti z-test. U tom slučaju test-veličina z se računa pomoću formule:

$$z = \frac{T_1 - \frac{n_1(n_1 + n_2 + 1)}{n_1}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

Odluka o testu se tada donosi usporedbom empirijskog z-omjera s teorijskim vrijednosti za z jedinične normalne distribucije (Šošić, 2006, str.345).

3. Opis skupa podataka

Skup podataka koji je analiziran u ovom radu sastoji se od 287 LibreOffice Calc datoteka. Svaka datoteka predstavlja mjerenje otpadnih voda na pojedinoj lokaciji određenog datuma. Lokacije na kojima su provedena mjerenja su nazvane Pročišćivač 1, Pročišćivač 2, Pročišćivač 3, Hotel, Prepumpna stanica i Stari kanal. Od tih lokacija su Pročišćivač 1, 2 i 3 stanice za pročišćavanje otpadnih voda i na njima je mjerena voda na ulazu i izlazu, tj. prije i poslije pročišćivanja. Mjerenja obuhvaćaju period od 2014. do 2017. godine. Mjerenja nisu slijedila pravila pa s toga za neke godine i lokacije imamo više mjerenja a za neke manje. U tablici ispod prikazan je broj i vrsta mjerenja po lokaciji i godini.

Tablica 1: Broj mjerenja po lokaciji, vrsti i godini

	Pročišćivač 1		Pročišćivač 2		Pročišćivač 3		Hotel	Prepumpna stanica	Stari kanal
	Ulaz	Izlaz	Ulaz	Izlaz	Ulaz	Izlaz	/	/	/
2014.	4	12	4	12	2	2	4	0	0
2015.	27	11	22	9	1	1	3	4	3
2016.	4	11	63	11	1	1	4	0	4
2017.	4	12	4	12	1	1	0	0	4

Kao što vidimo mjerenja nisu uvijek pratila stroga pravila, no primjećujemo da su za lokacije Pročišćivač 1 i 2 na ulazu mjerena u razmaku od 3 mjeseca osim u iznimnim slučajevima kad su mjerenja rađena svaki dan. Na izlazu istih lokacija su mjerenja rađena jednom mjesečno jer se time nadzirala ispravnost uređaja za pročišćavanje. Ostale lokacije su mjerene svaka 3 do 4 mjeseca osim lokacije Hotel za koju ne postoje mjerenja 2017. godine i Stari kanal za koji ne postoje mjerenja 2014. godine. Na lokaciji prepumpne stanice postoje mjerenja samo za 2015. godinu.

Podaci za svako mjerenje nalaze se u posebnoj LibreOffice Calc datoteci koja sadrži brojne informacije o provođenju analize, odgovornim osobama, porijeklu i količini uzorka, stanja vodopravnih dozvola i slično no za potrebe testiranja zanimljivi su samo fizikalno-kemijski parametri. Ispitivanja također nisu pratila stroga pravila o tome koji parametri će se testirati pa stoga imamo mnoge vrijednosti koje nedostaju. Parametri koje se testiralo zajedno sa mjernim jedinicama u zagradama su: pH-vrijednost (pH), temperatura vode (ºC),

ukupna suspendirana tvar (mg/l), boja, miris, taložive tvari (ml/l), električna vodljivost (µS/cm), biokemijska potrošnja kisika nakon 5 dana ili BPK_5 (mg O_2/l), kemijska potrošnja kisika – dikromatom ili KPKCr (mg O_2/l), amonijak (mg N/l), nitriti (mg N/l), nitrati (mg N/l), kloridi (mg/l), ukupni fosfor (mg/l), ortofosfati (mg/ P/l), ukupni dušik (mg/l), anionski detergenti (mg/l), ukupna ulja i masti (mg/l), ugljikovodici (mg/l), protok (l/s), cink (mg/l) i olovo (mg/l). Neke parametre nećemo uzimati u obzir jer nam nisu od interesa a neki su izuzetno rijetko mjereni pa s toga nije moguće povući valjan zaključak iz njih.

Imati podatke u 287 zasebnih datoteka nije praktično pa je s toga bilo potrebno okupiti sve podatke u jednoj datoteci. To je bio izazov jer je svaka datoteka bila drugačija i stoga nije bilo isplativo napraviti macro radnju u LibreOfficeu koja bi kopirala podatke u jednu datoteku. Datoteke također imaju specifičan format zbog kojeg ga nije bilo moguće jednostavno konvertirati u Excel datoteku. Taj problem je riješen pomoću metode gdje je korištena prazna LibreOffice Calc datoteka u koju je pomoću ćelija koje su korištene kao varijable da bi se kreiralo reference na pojedinačne datoteke bilo moguće dobiti podatke iz više datoteka odjednom. Na slici 5 vidimo primjenu te metode na primjeru dobivanja podataka za lokaciju Pročišćivač 2 za mjerenja 2016. godine. Princip te metode je bio isti za svaku lokaciju i godinu mjerenja.

2016 Procisciv ac 2 IZLAZ					
FOLDER 1	file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac				
FOLDER 2	file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/2016 2 dio/Procisciv ac				
EKSTENZIJA	.ods				
			#$Single.$AO23	$Single.$AQ3	#$Single.$AQ39
IME DOKUMENTA		REFERENCE	Datum	pH-v rijednost	Temperatura v ode
227 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/227 P	21/01/16	7.16	6.2
290 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/290 P	02/11/16	7.99	6.3
419 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/419 P	16/03/16	8.07	9.2
677 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/677 P	15/04/16	8.08	12.2
1267 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1267	16/06/16	7.66	19.1
1368 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1368	07/05/16	7.64	19.4
1520 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1520	08/07/16	7.71	15
1683 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/2016 2 dio/Procisciv ac2/1683 P	14/09/16	7.75	16.1
1815 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/2016 2 dio/Procisciv ac2/1815 P	10/12/16	7.78	13.1
2170 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/2016 2 dio/Procisciv ac2/2170 P	11/11/16	7.58	11.4
2287 Procisciv ac2-izlaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/2016 2 dio/Procisciv ac2/2287 P	12/09/16	7.95	10.8
2016 Procisciv ac 2 ULAZ					
			#$Single.$AO23	$Single.$AQ3	#$Single.$AQ39
IME DOKUMENTA		REFERENCE	Datum	pH-v rijednost	Ukupna suspendira
1366 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1366	07/05/16	7.71	2032
1367 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1367	07/06/16	7.55	1132
1381 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1381	07/07/16	7.7	295.2
1382 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1382	07/08/16	7.68	644.4
1390 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1390	07/09/16	7.34	114.2
1391 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1391	07/10/16	7.38	109.2
1392 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1392	07/11/16	7.7	238.6
1401 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1401	07/12/16	7.53	144.4
1413 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1413	13/07/16	7.66	124
1414 Procisciv ac2-ULAZ		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1414	14/07/16	7.55	108.4
1419 Procisciv ac2-ulaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1419	15/07/16	7.7	108.4
1429 Procisciv ac2-ulaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1429	16/07/16	7.62	108
1430 Procisciv ac2-ulaz		'file:///D:/Users/Nina/Desktop/Nina Perić/otp v ode 2016/ 2016 1 dio/Procisciv ac2/1430	17/07/16	7.78	166.8

Slika 5: Primjer postupka dobivanja podataka iz više LibreOffice Calc datoteka

Na prvi pogled možda djeluje kompliciranо ali na taj način je postupak kopiranja podataka iz više datoteka na uređen način uvelike ubrzan i olakšan. Metoda je funkcionirala na način da su u prvom redu ćelije u koje se kopira adresa datoteke u kojoj se nalazie podaci za pojedinu lokaciju jedne od godina te ekstenziju LibreOffice Calc datoteke. Te ćelije se dalje tretiraju kao konstante. Za svaku novu datoteku u kojoj su bile spremljene LibreOffice Calc datoteke samo se promijeni ta jedna ćelija. Na primjeru na slici su podaci spremljeni u dvije datoteke pa stoga postoje dvije varijable za lokaciju. Nakon toga napravljen je stupac u koji se unosi ime dokumenta u kojem se nalaze podaci. Potrebno je navesti točna imena svih datoteka iz kojih se dobivaju podaci. Na temelju tih podataka genera se referenca. Za svaki parametar se na desnoj strani napravi stupac. Iznad imena stupca nalazi se ćelija u kojoj se odredi na kojem listu i u kojoj ćeliji se nalazi podatak za taj parametar u LibreOffice Calc datoteci. Na kraju se u ćeliji u kojoj se žele prikazati podaci napiše formula koja je spajala gotovu referencu sa lokacijom na listu. Za mjerenja koja su mjerila različite parametre potrebno je prilagođavati stupce u kojima se prikazuju podaci i brojeve ćelija iz kojih se čitaju podaci.

Nakon što se gore opisanom metodom dobiju prikazi podataka parametara za svaku lokaciju po godini potrebno je objediniti ih u jedan dokument. Za to se kreira nova Excel datoteku koja ima novi list za svaku od lokacija. Podaci su na svakom listu odvojeni u zasebne tablice ovisno da li je mjerenje obavljeno na ulazu ili izlazu gdje je to primjenjivo. U tablicama stupci označavaju imena parametara a redovi su kronološki poredani prema datumima mjerenja. Potrebno je samo kopirati podatke is LibreOffice Calc datoteke, transponirati ih i staviti u odgovarajuću tablicu u Excel datoteci. Daljnje obrade podataka će se odraditi pomoću alata R. Za to je najpogodniji .csv format. Datoteke u tom formatu lako se dobiju pomoću export funkcije u Excelu. Na taj način može se pripremiti potrebne dijelove podataka za svaki test.

4. Provođenje analize i rezultati

Za analizu skupa podataka predviđeno je ukupno 9 skupina testiranja. Prije svakog od tih testiranja provodit će se test normalnosti kako bi se odredilo koji test se može provesti za testiranje hipoteze o aritmetičkoj sredini skupa podataka. Uspoređivat će se vrijednosti ukupne suspendirane tvari, BPK_5, KPKCr, amonijaka, klorida, fosfora, anionskih detergenata, masti i ulja te cinka i olova ovisno o dostupnosti podataka za pojedine lokacije. Testirat će se razlika aritmetičkih sredina uparenih uzoraka koji su uzeti na iste datume kod ulaza i izlaza iz stroja za pročišćivanje zasebno na lokacijama Pročišćivač 1, 2 i 3. Nakon toga će se analizirati sve uzorke koji su uzeti prije i poslije puštanja obalnog kolektora u rad. To će se napraviti zasebno za lokacije Pročišćivač 1 i 3 jer samo za njih imamo dovoljno podataka prije i poslije tog događaja. Za lokacije Hotel, Prepumpna stanica i Stari kanal će se napraviti zasebno jednosmjerne testove na jednom uzorku i uspoređivati ih s propisanim dozvoljenim količinama tvari za otpadne vode.

4.1. Usporedba uzoraka istih datuma

U skupu podataka postoje tri lokacije na kojima su obavljene analize na ulazu i izlazu iz istog uređaja za pročišćavanje istog datuma. Provođenjem prikladnih testova ćemo analizirati sve dostupne parametre prije i poslije pročišćivanja otpadne vode kako bi smo mogli provjeriti učinkovitost uređaja za pročišćavanje. To ćemo zaključiti na temelju usporedbe aritmetičkih sredina uzoraka otpadne vode prije i poslije pročišćavanja. Uzorci koji su uzeti na istoj lokaciji istog dana su zavisni pa ćemo to uzeti u obzir kod odabira prikladnog testa. Prva lokacija na kojoj imamo uzorke istih datuma je Pročišćivač 1 a nakon toga ćemo analizirati podatke s lokacija Pročišćivač 2 i 3.

4.1.1. Pročišćivač 1

Kao što je karakteristično za naš skup podataka, veći broj parametara je analiziran na izlazu nego na ulazu u stroj za pročišćavanje. Na lokaciji Pročišćivač 1 ukupno je 11 puta provedeno testiranje ulaza i izlaza istog datuma. S obzirom da se bavimo uspoređivanjem, analizirat ćemo samo one parametre koji su mjereni u oba slučaja. To su pH-vrijednost, temperatura vode, ukupna suspendirana tvar, BPK_5 i ukupna ulja i masti. Započet ćemo s analizom pH-vrijednosti. Potrebno je provjeriti normalnost distribucije skupa podataka. To ćemo napraviti Shapiro-Wilksovim testom koristeći alat R. Kod provođenja ovog i svakog sljedećeg Shapiro-Wilks testa odabrat ćemo razinu značajnosti 95% a hipoteze će biti sljedeće:

Tablica 2: Hipoteze kod Shapiro-Wilks testa

Hipoteza	Uvjet prihvaćanja
H_0 = ne postoji značajno odstupanje od normalne distribucije	$p > 0.05$
H_1 = postoji značajno odstupanje od normalne distribucije	$p \leq 0.05$

Tablica 2. Hipoteze kod provođena Shapiro-Wilks testa

U slučaju kad je broj parova manji od 30 s obzirom na to da su uzorci zavisni testira se normalnost distribucije razlike vrijednosti (y_i - x_i) a ne normalnost svakog od uzoraka. Razlike možemo lako izračunati pomoću Excela a rezultat izvezemo kao .csv datoteku kako bi podaci bili prikladni za rad s alatom R. Analizu u alatu provodimo tako da u konzolu u Rstudiu unesemo sljedeće linije koda:

```
prociscivac1_pH <- read.csv(choose.files())
shapiro.test(prociscivac1_pH$razlika)
```

U prvoj liniji koda prociscivac1_pH je ime varijable u koju spremamo našu .csv datoteku. Funkcija read.csv() se koristi za čitanje datoteka a argument choose.files() govori programu da ćemo odabrati .csv datoteku pomoću skočnog prozora Windows Explorera. Druga linija koda je naredba za izvođenje Shapiro-Wilksovog testa. Funkciji se prosljeđuje varijabla koja sadrži podatke. U našem slučaju postoje tri stupca u .csv datoteci pa je bilo potrebno naznačiti da se test provodi nad stupcem koji se zove razlika. Nakon provedenog testa dobili smo sljedeći ispis:

```
data: prociscivac1_pH$razlika
W = 0.97793, p-value = 0.9536
```

Vidimo da je rezultat testa p-vrijednost = 0.9536 što je značajno veće od 0.05. Iz tog razloga prihvaćamo hipotezu H_0 da ne postoji značajno odstupanje od normalne distribucije.

S obzirom na rezultat prethodnog testa odabrat ćemo parametarski test. Prikladan test za zavisne uzorke je t-test. Želimo provjeriti da li je Pročišćivač utjecao na pH vode na bilo koji način stoga je naš test dvosmjeran. H_0 hipoteza je da aritmetičke sredine oba nisu statistički značajno različite, a H_1 tvrdi suprotno. Test provodimo pomoću sljedeće linije koda:

```
t.test(prociscivac1_pH$ph_vrijednost_ulaz,
prociscivac1_pH$ph_vrijednost_izlaz, alternative="two.sided",
paired=TRUE, conf.level=0.95)
```

Funkcijom t.test pokrećemo test i prosljeđujemo joj argumente unutar zagrada. Prve dvije varijable sadržavaju podatke o naša dva uzorka, alternative="two.sided" određuje da je test dvosmjeran, paired je bool varijabla koja kad je postavljena na TRUE označava da se radi o zavisnim uzorcima a conf.level određuje razinu značajnosti kao 95%. Rezultat ovog testa koji vraća program je:

```
Paired t-test
```

```
data:  prociscivac1_pH$ph_vrijednost_ulaz and
prociscivac1_pH$ph_vrijednost_izlaz
t = -1.7025, df = 10, p-value = 0.1195
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.19939032 0.02666304
sample estimates:
mean of the differences
               -0.086364
```

Iz trećeg reda rezultata koje nam je program ispisao možemo vidjeti da je t-omjer = -1.7025, stupanj slobode df = 10 i p = 0.1195 što je veće od 0.05. Na temelju toga prihvaćamo hipotezu H_0 a to je da aritmetičke sredine oba skupa nisu statistički značajno različite. To ne znači da su potpuno jednake, nego da ne možemo sa dovoljnom razinom sigurnosti tvrditi da su značajno različite. Dalje vidimo da je aritmetička sredina razlika -0.086364. To znači da je pH vrijednost izmjerena na izlazu iz stroja za pročišćavanje prosječno za 0.086364 pH manja od one izmjerene na ulazu u stroj. Kakav značaj takva razlika ima u praktičnom smislu ćemo raspraviti u poglavlju diskusije.

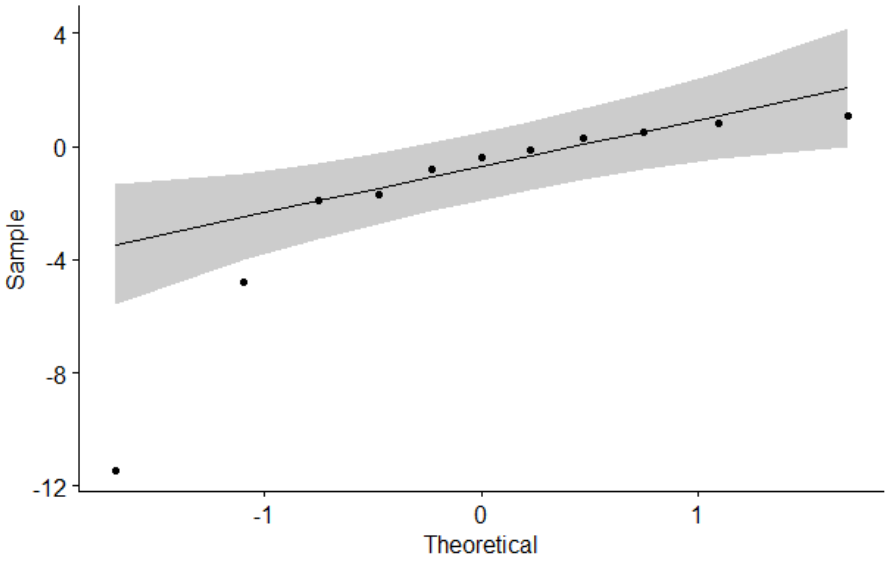
Sljedeća vrijednost koju ćemo analizirati na ovom uzorku je temperatura vode. Test započinjemo na isti način kao i u prethodnom primjeru. Testiramo normalnost distribucije razlika vrijednosti temperature na ulazu i izlazu iz Pročišćivača pomoću Shapiro-Wilks testa. Funkcije koje koristimo u programu za testiranje normalnosti gotovo su iste pa ih nećemo dalje posebno opisivati. Rezultat testa je:

```
data: prociscivac1_temp$razlika
W = 0.7187, p-value = 0.0008151
```

Vidimo da je p-vrijednost izuzetno mala pa stoga prihvaćamo H_1 hipotezu da postoji značajno odstupanje od normalne distribucije. Kako bi bolje razumjeli distribuciju napraviti ćemo Q-Q dijagram pomoću kojeg možemo vizualno prikazati distribuciju razlika. Alat R može generirati takav dijagram koristeći pakete ggplot2 i magtrittr. Naredba kojom alat generira željeni dijagram je:

```
ggqqplot(prociscivac1_temp$razlika)
```

Dijagram koji smo dobili prikazan je na slici 6. Vidimo da većina podataka ima normalnu distribuciju i nalazi se blizu središnjeg pravca, no imamo vrijednosti koje odudaraju.



Slika 6: Q-Q dijagram razlika temperatura mjerenih na lokaciji Pročišćivač 1 istih datuma

S obzirom da distribucija podataka nije normalna koristit ćemo primjereni neparametarski test. To je Wilcoxonov test na osnovi ekvivalentnih parova. Hipoteza H_0 je da ne postoji statistički značajna razlika aritmetičkih sredina vrijednosti temperatura između dva uzorka. Hipoteza H_1 je da takva razlika postoji. U alatu R navedeni test pokrećemo pomoću funkcije:

```
wilcox.test(prociscivac1_temp$ulaz_temp,
prociscivac1_temp$izlaz_temp, paired=TRUE)
```

Funkciji smo proslijedili varijable koje sadrže podatke o temperaturi prije i poslije pročišćavanja. Bool paired smo postavili na TRUE jer su naši uzorci zavisni. Rezultati koje program vraća su sljedeći:

```
Wilcoxon signed rank test

data:  prociscivac1_temp$ulaz_temp and prociscivac1_temp$izlaz_temp
V = 48, p-value = 0.2061
alternative hypothesis: true location shift is not equal to 0
```

Iz rezultata vidimo da je p-vrijednost = 0.2061 što je veće od 0.05. Prihvaćamo hipotezu H_0 da ne postoji statistički značajna razlika između medijana temperatura prije i poslije pročišćavanja. Aritmetička sredina razlika tih temperatura je -1.67545 ºC.

Analizu ukupne suspendirane tvari također ćemo započeti Shapiro-Wilks testom za normalnost distribucije. Rezultat testa je:

```
data:  prociscivac1_susp$razlika
W = 0.7159, p-value = 0.0007498
```

Vidimo da je p-vrijednost od 0.0007 značajno manja od 0.05 pa prihvaćamo hipotezu H_1 da distribucija podataka značajno odstupa od normalne distribucije. Provodimo prikladni neparametarski test, tj. Wilcoxonov test na osnovi ekvivalentnih parova. Koristimo istu funkciju kao i prošli put samo na novim varijablama. Rezultat testa je:

```
Wilcoxon signed rank test

data: prociscivac1_susp$ulaz and prociscivac1_susp$izlaz
V = 30, p-value = 0.8311
alternative hypothesis: true location shift is not equal to 0
```

P-vrijednost od 0.8311 je veća od 0.05 za razinu značajnosti od 95% pa prihvaćamo hipotezu H_0 da ne možemo zaključiti da postoji statistički značajna razlika između medijana ta dva skupa. Sljedeći parametar je BPK_5. Analizom Shapiro-Wilks testom dobiveni su sljedeći rezultati:

```
data:  prociscivac1_bpk$razlika
W = 0.9506, p-value = 0.6517
```

S obzirom da je p-vrijednost veća od 0.05 možemo prihvatiti hipotezu da podaci imaju normalnu distribuciju, pa ćemo odabrati t-test za zavisne uzorke. Rezultat tog testa je:

```
data:  prociscivac1_pH$ph_vrijednost_ulaz and
prociscivac1_pH$ph_vrijednost_izlaz
t = -1.7025, df = 10, p-value = 0.1195
mean of the differences
               -82.81818
```

P-vrijednost iznosi 0.1195 što je veće od 0.05. Prihvatiti hipotezu H_0 da ne postoji statistički značaja razlika između medijana ta dva skupa. Aritmetička sredina razlike prije i poslije pročišćavanja je -82.8182 mg O_2/l.

Zadnji parametar za lokaciju Pročišćivač 1 je ukupna ulja i masti. Tu se prvi put susrećemo s neodređenim vrijednostima. Metoda kojom je provedena analiza ne može detektirati prisutnost ulja i masti manju od 5 mg/l. Postoji mogućnost da odbacimo parove u kojima se javljaju takve vrijednosti ali u tom slučaju zanemarili bi gotovo polovicu parova i ne bi smo imali uzorak koji bi vjerno predstavljao stanje. Analizom takvog uzorka bi smo vjerojatno precijenili stvarnu količinu ulja i masti u otpadnoj vodi. Iz tog razloga ćemo sve takve vrijednosti zamijeniti sa 2.5 mg/l. To je vrijednost koja je na sredini intervala potencijalnih stvarnih vrijednosti. Dalje provodimo Shapiro-Wilks test za normalnost distribucije. Rezultati testa su:

```
data:  prociscivac1_masti$razlika
W = 0.38176, p-value = 6.002e-08
```

Iz rezultata vidimo da je p-vrijednost = 6.002×10^{-8} što je značajno manje od 0.05. Vidimo da podaci nemaju normalnu distribuciju. Provodimo Wilcoxonov test za ekvivalente parove. Rezultati tog testa su:

```
Wilcoxon signed rank test

data:  prociscivac1_masti$ulaz and prociscivac1_masti$izlaz
V = 52, p-value = 0.1016
alternative hypothesis: true location shift is not equal to 0
```

Iz rezultata vidimo da p-vrijednost od 0.1016 je veća od 0.05. Prihvaćamo hipotezu H_0 da ne možemo zaključiti da postoji statistički značajna razlika između medijana ta dva skupa.

4.1.2. Pročišćivač 2

Pročišćivač 2 je druga lokacija na kojoj imamo mjerene uzorke ulaza i izlaza iz stroja za pročišćavanje uzete istih datuma. U skupu podataka postoji ukupno 9 takvih parova mjerenja. Kao i na lokaciji Pročišćivač 1, na ulazu nisu mjereni svi parametri kao i na izlazu pa ćemo analizirati samo pH-vrijednost, temperatura vode, ukupna suspendirana tvar, BPK_5 i ukupna ulja i masti. S obzirom da su postupci provođenja testiranja jednaki kao i za lokaciju Pročišćivač 1 te će se samo ponavljati iste testove na drugim podacima, za lokaciju Pročišćivač 2 ćemo samo prikazati rezultate tih testiranja u tablici. U tablici ispod teksta će se također navoditi hipoteza koju ćemo prihvatiti na temelju rezultata i aritmetička sredina razlika parova.

Tablica 3: Rezultati testiranja mjerenja provedenih istih datuma na lokaciji Pročišćivač 2

Parametar	Shapiro-Wilks test	Test razlike vrijednosti	Rezultat
pH-vrijednost	W = 0.9784, p = 0.9556	T-test zavisnih uzoraka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 0.09444$
	Normalna distribucija podataka	t = 2.3922, df = 8, p = 0.0437	
Temperatura vode	W = 0.80496 p = 0.04585	Wilcoxonov test	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} =$ -0.92857
	Podaci nemaju normalnu distribuciju	V = 19.5 p = 0.3972	
Ukupna suspendirana tvar	W = 0.46257 $p = 2.319 \cdot 10^{-6}$	Wilcoxonov test	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} =$ -138.4
	Podaci nemaju normalnu distribuciju	V = 10 p = 0.1641	
BPK_5	W = 0.74613 p =0.004889	Wilcoxonov test	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} =$ -99.6667
	Podaci nemaju normalnu distribuciju	V = 34 p = 0.2031	
Ulja i masti	W = 0.70475 p = 0.004176	Wilcoxonov test	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} =$ -17.4557
	Podaci nemaju normalnu distribuciju	V = 13 p = 0.675	

4.1.3. Pročišćivač 3

Pročišćivač 3 je posljednja lokacija na kojoj su mjerenja na ulazu i izlazu iz pročišćivaća napravljena istog datuma. Napravljeno je ukupno 5 para takvih mjerenja. Parametri koje ćemo uspoređivati su isti kao i na prethodne dvije lokacije. Također ćemo vršiti ista testiranja kao i na prethodnim primjerima pa ćemo rezultate prikazati u tablici ispod teksta.

Tablica 4: Rezultati testiranja mjerenja provedenih istih datuma na lokaciji Pročišćivač 3

Parametar	Shapiro-Wilks test	Test razlike vrijednosti	Rezultat
pH-vrijednost	W = 0.9287 p = 0.5875	T-test zavisnih uzoraka	H_0= aritmetičke sredine skupova nisu statistički značajno različite $\bar{x} = 0.022$
	Normalna distribucija podataka	t = 0.12484, df = 4, p = 0.9067	
Temperatura vode	W = 0.80588 p = 0.09041	T-test zavisnih uzoraka	H_1= aritmetičke sredine skupova su statistički značajno različiti $\bar{x} = -0.42$
	Normalna distribucija podataka	t = -3.7717, df = 4, p = 0.01958	
Ukupna suspendirana tvar	W = 0.98884 p = 0.9755	T-test zavisnih uzoraka	H_0= aritmetičke sredine skupova nisu statistički značajno različiti $\bar{x} = 440.76$
	Normalna distribucija podataka	t = 1.7295, df = 4, p = 0.1558	
BPK_5	W = 0.90044 p =0.4123	T-test zavisnih uzoraka	H_0= aritmetičke sredine skupova nisu statistički značajno različite $\bar{x} = 18.2$
	Normalna distribucija podataka	t = 0.20953, df = 4, p = 0.8443	
Ulja i masti	W = 0.82849 p = 0.1355	T-test zavisnih uzoraka	H_0= aritmetičke sredine skupova nisu statistički značajno različiti $\bar{x} = -8.34$
	Normalna distribucija podataka	t = -0.03106, df = 4, p = 0.9767	

4.2. Usporedba prije i poslije obalnog kolektora

Izgradnja obalnog kolektora je projekt čiji je cilj bio sanacija kanalizacijske odvodnje u zaljev. Za realizaciju projekta utrošeno je približno 240 milijuna kuna. Neki od ispusta otpadne vode u zaljev su zatvoreni dok su ostali sanirani. Prije projekta u zaljev se ispuštala nepročišćena voda koja je stvarala talog na dnu zaljeva i zagađivala gradsku luku. Kolektor je krenuo s radom 24. ožujka 2015. godine (Hrvatske vode, 2015). Usporedit će se vrijednosti parametara izmjerenih prije i poslije tog datuma kako bi se donio zaključak o učinkovitosti kolektora. Analizirat će se mjerenja na izlazu iz uređaja za pročišćavanje otpadnih voda na lokacijama Pročišćivač 1 i 2 te na ispustu kod hotela. Za lokacije Pročišćivač 3, Prepumpna stanica i Stari kanal ne postoji dovoljan broj mjerenja prije i poslije datuma puštanja kolektora u rad da bi se mogao donijeti zaključak. Skupovi podataka na kojima se provode testovi nisu zavisni pa će se provoditi T-test za nezavisne uzorke ili Mann-Whitney-Wilcoxonov test ovisno o normalnosti distribucije koja će se ispitati Shapiro-Wilksovim testom.

4.2.1. Pročišćivač 1

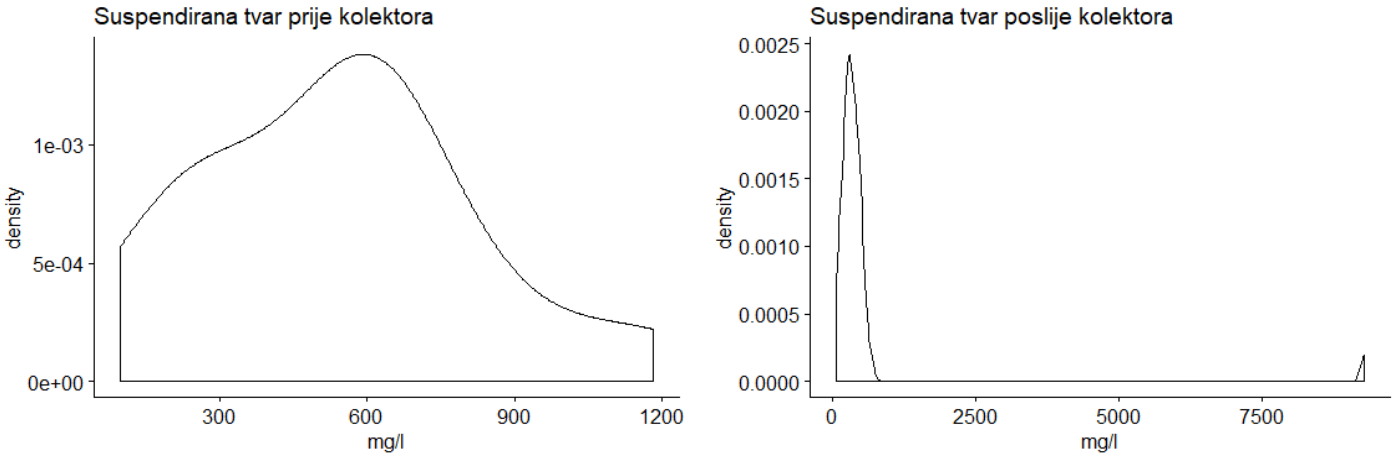
Na lokaciji Pročišćivač 1 u skupu podataka postoji 15 mjerenja prije dana puštanja obalnog kolektora u rad i 31 mjerenje poslije. Parametri koji će se analizirati su pH-vrijednost, temperatura vode, ukupna suspendirana tvar, BPK_5, KPKCr, amonijak, kloridi, ukupni fosfor, anionski detergenti, ukupna ulja i masti, cink i olovo. Za prvih par testova će biti opisani postupak testiranja a za daljnja testiranja koja prate isti princip će se samo navoditi rezultati.

Prvi parametar je pH-vrijednost. Prije odabira parametarskog ili neparametarskog testa potrebno je testirati normalnost distribucije podataka. Za ovo testiranje podaci nisu zavisni kao u prethodnim primjerima pa će se iz tog razloga Shapiro-Wilks test provoditi na svakom skupu pojedinačno. Uvjet za parametarski test je da oba skupa podataka od prije i poslije kolektora imaju normalnu distribuciju. Test se provodi na ranije opisan način pomoću alata R. Rezultat testa na skupu podataka pH-vrijednosti prije kolektora je p-vrijednost = 0.8424 na temelju čega prihvaćamo H_0 da ne postoji značajno odstupanje od normalne distribucije. Test na skupu podataka poslije kolektora je dao p-vrijednost = 0.4614 što ukazuje na to da i taj skup ima normalnu distribuciju podataka. Ispunjeni su uvijeti za parametarski T-test nezavisnih uzoraka. Vrijednosti varijanca nije potrebno posebno ispitivati jer alat R sam po potrebi radi korekcije koristeći Welchov T-test. Test pokrećemo naredbom

```
t.test(prociscivac1_pH$prije, prociscivac1_pH$poslije, paired =
FALSE, conf.level = 0.95)
```

Prvi argument naredbe je skup podataka prije kolektora, drugi element je skup podataka poslije, paired je zadano kao false jer su podaci nezavisni a razina značajnosti je određena sa 95%. Rezultat testa je t = -0.0103, df = 21.55 i p = 0.9919. Iz toga može se prihvatiti hipoteza H_0 da aritmetičke sredine skupova nisu statistički značajno različite. Rezultati pokazuju još i da je aritmetička sredina podataka prije 7.848667 a poslije 7.849677 što je približno jednako.

Drugi parametar je ukupna suspendirana tvar. Provodi se Shapiro-Wilksov test za normalnost distribucije na isti način. Rezultat za skup prije je p = 0.4577 a za skup poslije je $p = 1.13 \cdot 10^{-11}$. Za lakšu vizualizaciju moguće je napraviti i histogram koji prikazuje distribuciju podataka.



Slika 7: Histogrami distribucije podataka suspendirane tvari prije i poslije kolektora

Vidljivo je da skup prije ima normalnu distribuciju podataka ali skup poslije nema. U drugom skupu postoji ekstremna vrijednost izmjerena 21. siječnja 2016. godine koja narušava normalnost distribucije. Iz tog razloga ćemo koristiti Mann-Whitney-Wilcoxonov test. U alatu R se provodi pomoću funkcije

```
wilcox.test(prociscivac1_pH$prije, prociscivac1_pH$poslije, paired =
FALSE, conf.level = 0.95)
```

kojoj se proslijede isti parametri kao i u slučaju T-testa. Rezultati testa su W = 340 i p = 0.01216. Na temelju p-vrijednosti prihvaća se hipoteza H_1 da su medijani dva skupa statistički značajno različiti.

Na ostalim parametrima provode se testovi na isti način i radi preglednosti rezultati tih testova su prikazani u tablici 5.

Tablica 5: Rezultati testiranja prije i poslije obalnog kolektora na lokaciji Pročišćivač 1

<table>
<tr><th>Parametar</th><th colspan="2">Shapiro-Wilks test</th><th>Test razlike vrijednosti</th><th>Rezultat</th></tr>
<tr><td rowspan="3">pH-vrijednost</td><td>Prije:</td><td>p = 0.8424</td><td>T-test nezavisnih uzoraka</td><td rowspan="3">H_0 = aritmetičke sredine skupova nisu statistički značajno različite
$\bar{x}_1 = 7.84866$
$\bar{x}_2 = 7.84967$</td></tr>
<tr><td>Poslije:</td><td>p = 0.4614</td><td rowspan="2">t = -0.0103
df = 21.55
p = 0.9919</td></tr>
<tr><td colspan="2">Normalna distribucija podataka</td></tr>
<tr><td rowspan="3">Temperatura vode</td><td>Prije:</td><td>p = 0.05407</td><td>T-test nezavisnih uzoraka</td><td rowspan="3">H_0 = aritmetičke sredine skupova nisu statistički značajno različite
$\bar{x}_1 = 13.62$
$\bar{x}_2 = 13.3829$</td></tr>
<tr><td>Poslije:</td><td>p = 0.7367</td><td rowspan="2">t = 0.17293
df = 28.814
p = 0.8639</td></tr>
<tr><td colspan="2">Normalna distribucija podataka</td></tr>
<tr><td rowspan="3">Ukupna suspendirana tvar</td><td>Prije:</td><td>p = 0.4577</td><td>Mann-Whitney-Wilcoxonov test</td><td rowspan="3">H_1 = medijani skupova su statistički značajno različiti
$\bar{x}_1 = 542.1467$
$\bar{x}_2 = 612.0903$</td></tr>
<tr><td>Poslije:</td><td>p = $1.613 \cdot 10^{-11}$</td><td rowspan="2">W = 340
p = 0.01216</td></tr>
<tr><td colspan="2">Podaci nemaju normalnu distribuciju</td></tr>
<tr><td rowspan="3">BPK_5</td><td>Prije:</td><td>p = 0.9922</td><td>T-test nezavisnih uzoraka</td><td rowspan="3">H_1 = aritmetičke sredine skupova su statistički značajno različite
$\bar{x}_1 = 275.2667$
$\bar{x}_2 = 167.7742$</td></tr>
<tr><td>Poslije:</td><td>p = 0.7122</td><td rowspan="2">t = 3.3604
df = 18.298
p = 0.003422</td></tr>
<tr><td colspan="2">Normalna distribucija podataka</td></tr>
<tr><td rowspan="3">KPKCr</td><td>Prije:</td><td>p = 0.9023</td><td>T-test nezavisnih uzoraka</td><td rowspan="3">H_0 = aritmetičke sredine skupova nisu statistički značajno različite
$\bar{x}_1 = 754.1533$
$\bar{x}_2 = 629.6439$</td></tr>
<tr><td>Poslije:</td><td>p = 0.964</td><td rowspan="2">t = 1.5207
df = 18.919
p = 0.1449</td></tr>
<tr><td colspan="2">Normalna distribucija podataka</td></tr>
</table>

Amonijak	Prije:	p = 0.001507	Mann-Whitney-Wilcoxonov test	H_0 = medijani skupova nisu statistički značajno različiti $\bar{x}_1 = 55.84133$ $\bar{x}_2 = 44.51355$
	Poslije:	p = 0.01577	W = 311.5 p = 0.06583	
	Podaci nemaju normalnu distribuciju			
Kloridi	Prije:	p = 0.001054	Mann-Whitney-Wilcoxonov test	H_1 = medijani skupova su statistički značajno različiti $\bar{x}_1 = 1235.7246$ $\bar{x}_2 = 2398.2383$
	Poslije:	p = 0.00471	W = 106 p = 0.003151	
	Podaci nemaju normalnu distribuciju			
Ukupni fosfori	Prije:	p = 0.4839	T-test nezavisnih uzoraka	H_0 = aritmetičke sredine skupova nisu statistički značajno različite $\bar{x}_1 = 7.9480$ $\bar{x}_2 = 6.1473$
	Poslije:	p = 0.9996	t = 1.8061 df = 18.012 p = 0.08764	
	Normalna distribucija podataka			
Anionski detergenti	Prije:	p = 0.1069	T-test nezavisnih uzoraka	H_0 = aritmetičke sredine skupova nisu statistički značajno različite $\bar{x}_1 = 8.5733$ $\bar{x}_2 = 6.1473$
	Poslije:	p = 0.1712	t = 0.7706 df = 21.299 p = 0.4494	
	Normalna distribucija podataka			
Ulja i masti	Prije:	$p = 6.478 \cdot 10^{-5}$	Mann-Whitney-Wilcoxonov test	H_0 = medijani skupova nisu statistički značajno različiti $\bar{x}_1 = 32.856$ $\bar{x}_2 = 25.3671$
	Poslije:	$p = 1.754 \cdot 10^{-9}$	W = 252.5 p = 0.6448	
	Podaci nemaju normalnu distribuciju			
Cink	Prije:	p = 0.1256	Mann-Whitney-Wilcoxonov test	H_0 = medijani skupova nisu statistički značajno različiti $\bar{x}_1 = 0.1886$ $\bar{x}_2 = 0.3110$
	Poslije:	$p = 4.234 \cdot 10^{-5}$	W = 173 p = 0.1667	
	Podaci nemaju normalnu distribuciju			

Olovo	Prije:	$p = 1.987 \cdot 10^{-5}$	Mann-Whitney-Wilcoxonov test	H_0 = medijani skupova nisu statistički značajno različiti $\bar{x}_1 = 0.0134$ $\bar{x}_2 = 0.0218$
	Poslije:	$p = 2.2 \cdot 10^{-6}$	W = 194 p = 0.3526	
	Podaci nemaju normalnu distribuciju			

4.2.2. Pročišćivač 2

Za lokaciju Pročišćivač 2 testiraju se isti parametri kao i za lokaciju Pročišćivač 1. Postoji 15 mjerenja prije dana puštanja obalnog kolektora u rad i 29 mjerenje poslije. Princip provođenja testova je isti. Rezultati su prikazani u tablici 6.

Tablica 6: Rezultati testiranja prije i poslije obalnog kolektora na lokaciji Pročišćivač 2

Parametar	Shapiro-Wilks test		Test razlike vrijednosti	Rezultat
pH-vrijednost	Prije:	p = 0.1558	T-test nezavisnih uzoraka	H_0 = aritmetičke sredine skupova nisu statistički značajno različite $\bar{x}_1 = 7.892$ $\bar{x}_2 = 7.7682$
	Poslije:	p = 0.3628	t = 1.6527 df = 26.842 p = 0.11	
	Normalna distribucija podataka			
Temperatura vode	Prije:	p = 0.09858	T-test nezavisnih uzoraka	H_0 = aritmetičke sredine skupova nisu statistički značajno različite $\bar{x}_1 = 14.29333$ $\bar{x}_2 = 13.65517$
	Poslije:	p = 0.2104	t = 0.41995 df = 29.108 p = 0.6776	
	Normalna distribucija podataka			
Ukupna suspendirana tvar	Prije:	p = 0.00617	Mann-Whitney-Wilcoxonov test	H_1 = medijani skupova su statistički značajno različiti $\bar{x}_1 = 424.4133$ $\bar{x}_2 = 276.0827$
	Poslije:	p = 0.9684	W = 340 p = 0.002504	
	Podaci nemaju normalnu distribuciju			

BPK_5	Prije:	p = 0.001477	Mann-Whitney-Wilcoxonov test	H_1 = medijani skupova su statistički značajno različiti $\bar{x}_1 = 272{,}4$ $\bar{x}_2 = 158$
	Poslije:	p = 0.003897	W = 386 p = $3.182e \cdot 10^{-5}$	
	Podaci nemaju normalnu distribuciju			
KPKCr	Prije:	p = 0.01776	Mann-Whitney-Wilcoxonov test	H_1 = medijani skupova su statistički značajno različiti $\bar{x}_1 = 745{,}1733$ $\bar{x}_2 = 570{,}7717$
	Poslije:	p = 0.7121	W = 313 p = 0.01741	
	Podaci nemaju normalnu distribuciju			
Amonijak	Prije:	p = 0.9135	T-test nezavisnih uzoraka	H_0 = aritmetičke sredine skupova nisu statistički značajno različite $\bar{x}_1 = 57.12267$ $\bar{x}_2 = 43.51966$
	Poslije:	p = 0.3927	t = 2.0403 df = 17.314 p = 0.05688	
	Normalna distribucija podataka			
Kloridi	Prije:	p = $8.65 \cdot 10^{-5}$	Mann-Whitney-Wilcoxonov test	H_0 = medijani skupova nisu statistički značajno različiti $\bar{x}_1 = 2860{,}635333$ $\bar{x}_2 = 1514{,}205172$
	Poslije:	p = $2.63 \cdot 10^{-6}$	W = 205 p = 0.00769	
	Podaci nemaju normalnu distribuciju			
Ukupni fosfori	Prije:	p = 0.4839	T-test nezavisnih uzoraka	H_0 = aritmetičke sredine skupova nisu statistički značajno različite $\bar{x}_1 = 7.948067$ $\bar{x}_2 = 6.147323$
	Poslije:	p = 0.9996	t = 1.8061 df = 18.012 p = 0.08764	
	Normalna distribucija podataka			
Anionski detergenti	Prije:	p = 0.02083	Mann-Whitney-Wilcoxonov test	H_0 = medijani skupova nisu statistički značajno različiti $\bar{x}_1 = 9{,}9793$ $\bar{x}_2 = 7{,}7674$
	Poslije:	p = 0.01659	W = 237 p = 0.6418	
	Podaci nemaju normalnu distribuciju			

<table>
<tr><td rowspan="3">Ulja i masti</td><td>Prije:</td><td>p = 1.683· 10^{-6}</td><td>Mann-Whitney-Wilcoxonov test</td><td rowspan="3">H_0 = medijani skupova nisu statistički značajno različiti
$\bar{x}_1 = 45{,}3066$
$\bar{x}_2 = 37{,}4165$</td></tr>
<tr><td>Poslije:</td><td>p = 3.492 · 10^{-9}</td><td rowspan="2">W = 224
p = 0.8812</td></tr>
<tr><td colspan="2">Podaci nemaju normalnu distribuciju</td></tr>
<tr><td rowspan="3">Cink</td><td>Prije:</td><td>p = 0.2611</td><td>T-test nezavisnih uzoraka</td><td rowspan="3">H_0 = aritmetičke sredine skupova nisu statistički značajno različite
$\bar{x}_1 = 0.1989$
$\bar{x}_2 = 0.3051$</td></tr>
<tr><td>Poslije:</td><td>p = 0.1289</td><td rowspan="2">t = - 2.0348
df = 28.602
p = 0.05124</td></tr>
<tr><td colspan="2">Normalna distribucija podataka</td></tr>
<tr><td rowspan="3">Olovo</td><td>Prije:</td><td>p = 0.005467</td><td>Mann-Whitney-Wilcoxonov test</td><td rowspan="3">H_0 = medijani skupova nisu statistički značajno različiti
$\bar{x}_1 = 0{,}0074$
$\bar{x}_2 = 0{,}0184$</td></tr>
<tr><td>Poslije:</td><td>p = 4.73 · 10^{-7}</td><td rowspan="2">W = 185.5
p = 0.4187</td></tr>
<tr><td colspan="2">Podaci nemaju normalnu distribuciju</td></tr>
</table>

4.2.3. Hotel

Posljednja lokacija za koju postoji dovoljno podataka za testiranje prije i poslije puštanja obalnog kolektora u rad je ispust kod hotela. Testirat će se parametri pH-vrijednost, temperatura vode, ukupna suspendirana tvar, BPK_5 i KPKCr. Postoji 5 mjerenja prije i 6 poslije 24.3.2015. Rezultati testiranja prikazani su u tablici 7.

Tablica 7: Rezultati testiranja prije i poslije obalnog kolektora na lokaciji Hotel

<table>
<tr><th>Parametar</th><th colspan="2">Shapiro-Wilks test</th><th>Test razlike vrijednosti</th><th>Rezultat</th></tr>
<tr><td rowspan="3">pH-vrijednost</td><td>Prije:</td><td>p = 0.5609</td><td>T-test nezavisnih uzoraka</td><td rowspan="3">H_0 = aritmetičke sredine skupova nisu statistički značajno različite
$\bar{x}_1 = 7.382$
$\bar{x}_2 = 7.2366$</td></tr>
<tr><td>Poslije:</td><td>p = 0.301</td><td rowspan="2">t = 0.72165
df = 4.1911
p = 0.5087</td></tr>
<tr><td colspan="2">Normalna distribucija podataka</td></tr>
<tr><td rowspan="3">Temperatura vode</td><td>Prije:</td><td>p = 0.4772</td><td>T-test nezavisnih uzoraka</td><td rowspan="3">H_0 = aritmetičke sredine skupova nisu statistički značajno različite
$\bar{x}_1 = 14.98000$
$\bar{x}_2 = 15.81667$</td></tr>
<tr><td>Poslije:</td><td>p = 0.3984</td><td rowspan="2">t = -0.33388
df = 8.597
p = 0.7465</td></tr>
<tr><td colspan="2">Normalna distribucija podataka</td></tr>
</table>

Ukupna suspendirana tvar	Prije:	p = 0.01853	Mann-Whitney-Wilcoxonov test	H_0 = medijani skupova nisu statistički značajno različiti $\bar{x}_1 = 556{,}78$ $\bar{x}_2 = 63{,}8167$
	Poslije:	$p = 6.808 \cdot 10^{-5}$	W = 25 p = 0.08004	
	Podaci nemaju normalnu distribuciju			
BPK_5	Prije:	p = 0.4002	Mann-Whitney-Wilcoxonov test	H_1 = medijani skupova su statistički značajno različiti $\bar{x}_1 = 159$ $\bar{x}_2 = 6{,}7$
	Poslije:	p = 0.0001408	W = 28 p = 0.01732	
	Podaci nemaju normalnu distribuciju			
KPKCr	Prije:	p = 0.3149	Mann-Whitney-Wilcoxonov test	H_1 = medijani skupova su statistički značajno različiti $\bar{x}_1 = 355{,}482$ $\bar{x}_2 = 3{,}45$
	Poslije:	$p = 2.073 \cdot 10^{-5}$	W = 28 p = 0.01669	
	Podaci nemaju normalnu distribuciju			

4.3. Usporedba s propisanim dozvoljenim količinama

Ministarstvo poljoprivrede propisuje maksimalne dozvoljene količine tvari u otpadnim vodama koje se ispuštaju u površinske vode kao što je to slučaj u zaljevu primorskog grada. Količine su određene pravilnikom o graničnim vrijednostima emisija otpadnih voda (Ministarstvo Poljoprivrede, 2013.). Parametri preostalih lokacija za koje ne postoje mjerenja prije i poslije pročišćavanja će biti uspoređivani s propisanim maksimalnim dozvoljenim količinama. Lokacije za koje će se provoditi testovi su Hotel, Prepumpna stanica i Stari kanal. Propisane vrijednosti s kojima ćemo uspoređivati skupove podataka su temperatura vode do 30 ºC, ukupna suspendirana tvar do 35 mg/l, BPK_5 do 25 mg/l, KPKCr do 125 mg/l, ukupna ulja i masti do 20 mg/l, amonijak do 10 mg/l i ukupni fosfor 2 mg/l. Za pH-vrijednost vode je propisano da ne smije premašivat interval od 6,5 do 9,0 no za potrebne našeg testiranja koristit ćemo vrijednost 7 pH jer u toj vrijednosti voda nije ni kisela ni bazična. Provodit će se testovi nad jednim uzorkom. Ukoliko su ispunjeni uvjeti za parametarski test koristit će se T-test a ukoliko nisu koristit će se test predznaka.

4.3.1. Hotel

Na lokaciji ispusta kod hotela postoji ukupno 11 mjerenja koja će se uspoređivati s propisanim vrijednostima. Parametri koji su mjereni su pH-vrijednost, temperatura vode, ukupna suspendirana tvar, BPK_5, KPKCr i ukupna ulja i masti. Postupak provođenja T-testa za jedan uzorak i testa predznaka će biti opisan na jednom primjeru a dalje će se samo navoditi rezultati u tablicama. Koristi se alat R kao i u prethodnim primjerima.

Prvi parametar koji se testira je pH-vrijednost. Potrebno je testirati normalnost distribucije podataka Shapiro-Wilks testom. Rezultat tog testa je p-vrijednost = 0.6637 na temelju čega možemo zaključiti da podaci imaju normalnu distribuciju. Zatim se provodi parametarski t-test na jednom uzorku gdje će se aritmetička sredina uzorka uspoređivati s vrijednošću 7. U alatu R to možemo s naredbom

```
t.test(hotel$pH, mu=7, conf.level=0.95)
```

U kojoj je hotel$pH varijabla u kojoj je pohranjen skup podataka, mu=7 određuje vrijednost s kojoj se uspoređuje a conf.level=0.95 određuje razinu značajnosti kao 95%. Rezultati testa su t = 3.388, df = 10 i p-vrijednost = 0.006909 čime vidimo da se može prihvatiti hipoteza H_1 da je aritmetička sredina vrijednosti statistički značajno različita. Kao dio rezultata dobivena je i aritmetička sredina vrijednosti iz skupa koja iznosi 7.302727.

Sljedeći parametar je ukupna suspendirana vrijednost. Shapiro-Wilksovim testom dobiva se p-vrijednost koja iznosi $4.063 \cdot 10^{-6}$. Podaci nemaju normalnu distribuciju pa će se koristiti test predznaka. Za taj test je u alatu R potrebno instalirati paket DSBA. Test se zatim pokreće naredbom

```
SIGN.test(hotel$susp, md=35, conf.level=0.95)
```

gdje je hotel$susp skup podataka a md=35 vrijednost medijana s kojim uspoređujemo skup podataka. Rezultat testa je s = 4 i p-vrijednost = 0.5488. Na temelju rezultata prihvaća se hipoteza H_0 da medijani skupova nisu statistički značajno različiti. Test računa i procjenu medijana skupa koja iznosi 20.1. Rezultati ostalih testova zajedno s graničnim vrijednostima označenih s G prikazani su u tablici 8.

Tablica 8: Rezultati testiranja usporedbe parametara s propisanim vrijednostima na lokaciji Hotel

Parametar	Shapiro-Wilks test	Test razlike vrijednosti	Rezultat
pH-vrijednost	W = 0.95154 p = 0.6637	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 7.302727$ 6,5 ≤ G ≤ 9,0
	Normalna distribucija podataka	t = 3.388, df = 10, p = 0. 006909	
Temperatura vode	W = 0.90991 p =0.2433	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 15.43636$ G ≤ 30
	Normalna distribucija podataka	t = -12.244, df = 10, p = $2.417 \cdot 10^{-7}$	
Ukupna suspendirana tvar	W = 0.53532 p = $4.063 \cdot 10^{-6}$	Test predznaka	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} = 287.8909$ M = 20.1 G ≤ 35
	Podaci nemaju normalnu distribuciju	s = 4 p = 0.5488	
BPK_5	W =0.64661 p = $9.761 \cdot 10^{-6}$	Test predznaka	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} = 75.92$ M = 6 G ≤ 25
	Podaci nemaju normalnu distribuciju	s = 3 p = 0.3438	
KPKCr	W = 0.63935 p = $7.906 \cdot 10^{-5}$	Test predznaka	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} = 217.5818$ M = 40.9 G ≤ 125
	Podaci nemaju normalnu distribuciju	s = 3 p = 0.2266	
Ukupna ulja i masti	W = 0.35234 p =$2.729 \cdot 10^{-8}$	Test predznaka	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} = 163.4645$ M = 2.5 G ≤ 20
	Podaci nemaju normalnu distribuciju	s = 2 p = 0.06543	

4.3.2. Prepumpna stanica

Na lokaciji Prepumpna stanica postoji samo 4 mjerenja koja će se uspoređivati s propisanim vrijednostima. Parametri koji će se uspoređivati su pH-vrijednost, temperatura vode, ukupna suspendirana tvar, BPK_5, KPKCr, amonijak i ukupni fosfor. Testovi se provode na prethodno opisani način. Rezultati testiranja nalaze se u tablici 9.

Tablica 9: Rezultati testiranja usporedbe parametara s propisanim vrijednostima na lokaciji Prepumpna stanica

Parametar	Shapiro-Wilks test	Test razlike vrijednosti	Rezultat
pH-vrijednost	W = 0.97589 p = 0.8775	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 7.99$ 6,5 ≤ G ≤ 9,0
	Normalna distribucija podataka	t = 6.6295, df = 3, p = 0.006991	
Ukupna suspendirana tvar	W = 0.89411 p = 0.4024	T-test jednog uzorka	H_0= aritmetičke sredine skupova nisu statistički značajno različite $\bar{x} = 418.4$ G ≤ 35
	Normalna distribucija podataka	t = 2.0055, df = 3, p = 0.1386	
BPK_5	W = 0.9613 p = 0.7871	T-test jednog uzorka	H_0= aritmetičke sredine skupova nisu statistički značajno različite $\bar{x} = 229.75$ G ≤ 25
	Normalna distribucija podataka	t = 2.7831, df = 3, p = 0.06881	
KPKCr	W =0.69555 p = 0.01029	Test predznaka	H_0= medijani skupova nisu statistički značajno različiti $\bar{x} = 1055.2244$ M = 665.8988 G ≤ 125
	Podaci nemaju normalnu distribuciju	s = 4 p = 0.125	
Amonijak	W = 0.78307 p = 0.07517	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 47.185$ G ≤ 10
	Normalna distribucija podataka	t = 13.896, df = 3, p = 0.0008069	
Ukupni fosfor	W = 0.93386 p = 0.6172	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 5.26975$ G ≤ 2
	Normalna distribucija podataka	t = 6.7936, df = 3, p = 0.006521	

4.3.3. Stari kanal

Na starom kanalu izvršeno je 11 mjerenja koja će se uspoređivati s propisanim vrijednostima. Parametri koji će se uspoređivati su pH-vrijednost, ukupna suspendirana tvar, BPK_5, KPKCr te ukupna ulja i masti. Rezultati provedenih testova vidljivi su u tablici 10.

Tablica 10: Rezultati testiranja usporedbe parametara s propisanim vrijednostima na lokaciji Stari kanal

Parametar	Shapiro-Wilks test	Test razlike vrijednosti	Rezultat
pH-vrijednost	W = 0.92332 p = 0.3473	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 7.5972$ 6,5 ≤ G ≤ 9,0
	Normalna distribucija podataka	t = 6.7082, df = 10, p = $5.31 \cdot 10^{-5}$	
Temperatura vode	W = 0.9583 p = 0.7503	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 14.6272$ G ≤ 30
	Normalna distribucija podataka	t = -11.668, df = 10, p = $3.802 \cdot 10^{-7}$	
Ukupna suspendirana tvar	W = 0.94846 p = 0.6243	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 106.9273$ G ≤ 35
	Normalna distribucija podataka	t = 4.5421, df = 10, p = 0.001071	
BPK_5	W = 0.95765 p = 0.742	T-test jednog uzorka	H_1= aritmetičke sredine su statistički značajno različite $\bar{x} = 78$ G ≤ 25
	Normalna distribucija podataka	t = 4.7418, df = 10, p = 0.0007899	
KPKCr	W = 0. 98146 p = 0. 9737	T-test jednog uzorka	H_1= aritmetičke sredine skupova su statistički značajno različite $\bar{x} = 322.8091$ G ≤ 125
	Normalna distribucija podataka	t = 4.6737, df = 10, p = 0.0008759	
Ulja i masti	W = 0.83873 p = 0.03039	Test predznaka	H_1= medijani skupova su statistički značajno različiti $\bar{x} = 4.945$ M = 5.46 G ≤ 20

5. Diskusija

Nakon provedenih testova potrebno je protumačiti od kakvog su značaja rezultati i što se iz njih može zaključiti. Na testovima gdje je utvrđena statistički značajna razlika ne mora nužno značiti da je ta razlika značajna u praksi. Također je zanimljivo da je i nekim testovima utvrđeno da ne postoji statistički značajna razlika gdje bi ju se možda očekivalo. Prodiskutirat će se i zašto bi se to moglo dogoditi na pojedinim slučajevima.

5.1. Analiza testova na uzorcima istih datuma

Prvi niz testiranja proveden je na uzorcima s lokacija Pročišćivač 1, 2 i 3. Uspoređivani su uzorci otpadne vode na ulazu i izlazu iz uređaja za pročišćavanje koji su uzeti istog dana. Na taj način se eliminira mogućnost nekih drugih utjecaja na kvalitetu vode poput različitog opterećenja različitih dana. Ovakvom usporedbom može se zaključiti koliko učinkovito se odvijalo pročišćivanje na pojedinim uređajima za svaki parametar kroz period od 2014. do 2017. godine. Analizom pH-vrijednosti na sve tri lokacije samo je na lokaciji Pročišćivača 2 se pokazalo da postoji statistički značajna razlika. Računanjem aritmetičke sredine razlike dvaju uzoraka vidimo da ona iznosi 0.0944 pH što nije značajno u praksi. Možemo zaključiti da Pročišćivači nemaju veliki utjecaj na pH-vrijednost vode koju pročišćuju.

Analizom vrijednosti temperature samo u na lokaciji pročiščivaća 3 možemo zaključiti da postoji statistički značajna razlika u vrijednostima aritmetičkih sredina. Ako izračunamo aritmetičku sredinu razlika za sve uzorke vidimo sa je otpadna voda nakon pročišćavanja prosječno 1ºC hladnija. To ne predstavlja veliku razliku u praksi pogotovo ako uzmemo u obzir da otpadna voda ima negativni utjecaj na okoliš u slučaju da je značajno toplija.

Ukupna suspendirana tvar i BPK_5 su dobar pokazatelj kvalitete otpadne vode i poželjno je da te vrijednosti budu što niže. Ukupna suspendirana tvar pokazuje koliko je ukupne neotopive tvari u otpadnoj vodi a BPK_5 ili biokemijska potreba kisika u roku od 5 dana pokazuje koliko je voda opterećena s organskom tvari (Hrvatska Enciklopedija, 2018). Analizom ukupne suspendirane tvari pokazano je da ne postoji statistički značajna razlika aritmetičkih sredina na promatranim lokacijama. Pogledom na sirove podatke vidljivo je da su razine suspendirane tvari nakon pročišćivanja ponekad niže a ponekad čak i više od vrijednosti na ulazu u stroj. Ipak može se primijetiti da kod ulaza često se javljaju ekstremne vrijednosti koje ne postoje na izlazima. Na temelju toga može se zaključiti da uređaj za pročišćivanje ima granicu ispod koje više ne postiže jednaku učinkovitost u uklanjanju suspendirane tvari kao i na ekstremnim vrijednostima.

Analizom BPK_5 nije pokazana statistički značajna razlika samo na ni jednoj od lokacija. Aritmetička sredina razlika jednaka na lokaciji Pročišćivač 1 je -82.8182 O_2/l što ipak ukazuje na trend smanjivanja količine. Sirovi podaci pokazuju slična svojstva kao što je slučaj i za ukupnu suspendiranu tvar pa možemo zaključiti da postoji slična granica u mogućnostima stroja za pročišćavanje kod nižih ulaznih razina BPK_5.

Analizom ukupnih ulja i masti pokazano je da ne postoji statistički značajna razlika aritmetičkih sredina vrijednosti prije i poslije pročišćivanja na ni jednoj od analiziranih lokacija. Možemo pretpostaviti da uređaji nisu strogo predviđeni za filtriranje masti i ulja iz otpadnih voda.

5.2. Analiza testova učinkovitosti obalnog kolektora

Obalni kolektor pušten je u rad 24. ožujka 2015. godine. Provedeni su testovi razlika aritmetičkih sredina i medijana na podacima uzoraka prije i poslije tog datuma za lokacije Pročišćivač 1, 2 i Hotel. Analizom podataka prije i poslije moći će se vidjeti postoji li značajna razlika u svojstvima otpadne vode koja se ispušta u zaljev nakon tog projekta.

Analizom vrijednosti svih parametara statistički značajna razlika se javlja samo u parametrima ukupna suspendirana tvar, BPK_5, KPKCr i Kloridi. Razine ukupne suspendirane tvari su vidljivo smanjene na lokaciji Pročišćivač 2 na kojoj je izmjereno prosječno 150 mg/l manje suspendirane tvari na izlazu nakon ugradnje obalnog kolektora u odnosu na prije. Na lokaciji Pročišćivač 1 također je utvrđena razlika u medijanima skupova no razlika u aritmetičkim sredinama uzoraka prije i poslije ukazuje na to da je prosječna razina suspendirane tvari porasla od ugradnje kolektora. To može biti posljedica dodataka koji se dodaju kako bi se dezinficirala voda poput klora. Razina suspendirane tvari na lokaciji Hotel nije statistički značajno različita prije i poslije ugradnje kolektora. Za tu lokaciju ne postoji veliki broj mjerenja prije i poslije ugradnje kolektora a i vrijednosti su većinom ekstremne pa je teško donijeti valjani zaključak. Statistički značajna razlika aritmetičkih sredina vrijednosti BPK_5 je utvrđena na sve tri lokacije. Na lokacijama Pročišćivač 1 i 2 prosječna razina BPK_5 nakon pročišćavanja je gotovo prepolovljena, a na lokaciji Hotel je BPK_5 gotovo eliminiran.

KPKCr je pokazatelj sličan BPK5 osim što pokazuje podložnost oksidaciji uz prisutnost jakog oksidacijskog reagensa (Hanna Instruments, 2018). Na lokacijama Pročišćivač 2 i Hotel postoji statistički značajna razlika aritmetičkih sredina vrijednosti KPKCr prije i poslije ugradnje obalnog kolektora. U oba slučaja je razlika u približno 100 mg/l što predstavlja značajnu razliku i u praksi.

Razine klorida mjerene su samo na lokacijama Pročišćivač 1 i 2. Na lokaciji Pročišćivač 1 postoji značajna razlika u količini klora nakon obalnog kolektora. Aritmetičke sredine skupova prije i poslije kolektora nam ukazuju da se nakon ugradnje kolektora u vodi

nalazi približno dvostruko veća količina klora. Može se pretpostaviti da klor dolazi iz uređaja za pročišćavanje jer se klor često koristi za dezinfekciju vode.

Za parametre temperature vode, pH-vrijednosti, amonijaka, ukupnog fosfora, anionski detergenata te ukupnih ulja i masti ne postoji statistički značajna razlika aritmetičkih sredina i medijana prije i poslije obalnog kolektora. Ipak, ako se pogledaju vrijednosti aritmetičkih sredina skupova za svaki od navedenih parametara, može se uočiti da su uvijek za određeni postotak manje nakon ugradnje kolektora u odnosu na prije. Jedina iznimka su olovo i cink čija je aritmetičkih sredina skupa u oba slučaja veća nakon ugradnje kolektora. To može biti posljedica uređaja za pročišćavanje ili većeg opterećenja uređaja cinkom i olovom nakon 2015. godine. Nažalost skup podataka na kojem je vršena analiza ne daje odgovor jer razine cinka i olova nisu mjerene na ulazu u stroj za pročišćavanje.

5.3. Analiza testova uspoređivanja s propisanim vrijednostima

Posljednji niz testiranja je bio u svrhu uspoređivanja vrijednosti izmjerene na lokacijama Hotel, Prepumpna stanica i Stari kanal s propisanim graničnim vrijednostima za otpadne vode koje se ispuštaju u površinske vode koje je donijelo Ministarstvo poljoprivrede. Na taj način može se prokomentirati kakvoća otpadne vode koja se ispušta u zaljev.

Analizom pH-vrijednosti na sve tri lokacije utvrđeno je da se aritmetička sredina statistički značajno razlikuje od 7 pH na kojoj bi voda bila neutralna. Aritmetičke sredine pojedinih uzoraka ukazuje na to da je na sve tri lokacije voda kiselija od te vrijednosti u rasponu od približno 7,3 do 8 pH što je unutar granice do 9 pH. Na temelju toga može se pretpostaviti da kiselost otpadne vode nema snažan negativan ekološki utjecaj na zaljev.

Temperature na obje analizirane lokacije također su značajno različite od granične vrijednosti koja iznosi 30ºC. Ako se uzme u obzir aritmetičke sredine analiziranih uzoraka vidimo da su prosječne temperature približno 15ºC što je značajno niže od propisane maksimalne količine. Hladnija otpadna voda nema značajan negativan utjecaj na okoliš kao što ima topla voda jer ne potiče rast algi.

Analizom ukupne suspendirane tvari postoji statistički značajna razlika aritmetičke vrijednosti i propisane količine na lokaciji Stari kanal. Ako se uzme u obzir aritmetička sredina izmjerenih vrijednosti količina suspendirane tvari u kanalu vidimo da se ta vrijednost od 106.93 mg/l značajno razlikuje od propisane količine koja iznosi 35 mg/l. Ovakva količina suspendirane tvari je previsoka i pogledom na podatke iz skupa vidljivo je da pokazuje trend rasta unatoč tome što je 2016. godine kanal prenamijenjen samo za oborinske vode.

Analizom BPK_5 samo na lokaciji Stari kanal je pokazana statistički značajna razlika u odnosu na propisanu vrijednost od 25 mg/l. Aritmetička sredina uzorka pokazuje prosječnu

količinu BPK_5 od 78 mg/l u razdoblju od 2015. do 2017. godine. To nam govori da je kanal previše opterećen organskom tvari unatoč prenamjeni.

Razine KPKCr su također statistički značajno različite od propisane na lokaciji Stari kanal. Propisana maksimalna količina je 125 mg/l dok je aritmetička sredina količine KPKCr u kanalu 322.8 mg/l. Ovo je još jedan pokazatelj koji nam govori da je Stari kanal previše opterećen.

Ukupna ulja i masti analizirana su samo na lokacijama ispusta kod hotela i starog kanala. Statistički značajna razlika u odnosu na propisanu vrijednost pokazana je samo kod kanala. Ipak, za razliku od prethodna dva pokazatelja prosječna količina ulja i masti u kanalu je značajno ispod dozvoljene granice. Na lokaciji hotela pogledom na podatke možemo vidjeti da su razine ulja i masti na svakom mjerenju bile ispod granice osim u jednom ekstremnom slučaju.

Razine amonijaka i fosfora mjerene su samo na lokaciji Prepumpna stanica i u oba slučaja postoji statistički značajna razlika u odnosu na propisanu vrijednost. Propisane vrijednosti iznose 2 mg/l za fosfor i 10 mg/l za amonijak. Fosfor je u otpadnoj vodi na toj lokaciji prosječno prisutan u 2 puta većoj količini a količina amonijaka je prosječno približno 5 puta veća od propisane količine. Ovakve količine amonijaka mogu biti otrovne za neke vrste riba (Oregon Department of Human Services, 2000). Količine fosfora koje su prisutne u otpadnoj vodi na lokaciji Prepumpna stanica mogu snažno poticati rast algi koje onečišćuju zaljev (Oram, 2014).

6. Zaključak

Predmet ovog rada bila je statistička analiza skupa podataka koji sadrži vrijednosti parametrima otpadne vode na području primorskog grada. Analizirano je ukupno 287 mjerenja sa 12 parametara za koje se procijenilo da bi mogli biti od interesa. Mjerenja su provođena na lokacijama Pročišćivač 1, Pročišćivač 2, Pročišćivač 3, ispusta kod hotela, Prepumpna stanica i Stari kanal. Korištene su opisane metode inferencijalne statistike kojima smo prihvaćali ili odbacivali hipoteze o aritmetičkim sredinama i medijanima skupova podataka o izmjerenim parametrima. Provedeni su parametarski i neparametarski testovi na zavisnim i nezavisnim skupovima te na pojedinačnim uzorcima koje smo uspoređivali s konstantnim vrijednostima. Za provođenje testova korišten je alat R, a za obradu skupa podataka LibreOffice i Excel.

Proces testiranja smo podijelili u tri skupine prema svrsi i vrsti testiranja. Prvi niz testova uspoređivao je mjerenja prije i poslije prolaska kroz uređaj za pročišćavanje izmjerenih istih datuma. Pokazalo se da uređaj za pročišćavanje u praksi nije imao značajan utjecaj na pH-vrijednost, temperaturu vode i količinu masti i ulja. Na količine ukupne suspendirane tvari i BPK_5 je uređaj utjecao do određene granice na način da je velike količine tvari uspješno uklanjao no kada su vrijednosti tvari na ulazu u st roj bile niske nije se vidjela velika promjena.

Drugi niz testova uspoređivao je vrijednosti parametara prije i poslije projekta ugradnje obalnog kolektora. Testovi su se provodili samo na lokacijama na kojima je za to bilo dovoljno podataka a to su Pročišćivač 1, 2 i 3. Pokazano je da je nakon projekta postignuta pozitivna promjena na količine ukupne suspendirane tvari, BPK_5 i KPKCr što nam govori da je smanjena količina neotopljene, organske tvari i tvari sklonoj oksidaciji. To je značajna pozitivna promjena jer su to glavni pokazatelji zagađenosti otpadne vode. Postoji i promjena u razinama klora koja je veća u većini slučajeva. Razlog tome je najvjerojatnije klor koji se koristi u svrhu dezinfekcije vode. Ugradnja obalnog kolektora nije pokazala veliki utjecaj na ostale parametre.

Posljednji niz testova je uspoređivao vrijednosti parametara s graničnim vrijednostima koje propisuje Ministarstvo poljoprivrede kako bi se mogao donjeti zaključak o kvaliteti otpadnih voda koje se ispuštaju u zaljev. Pokazano je da se vrijednosti analiziranih lokacija većinom ne razlikuju od propisanih vrijednosti osim na lokaciji Stari kanal gdje postoji značajno veća količina suspendirane tvari, BPK_5 i KPKCr. Problem također postoji i na lokaciji Prepumpna stanica gdje postoje značajno povišene količine amonijaka i fosfora.

7. Popis literature

T Test (Student's T-Test): Definition and Examples (25. 6. 2018). Preuzeto 7. 9. 2018. iz S tatistics How To: http://www.statisticshowto.com/probability-and-statistics/t-test/

Bacarea, V., & Marusteri, M. (2010). *Biochemia Medica*. Preuzeto 27. 8 2018 iz http://www.biochemia-medica.com/content/marius-marusteri-vladimir-bacarea-kako-odabrati-pravi-test-za-procjenu-statisticke-znacajnos

Biokemijska potrošnja kisika. (n.d.). Preuzeto 14. 9 2018 iz Hrvatska Encikopedija: http://www.enciklopedija.hr/natuknica.aspx?id=7750

DZS. (2011). *Državni zavod za statistiku.* Preuzeto 14. 8 2018 iz https://www.dzs.hr/Hrv/censuses/census2011/results/htm/H01_06_01/H01_06_01.html

Eterović, P. d. (2011). *Statističko zaključivanje.* Preuzeto 4. 9 2018 iz https://www.google.hr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=2ahUKEwjK4pH6g6HdAhVCJFAKHdaKDQYQFjABegQICRAC&url=http%3A%2F%2Fneuron.mefst.hr%2Fdocs%2Fgraduate%2520school%2Febm%2FPredmeti%2Fklinicka%2520biostatistika%2Fnastavni%2520materijali%2FFilozo

Hrvatske vode. (24. 3 2015). *Svečano otvaranje obalnog kolektora*. Preuzeto 9. 9 2018 iz Hrvatske vode: https://goo.gl/QVoNiC

Kemijska potrošnja kisika (KPK). (2018). Preuzeto 14. 9 2018 iz Hanna Instruments: http://hannainst.hr/kemijska-potrosnja-kisika-kpk/

Ministarstvo Poljoprivrede. (2013.). *Pravilnik o graničnim vrijednostima emisija otpadnih voda.* Preuzeto 29. 8 2018 iz Narodne novine: https://narodne-novine.nn.hr/clanci/sluzbeni/2013_06_80_1681.html

Oram, B. (2014). *Phosphates in the Enviroment*. Preuzeto 14. 9 2018 iz Water Research Center: https://www.water-research.net/index.php/phosphates

Oregon Department of Human Services. (1 2000). *Health effects of ammonia.* Preuzeto 14. 9 2018 iz Oregon.gov: https://www.oregon.gov/oha/PH/HealthyEnvironments/DrinkingWater/Monitoring/Documents/health/ammonia.pdf

Poslovni Dnevnik. (26. 3 2015). Preuzeto 14. 8 2018 iz https://goo.gl/3PiyWN

Razali, N. M., & Yap, B. W. (1 2011). *ResearchGate.* Preuzeto 27. 8 2018 iz https://www.researchgate.net/publication/267205556_Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirnov_Lilliefors_and_Anderson-Darling_Tests

Šošić, I. (2006). *Primjenjena Statistika.* Zagreb: Školska knjiga.

Statistics How To. (11. 3 2015). Preuzeto 12. 9 2018 iz http://www.statisticshowto.com/welchs-test-for-unequal-variances/

Zaionts, C. (2013). *Real statistics using Excel*. Preuzeto 2. 9 2018 iz Real statistics: http://www.real-statistics.com/tests-normality-and-symmetry/statistical-tests-normality-symmetry/shapiro-wilk-test/

Zaiontz, C. (2013.). *Real statistics*. Preuzeto 2. 9 2018 iz Shapiro-Wilks Tables: http://www.real-statistics.com/statistics-tables/shapiro-wilk-table/

8. Popis slika

Slika 1: Jednosmjerna hipoteza na lijevu stranu (autorski rad prema Bacarea & Marusteri, 2010) 3

Slika 2: Jednosmjerna hipoteza na desnu stranu (autorski rad prema Bacarea & Marusteri, 2010) 3

Slika 3: Dvosmjerna hipoteza (autorski rad prema Bacarea & Marusteri, 2010) 3

Slika 4: Gaussova distribucija (autorski rad) 4

Slika 5: Primjer postupka dobivanja podataka iz više LibreOffice Calc datoteka 10

Slika 6: Q-Q dijagram razlika temperatura mjerenih na lokaciji Pročišćivač 1 istih datuma ...15

9. Popis tablica

Tablica 1: Broj mjerenja po lokaciji, vrsti i godini 9

Tablica 2: Hipoteze kod Shapiro-Wilks testa.....13

Tablica 3: Rezultati testiranja mjerenja provedenih istih datuma na lokaciji Pročišćivač 2 17

Tablica 4: Rezultati testiranja mjerenja provedenih istih datuma na lokaciji Pročišćivač 3 18

Tablica 5: Rezultati testiranja prije i poslije obalnog kolektora na lokaciji Pročišćivač 1 21

Tablica 6: Rezultati testiranja prije i poslije obalnog kolektora na lokaciji Pročišćivač 2 23

Tablica 7: Rezultati testiranja prije i poslije obalnog kolektora na lokaciji Hotel.....25

Tablica 8: Rezultati testiranja usporedbe parametara s propisanim vrijednostima na lokaciji Hotel.....28

Tablica 9: Rezultati testiranja usporedbe parametara s propisanim vrijednostima na lokaciji Prepumpna stanica 29

Tablica 10: Rezultati testiranja usporedbe parametara s propisanim vrijednostima na lokaciji Stari kanal 30

10. Prilozi

[1] Excel dokument uređenog skupa podataka „Otpadne vode - podaci.xlsx“