

Prikaz i analiza podataka

doc. dr Nadica Miljković

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite. Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. _____ Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M.M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'oeil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

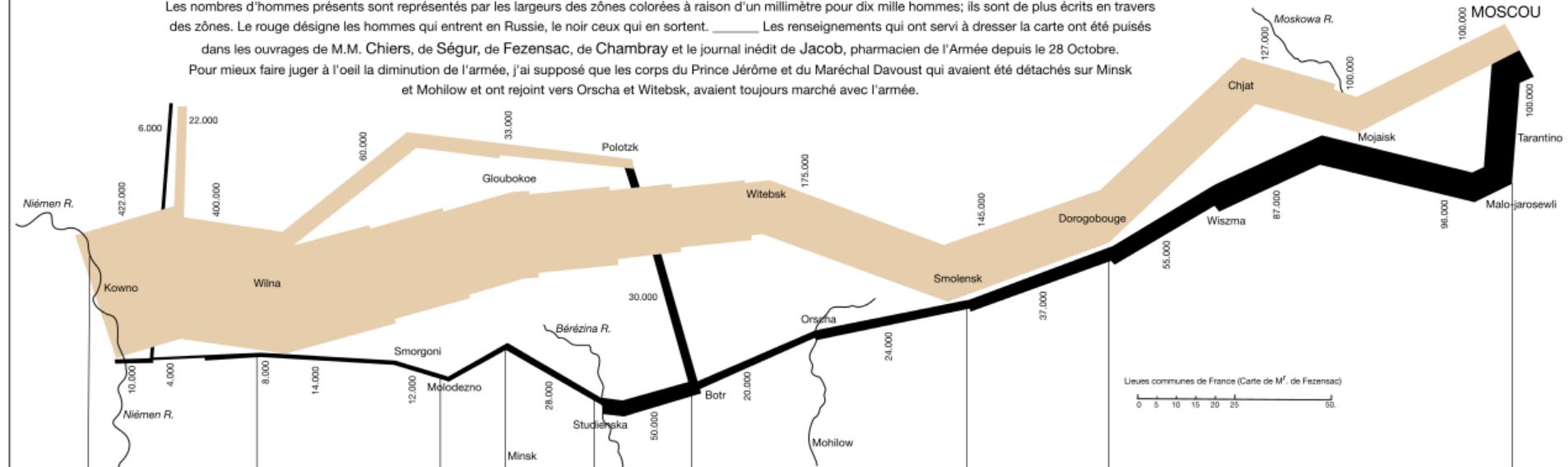


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosacques passent au galop
le Niémen gelé.

- 26.[°] le 7 X^{rme}

- 30.[°] le 6 X^{rme}

- 24.[°] le 1^{er} X^{rme}

- 20.[°] le 28 g^{me}

- 11.[°]

- 21.[°] le 14 g^{me}

- 9.[°] le 9 g^{me}

Zéro le 18 8^{me}
- 5
- 10
- 15
- 20
- 25
- 30

[Vectorization CC-BY-SA martingrandjean.ch 2014]

Imp. Lith. Regnier et Dourdet.

By Martin Grandjean - Own work <http://www.martingrandjean.ch/historical-data-visualization-minard-map/>, CC BY-SA 3.0,
<https://commons.wikimedia.org/w/index.php?curid=32985771>

Sve po redu ...



Sve po redu ...

Prvo je potrebno izmeriti signale.



Merenje započinje poznavanjem informacije od interesa!

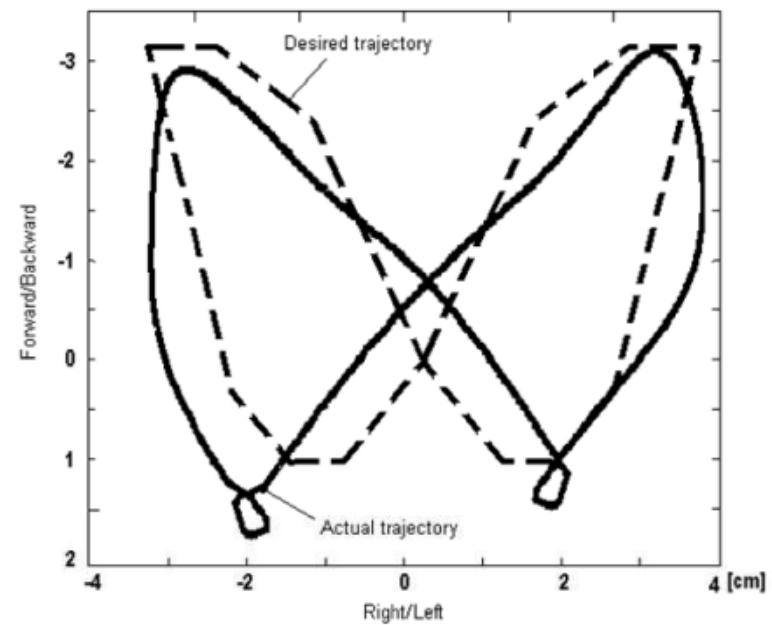
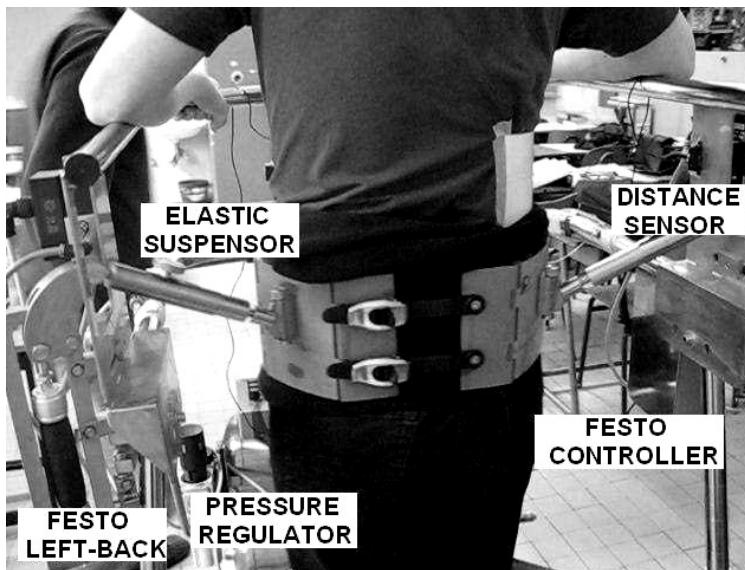
1. Koju fizičku veličinu je potrebno izmeriti ili koji deo fizičke veličine predstavlja informaciju od interesa?
2. U kom opsegu je fizička veličina od značaja za merenje?
3. Pod kojim uslovima se meri fizička veličina? Protokol?
4. Koje su karakteristike senzora koji se koristi?

Primer (projekat iz predmeta Neuralno inženjerstvo 2015. godine pod nazivom **“Upotreba IR senzora za merenje rastojanja”** studentkinje S. Poledice):

1. Meri se rastojanje – pomeranje kuka u transverzalnoj i sagitalnoj osi.
2. Fizička veličina je u opsegu do maksimalno 10 cm.
3. Meri se tokom hoda u Hodalici. Protokol omogućava merenje rastojanja sa površina različitih karakteristika.
4. Proizvođač? Ispitati (modelovanje karakteristike senzora i analiza uticaja vrste refleksione površine)?

Cilj merenja?

Cilj merenja je odrediti trajektoriju kuka tokom hoda u mehanizovanoj Hodalici u transverzalnoj i sagitalnoj ravni. Važno za upravljanje Hodalicom prilikom rehabilitacije pacijenata nakon moždanog udara i tokom njihovog vežbanja hoda.



Slike su preuzete i prilagođene iz rada: J. F. Veneman, S. Došen, N. Miljković, N. Jovičić, A. Veg, D. B. Popović, Member IEEE, and T. Keller, Member IEEE. A device for active posture assistance during over ground gait training. *Proc of the 1st International Conference on Applied Bionics and Biomechanics, ICABB-2010, CD, pp. 1-5, Venice, Italy, 2010.*

Protokoli

1. Ispitivanje karakteristika senzora

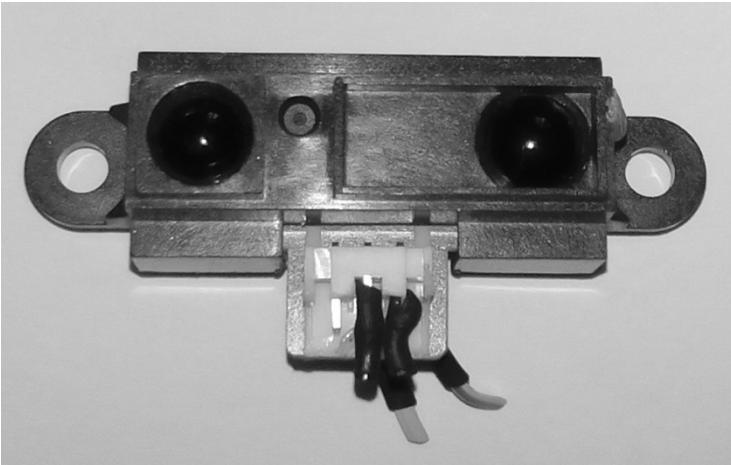
U eksperimentu su kao reflektujuce površine (površine koje će se naći na putu IR zraka) korišćeni: beli papir, crni papir, tamno siva površina sa rupicama-plastika (zvučnik) i plastična podloga bledo sive boje. Izmeren je izlaz za objekte postavljene na rastojanjima od 5 cm, 10 cm, 20 cm, 30 cm, 40 cm, 50 cm, 60 cm, 70 cm, 80 cm, 90 cm i 100 cm u trajanju od nekoliko sekundi.

2. Merenje pokreta kuka

Senzori su postavljeni na Hodalicu i tekstura ravne i nereflektujuće površine je postavljena na telo ispitanika. Pomeranje se meri u transverzalnoj i sagitalnoj ravni. Merenje se radi na modelu, zdravom ispitaniku i pacijentu¹.

1. J. F. Veneman, S. Došen, N. Miljković, N. Jovičić, A. Veg, D. B. Popović, Member IEEE, and T. Keller, Member IEEE. A device for active posture assistance during over ground gait training. Proc of the 1st International Conference on Applied Bionics and Biomechanics, ICABB-2010, CD, pp. 1-5, Venice, Italy, 2010

Senzor



Karakteristike senzora iz familije Sharp GP2D12 (preuzete iz *Data Sheet*-a senzora, na slici je prikazana nešto drugačija verzija iste familije senzora Sharp 2DI20X):

Opseg rastojanja: od 10 cm do 80 cm

Frekvencija/period ažuriranja: 25 Hz / 40 ms

Maksimalan ugao rasipanja svetlosti: >40°

Primenjen napona napajanja: od 4.5 V do 5.5 V

Šum na analognom izlazu: <200 mV

Prosečna vrednost struje: 33 mA

Rezultati

BELA POVRŠINA

d (cm)	5	10	20	30	40	50	60	70	80	90	100
sr. vr.	1.47	2.63	1.48	1.03	0.80	0.65	0.56	0.49	0.43	0.40	0.37
devijacija	0.03	0.04	0.03	0.02	0.03	0.02	0.03	0.02	0.02	0.03	0.03

CRNA POVRŠINA

d (cm)	5	10	20	30	40	50	60	70	80	90	100
sr. vr.	1.73	2.62	1.46	1.01	0.81	0.63	0.53	0.47	0.42	0.38	0.38
devijacija	0.03	0.03	0.02	0.03	0.09	0.03	0.04	0.03	0.03	0.03	0.03

TAMNOSIVA POVRŠINA SA RUPICAMA

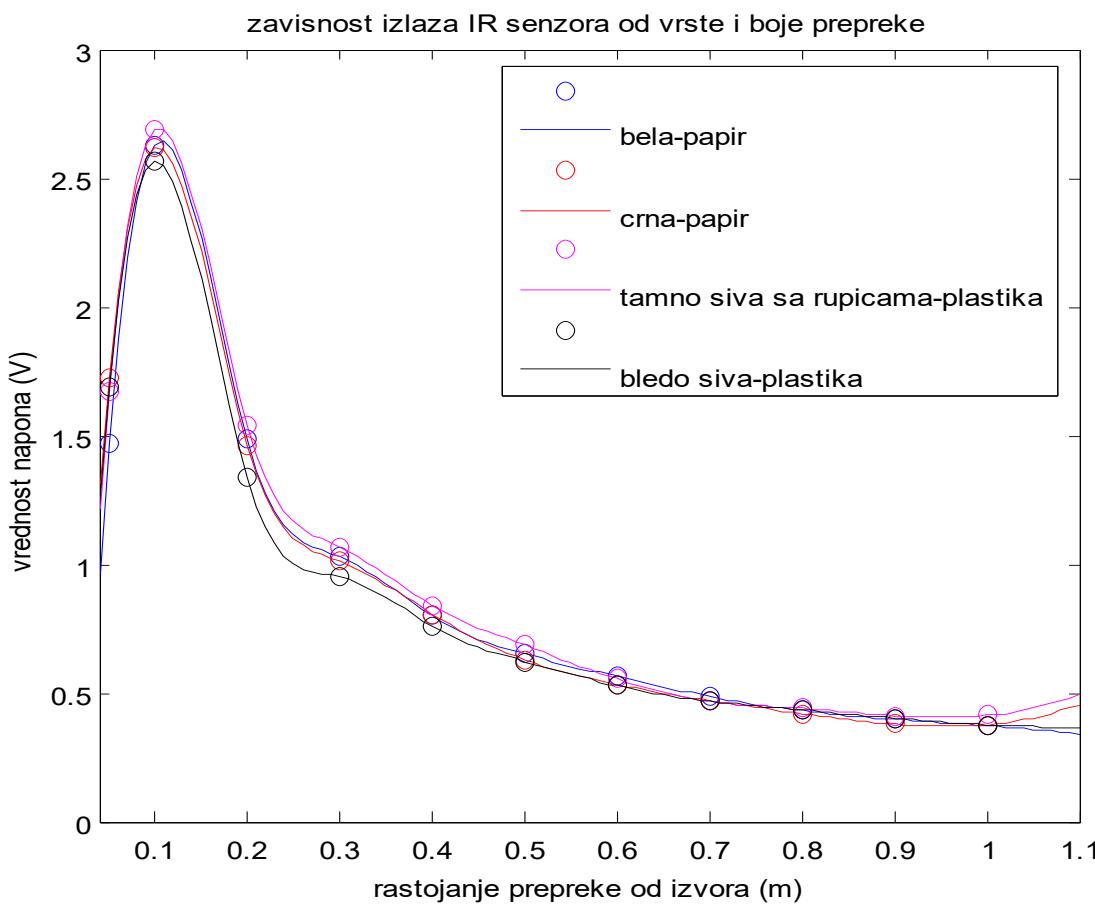
d (cm)	5	10	20	30	40	50	60	70	80	90	100
sr. vr.	1.67	2.69	1.54	1.07	0.84	0.68	0.55	0.47	0.44	0.41	0.41
devijacija	0.03	0.04	0.03	0.02	0.03	0.03	0.03	0.04	0.04	0.03	0.03

PLASTIKA BLEDOSIVE BOJE

d (cm)	5	10	20	30	40	50	60	70	80	90	100
sr. vr.	1.69	2.57	1.33	0.95	0.76	0.62	0.53	0.47	0.43	0.40	0.38
devijacija	0.03	0.05	0.03	0.03	0.03	0.03	0.03	0.04	0.02	0.04	0.04

Srednje vrednosti i standardne devijacije merenih izlaza senzora za različite tipove refleksionih površina i rastojanja senzora od izvora, date su u tabeli. Cilj ovog merenja je bio da se odredi karakteristika senzora.

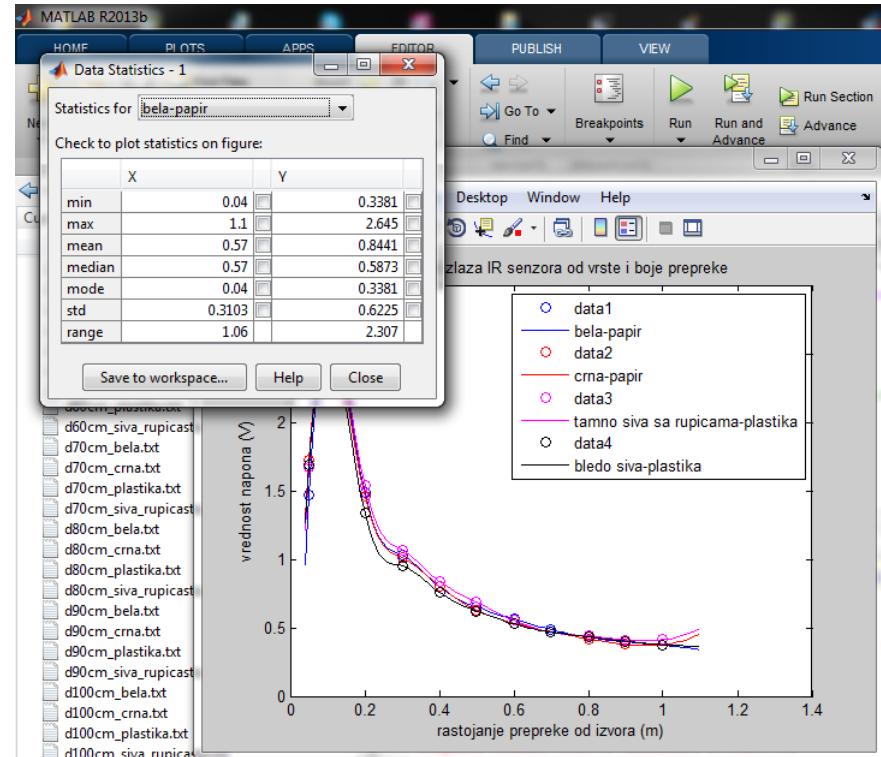
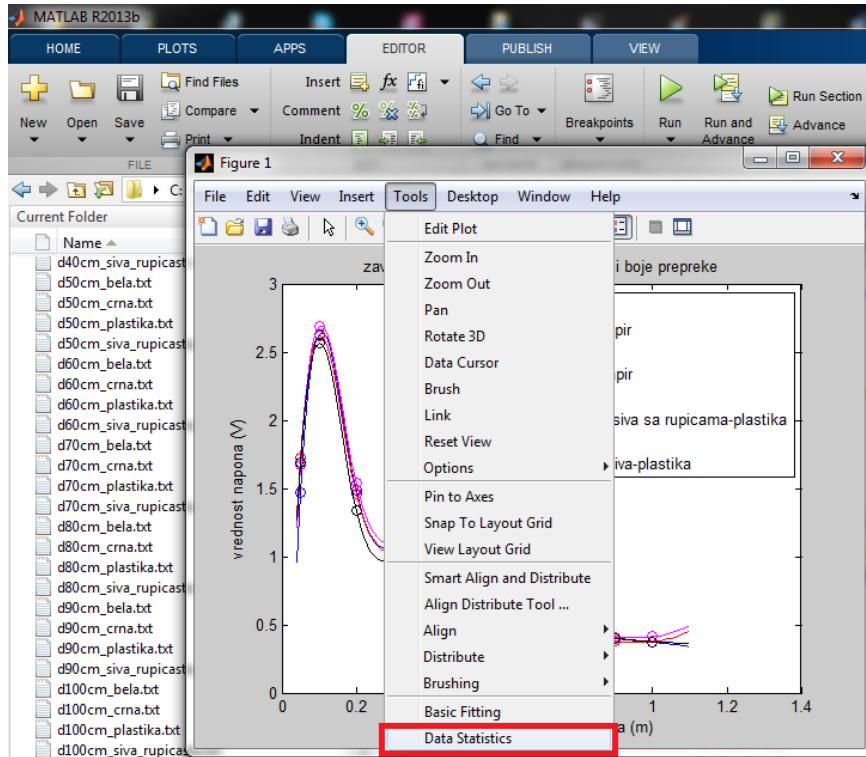
Karakteristika senzora



Interpolirana karakteristika izlaza senzora za različite tipove refleksionih površina dobijena na osnovu merenja izlaza senzora u konačnom broju rastojanja

Uporediti ove karakteristike sa karakteristikom koju je proizvođač dao u uputstvu za GP2D12 senzor, https://engineering.purdue.edu/ME588/SpecSheets/sharp_gp2d12.pdf.

Rezultati – deskriptivna statistika



Pokretanjem programa *interpolacija.m* dobija se grafik, kao na slici levo (ovaj prikaz je realizovan kao primer za Matlab R2013b). Odabirom opcija *Tools/Data Statistics* otvara se prozor kao na slici desno.

Zadatak – interpolacija

Dopuniti program *interpolacija* prema uputstvu koje je dato u komentarima u kodu (snimiti ga pod nazivom *interpolacijalmePrezime* na Desktopu računara).

Programski jezici za analizu

- Na KLI predmetu, studentkinje i studenti mogu koristiti i uz konsultacije sa dežurnim/om instalirati programe po izboru za merenje i analizu signala.
- Polinomijalno fitovanje u Python-u se najčešće obavlja korišćenjem funkcije numpy.polyfit, a više na <https://docs.scipy.org/doc/numpy/reference/generated/numpy.polyfit.html>
- Pored Matlaba, polinomijalna interpolacija je zastupljena i u GNU Octave-u, a više na <https://octave.org/doc/v4.4.1/Polynomial-Interpolation.html>
- Kratko uputstvo za fitovanje polinoma u R-u, može se naći na <https://datascienceplus.com/fitting-polynomial-regression-r/>



PODSETNIK: PREDSTAVLJANJE REZULTATA MERENJA

Za ovaj deo prezentacije korišćena je literatura:

N. Miljković, Metode i instrumentacija za električna merenja, Elektrotehnički fakultet u Beogradu, ISSN: 978-86-7225-057-2, 2016, doi: [10.5281/zenodo.1335249](https://doi.org/10.5281/zenodo.1335249).

Rezultat merenja

Merna nesigurnost je interval u kome se nalazi merna veličina sa određenom verovatnoćom.

Merna nesigurnost:

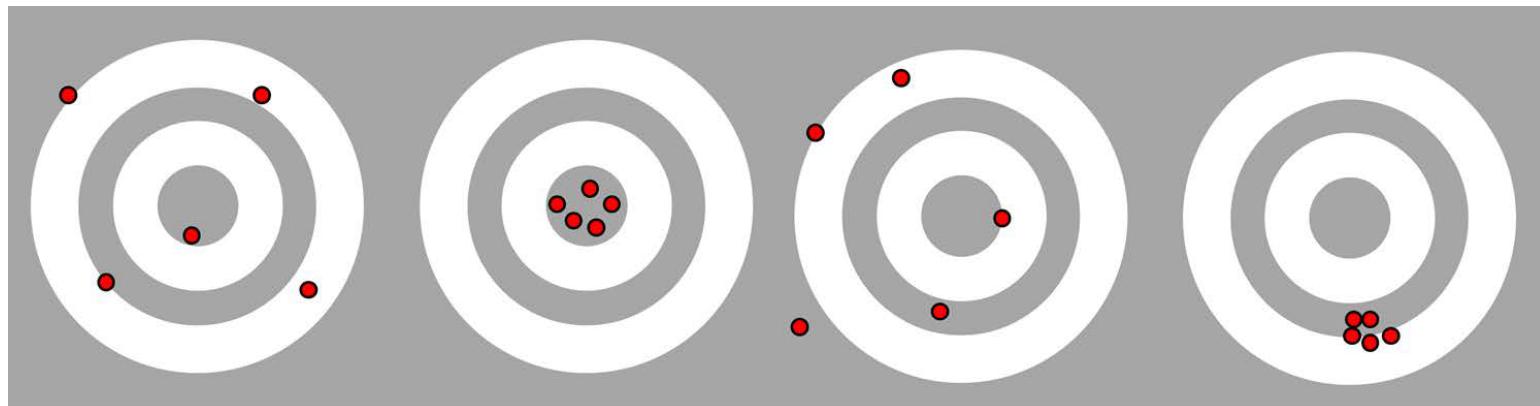
tipa A – određuje se statističkom analizom rezultata merenja (u_A)

tipa B – procenjuje se na osnovu znanja o mernoj metodi i postupku merenja, o karakteristikama instrumenta i svim ostalim podacima, osim uračunate merne nesigurnosti kroz postupak za tip A (u_B)

Kombinovana merna nesigurnost ili ukupna merna nesigurnost (u_C):

$$u_C = \sqrt{u_A^2 + u_B^2}$$

Tačnost i preciznost

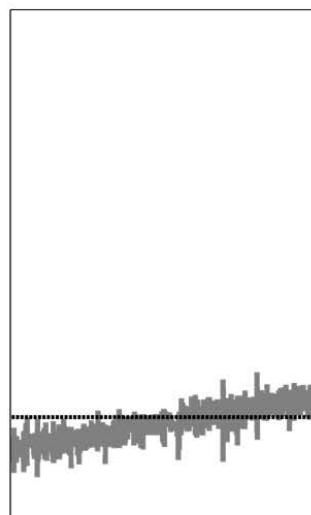


neprecizno i tačno

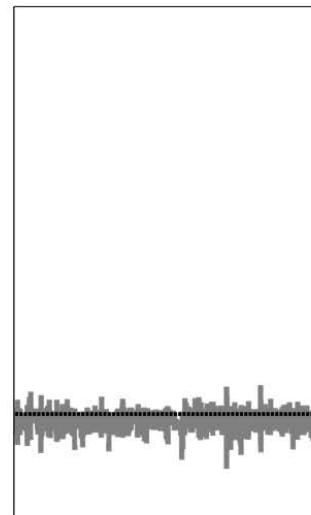
precizno i tačno

neprecizno i netačno

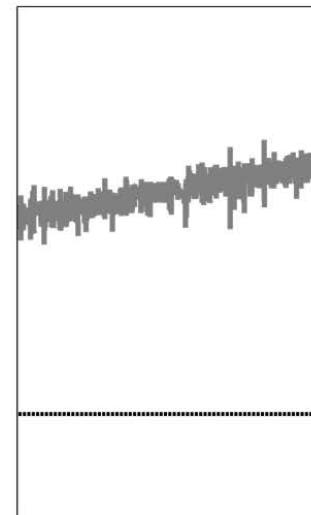
precizno i netačno



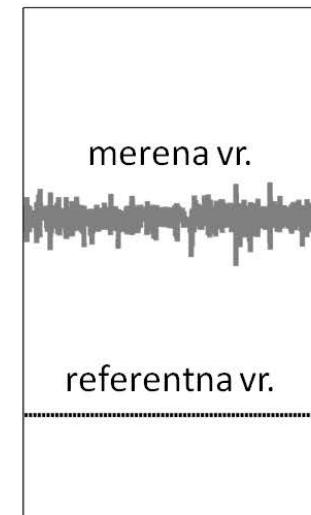
nestabilno i tačno



stabilno i tačno



nestabilno i netačno



merena vr.

referentna vr.

stabilno i netačno

Merna nesigurnost tipa A

Kako je srednja vrednost samo procena tačne vrednosti x_0 , to i dalje postoji neko ε koje se ne može zanemariti i koje predstavlja razliku procenjene i tačne srednje vrednosti.

$$x_0 = \bar{x} \pm \varepsilon$$

Standardna devijacija s (https://en.wikipedia.org/wiki/Standard_deviation) opisuje kako su greške raspodeljene (u opsegu oko srednje vrednosti).

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



otkud kvadrat, koren i $(n-1)$?

Merna nesigurnost tipa A – u_A

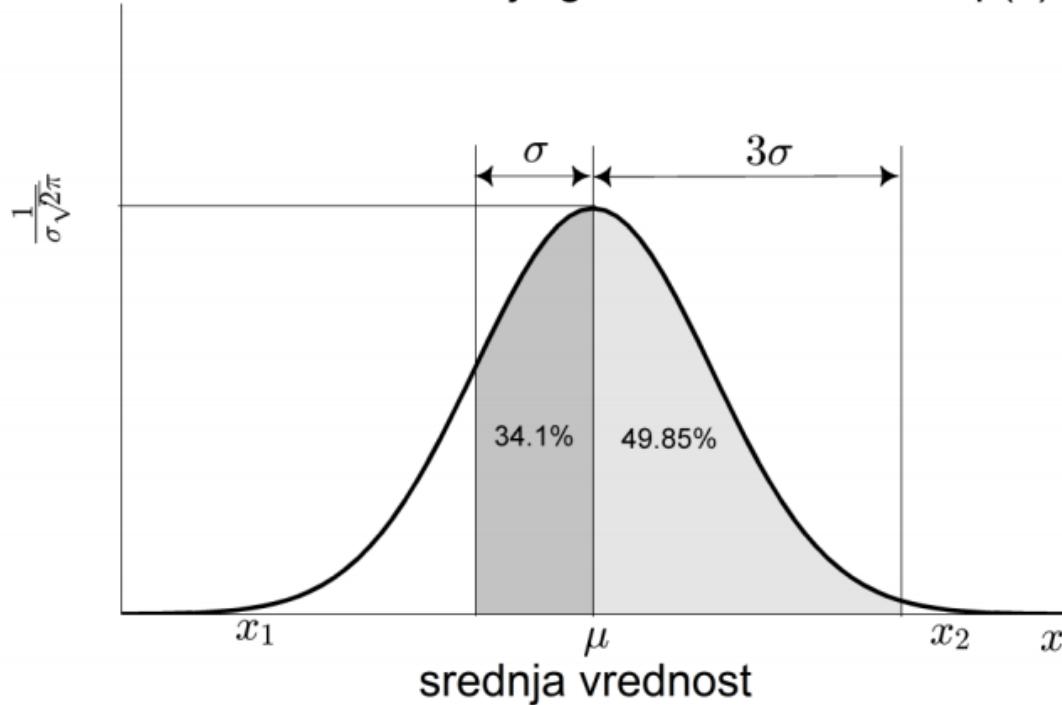
$$u_A = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}}$$

Kako “stvarne” greške prilikom merenja mogu sa podjednakom verovatnoćom da imaju i pozitivan i negativan predznak, potrebno je uvesti način da one sa istim težinskim faktorom budu uračunate u ukupnu mernu nesigurnost.

Umesto n^2 , koristi se $n(n-1)$ zbog [Beselove korekcije](#). Ovakva korekcija doprinosi većoj mernoj nesigurnosti, ali zato procena merne nesigurnosti ne zavisi od procene srednje vrednosti (za slučaj da nije unapred poznato odstupanje srednje vrednosti od populacije, već od uzorka).

Merna nesigurnost tipa A

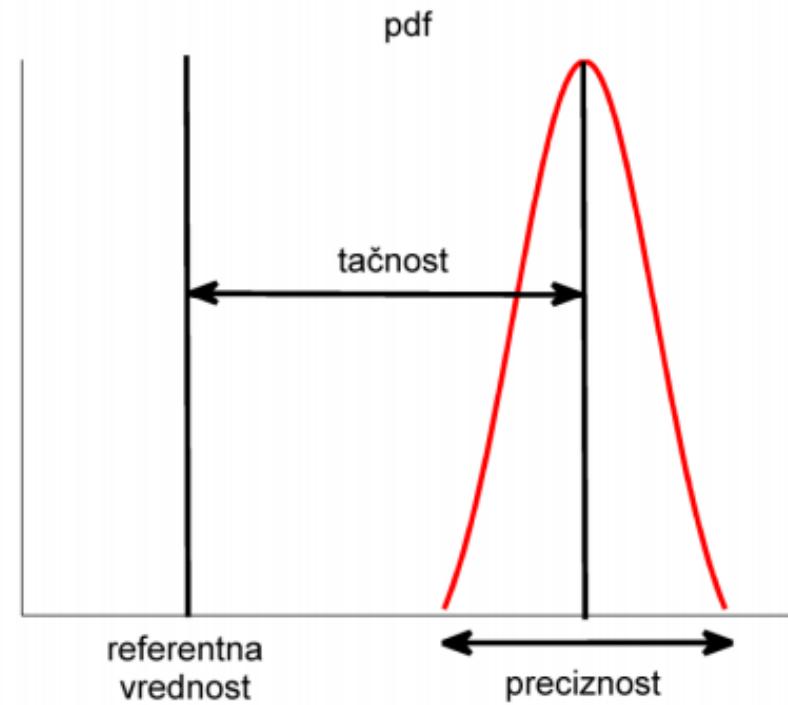
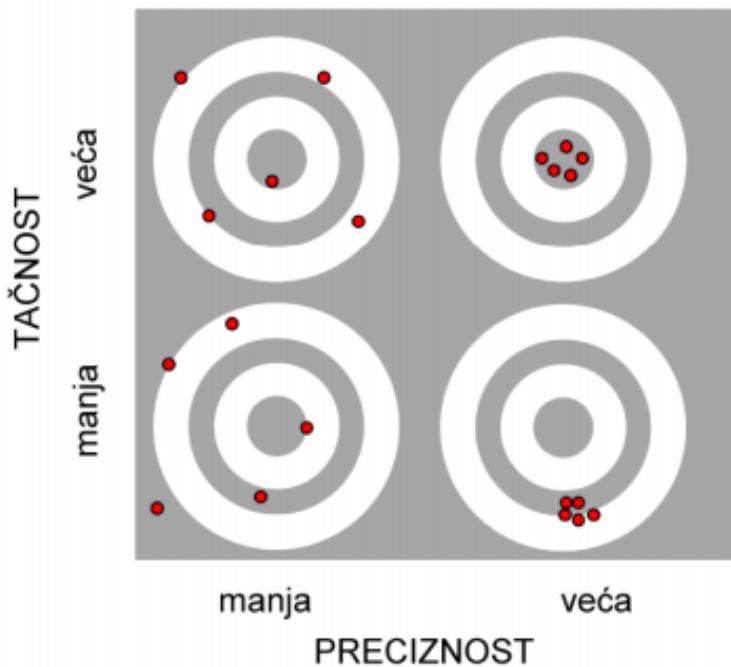
Gausova funkcija gustine verovatnoće $p(x)$



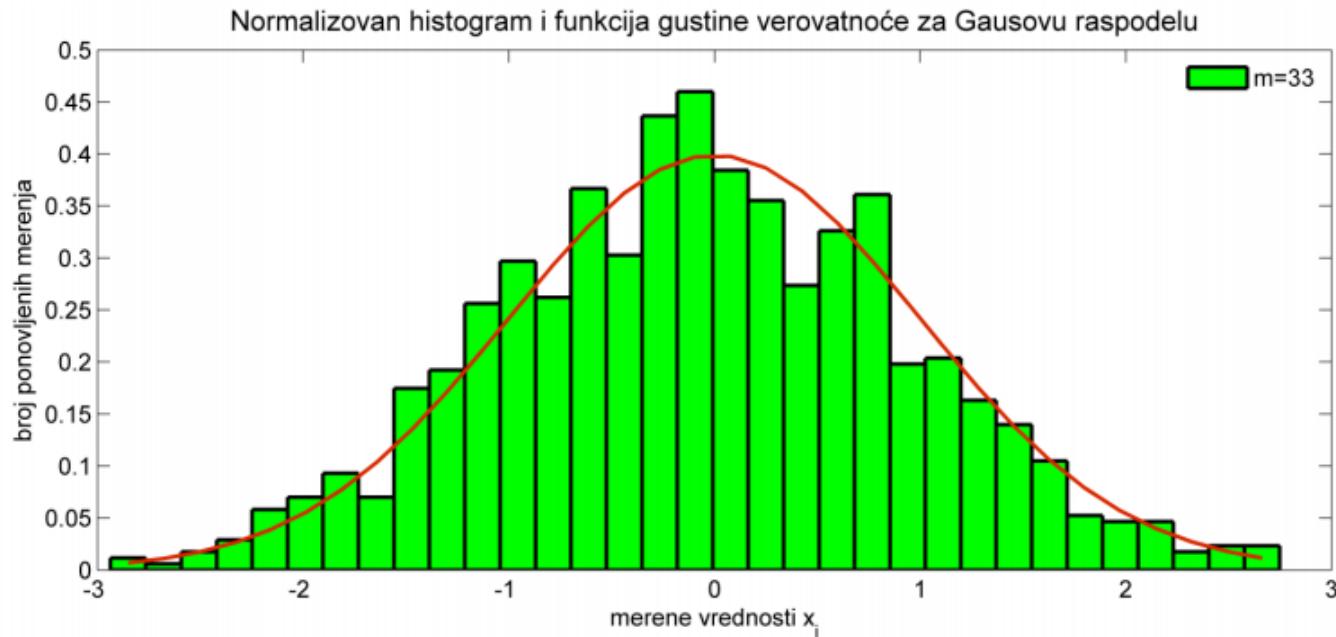
Mernoj nesigurnosti tipa A odgovara verovatnoća od 68%, tačnije 68.2.

Ukoliko se zahteva da verovatnoća bude 95%, računa se proširena merna nesigurnost sa faktorom proširenja 2 (verovatnoća od 95% odgovara 2σ). Za faktor proširenja 3, verovatnoća je 99.7% (analogno 3σ).

Tačnost i preciznost



Prikaz rezultata ponovljenih merenja



n – broj ponovljenih merenja

m – broj intervala histograma

u opsegu (x_{min}, x_{max}) , preporučljivo* je:

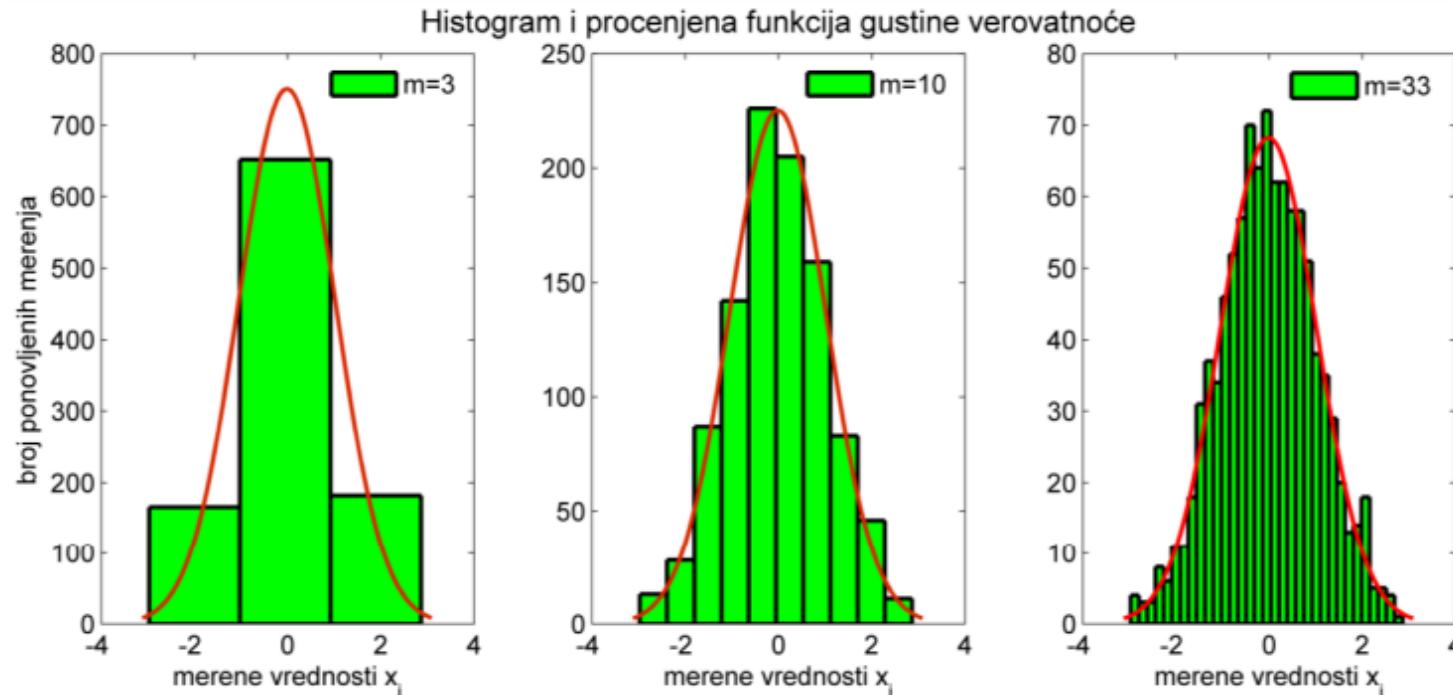
$$m \approx \sqrt{n} + 1$$

n_i je broj ishoda u intervalu opsega i , pa važi:

$$\sum_{i=1}^m n_i = n$$

*Za m koje je relativno malo, histogram "grubo" prikazuje raspodelu rezultata merenja, a ukoliko je m relativno veliko neki od intervala ostaju "prazni" tj. javljaju se prekidi u prikazu raspodele.

Prikaz rezultata ponovljenih merenja - histogram



p_i je relativan broj
ishoda u intervalu
opsega i

$$\sum_{i=1}^m p_i = 1$$

$$p_i = \frac{n_i}{n}$$

Broj ishoda u intervalu n_i se naziva i učestanost intervala.

Relativna učestanost intervala (p_i) se može tumačiti i kao verovatnoća intervala.

Gustina verovatnoće intervala se dobija kada se relativna učestanost podeli sa širinom intervala -> normalizovan histogram (kao na slici gore).

Indirektno merenje

$$u_y = \sqrt{\sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} u_{xi} \right)^2}$$

$$y = f(x_1, x_2, \dots, x_N)$$

y – indirektno/posredno merena veličina (npr. otpornost)

x_1, x_2, \dots, x_N – direktno merene veličine (npr. x_1 – struja merena ampermetrom i x_2 – napon meren voltmetrom)

merena veličina \equiv veličina procenjena merenjem

Izraz za u_y važi pod uslovom da su x_1, x_2, \dots, x_N nekorelisane

u_{xi} je kombinovana merna nesigurnost procenjena prilikom merenja merne veličine x_i

Kako predstavljati podatke?

- Svakako iz finalni rezultat treba prikazati i mernu nesigurnost, ali i opisati proceduru kako je računata merna nesigurnost, proceduru kako je realizovano merenje, ...
- Naravno, ovo se ne radi ako se merenje sprovodi rutinski i ako nije u fokusu: npr. merenje otpornosti otpornika primenom unimera prilikom realizacije električnog kola nije merenje od značaja. Računati mernu nesigurnost za takvo merenje je čisto gubljenje vremena.
- Sa druge strane, za merenja koja su od suštinskog značaja (ranije je to bila provera etalona, ali sa novim SI sistemom, pogledati https://en.wikipedia.org/wiki/2019_redefinition_of_SI_base_units, to bi mogao biti neki drugi primer) potrebno je primeniti što kompletniju tj. kompleksniju proceduru za prikaz rezulata.
- O novim usvojenim definicijama SI jedinica, dostupni su odgovarajući materijali i vesti na sajtu Direkcije za mere i dragocene metale (<http://www.dmdm.rs/lt/Vesti2018.php>).

(Ne tako) savremeni trendovi

- Nedavno je prof. Irizari sa Harvard Univerziteta poslao otvoreno pismo urednicima naučnih časopisa (<https://simplystatistics.org/2019/02/21/dynamite-plots-must-die/>, pristupljeno 07.03.2019.) koje grafike treba, a koje ne treba koristiti.
- Ideja je da grafik prikaže što više podataka na što precizniji način.

Obrada signala započinje informacijom...

Za uspešnu primenu tehnika procesiranja signala, potrebno je što više informacija o mernoj veličini i mernoj metodi.

Poznavanje prirode tj. porekla merene veličine i smetnji za korišćenu mernu metodu je od primarnog značaja. U cilju primene najbolje metode, korisnik mora da zna cilj procesiranja i karakteristike merenog signala.

Obrada signala ima za cilj da izdvoji relevantne informacije iz signala. Nekad je relevantna informacija određena fizička veličina, a nekad je ona deo fizičke veličine.

Deterministički i stohastički signali

Deterministički su oni signali koji se mogu opisati matematičkim izrazom (npr. EKG signal i karakteristike QRS kompleksa su blisko deterministički signali).

Stohastički ili slučajni signal se ne može opisati "formulom", već samo statistički koristeći verovatnoću (npr. varijabilnost srčanog ritma tj. promene RR intervala).

Za stohastički ansambl koji čine slučajni procesi $s_1(t), s_2(t), \dots, s_n(t)$ može se definisati združena funkcija raspodele verovatnoće (*joint probability density function - JPDF*):

$$P[s_1(t) \leq s_1, s_2(t) \leq s_2, \dots, s_n(t) \leq s_n] = P(s_1, s_2, \dots, s_n)$$

Stohastički signal karakteristiše i gustina funkcije verovatnoće (*probability density function – PDF*), a ona se definiše kao N -ti parcijalni izvod funkcije JPDF:

$$p(s_1, s_2, \dots, s_n) = \frac{\partial^n}{\partial s_1 \partial s_2 \dots \partial s_n} [P(s_1, s_2, \dots, s_n)]$$

Filtriranje

Osnovni alat za procesiranje signala je filtriranje.

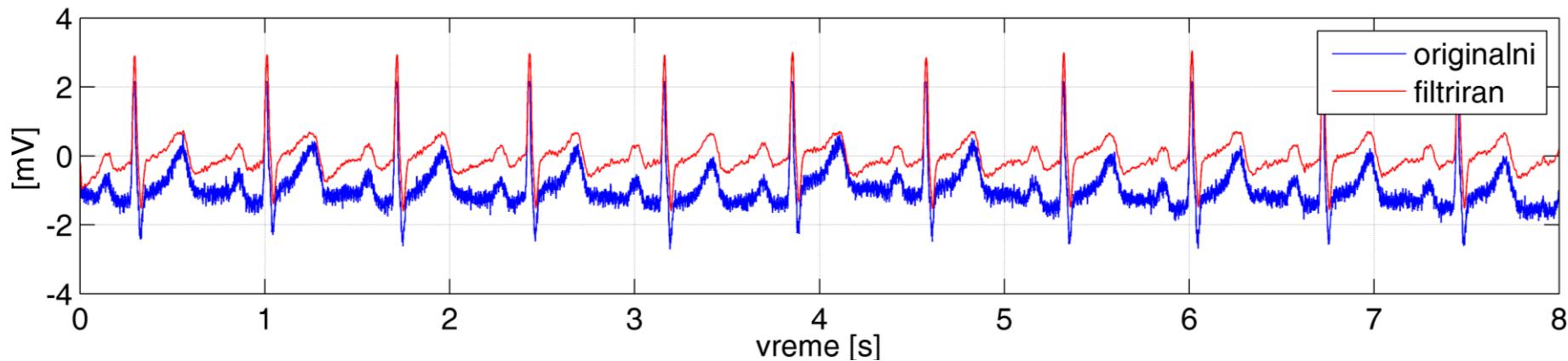
Filtriranje se može primeniti i u vremenskom i u frekvencijskom domenu.

Najčešće se primenjuje u frekvencijskom domenu.

Filtriranje se koristi u obradi determinističkih i stohastičkih signala.

Pri procesiranju stohastičkih signala, Furijeova transformacija se primenjuje na autokorelacionu funkciju pre nego na samu funkciju. Tada se određuje gustina spektralne snage.

Metode filtriranja



Kako bi izgledale promene na FFT-u kod filtriranog signala koji je prikazan na slici izna?

Linearno filtriranje

Optimalno filtriranje

- Wiener-ov filter (minimizacija srednje kvadratne greške)
- “prilagođen” filter (maksimizacija odnosa signal/šum)

Adaptivno filtriranje

Pored gore navedenih vrsta filtara, filtri mogu biti i:

- digitalni i analogni filtri
- IIR i FIR filtri
- linearni i nelinearni
- ...

Analiza signala u vremenskom domenu

Uobičajene metode za redukciju šuma su:

- filtriranje u vremenskom domenu,
- sinhronizovano usrednjavanje i
- druge.

Uobičajene metode za ekstrakciju informacija su:

- auto i kros-korelaciona funkcija,
- usrednjavanje (*averaging*),
- procena amplitude,
- prolazak kroz nulu (*zero-crossing*),
- procena talasne kompleksnosti,
- rekurentna analiza i
- druge.

Filtriranje signala u vremenskom domenu

Prednosti u odnosu na filtriranje u frekvencijskom domenu su:

- može da bude brže u većini slučajeva,
- nije neophodno direktno određivanje spektralnih karakteristika signala i
- nije neophodno određivanje spektralnih karakteristika šuma.

Služi za potiskivanje "slučajnog" šuma (*šum*) za ponovljene sekvence signala (*n* sekvenci *sig* signala), kada se spektri signala i šuma preklapaju.

$$\frac{1}{n} \sum_{k=1}^n sigSaŠumom = \frac{1}{n} \sum_{k=1}^n sig + \frac{1}{n} \sum_{k=1}^n šum$$

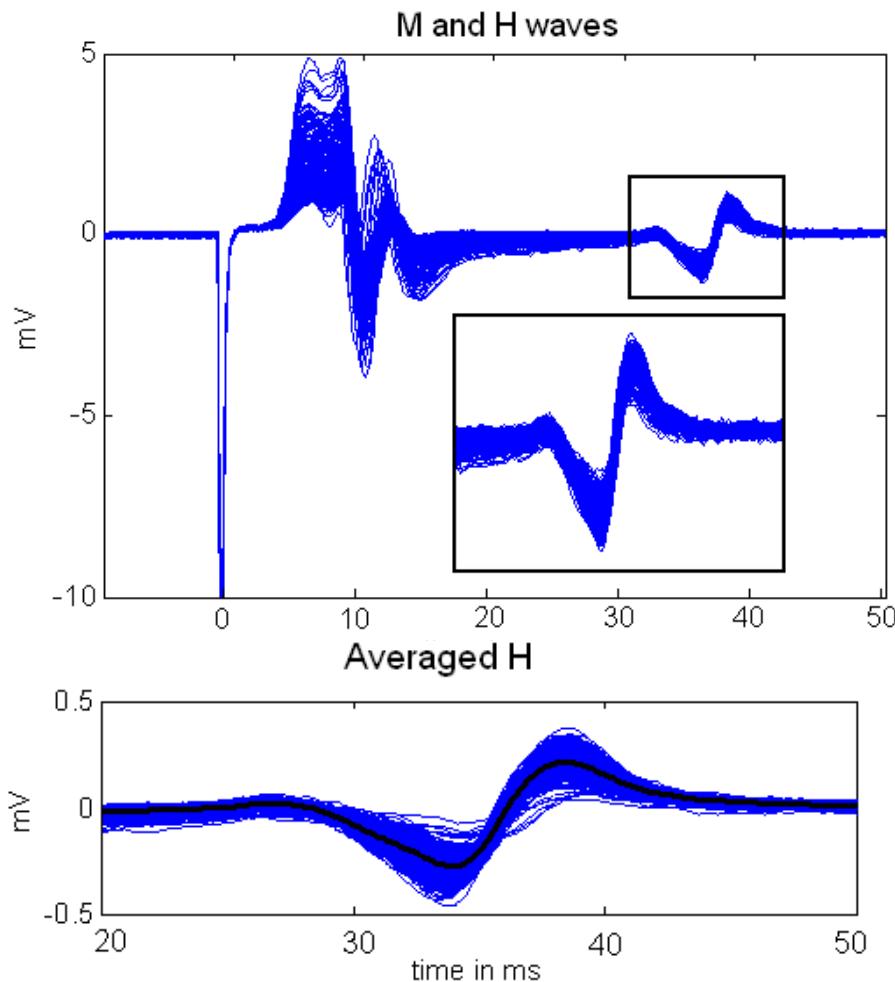
Postoje i ograničenja u primeni ove metode:

- Neophodno je obezrediti odgovarajuću sinhronizaciju i triger signala.
- Neophodno je da postoje trigerovane sekvence signala iste dužine.

Primenjuje se npr. kod merenja brzine provođenja senzornih nerava.

Sinhronizovano usrednjavanje

Primer: H-refleks



Triger signala obeležen je na slici 0.

Sinhronizovane sekvence M i H talasa su date na slici.

Brzina provođenja se računa po dodejivanjem mesta karakterističnih tačaka H talasa (minimuma i maksimuma)

Na donjoj slici izdvojene su ponovljene sekvence H-refleksa sa usrednjjenim rezultatom.

Vremensko usrednjavanje

Koristi se kada je potrebno odrediti tehniku procesiranja kada postoji samo jedna epoha signala ili samo jedan događaj od interesa.

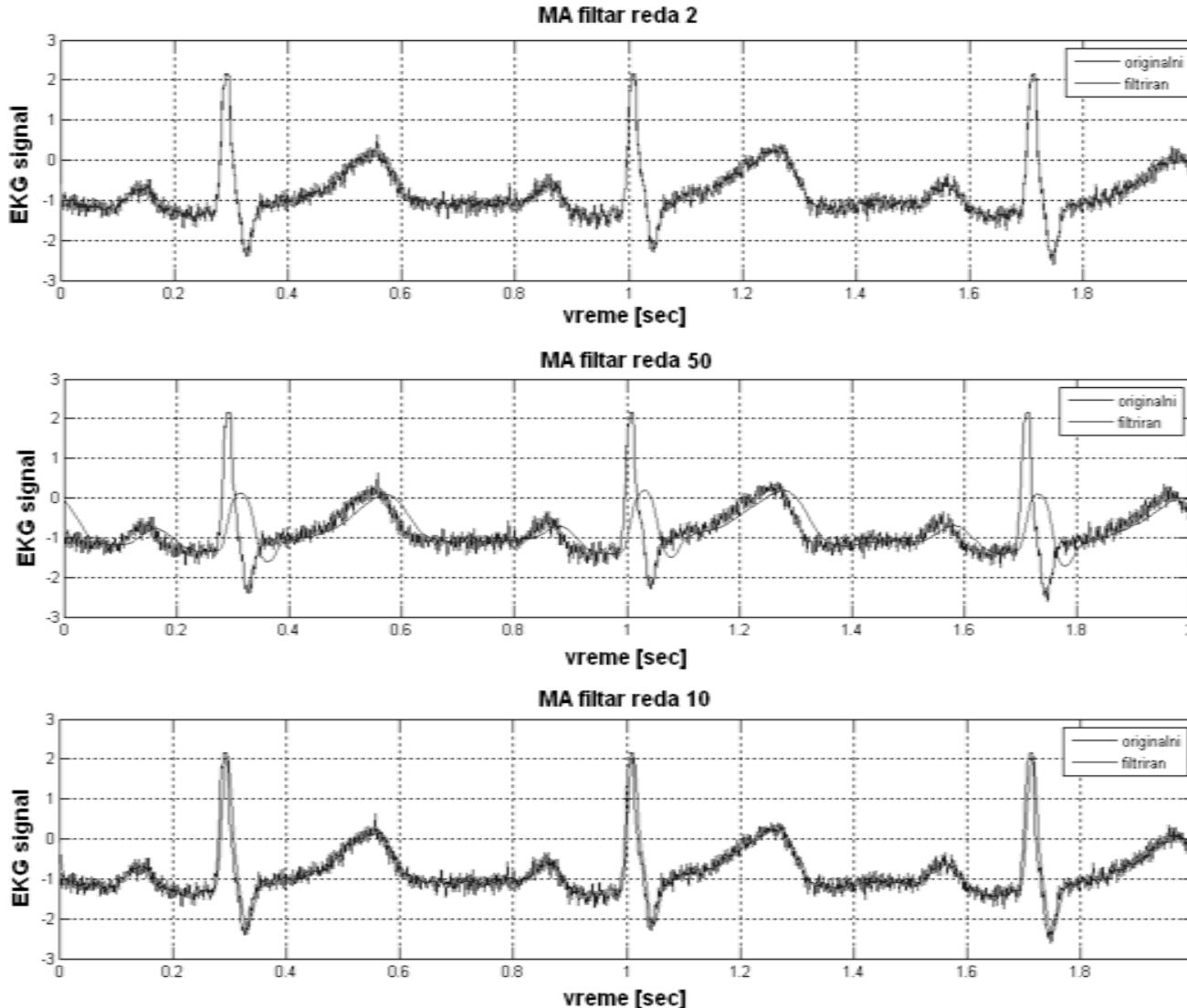
Za razliku od sinhronog usrednjavanja, ovde statistici podleže proces umesto ansambla: *Moving Average* (MA) filtri ili tehnika usrednjavanja sa pokretnim prozorom.

$$y(n) = \sum_{k=0}^N b_k x(n - k)$$

U prethodnoj relaciji sa y je označen izlaz filtra, sa x ulaz filtra, sa N red filtra, sa b_k koeficijenti filtra.

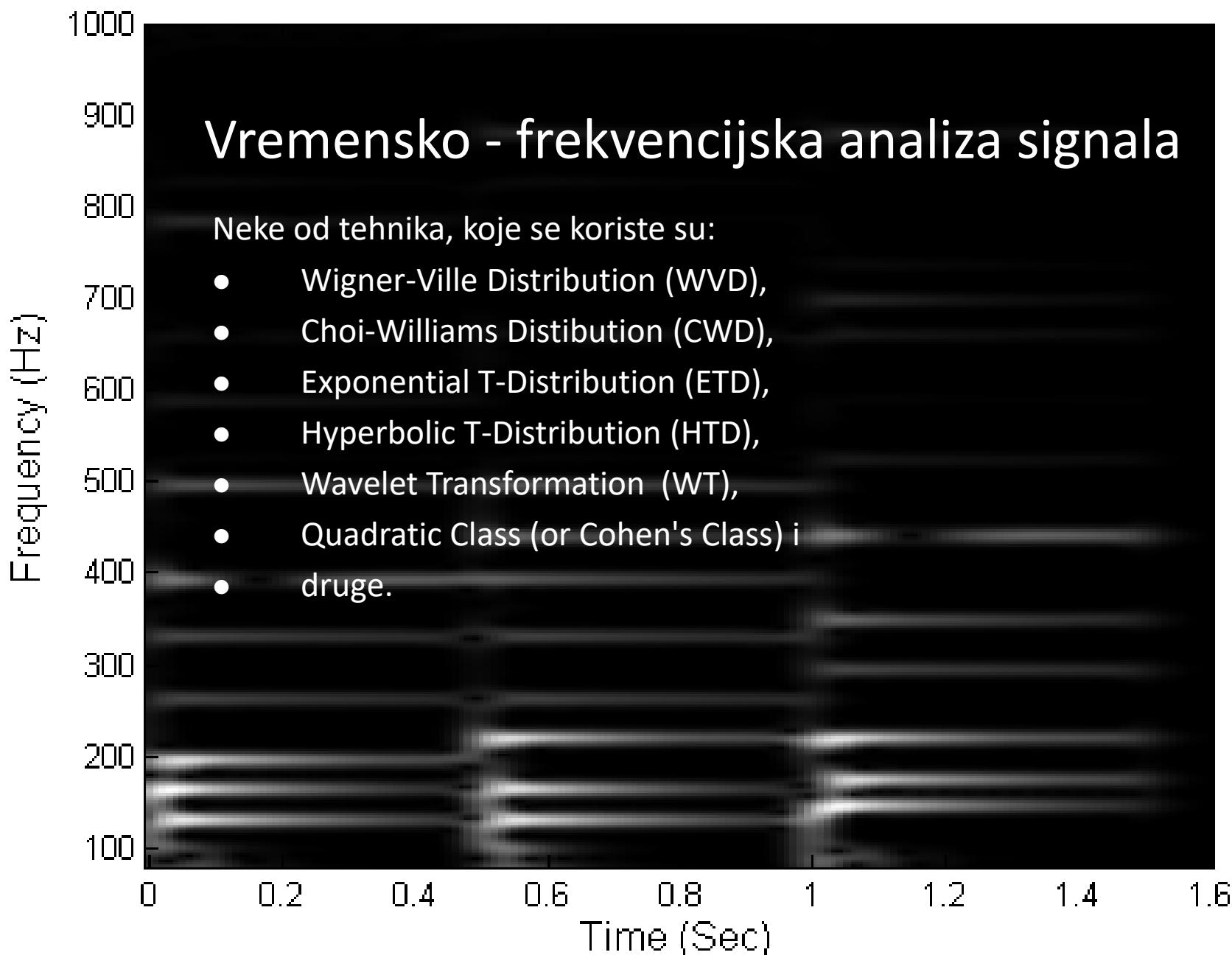
MA filter "liči" na VF filtriranje. MA filter je linearan filter.

Potiskivanje šuma MA filtrom u EKG signalu

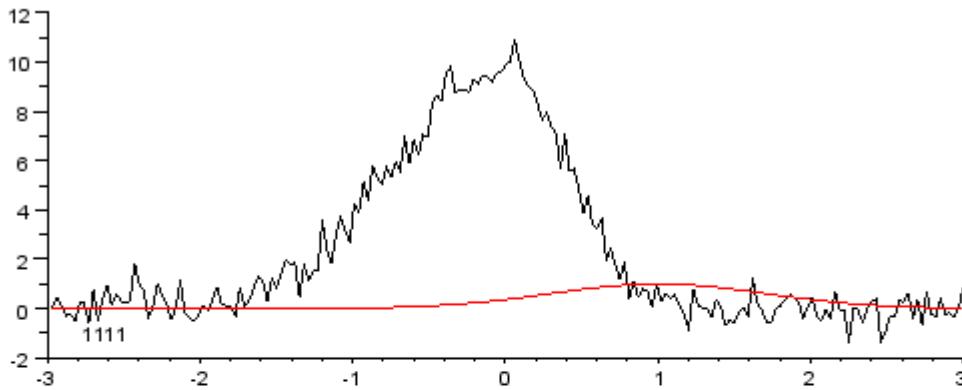


Prikazani su podaci koji su korišćeni na predmetu [Metode analize elektrofizioloških signala](#).

Gabor Transform



Zadatak – prikaz rezultata



Modifikovana slika: By Cdang (Christophe Dang Ngoc Chan) - Own work, CC BY-SA 3.0,
<https://commons.wikimedia.org/w/index.php?curid=23019015>.

Simulirati signale u programskom paketu po izboru primenom pseudo slučajnog generatora brojeva. Predstaviti rezultate u formi histograma. Predstaviti rezultate u formi box plot-a (https://en.wikipedia.org/wiki/Box_plot), izračunati srednju vrednost i standardnu devijaciju rezultata merenja.

DODATNI ZADATAK:

Prikazati izmerene odbirke kontinualno, diskretno i prikazati signal koji je fitovan polinomom 4. reda.