

Obrada informacija: Prva laboratorijska vježba

1. Učitajte biblioteke: NumPy, matplotlib.pyplot, scipy.fft, signal iz scipy te pywt.

Rijeke imaju veliki gospodarski značaj, od termoelektrana i nuklearnih elektrana kojima su rijeke nužne za rad, javne vodoopskrbe do poljoprivrede i transporta. Premda ljudi danas uvelike upravljaju vodotocima, uslijed vremenskih prilika ipak može doći do poplava, erozija tla i sličnih prirodnih nepogoda. Kako bismo se mogli na vrijeme pripremiti za potencijalne probleme, potrebno je istražiti ponašanje rijeka na temelju višegodišnjih praćenja vodostaja. U ovoj laboratorijskoj vježbi ćete analizirati vodostaje rijeke Save kroz period od 25 godina.

Unutar komprimiranog direktorija u kojem se nalazi ova bilježnica, nalazi se i .mat datoteka s podacima o vodostaju rijeke Save u periodu od 1.1.1982. do 31.12.2007. Svaki podatak označava jedan dan mjerenja vodostaja.

Vježba se izvodi u Pythonu/Google Colabu, a sve naredbe potrebne za provedbu vježbe te njihova objašnjenja dana su u predavanju. Ova laboratorijska vježba nosi 3 boda.

Kad ste gotovi s vježbom, na Moodle postavite .pdf izvješće s vježbe (.pdf izvješće možete generirati s naredbom File-Print-Location: PDF).

```
# Ovo je mjesto na kojem možete izvoditi svoj kod.
import numpy as np
import matplotlib.pyplot as plt
from scipy import signal
import pywt
```

2. Učitajte podatke o vodostaju iz .mat datoteke koja se nalazi unutar istog komprimiranog direktorija. Učitani podaci su spremljeni u rječnik. Izdvojite ključ i pripadne podatke o vodostajima te po potrebi smanjite dimenziju.

Savjet: .mat datoteku učitajte sa svog Google Drivea (unutar mape Colab Notebooks na svom Google Driveu postavite .mat datoteku). Ostali savjeti nalaze se u sljedećem programskom odsječku.

```
# Ovo je mjesto na kojem možete izvoditi svoj kod.
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
# Nakon što pokrenete ovaj odsječak koda, pojavit će se link u prozoru ispod. Kliknite na link, prijavite se sa svojim
# dobiti kod koji trebate kopirati u izlaz ispod ovog prozora (u prozor "Enter your authorization code: "). Nakon što
# pojavit će se poruka "Mounted at /content/drive".
# Ako unutar takvog Colab Notebook direktorija imate spremljenu datoteku "OIkobas.mat", možete joj pristupiti sljedeći
#from scipy import io
#kobas_iz_matlaba=io.loadmat("drive/My Drive/Colab Notebooks/OIkobas.mat")
#print(kobas_iz_matlaba)
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", forc
```



3. Ispišite: broj dana za koje postoje podaci o vodostajima, srednju vrijednost vodostaja, standardnu devijaciju, minimalnu vrijednost, maksimalnu vrijednost te median.

```
# Ovo je mjesto na kojem možete izvoditi svoj kod.
```

```
from scipy import io
kobas_iz_matlaba=io.loadmat("/content/drive/My Drive/Colab Notebooks/OIkobas.mat")
#print(kobas_iz_matlaba)
```

```

podatci = kobas_iz_matlaba.get('kobas')
#print(podatci)

i = 0
for x in podatci:
    if (np.dtype(x[0]) == 'uint16' and x[0] != 0):
        i += 1

print("Broj dana za koje postoje podatci o vodostajima je:", i)
print("Srednja vrijednost vodostaja je:", np.round(podatci.mean(), 3))
print("Standardna devijacija podataka je:", np.round(podatci.std(), 3))
print("Najmanji zabilježeni vodostaj je:", podatci.min())
print("Najveći zabilježeni vodostaj je:", podatci.max())
print("Medijan podataka o vodostajima je:", np.median(podatci))

```

```

Broj dana za koje postoje podatci o vodostajima je: 9496
Srednja vrijednost vodostaja je: 294.258
Standardna devijacija podataka je: 183.092
Najmanji zabilježeni vodostaj je: 17
Najveći zabilježeni vodostaj je: 878
Medijan podataka o vodostajima je: 255.0

```

4. Nacrtajte vodostaj. Obilježite x i y os te naslov slike.

```

# Ovo je mjesto na kojem možete izvoditi svoj kod.
import numpy as np
from scipy.fft import fft, ifft
import matplotlib.pyplot as plt

podatci2 = []
for z in podatci:
    podatci2.append(z[0])
#print(podatci2)

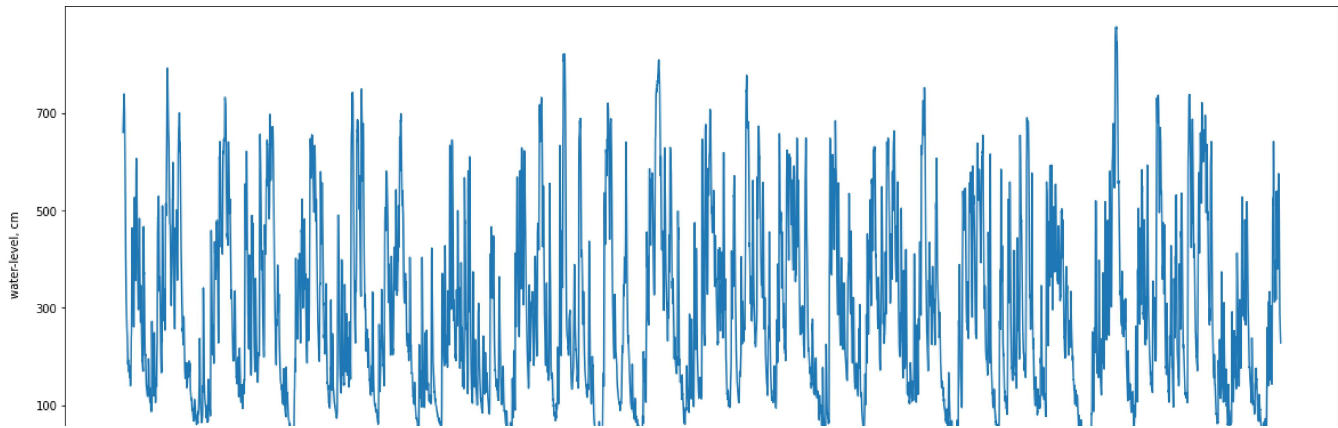
f = plt.figure().set_figwidth(20)
#plt.stem(podatci2) # ova opcija je teža za renderanje, duže traje
plt.plot(np.arange(0, 9496, 1), podatci2)

plt.xlabel("days")
plt.ylabel("water-level, cm")

plt.yticks(np.arange(-100, 900, step=200))
plt.xticks(np.arange(0, 9500, 365))

plt.show()

```



5. Izračunajte diskretnu Fourierovu transformaciju zadanog vodostaja. Ispišite dobivene vrijednosti. Nacrtajte apsolutnu vrijednost dobivenih koeficijenata.

Ovo je mjesto na kojem možete izvoditi svoj kod.

```
X = fft(podatci2)
```

```
print(X)
```

```
plt.figure().set_figwidth(13)
```

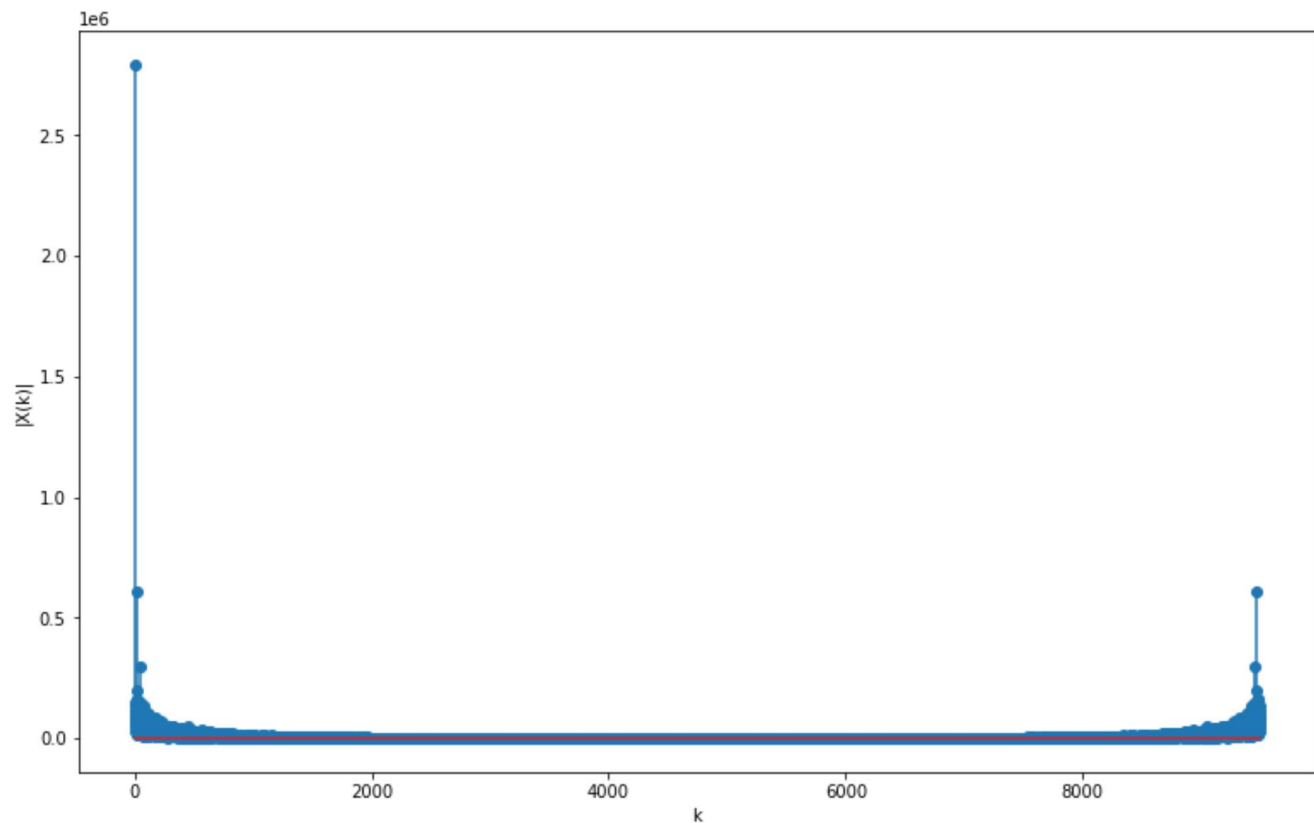
```
plt.stem(abs(X)) # ovdje sam izabrao .stem() jer je takav prikaz i u ppt
```

```
plt.xlabel("k")
```

```
plt.ylabel("|X(k)|")
```

```
plt.show()
```

```
[2794270.          -0.j          -9312.52756202+49784.33987081j
 80617.82235263-75905.53826228j ... -75144.8689581 -18406.78122837j
 80617.82235263+75905.53826228j  -9312.52756202-49784.33987081j]
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:7: UserWarning: In Matplotlib 3.3 individu
import sys
```



6. Izračunajte Fourierovu transformaciju na vremenskom otvoru zadanog vodostaja koristeći pravokutni otvor širine 2 godine. Prikažite rezultat pomoću pcolormesh. Odgovorite: koje frekvencije su vidljive u vodostaju? Koje godine se javljaju?

```
# Ovo je mjesto na kojem možete izvoditi svoj kod.
T = 365
fs = 365
window = 'boxcar'
f, t, Zxx = signal.stft(podatci2, fs, window, nperseg = 2*T)

#print(f)
#print(t)
#print(np.ndarray.flatten(Zxx))

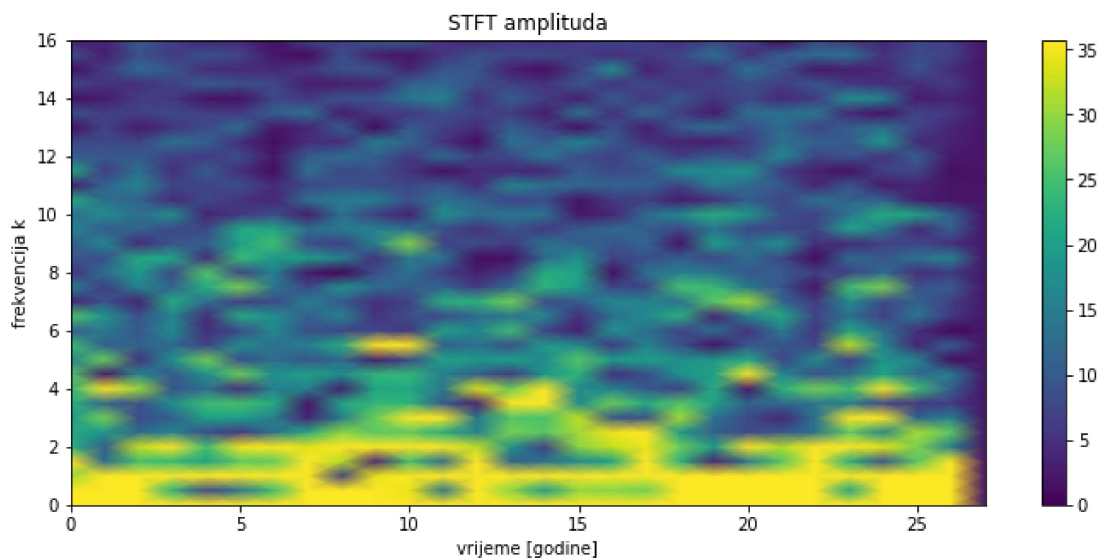
fig = plt.figure()
fig.set_figheight(5)
fig.set_figwidth(12)

plt.pcolormesh(t, f, np.abs(Zxx), vmin = 0, vmax = np.abs(Zxx.max()) / 10, shading = 'gouraud')
plt.colorbar()

plt.title("STFT amplituda")
plt.ylabel("frekvencija k")
plt.xlabel("vrijeme [godine]")

plt.ylim(top = 16)    # restrict na prikaz vrijednosti grafa

plt.show()
```



ODGOVOR: Vidljive su frekvencije:

- **0 godina** (nema smisla previše govoriti o tome - to bi bila kontinuirana promjena)
- **1 godina** - približno ponavljajući uzorak svake godine
- **2 godine** - malo manje izražen i pravilan kao prethodni, ali vidljiv uzorak kod frekvencije od 2 godine
- ostale su u djelovima ponavljajuće, ali nedovoljno da bih mogao zaključiti da su prisutne kroz sav vremenski interval od 26 godina mjerenja

7. Upotrijebite drugu širinu otvora (po izboru) i ponovite prethodni zadatak. Ispišite koju širinu ste upotrijebili. Objasnite razlike

```
# Ovo je mjesto na kojem možete izvoditi svoj kod.
T = 365
fs = 365
window = 'boxcar'
f, t, Zxx = signal.stft(podatci2, fs, window, nperseg = 8*T)

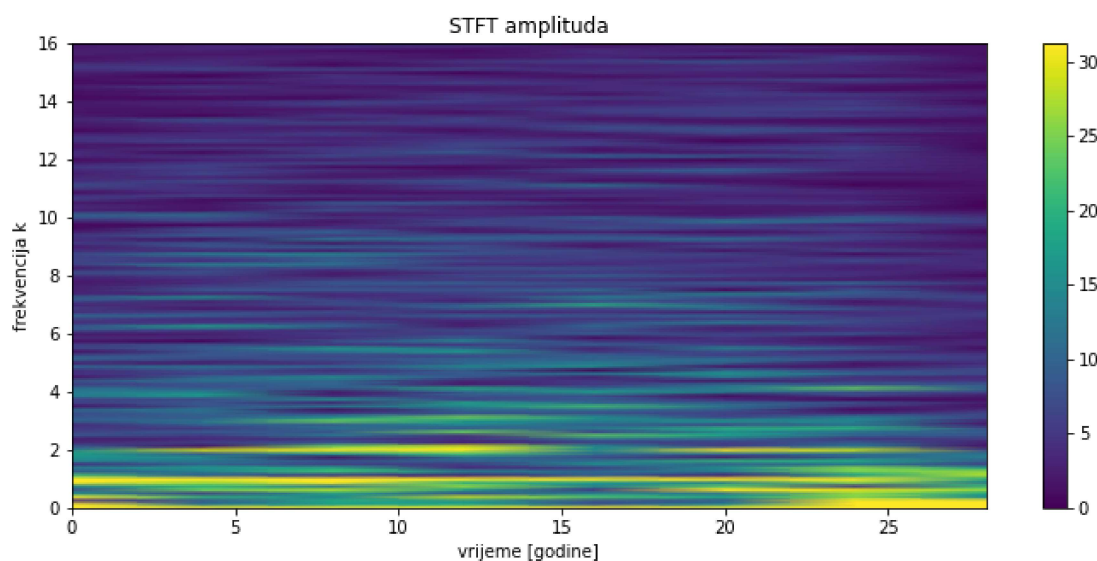
fig = plt.figure()
fig.set_figheight(5)
fig.set_figwidth(12)

plt.pcolormesh(t, f, np.abs(Zxx), vmin = 0, vmax = np.abs(Zxx.max()) / 10, shading = 'gouraud')
plt.colorbar()

plt.title("STFT amplituda")
plt.ylabel("frekvencija k")
plt.xlabel("vrijeme [godine]")

plt.ylim(top = 16)

plt.show()
```



ODGOVOR: Promjenom širine otvora (u konkretnom sam slučaju izabrao povećanje okvira na 8 godina) događa se sljedeće:

Povećala se jasnoća (preciznost, oštrina) u *vertikalnom smjeru*, odnosno jasnije se razaznaju bitne frekvencije - npr. iz ovakvog je grafa vidljivo da je i bitna frekvencija od 0,5 godina (u prethodnom se grafu "stopila" s frekvencijama od 0 i 1 godine).

Smanjila se jasnoća u *horizontalnom smjeru* - to se može interpretirati kao nepreciznost u trajanju utjecaja zadane frekvencije na vodostaj rijeke

Zaključak: Mijenjanjem širine otvora radimo "tradeoff" između 2 vrste preciznosti - vremenske i frekvencijske. U drugim se područjima znanosti (npr. kvantna mehanika) ovaj efekt naziva Heisenbergov princip neodređenosti.

8. Ispišite sve obitelji kontinuiranih valića koje se nalaze u PyWavelets biblioteci. Na istoj slici, koristeći subplot naredbu, nacrtajte dvije valićne funkcije po izboru. U naslovu svake slike napišite o kojim valićima se radi.

```
# Ovo je mjesto na kojem možete izvoditi svoj kod.
import pywt
import pylab
```

```

print("Imena obitelji kontinuiranih valića iz PyWavelets biblioteke:")
aaa = pywt.wavelist(kind='continuous')
for i in aaa:
    print(i, end=", ") if (i != aaa[-1] ) else print(i)
print("\n")
##### FUNCTIONS #####

wave1 = pywt.ContinuousWavelet('mexh')
#print(wave1)
[psi_1, x_1] = wave1.wavefun(level=20)

wave2 = pywt.ContinuousWavelet('shan')
#print(wave2)
[psi_2, x_2] = wave2.wavefun(level=20)

##### PLOTTING #####
plt.figure(1).set_figwidth(20)
plt.subplot(1, 2, 1)
pylab.plot(x_1, psi_1)
plt.title('mexican hat function')

plt.subplot(1, 2, 2)
pylab.plot(x_2, psi_2)
plt.title('shannon function')

plt.show()

```

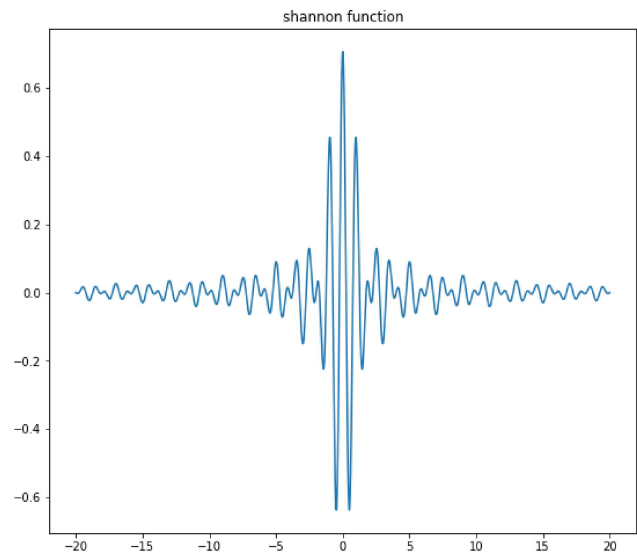
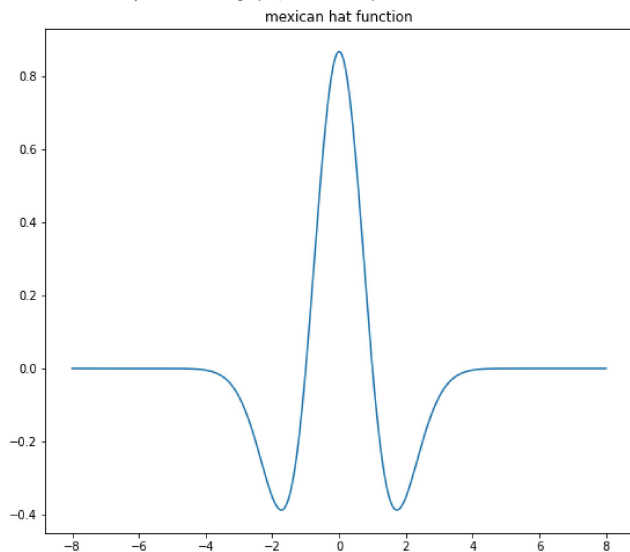
Imena obitelji kontinuiranih valića iz PyWavelets biblioteke:

cgau1, cgau2, cgau3, cgau4, cgau5, cgau6, cgau7, cgau8, cmor, fbsp, gaus1, gaus2, gaus3, gaus4, gaus5,

```

/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:16: FutureWarning: Wavelets from the famil
app.launch_new_instance()
/usr/local/lib/python3.7/dist-packages/matplotlib/cbook/__init__.py:1317: ComplexWarning: Casting compl
return np.asarray(x, float)

```



9. Koristeći Morlet valić, odredite kontinuiranu valićnu transformaciju zadanog vodostaja. Nacrtajte apsolutnu vrijednost dobivenih koeficijenata. Odgovorite: koje frekvencije su vidljive u vodostaju? Koje godine se javljaju?

```
# Ovo je mjesto na kojem možete izvoditi svoj kod.
```

```
step = 1
```

```
scale = np.arange(1, 4*365, step) # u odnosu na ppt promijenio sam širinu otvora
```

```
coef, freqs = pywt.cwt(podatci2, scale, 'morl')
```

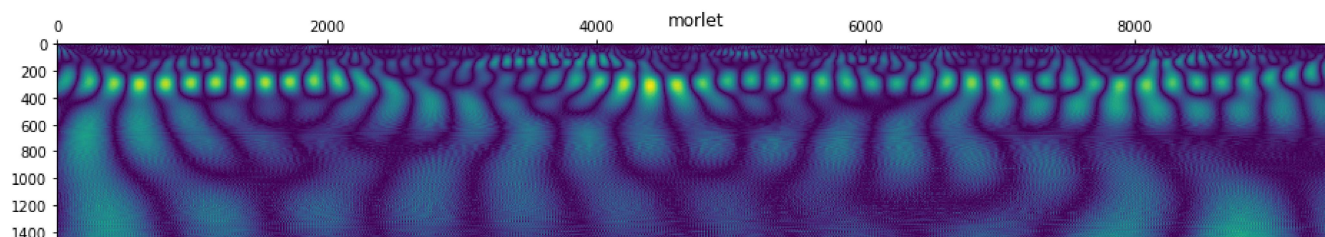
```
plt.figure().set_figheight(10)
```

```
plt.matshow(abs(coef))
```

```
plt.title('morlet')
```

```
plt.show()
```

<Figure size 864x720 with 0 Axes>



ODGOVOR:

Budući da su najsvjetliji dijelovi grafa oko pravca **x=365**, čini se da je najnaglašenija frekvencija upravo ta - promjena svakih **godinu dana**. Prema grafu se da naslutiti i periodičnost u vodostaju svakih 1300 dana ~ **4 godine** Od frekvencija većih od 1 godine - iako se one na ovom konkretnom grafu ne vide bajno, postoje i tamo pravilnosti: za periode od 150 dana (~**6 mjeseci**) i oko 90 i 100 dana (~**3 mjeseca**), ali ne toliko izražene kao ona od godine dana

Prvih 10ak godina vodostaj je u prosjeku bio relativno visok, dok su kasnije uslijedile sušne godine. Zatim se (ne suviše pravilno) izmjenjuju u prosjeku vlažnije i sušnije godine

10. Odgovorite: objasnite razliku u dobivenim rezultatima STFT i CWT.

ODGOVOR:

Osobno mislim da je za konkretne podatke prikladnije koristiti graf STFT jer jasnije prikazuju periodičnosti u manjim vremenskim intervalima. STFT je (naravno, ovisno o onom "tradeoffu") jasnije pokazivao o kojim se točno frekvencijama radi i/ili od kad do kada je trajao njihov utjecaj.

CWT bi možda bio jasniji i bolji izbor od dva kada bi se radilo o drukčijem signalu, npr. onome koji se radikalnije mijenja u vremenu