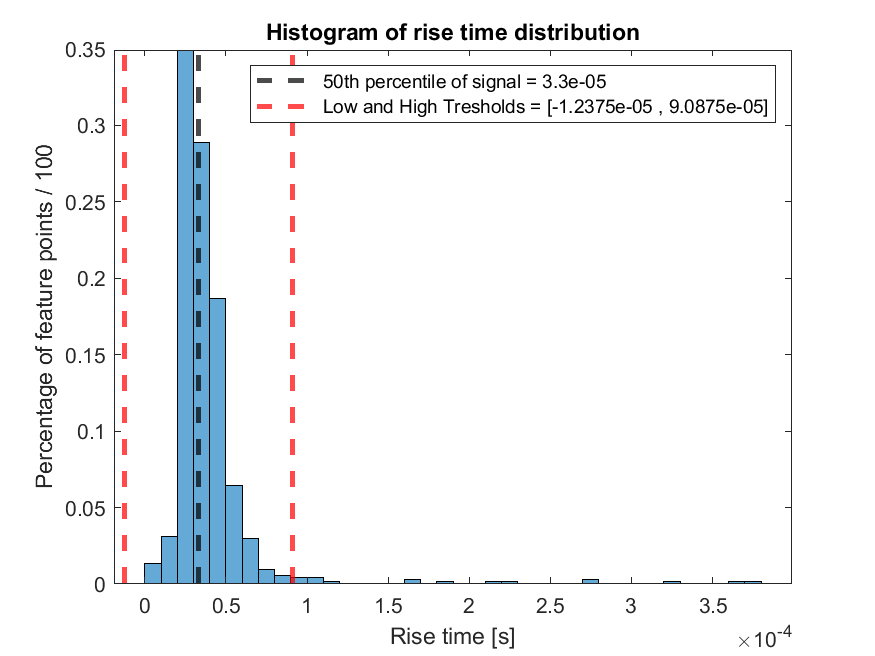
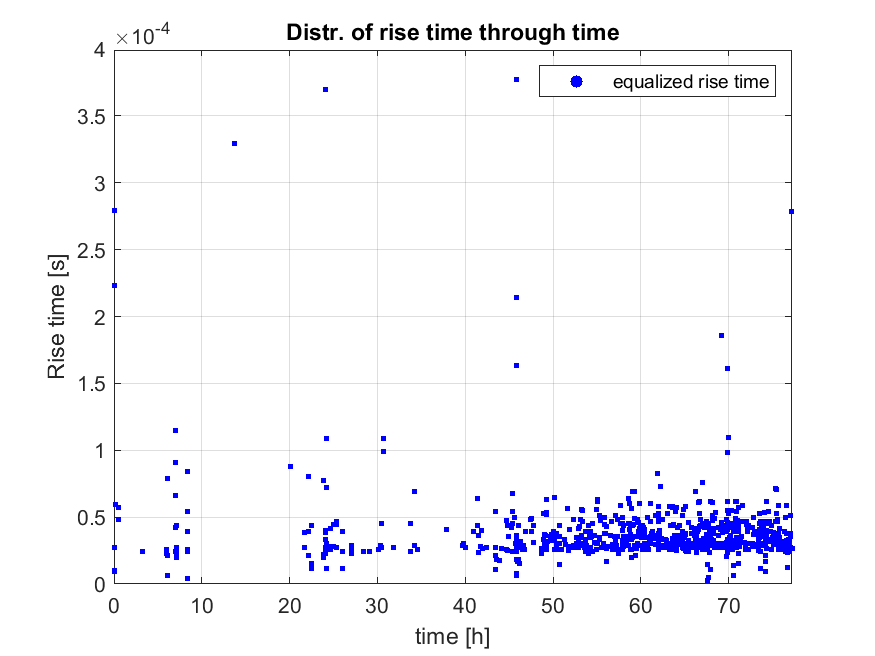
# Primjer izbacivanja outliera i normalizacije za manji dataset

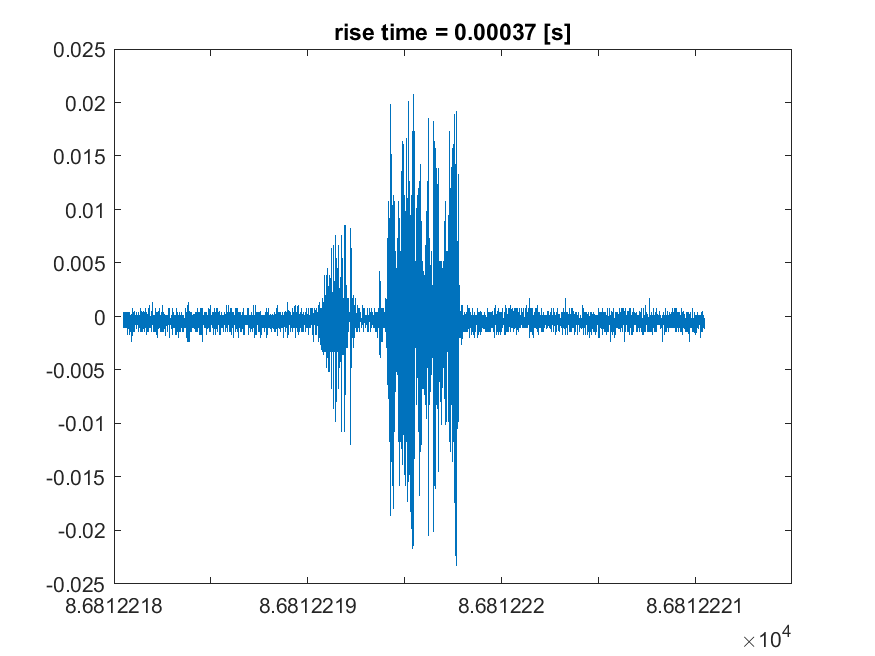
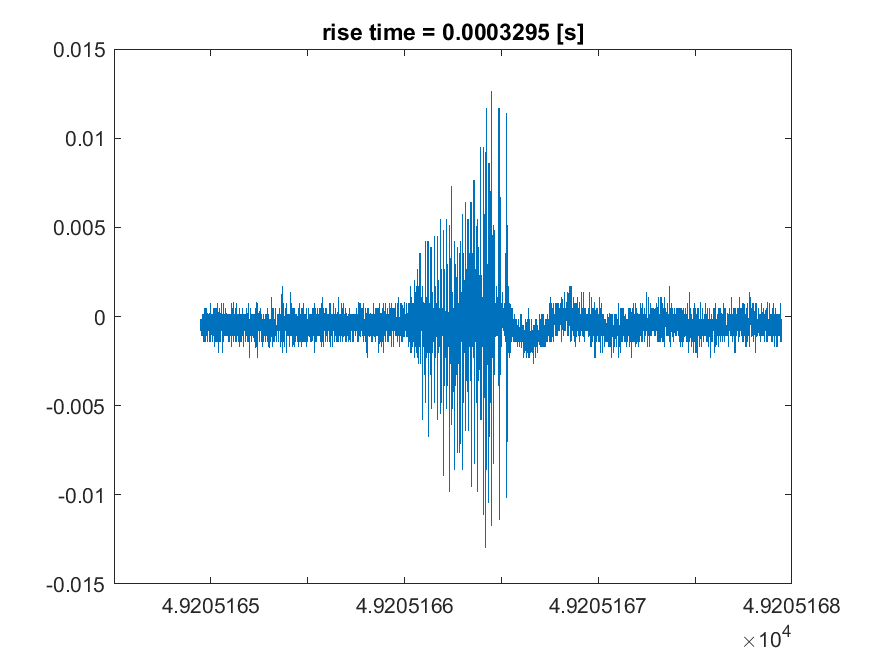
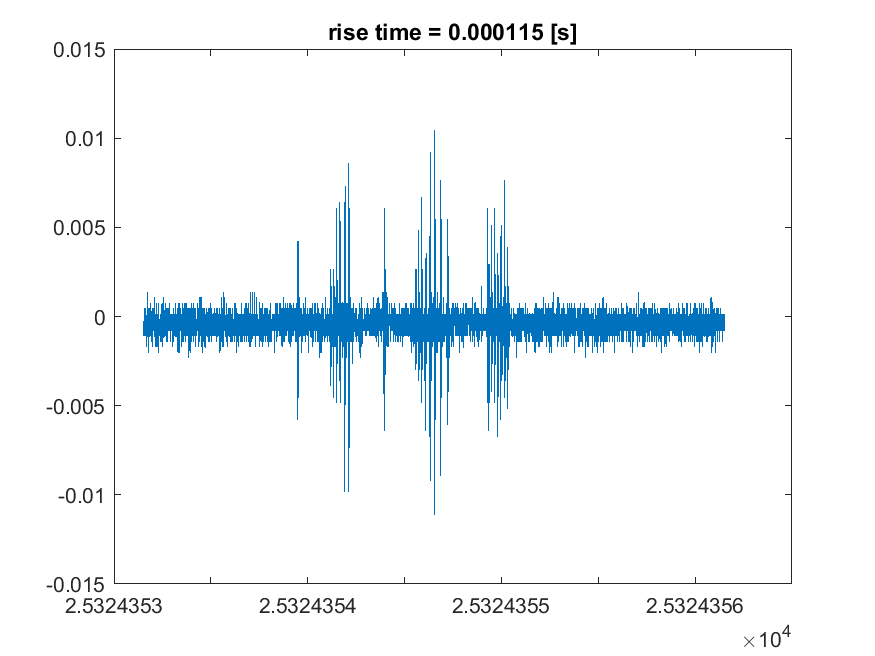
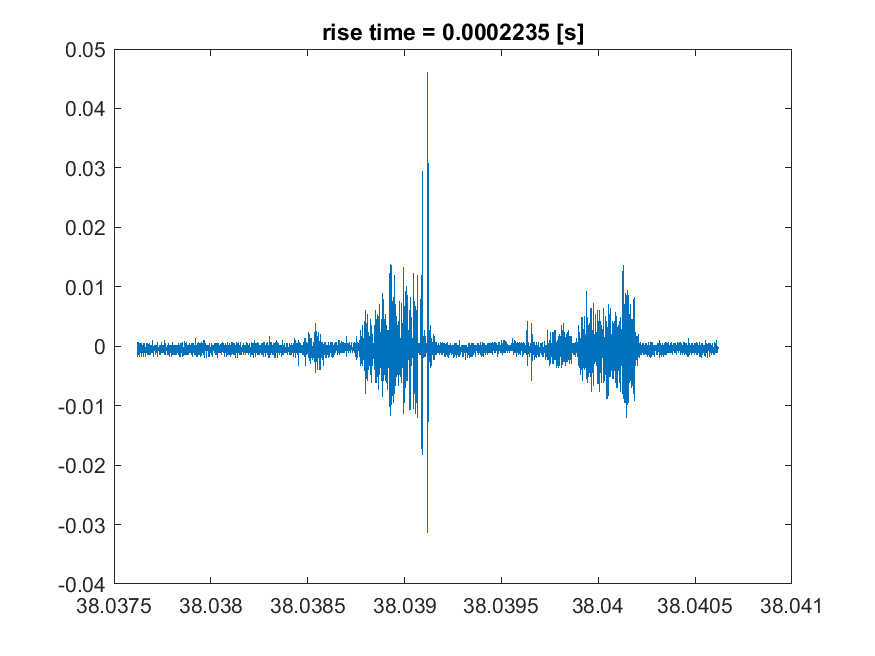
**Za potencijalne outliere određenih emisija će se tražiti vrijednosti određenih značajki iznad ili ispod granica određenih pomoću IQR mjere. Za IQR vrijednost je uzeta razlika između 95. percentila i 5. percentil pojedine značajke pomnožena s koeficijentom 1.25. Zatim je određena gornja granica outliera kao IQR + 95. percentil te donja granica outliera kao IQR – 5. percentil. Navedene granice vidimo na slici označene kao crvene iscrtkane linije. Potencijalni outlieri postaju pravi outlieri ako prepoznamo grešku kod odvajanja signala od šuma pomoću binarne omotnice ili ako prepoznamo da emisije unose previše šuma i nisu prava reprezentacija skupa podataka.**

**RISE\_TIME:**

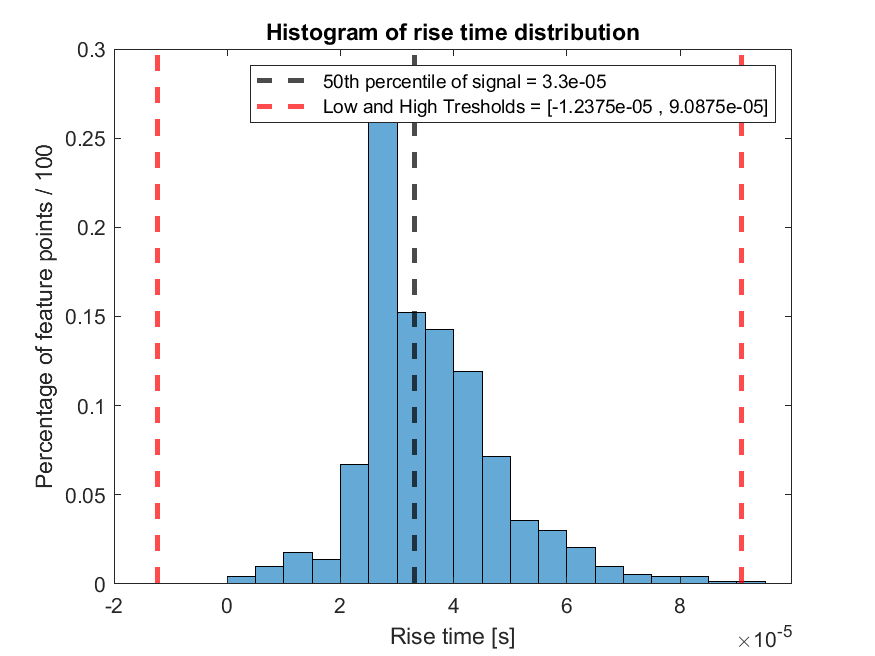
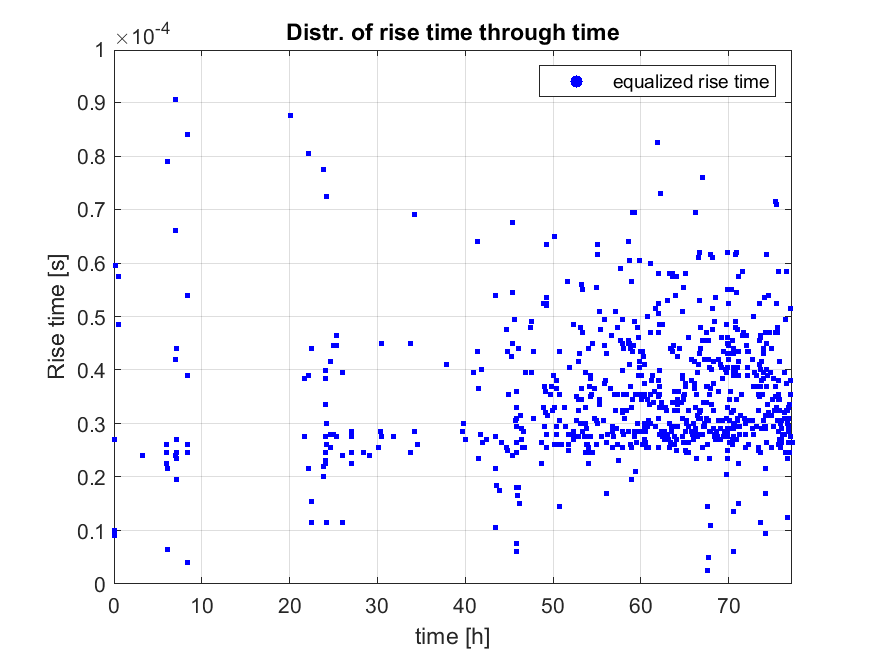


**Potencijalni outlieri iznad 9.0875e-05 s i ispod -1.2375e-05 s RISE TIME:**

**To je više emisija neposredno jedna nakon druge pa je binarna omotnica uzela predugi dio signala kao emisiju.**



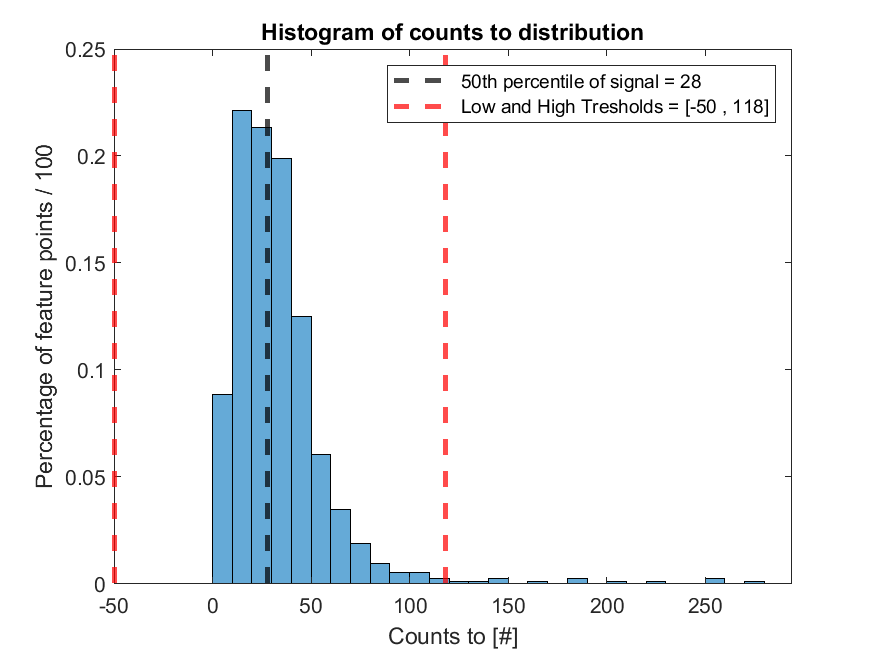
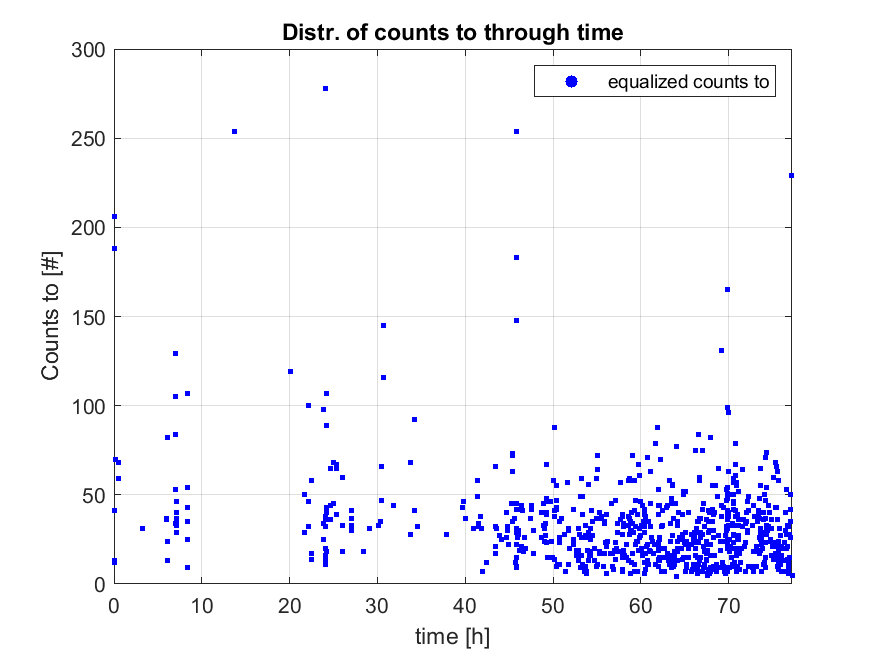
**Nakon izbacivanja outlieri iznad 9.0875e-05 s i ispod -1.2375e-05 s RISE TIME dobivamo slijedeću distribuciju:**



**RISE TIME outliers are emissions with value more than 9.0875e-05 [s]**

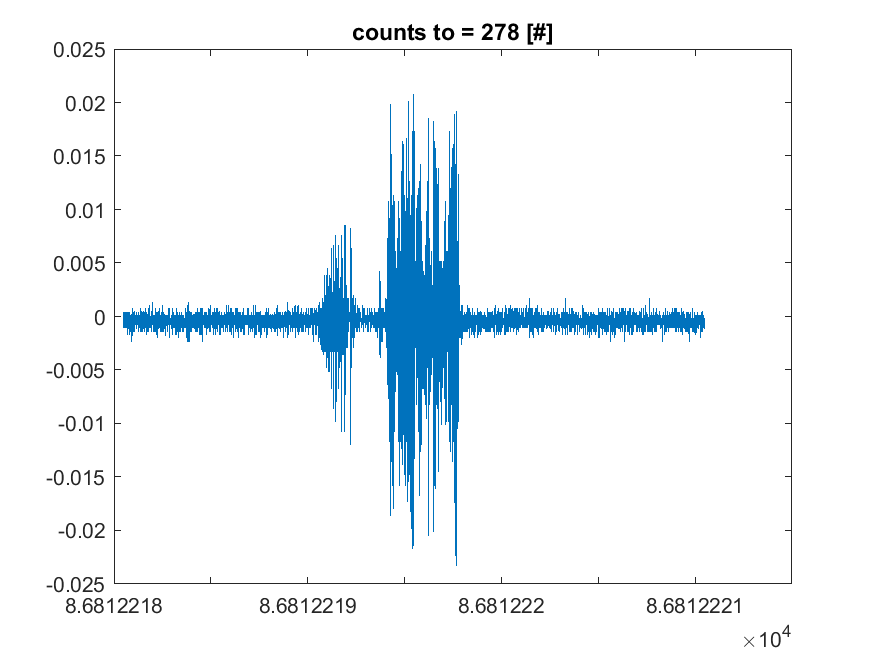
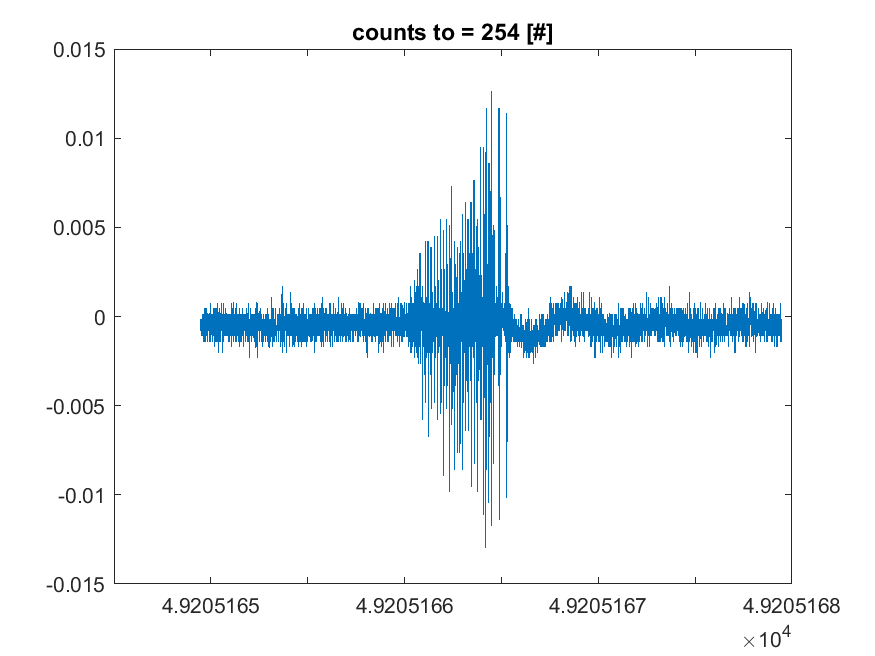
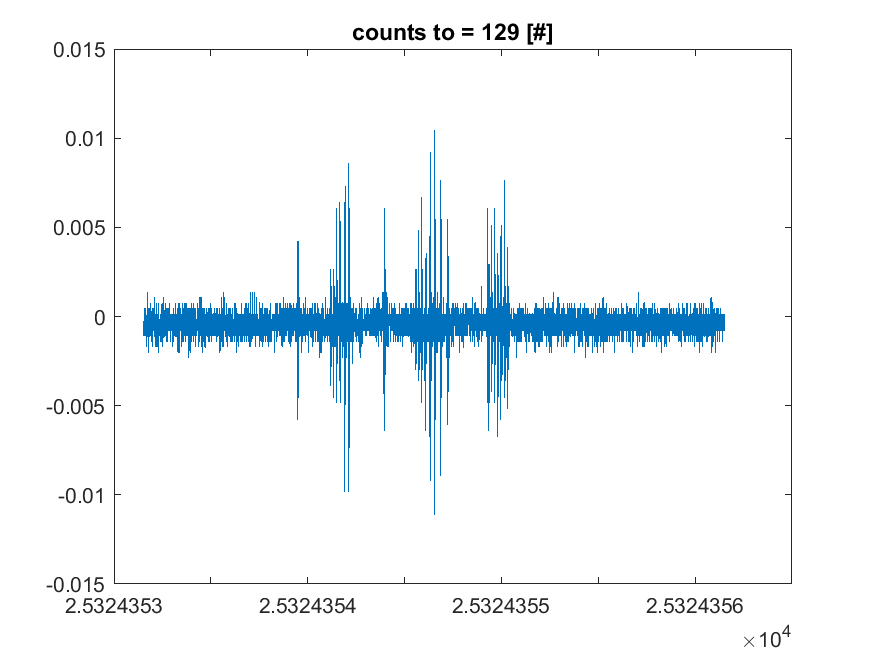
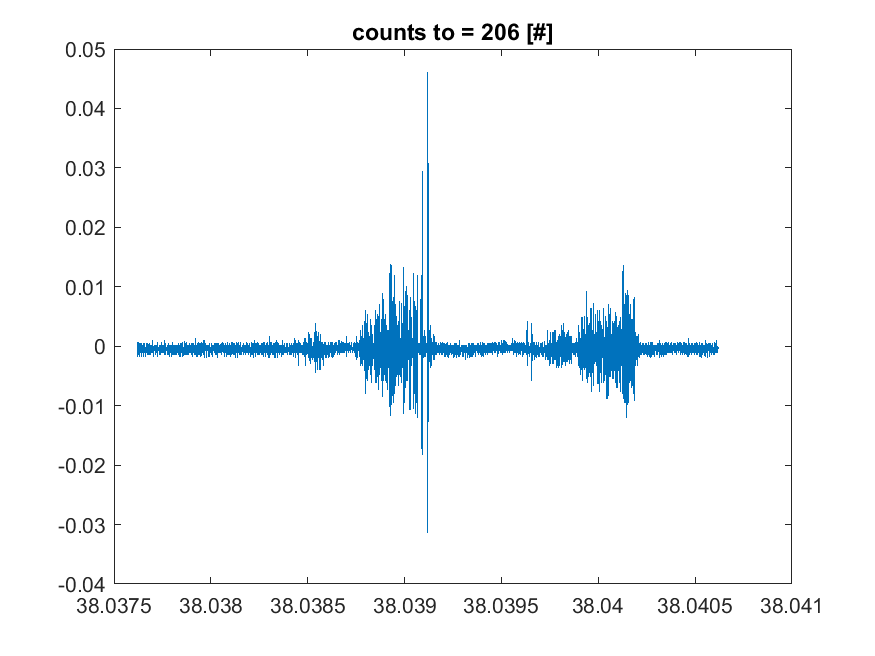
**Percentage of outliers for RISE TIME: 16/745 = 2.1477 %**

**COUNTS\_TO:**



**Potencijalni outlieri iznad 118 # i ispod -50 # COUNTS TO:**

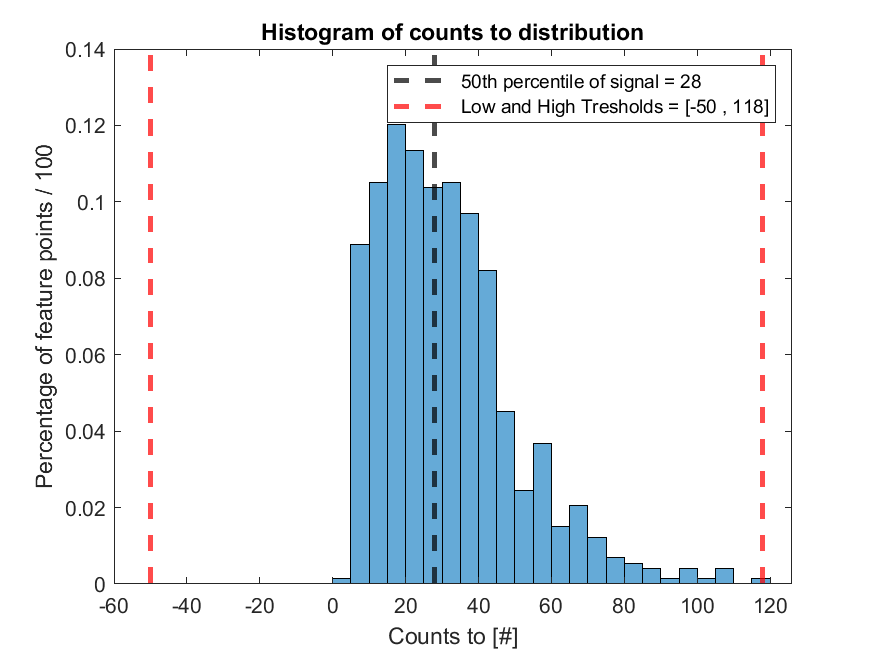
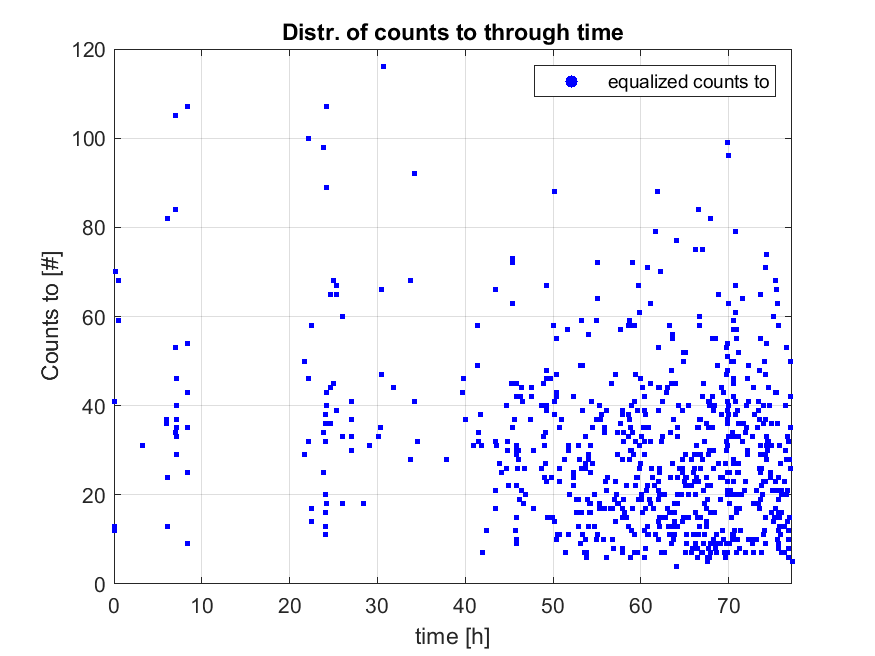
**Isti razlog i emisije kao kod RISE TIME.**



**COUNTS TO outliers are emissions with value more than 118 [#]**

**Percentage of outliers for COUNTS TO: 13/745 = 1.745 %**

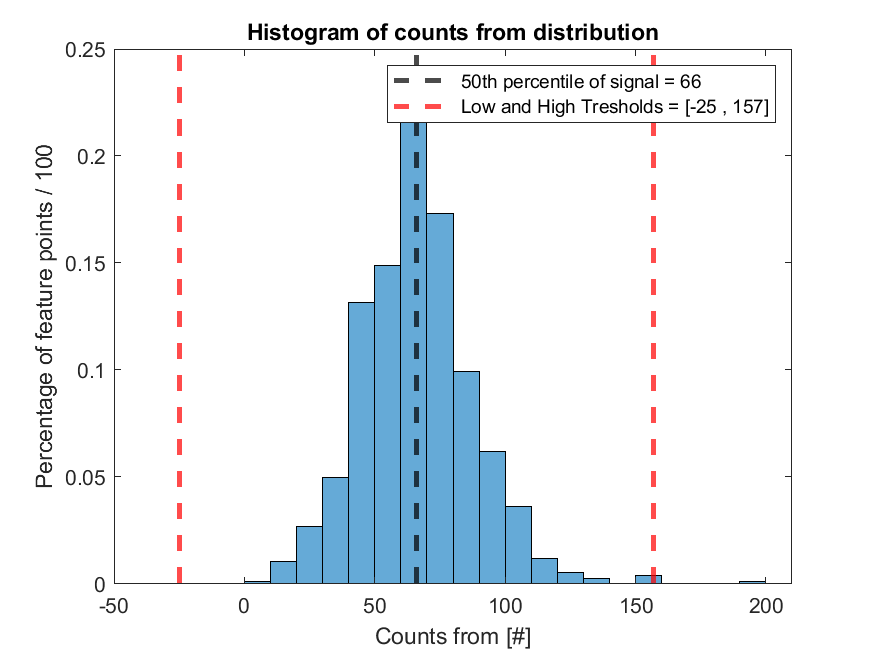
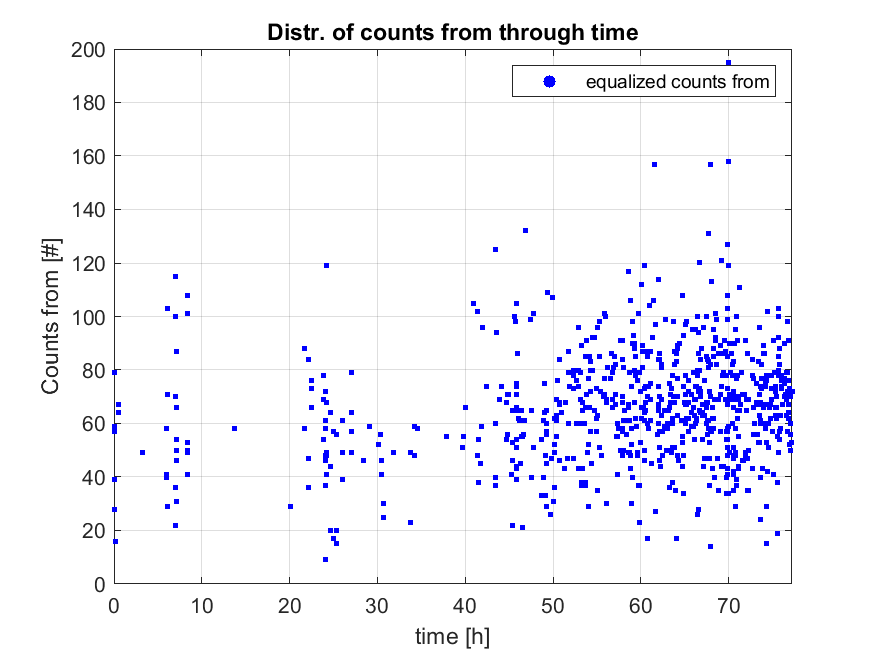
**Nakon izbacivanja outlieri iznad 118 # i ispod -50 # COUNTS TO dobivamo slijedeću distribuciju:**



**COUNTS TO outliers are emissions with value more than 118 [#]**

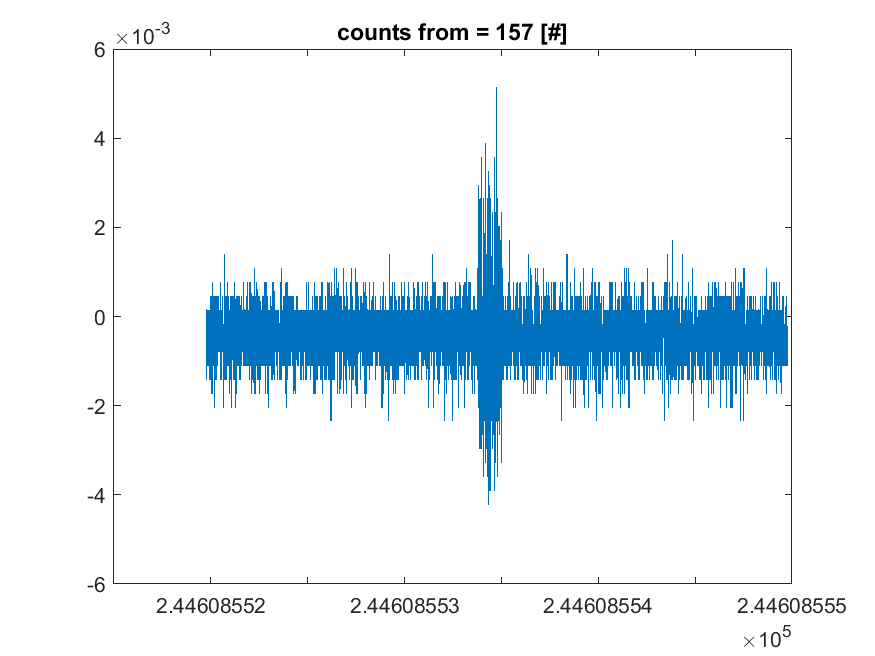
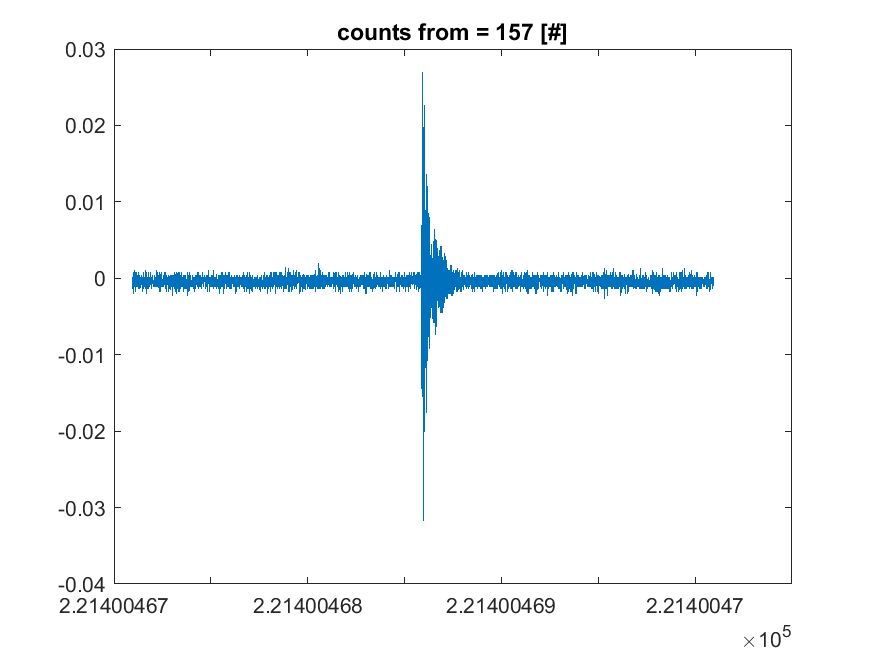
**Percentage of outliers for COUNTS TO: 13/745 = 1.745 %**

**COUNTS\_FROM:**

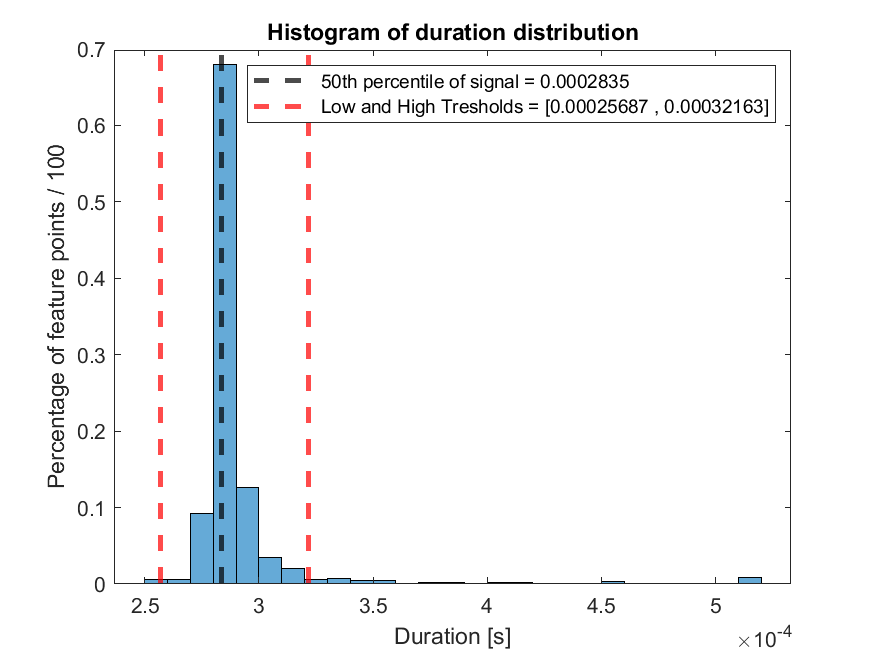
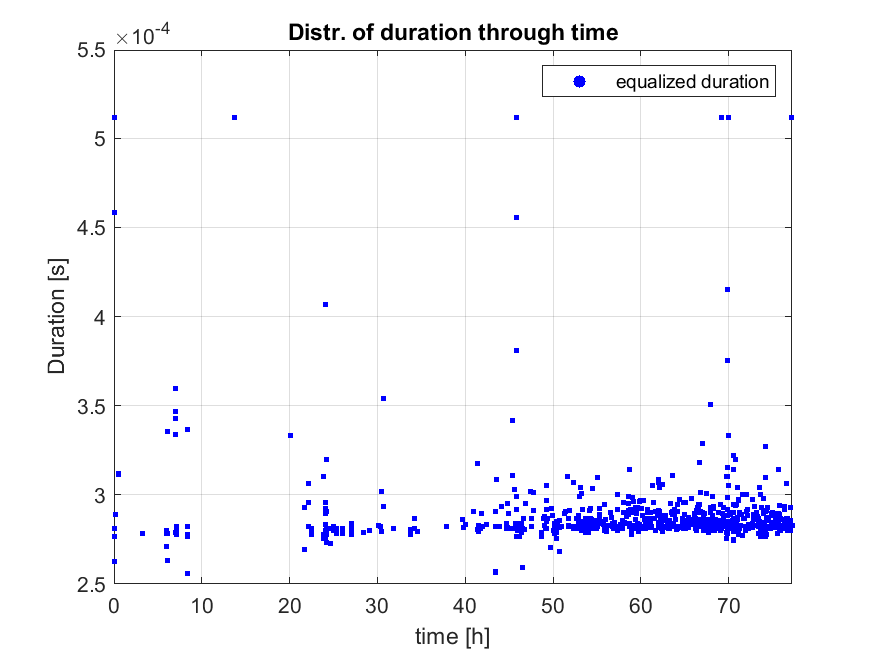


**Potencijalni outlieri iznad 157 # i ispod -25 # COUNTS FROM:**

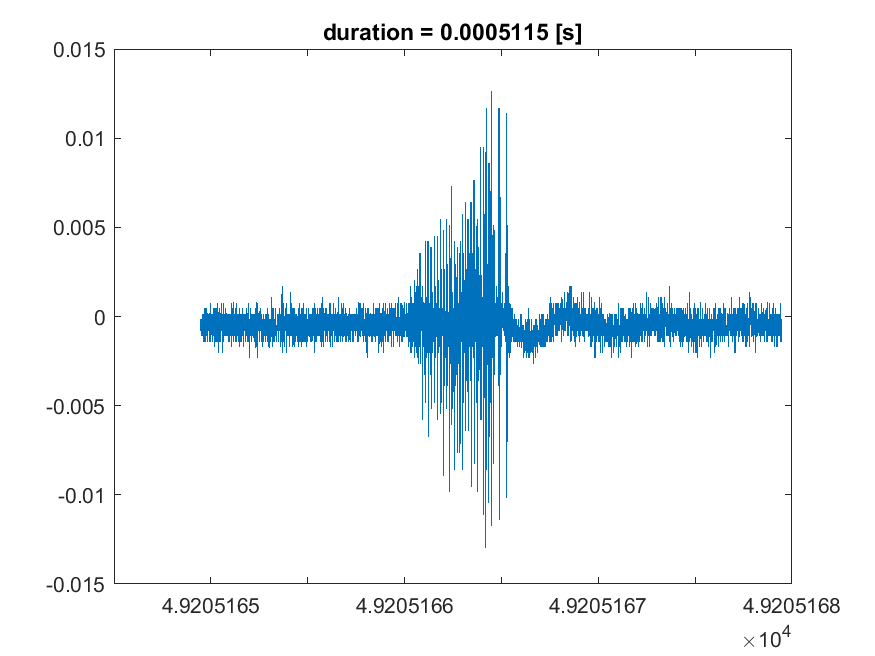
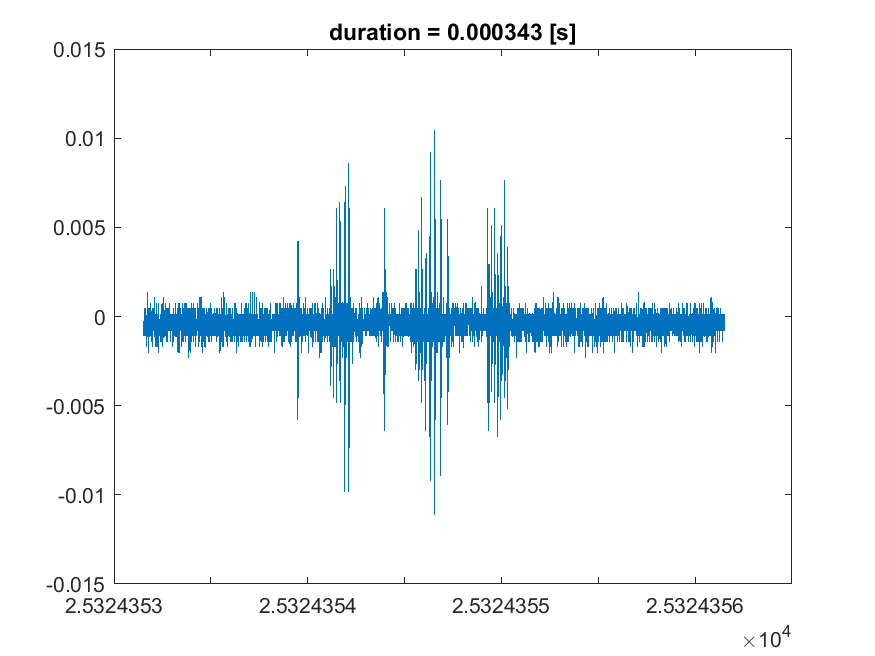
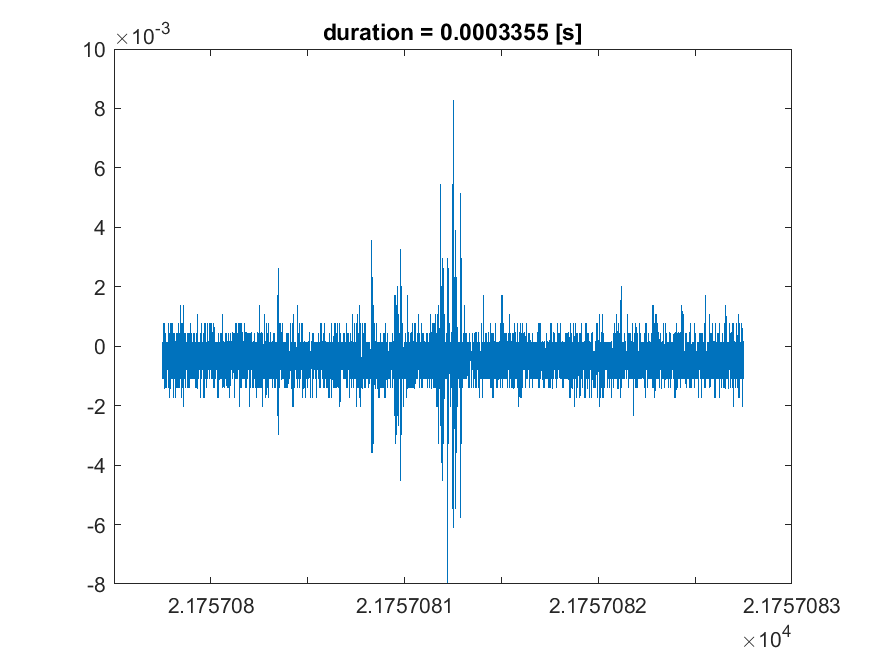
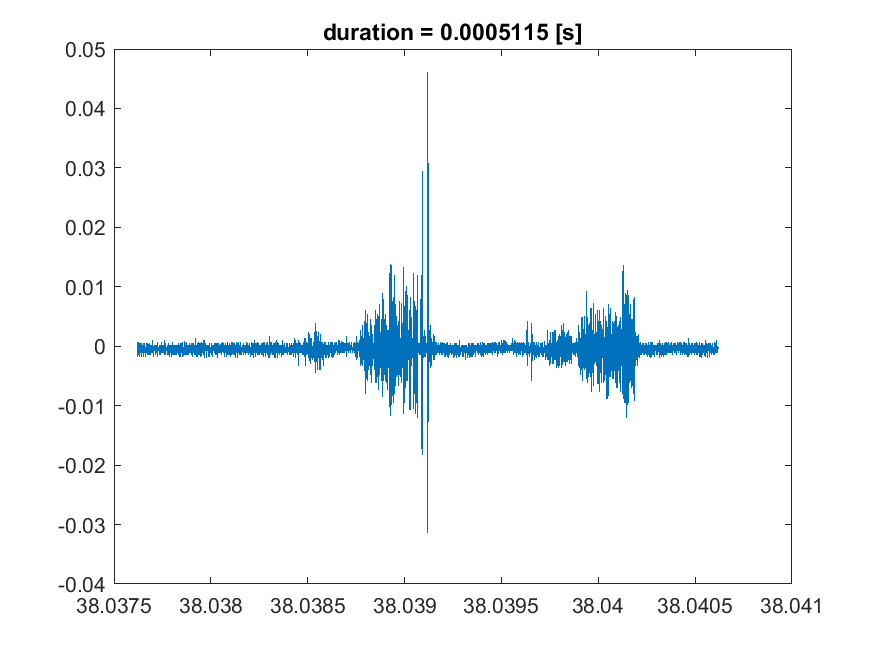
**Postoje validne emisije pa nisu pravi outlieri.**



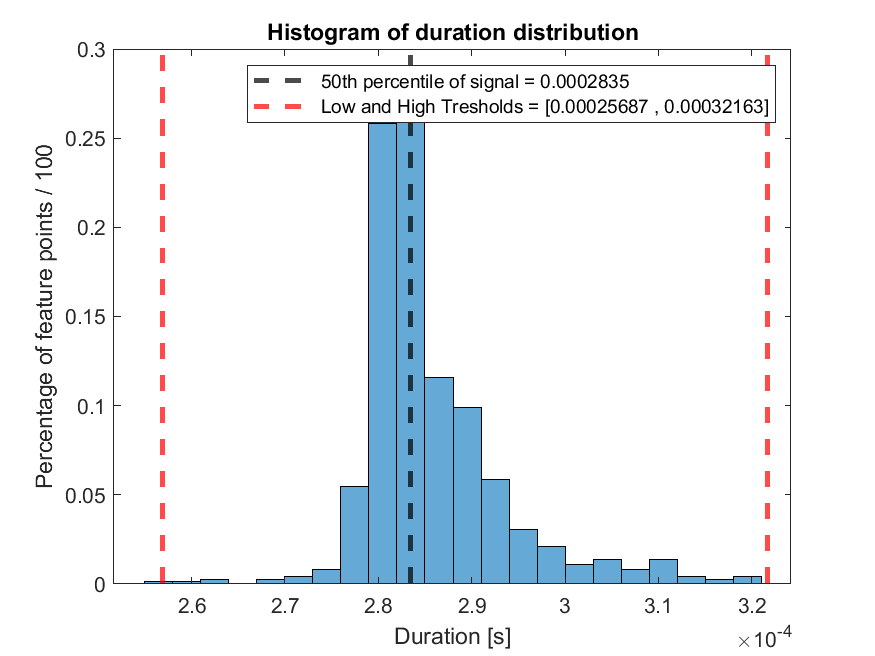
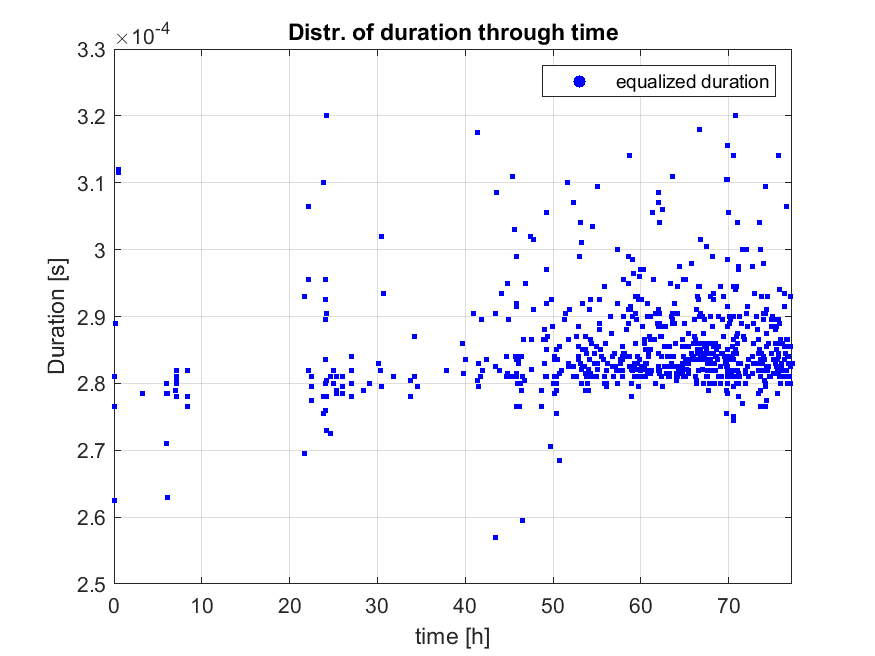
**DURATION:**



**Potencijalni outlieri iznad 0.00032163 s i ispod 0.00025687 s DURATION:**



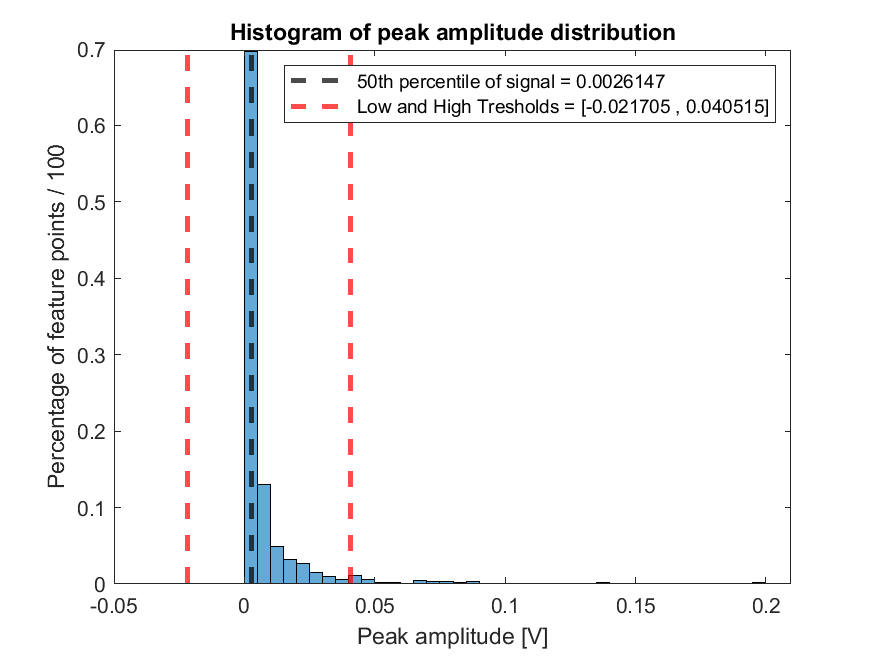
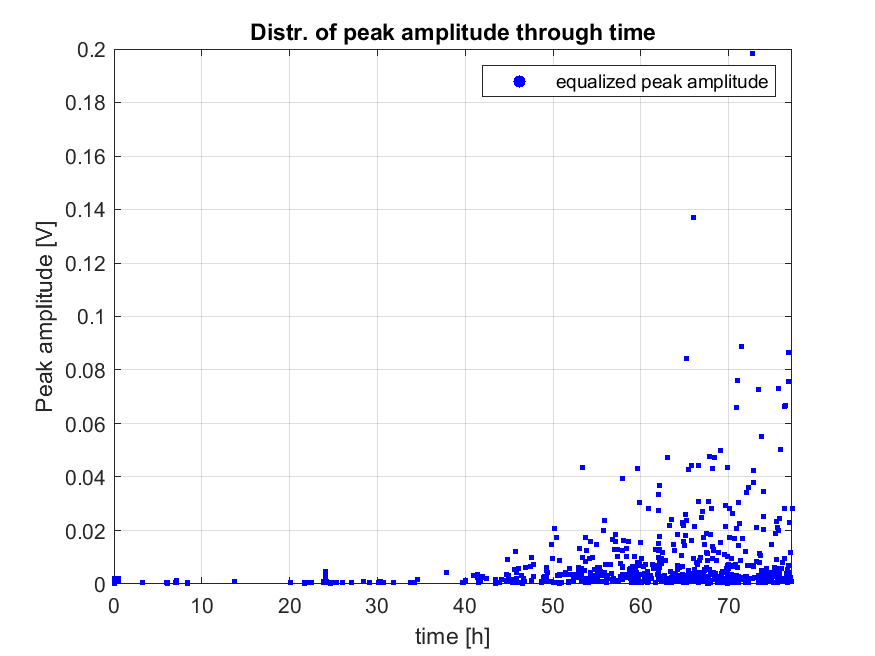
**Nakon izbacivanja outlieri iznad 0.00032163 s i ispod 0.00025687 s DURATION dobivamo slijedeću distribuciju:**



**DURATION outliers are emissions with value more than 0.00032163 [s]**

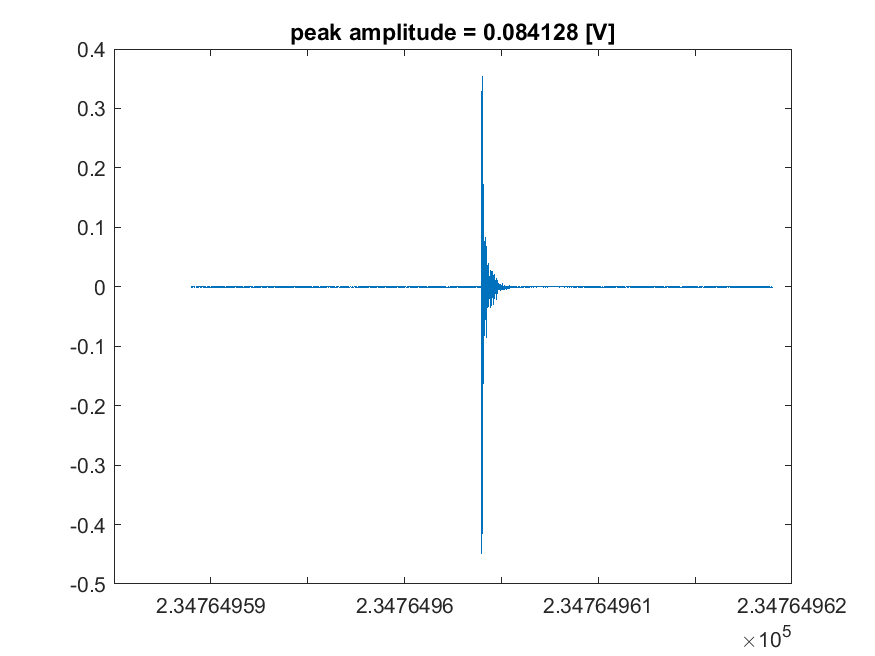
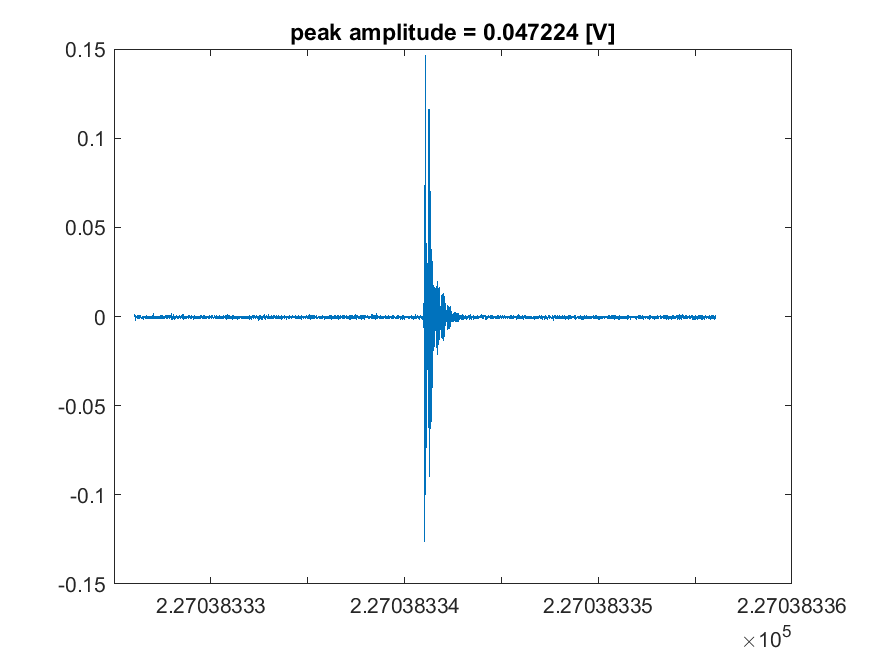
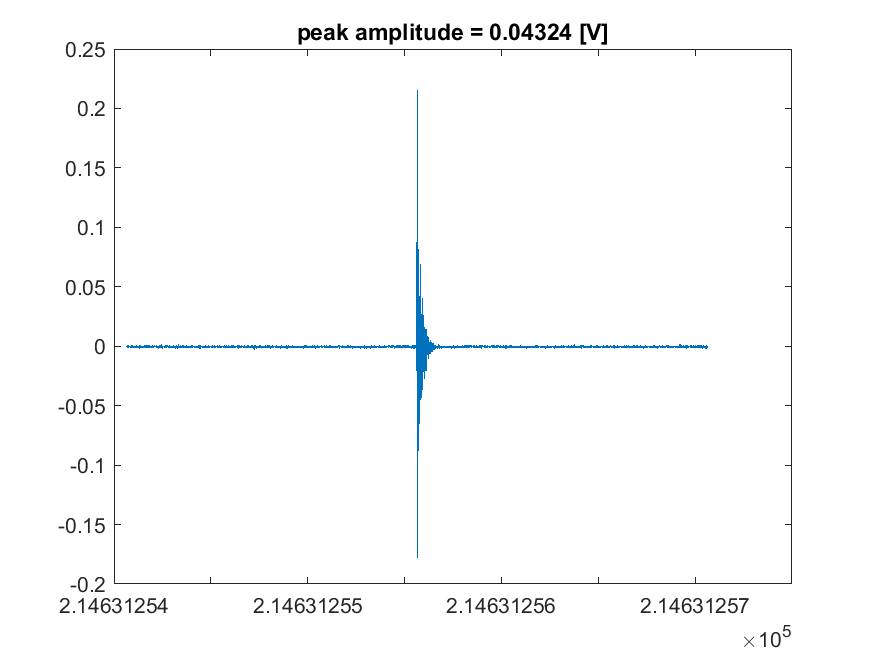
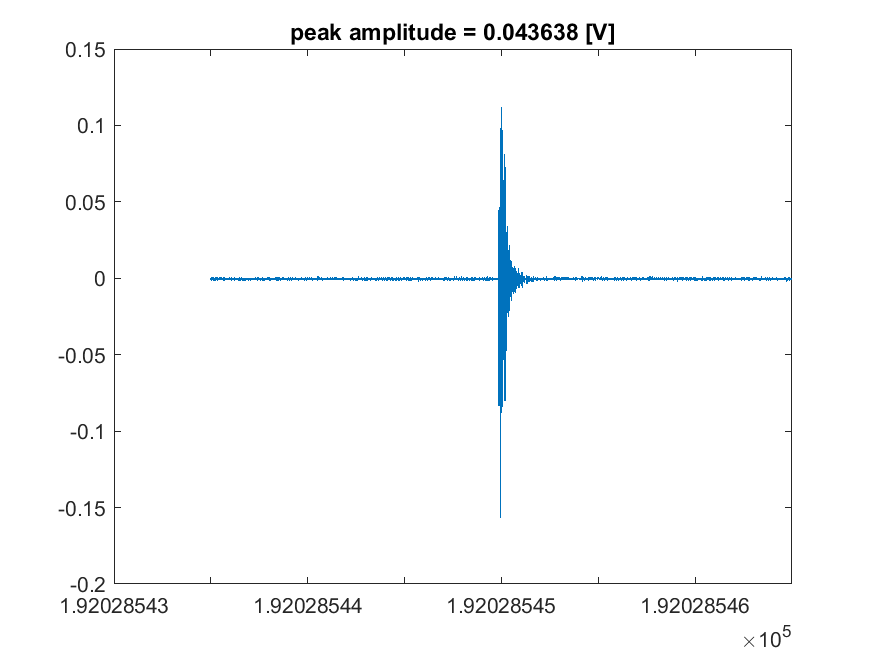
**Percentage of outliers for DURATION: 28/745 = 3.7584 %**

**PEAK\_AMPLITUDE:**

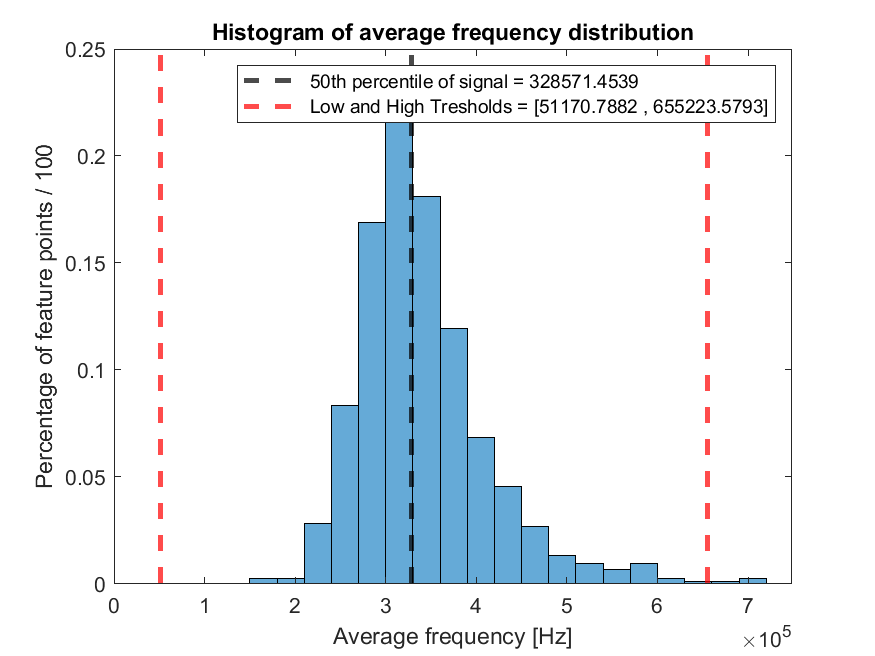
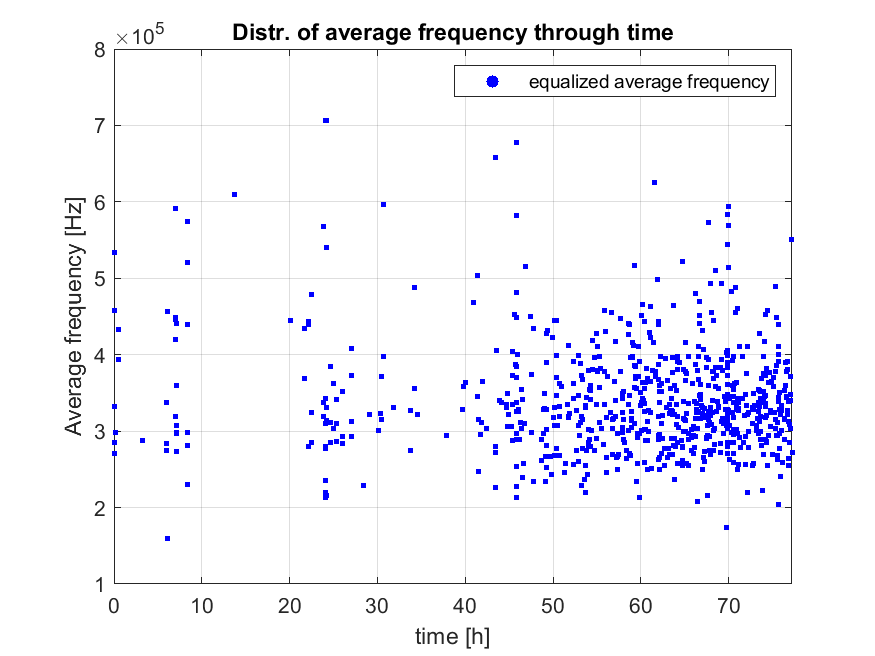


**Potencijalni outlieri iznad 0.040515 V i ispod -0.021705 V PEAK AMPLITUDE:**

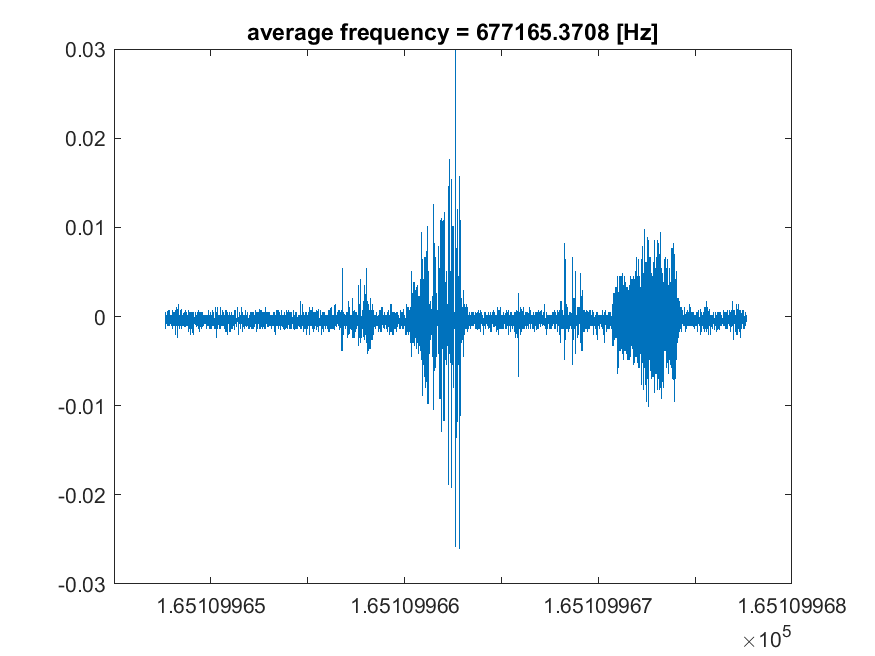
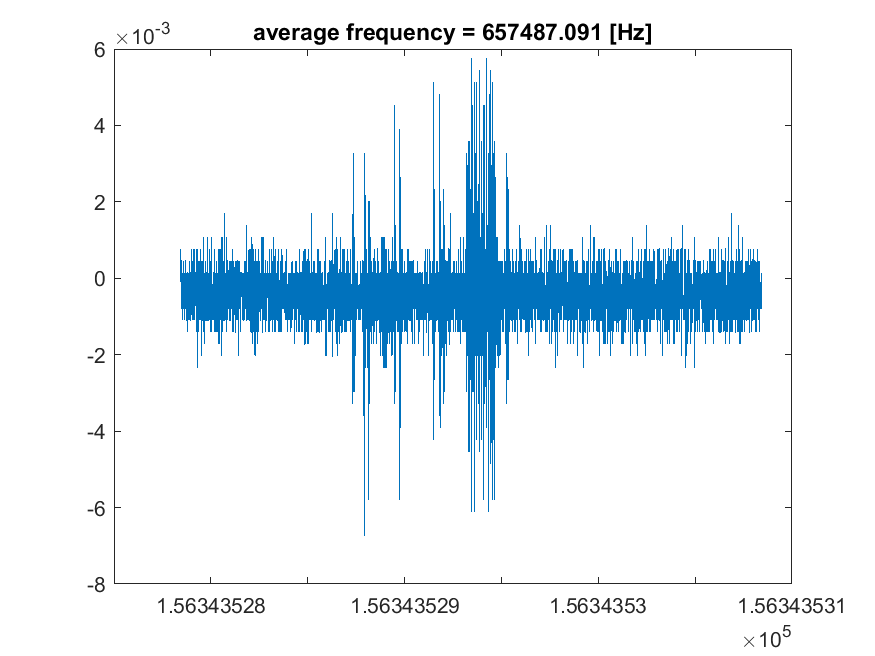
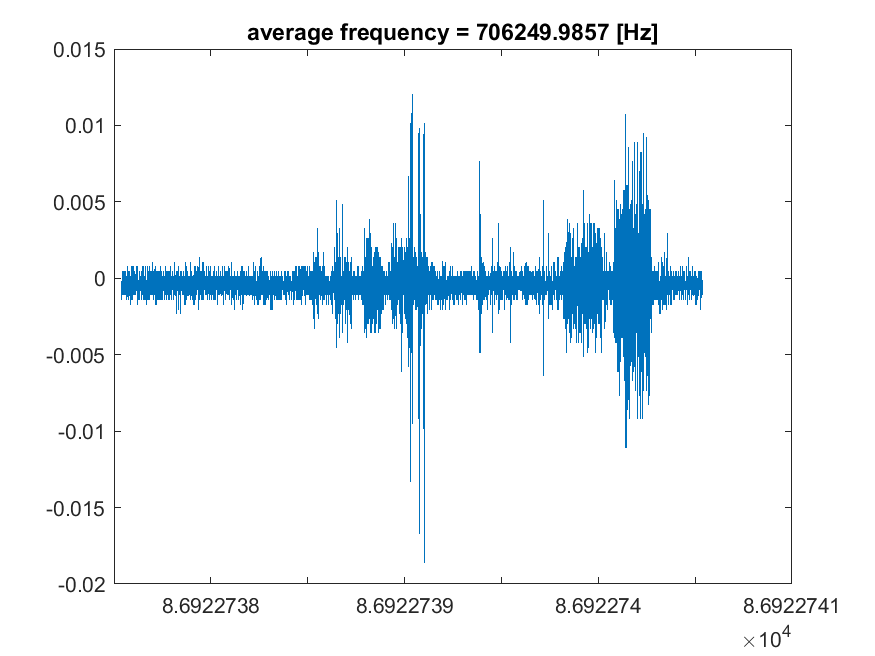
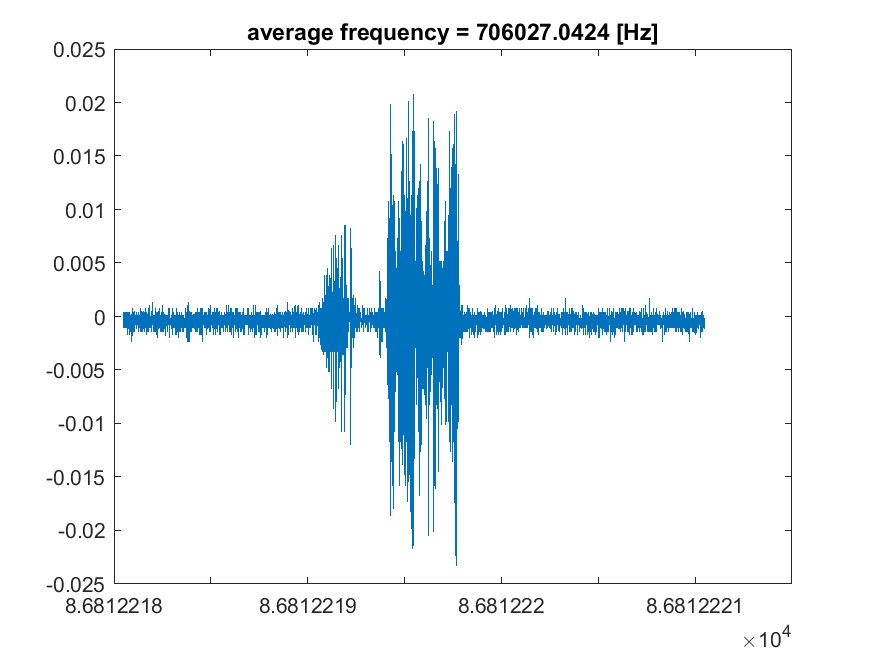
**Postoje validne emisije pa nisu pravi outlieri.**



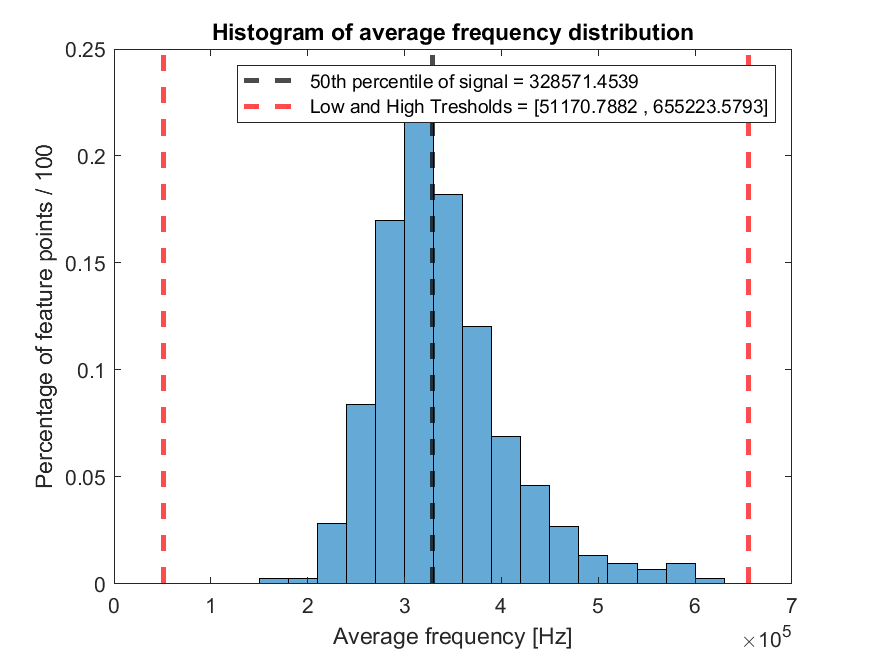
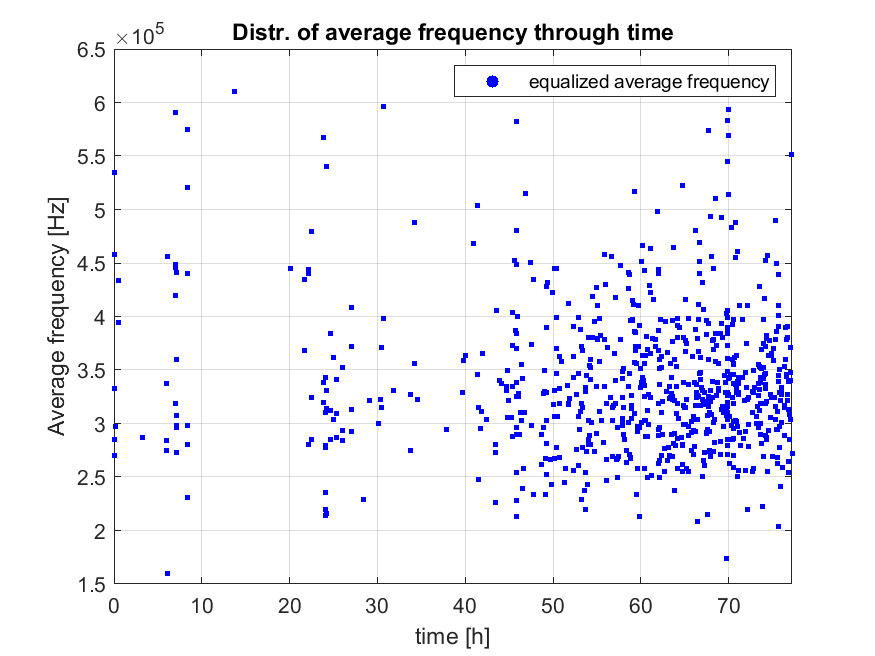
**AVERAGE\_FREQUENCY:**



**Potencijalni outlieri iznad 655223.5793 Hz i ispod 51170.7882 Hz AVERAGE FREQUENCY:**



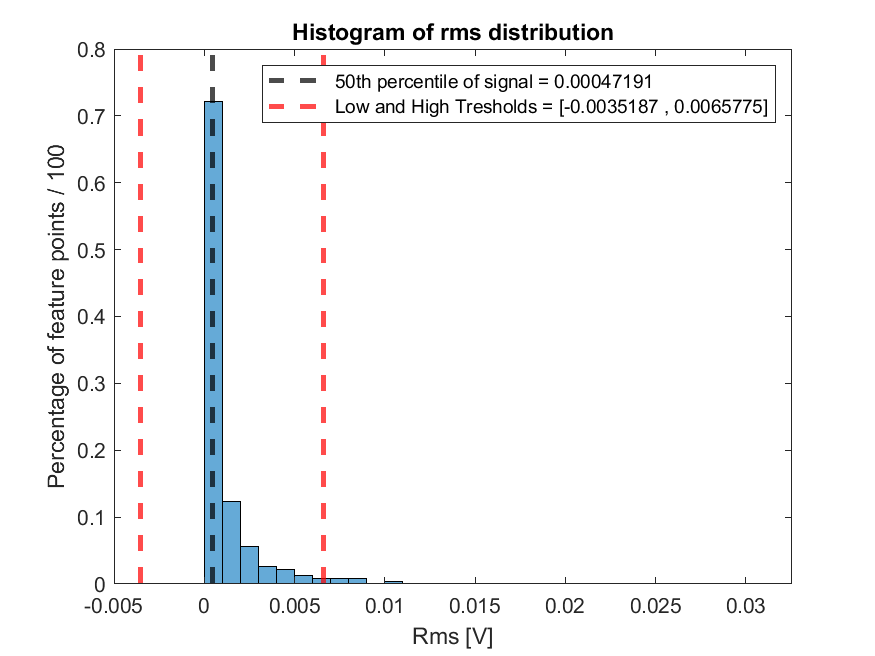
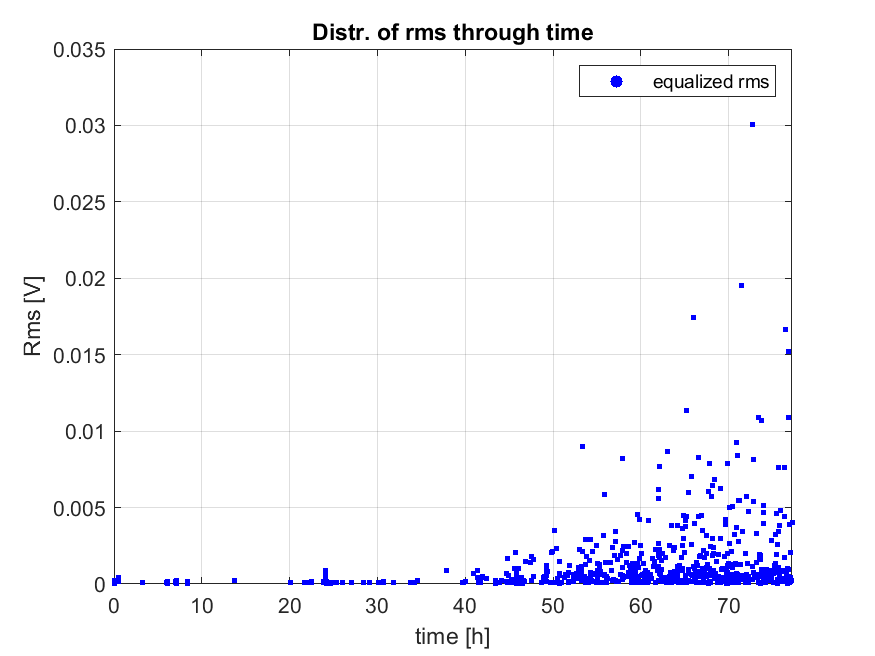
**Nakon izbacivanja outlieri iznad 655223.5793 Hz i ispod 51170.7882 Hz AVERAGE FREQUENCY dobivamo slijedeću distribuciju:**



**AVERAGE FREQUENCY outliers are emissions with value more than 655223.5793 [Hz]**

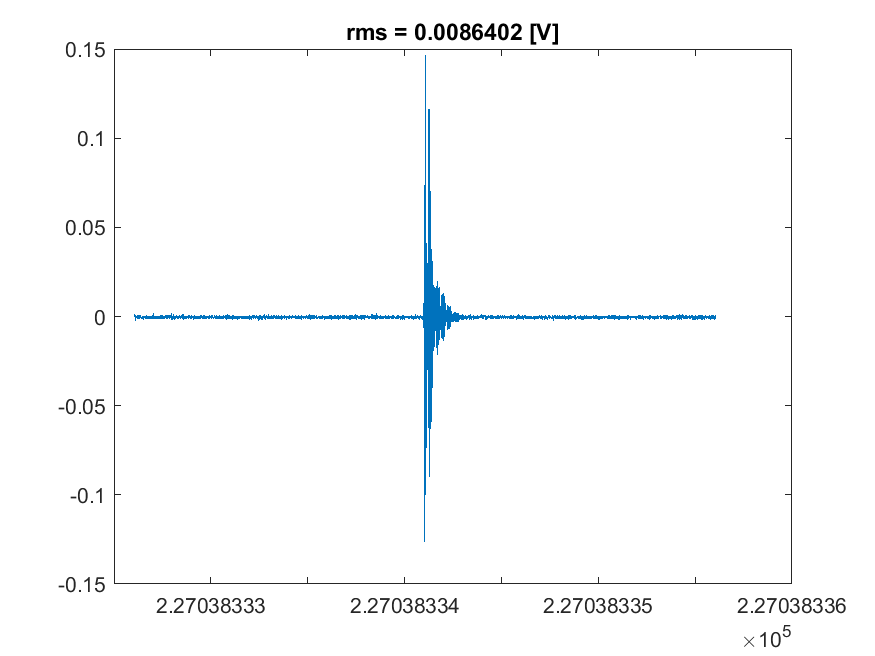
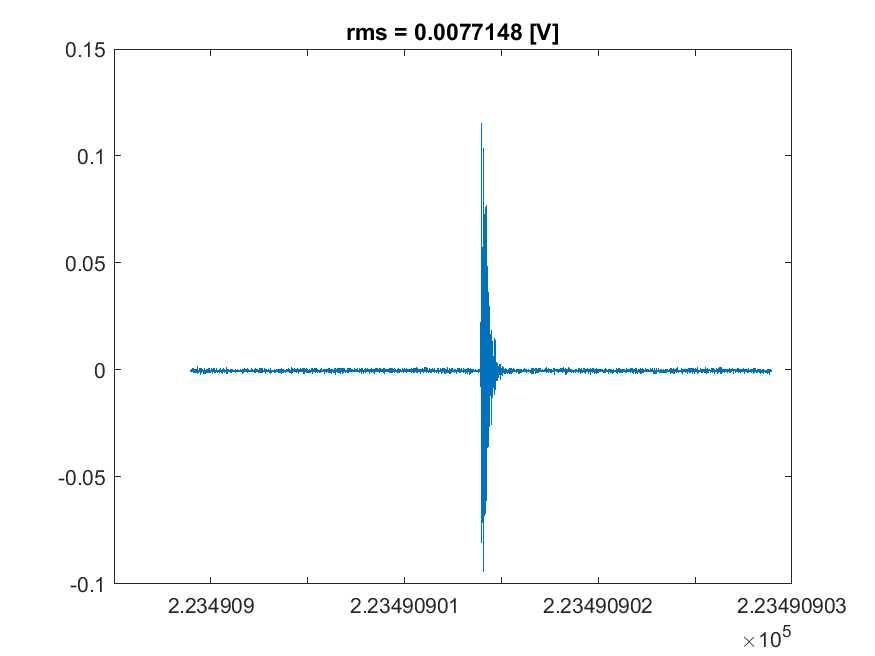
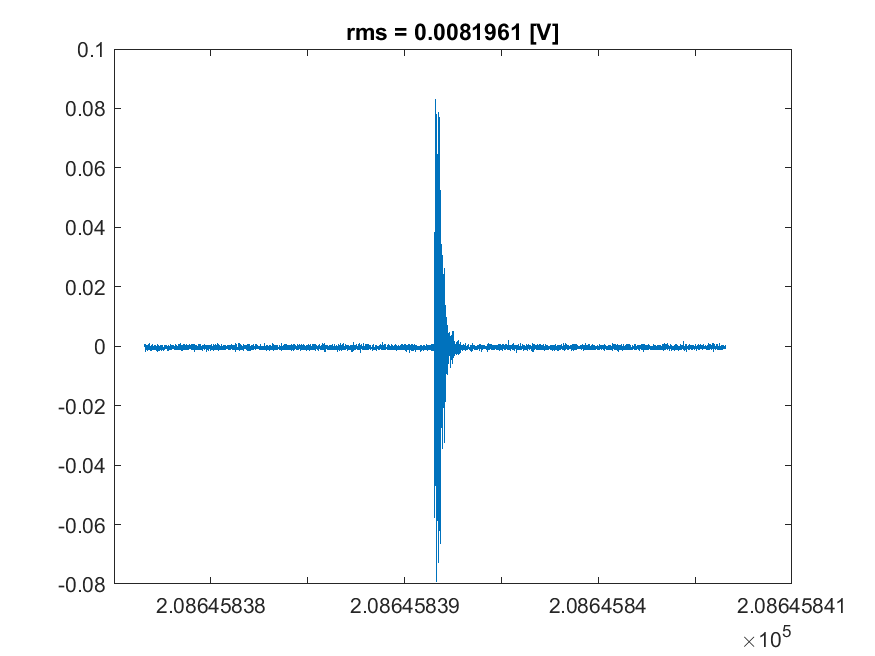
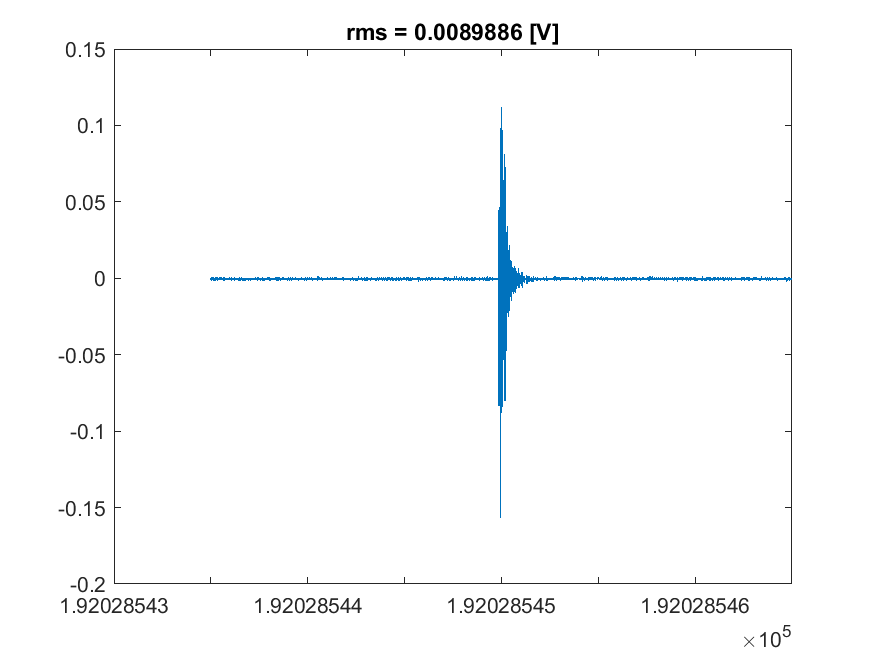
**Percentage of outliers for AVERAGE FREQUENCY: 4/745 = 0.53691 %**

**RMS:**

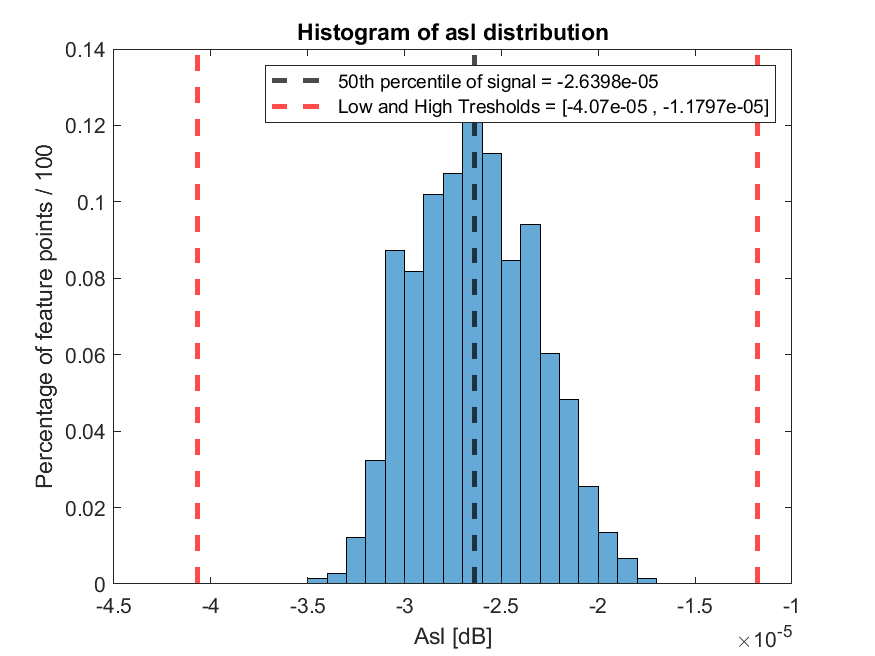
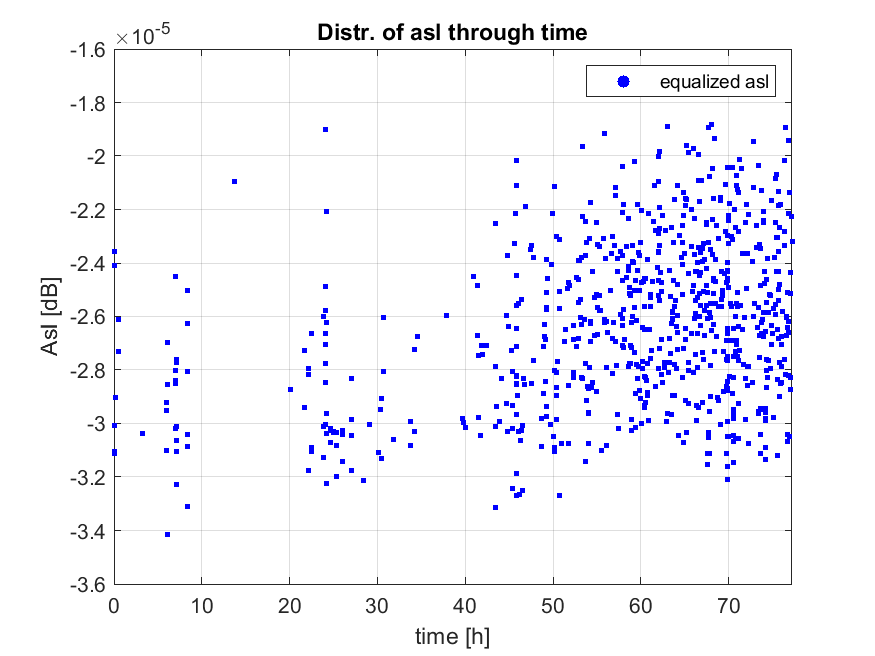


**Potencijalni outlieri iznad 0.0065775 V i ispod -0.0035187 V RMS:**

**Postoje validne emisije pa nisu pravi outlieri.**

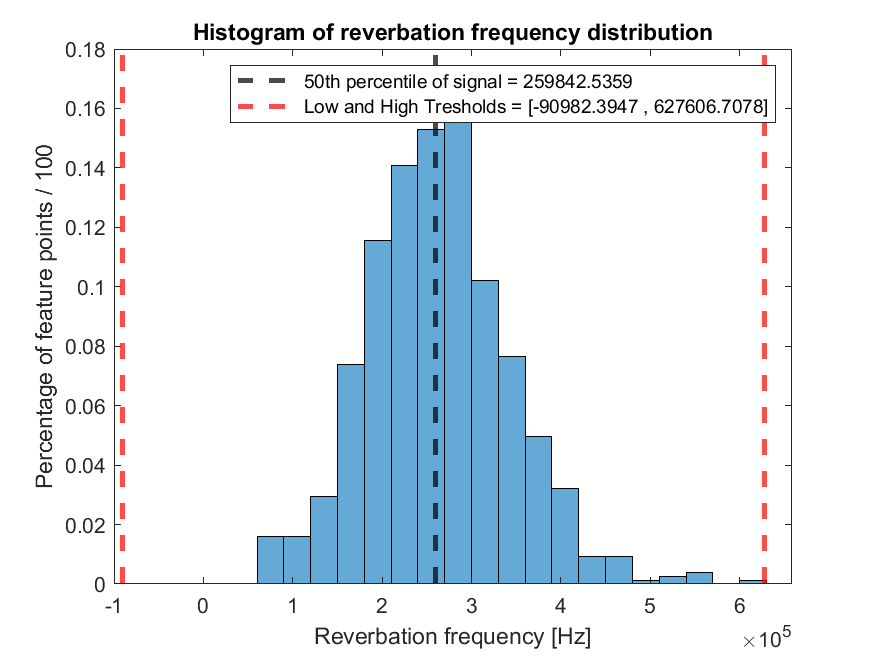
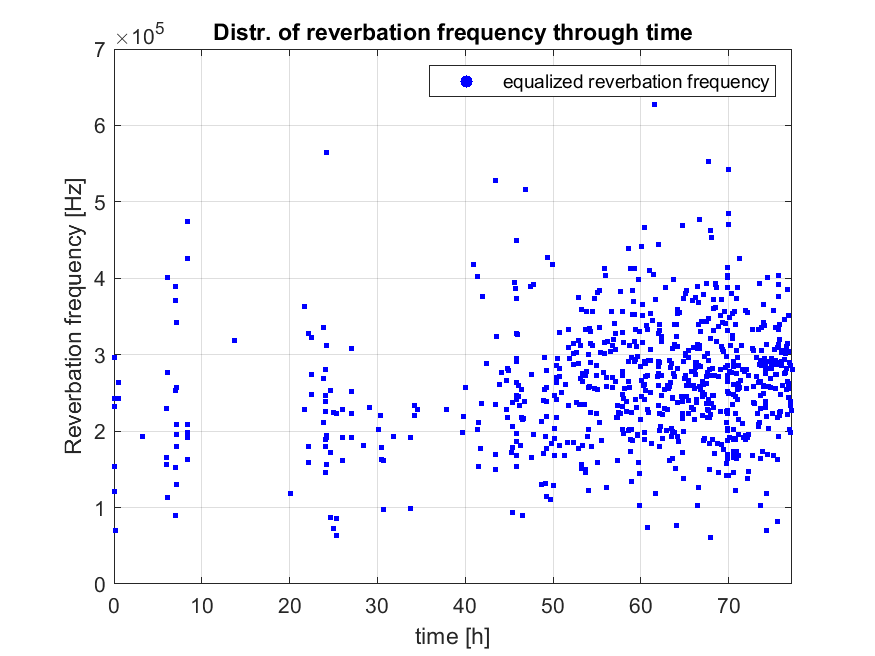


**ASL:**



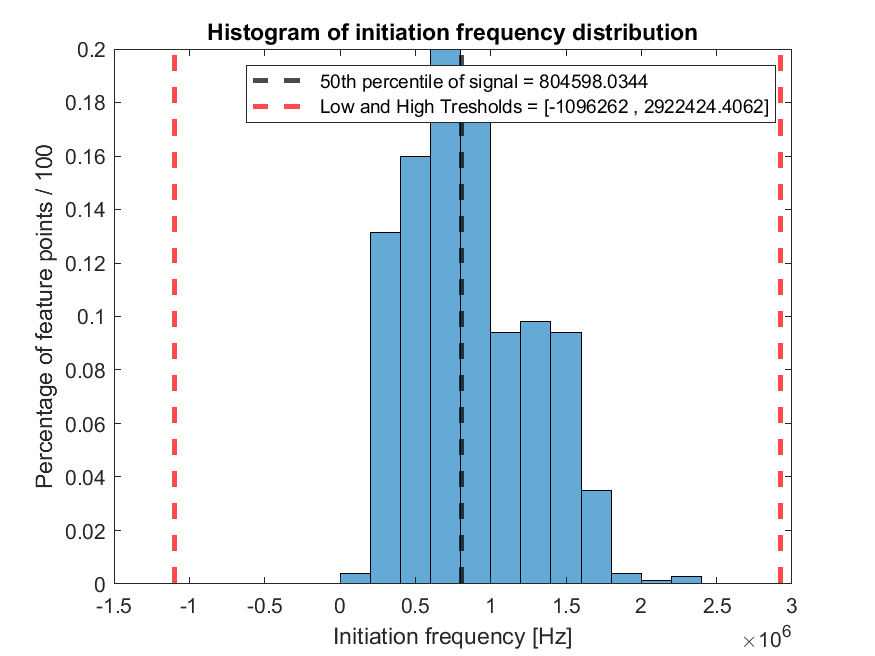
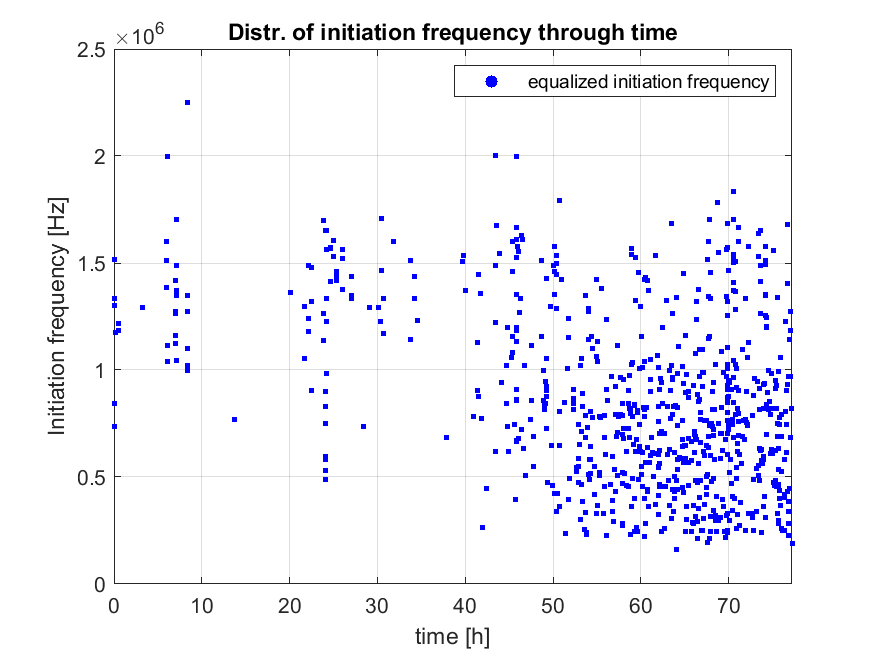
**Ne postoje potencijalni outlieri iznad -1.1797e-05 dB i ispod -4.07e-05 dB ASL.**

**REVERBATION\_FREQUENCY:**



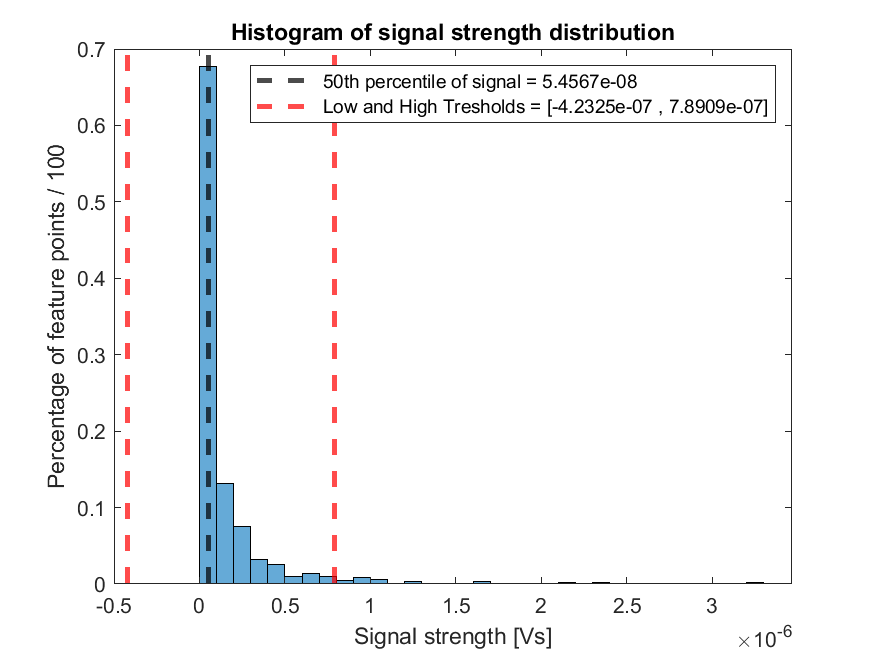
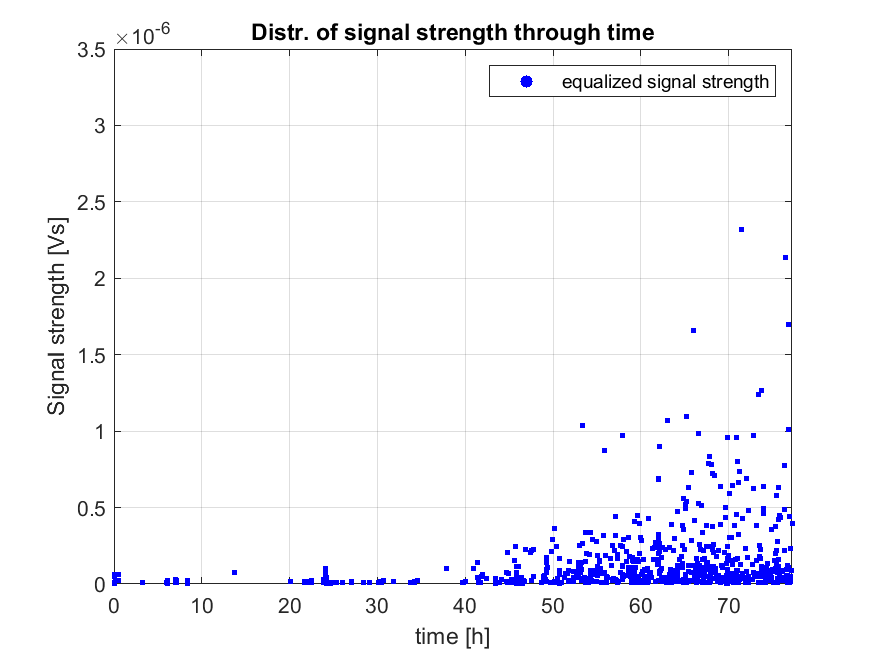
**Ne postoje potencijalni outlieri iznad 627606.7078 Hz i ispod -90982.3947 Hz REVERBATION FREQUENCY.**

**INITIAL\_FREQUENCY:**



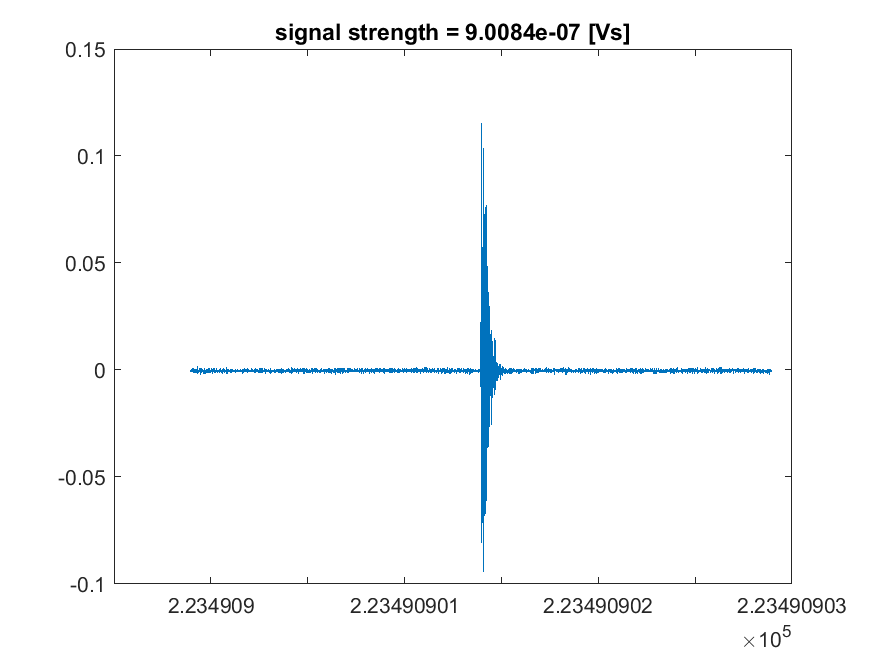
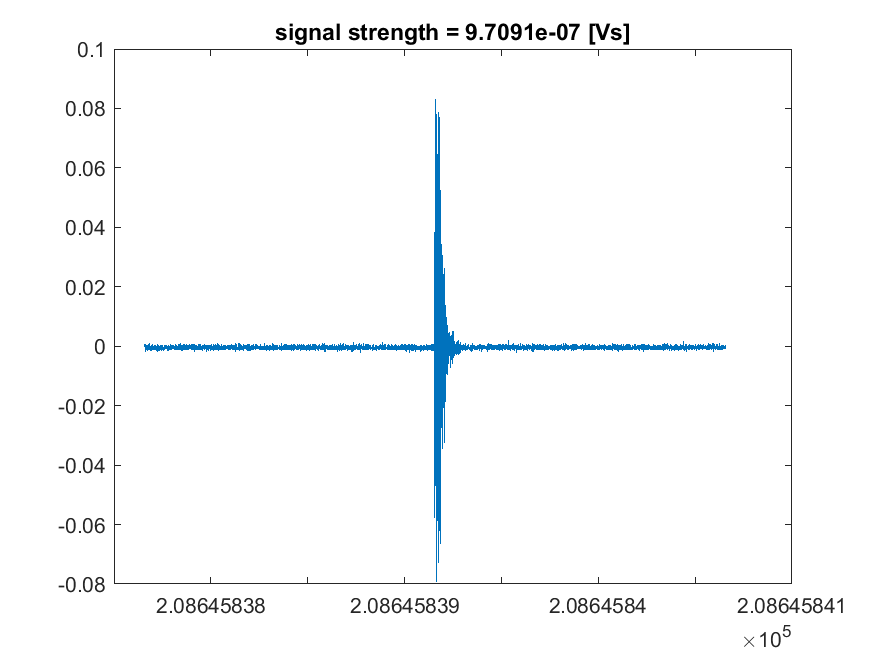
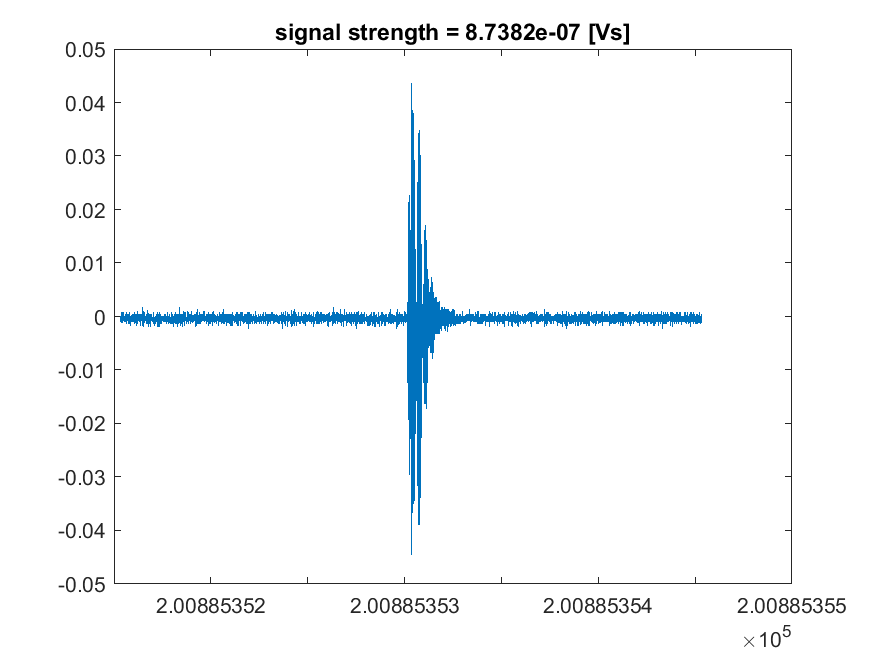
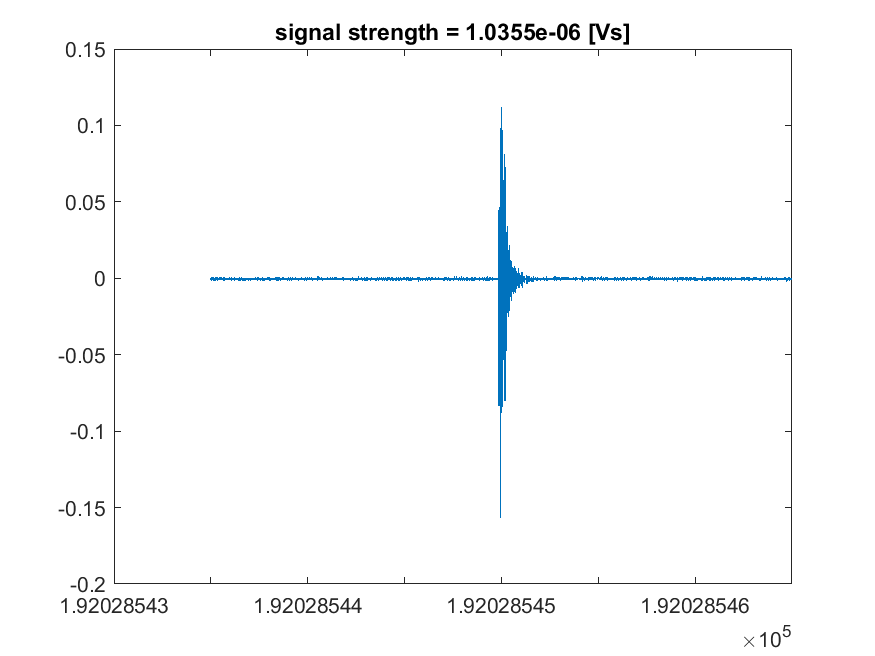
**Ne postoje potencijalni outlieri iznad 2922424.4062 Hz i ispod -1096262 Hz INITIATION FREQUENCY.**

**SIGNAL\_STRENGTH:**

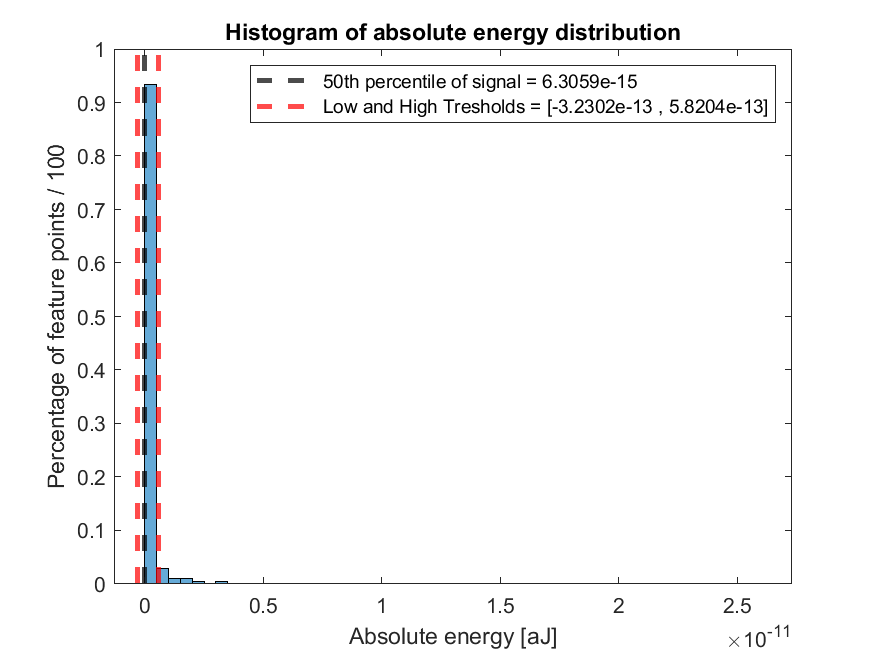
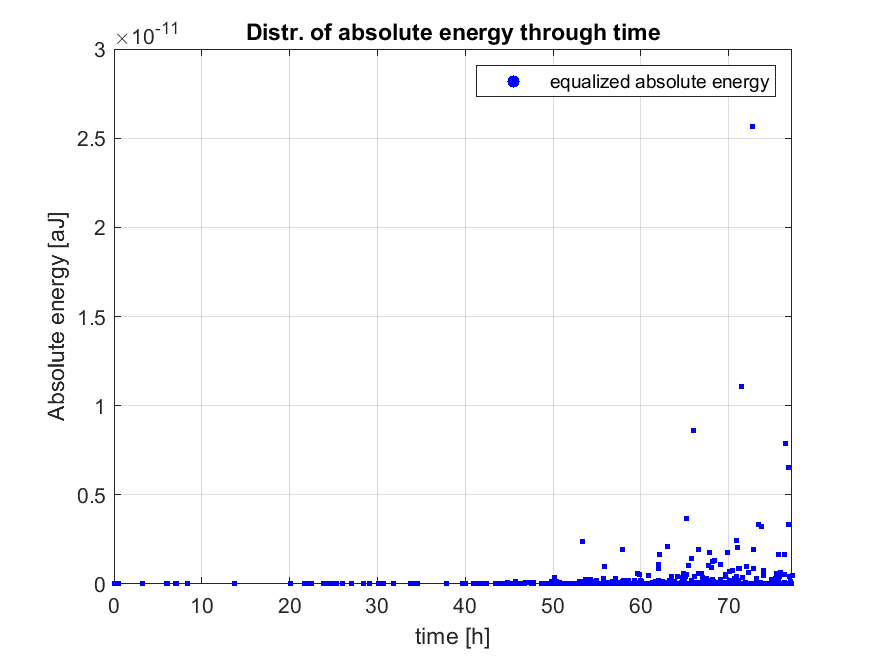


**Potencijalni outlieri iznad 7.8909e-07 Vs i ispod -4.2325e-07 Vs SIGNAL STRENGTH:**

**Postoje validne emisije pa nisu pravi outlieri.**

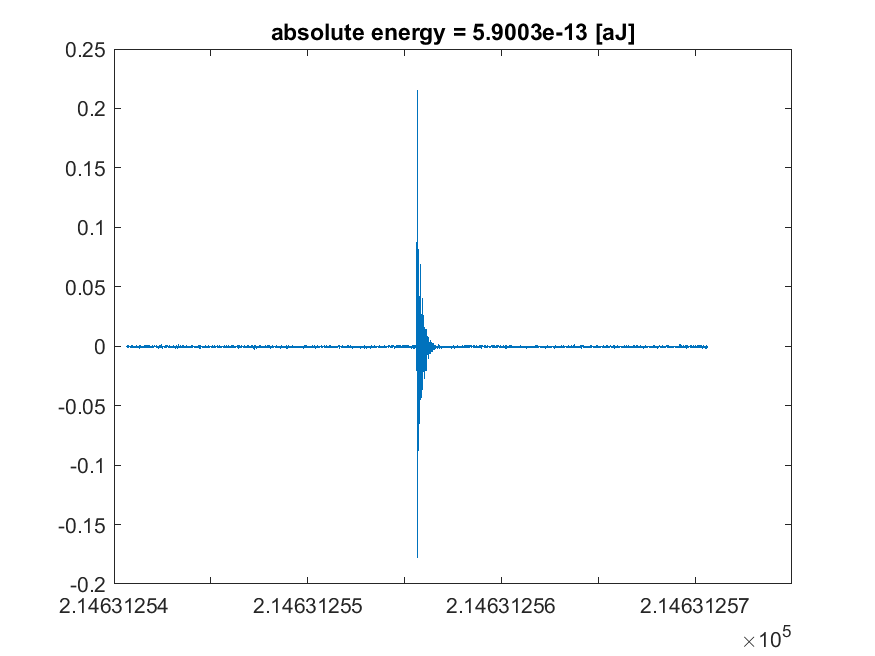
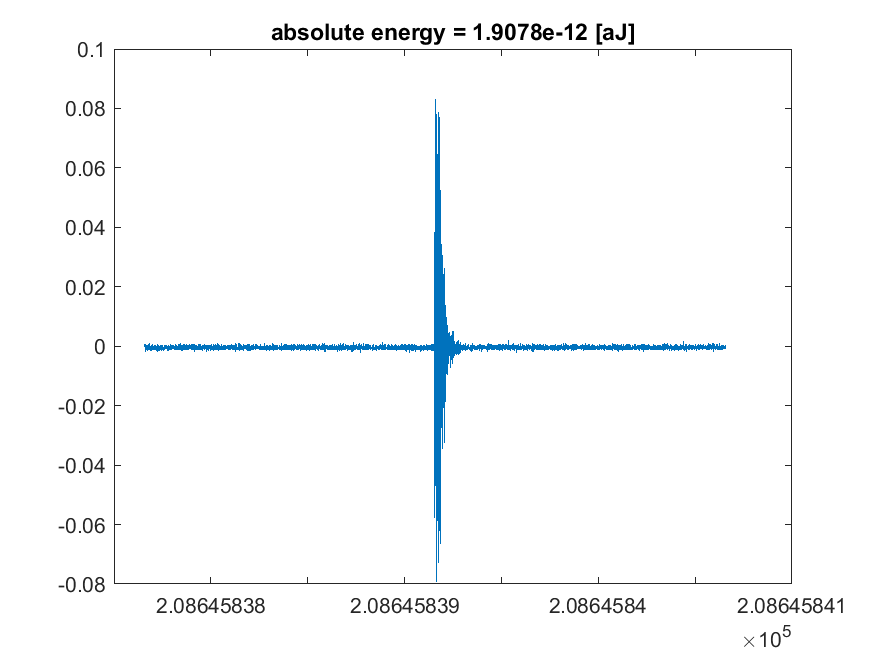
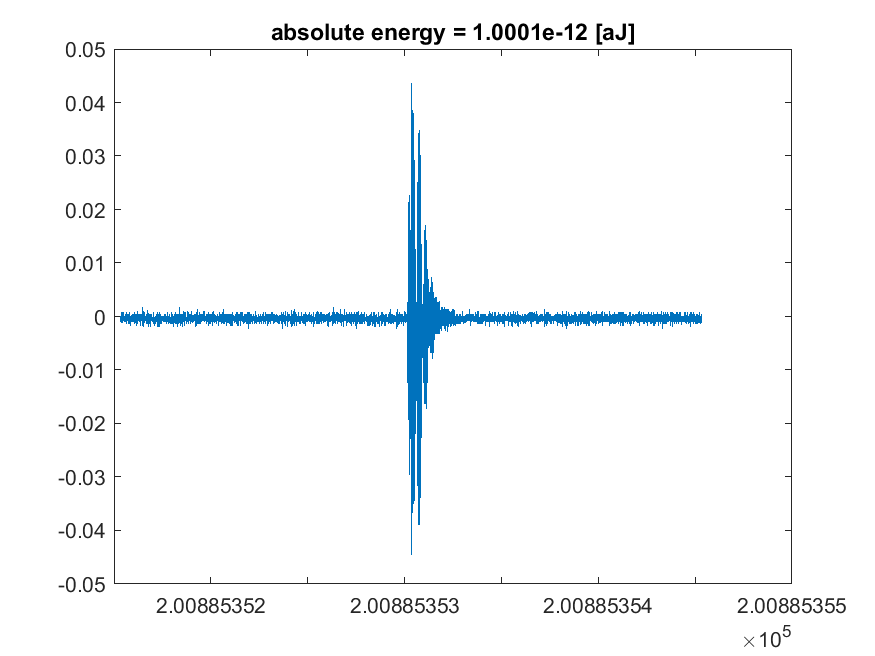
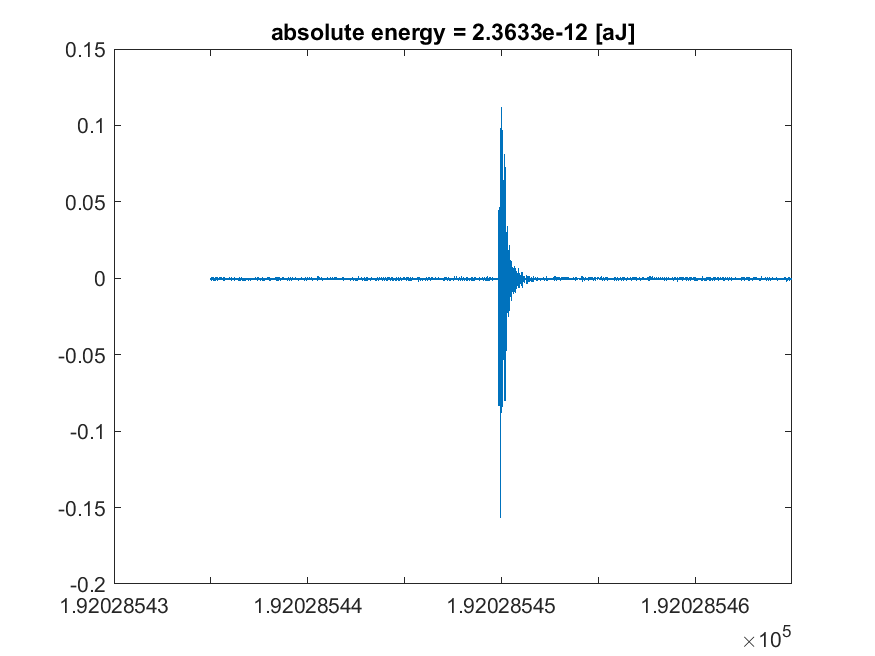


**ABSOLUTE\_ENERGY:**

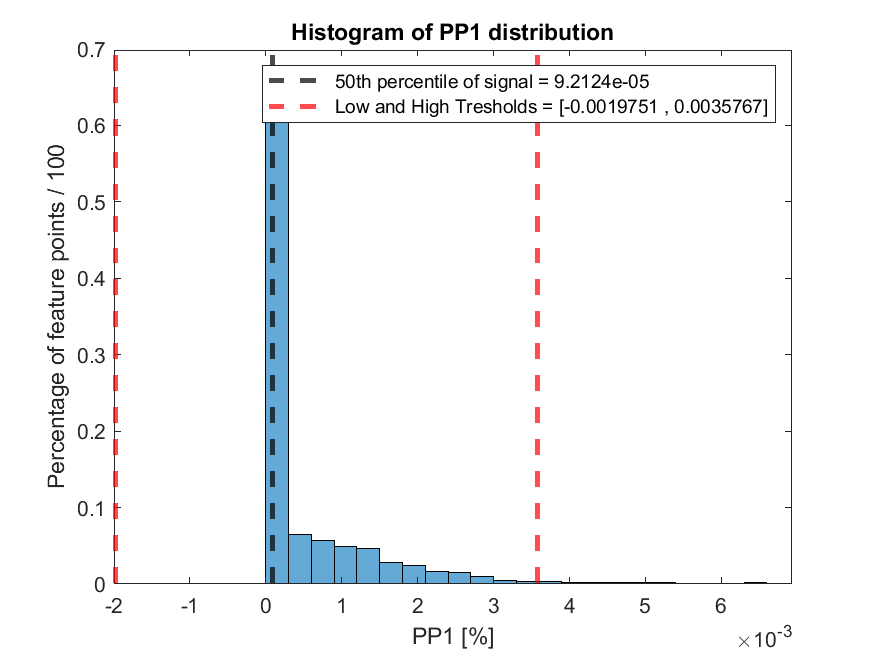
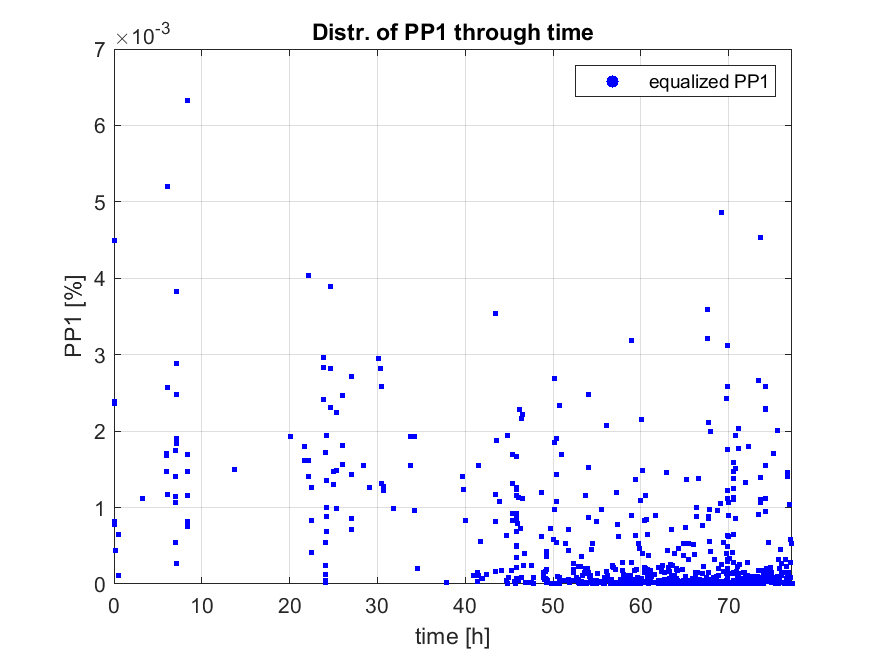


**Potencijalni outlieri iznad 5.8204e-13 aJ i ispod -3.2302e-13 aJ ABSOLUTE ENERGY:**

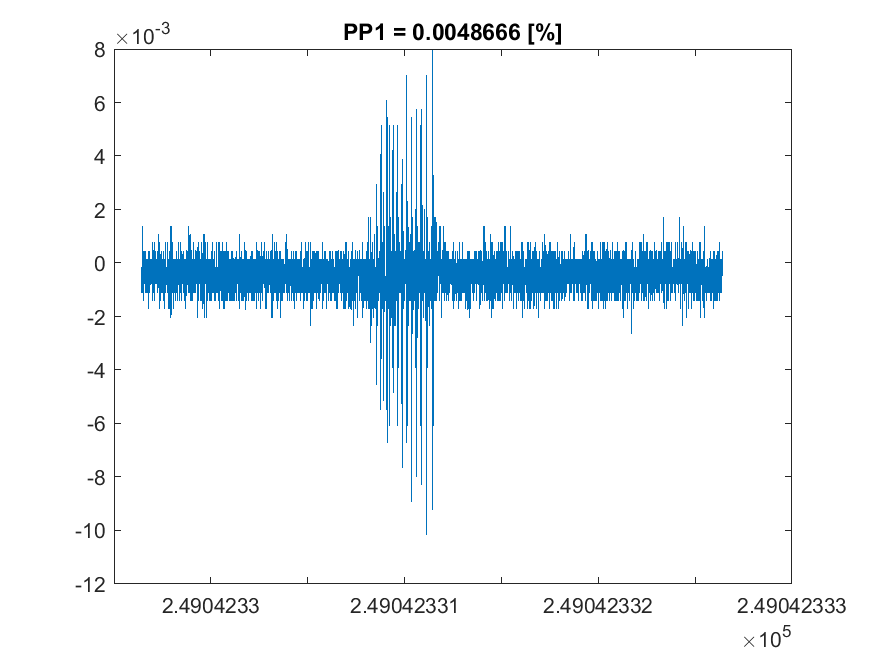
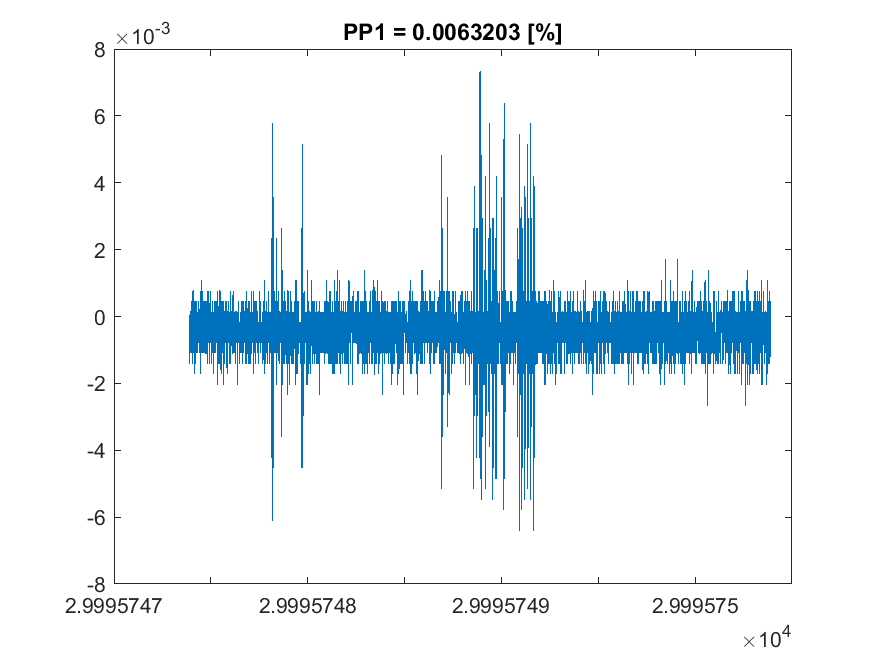
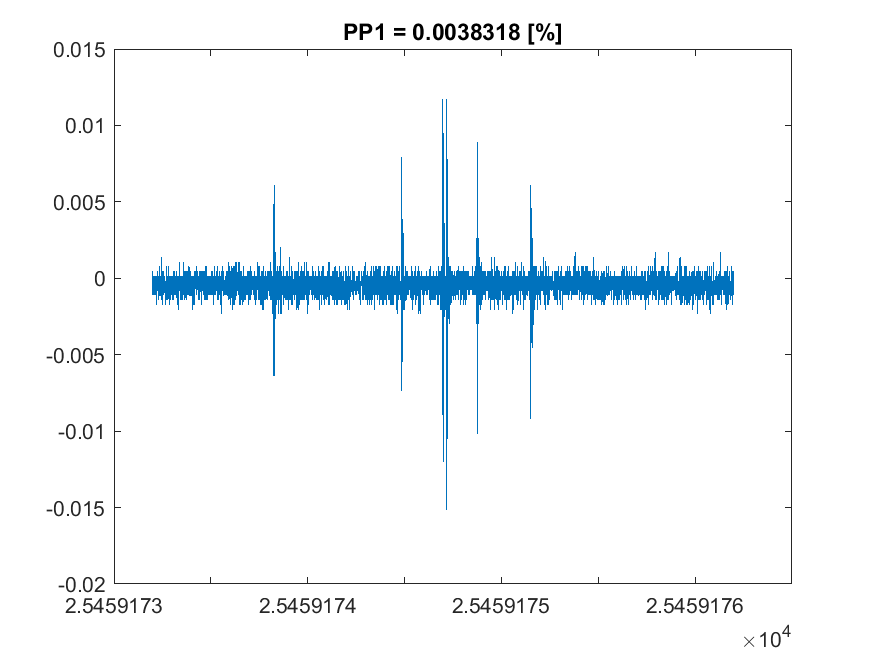
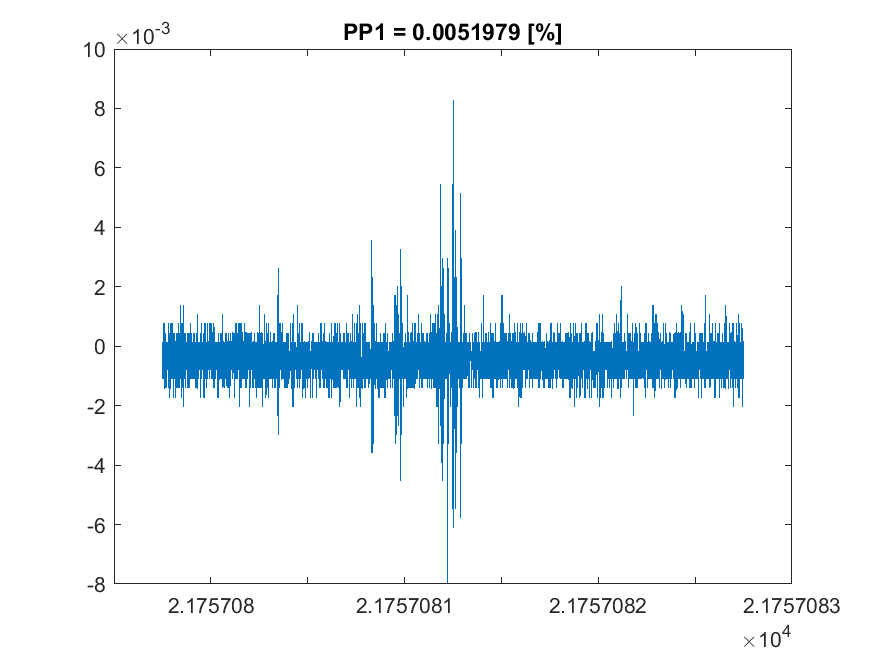
**Postoje validne emisije pa nisu pravi outlieri.**



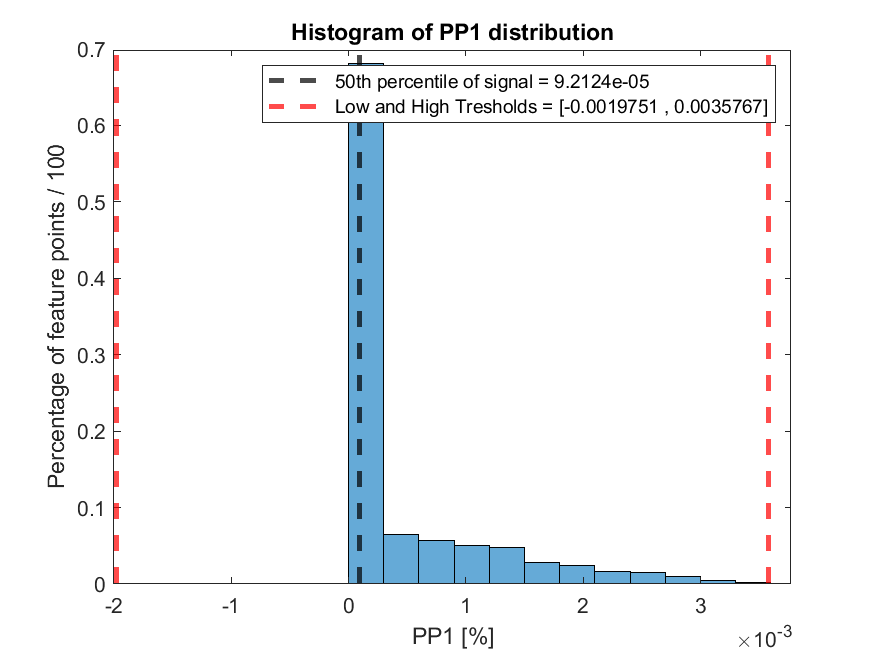
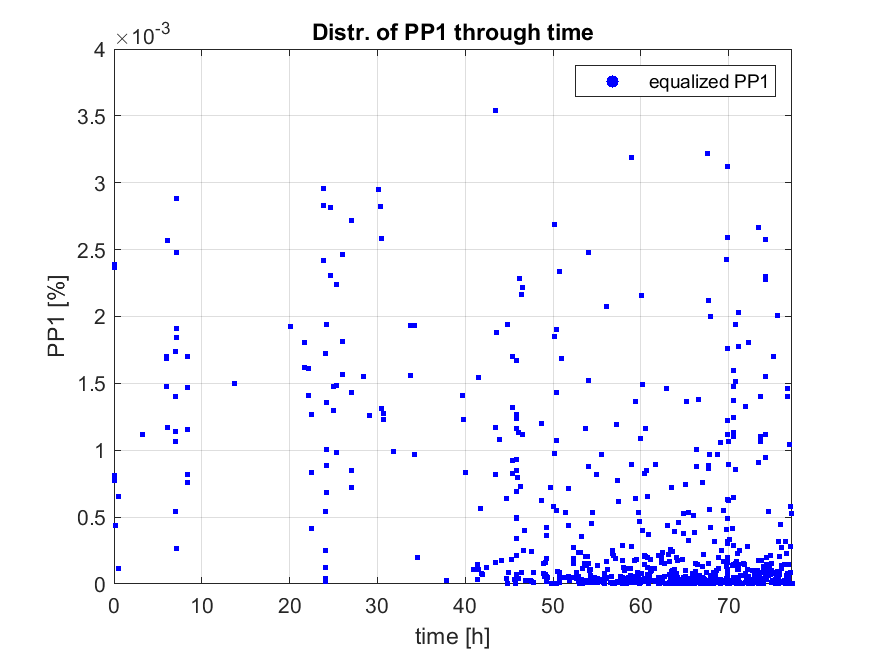
**PP1:**



**Potencijalni outlieri iznad 0.0035767 % i ispod -0.0019751 % PP1:**



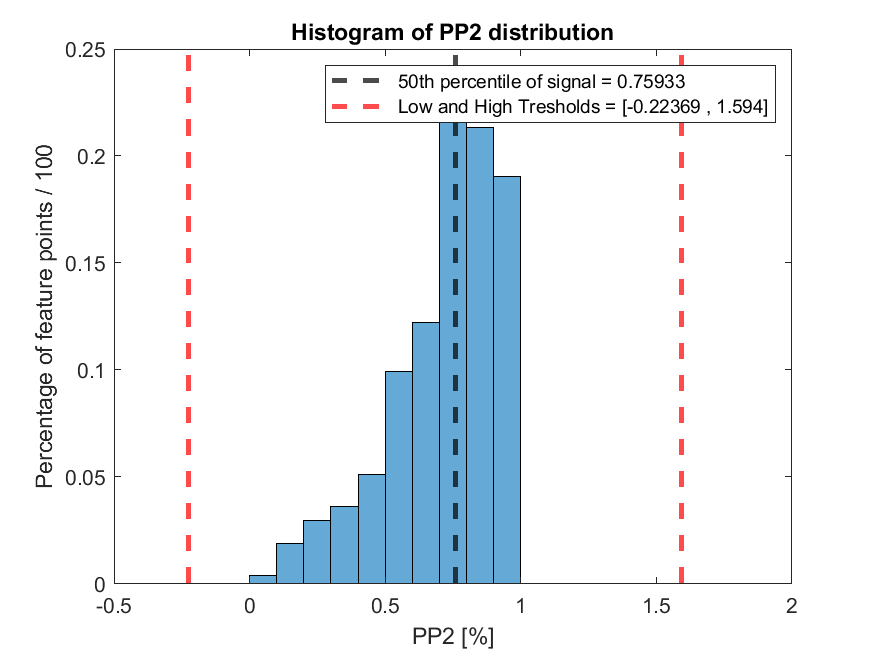
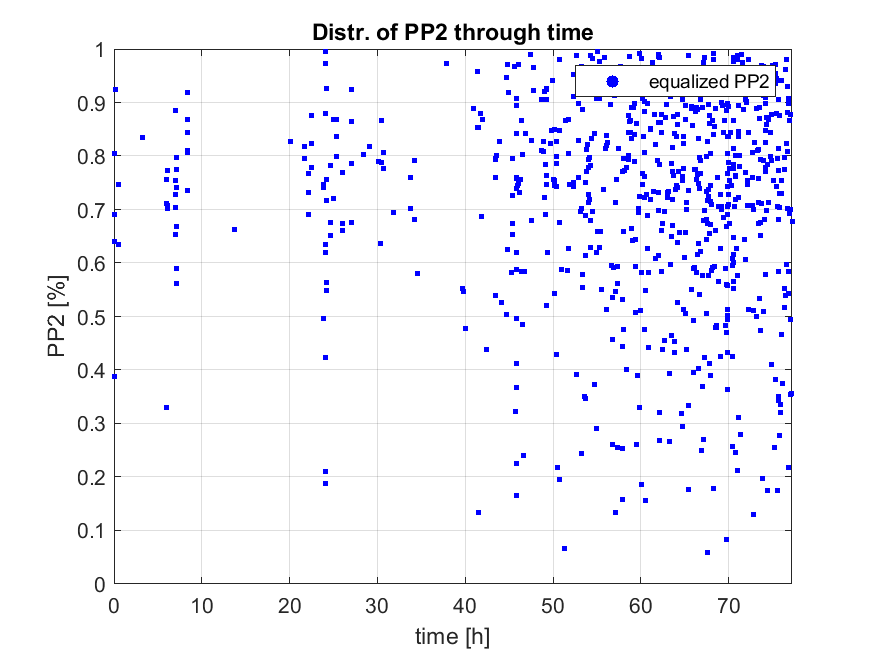
**Nakon izbacivanja outlieri iznad 0.0035767 % i ispod -0.0019751 % PP1 dobivamo slijedeću distribuciju:**



**PP1 outliers are emissions with value more than 0.0035767 [%]**

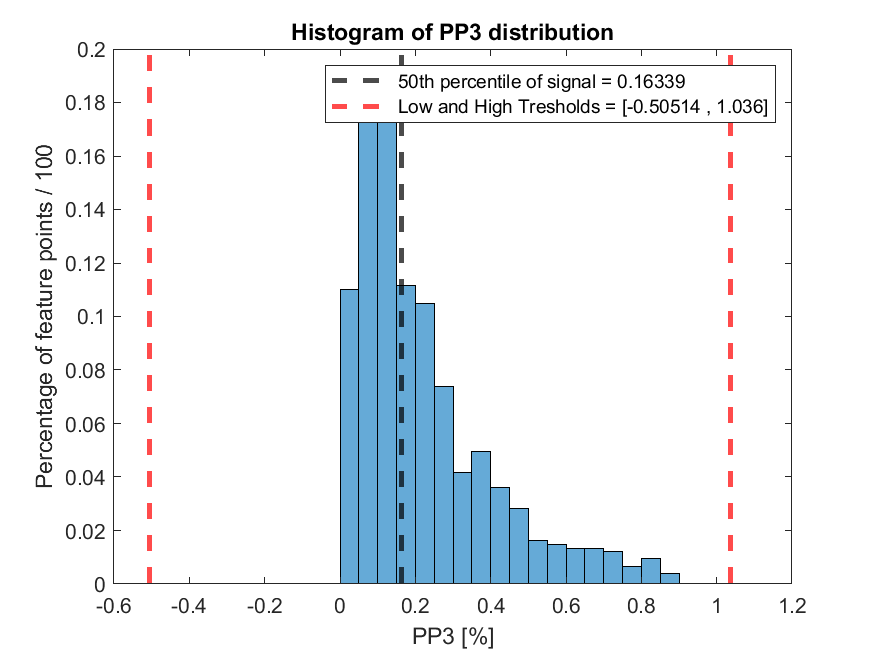
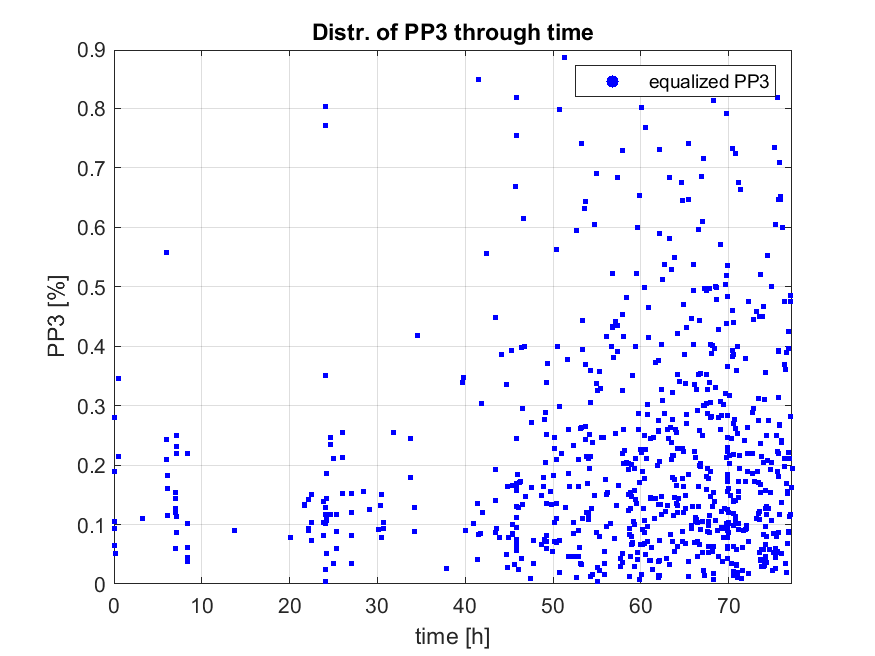
**Percentage of outliers for PP1: 9/745 = 1.2081 %**

**PP2:**



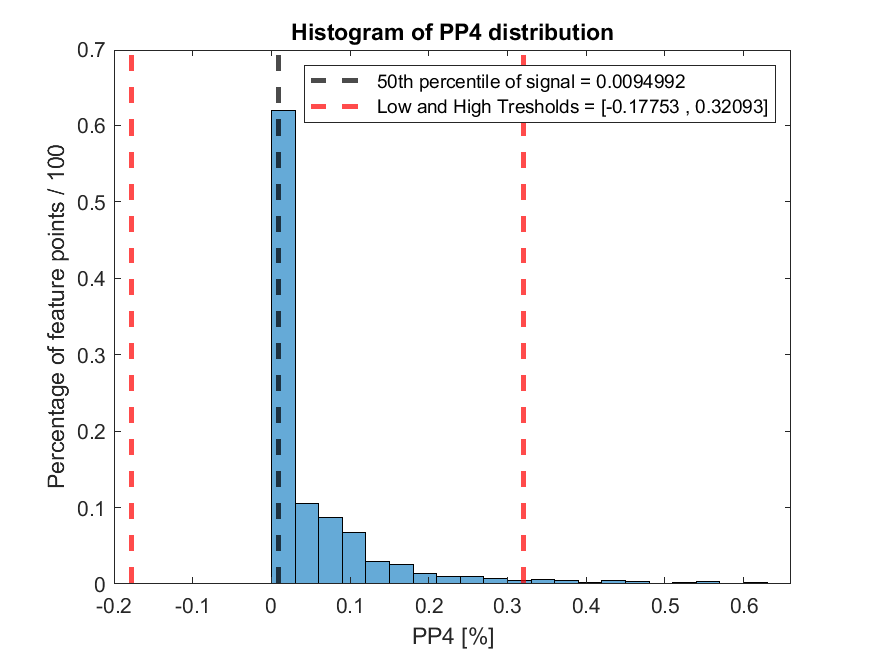
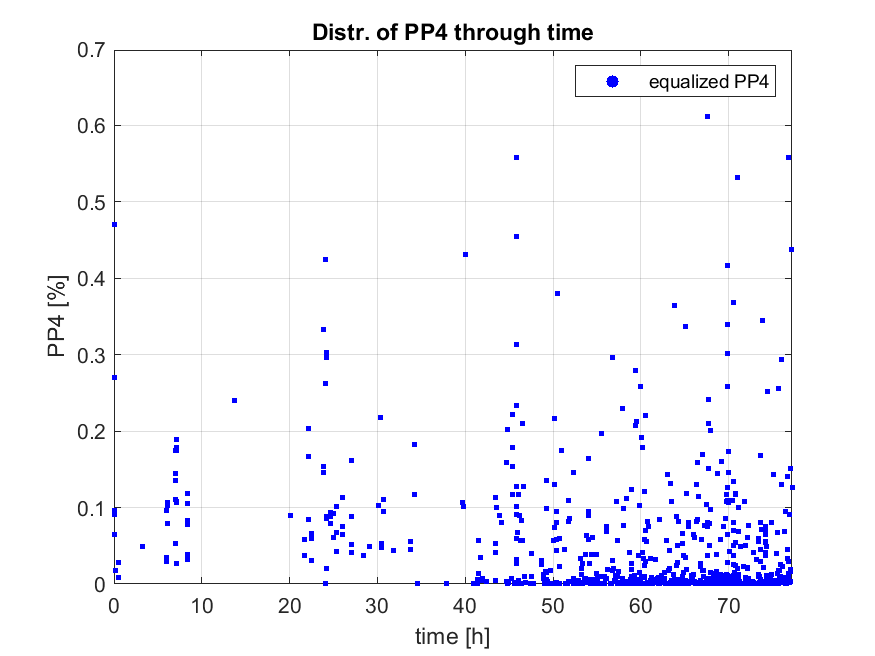
**Nema potencijalni outlieri iznad 1.594 % i ispod -0.22369 % PP2:**

**PP3:**



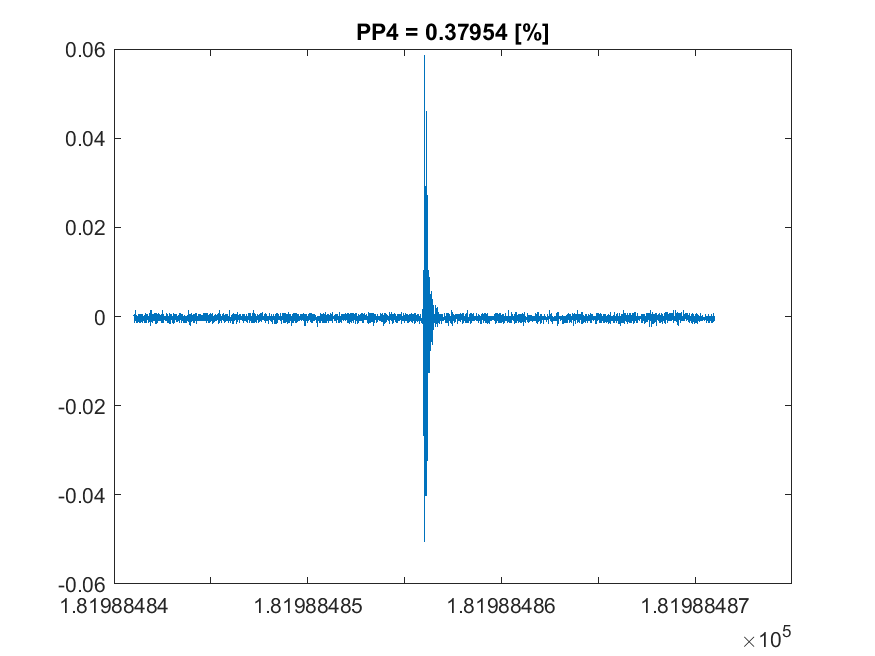
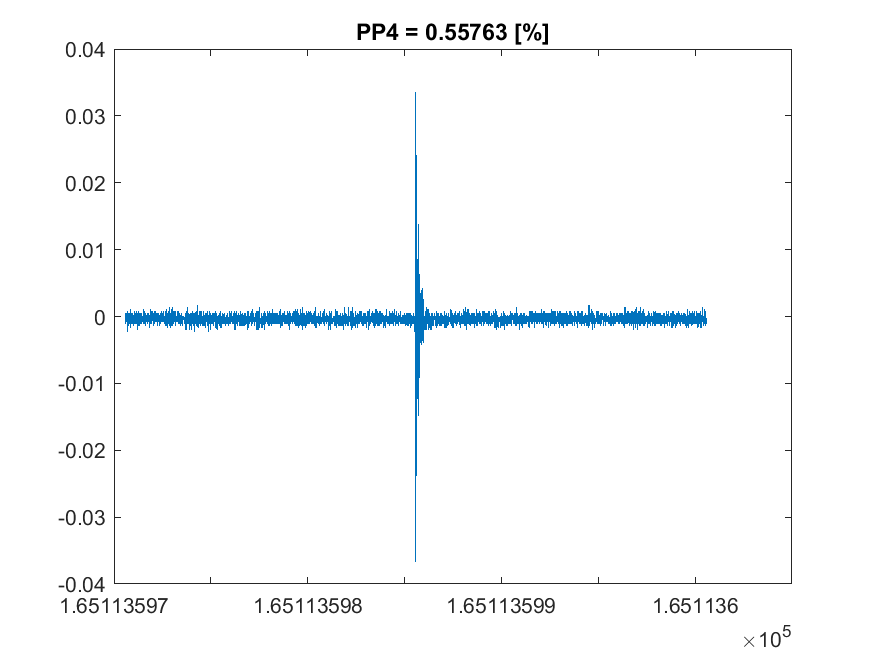
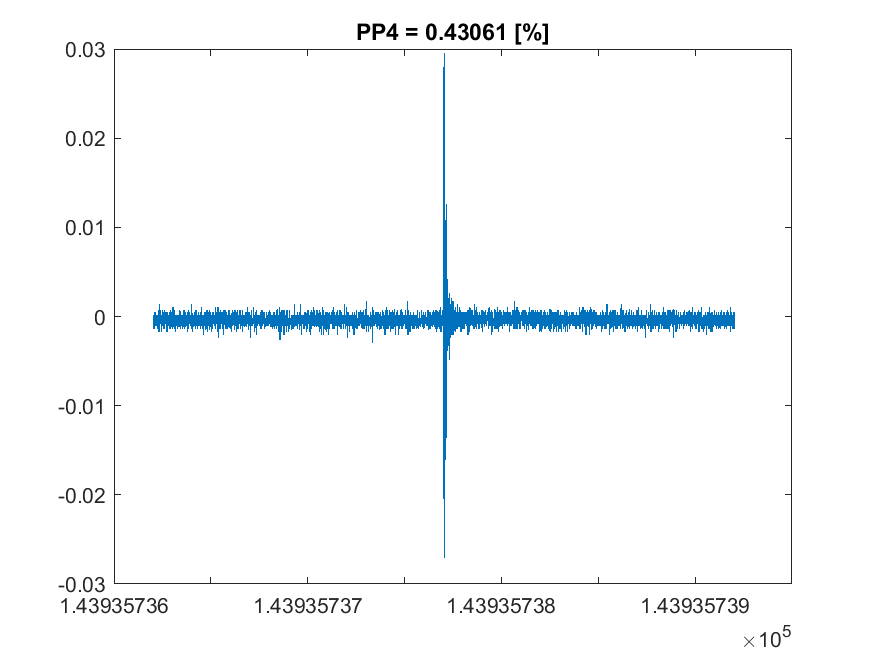
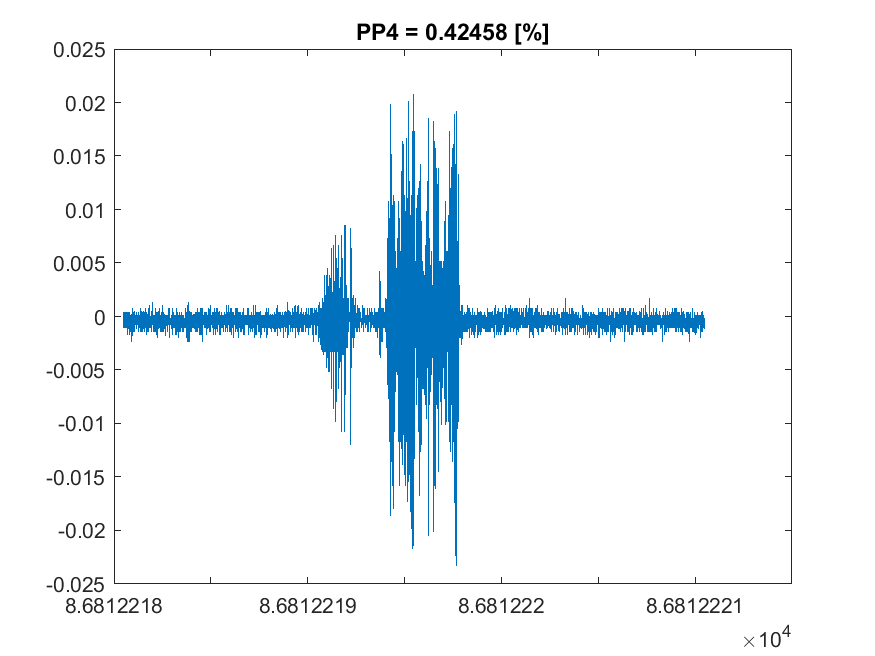
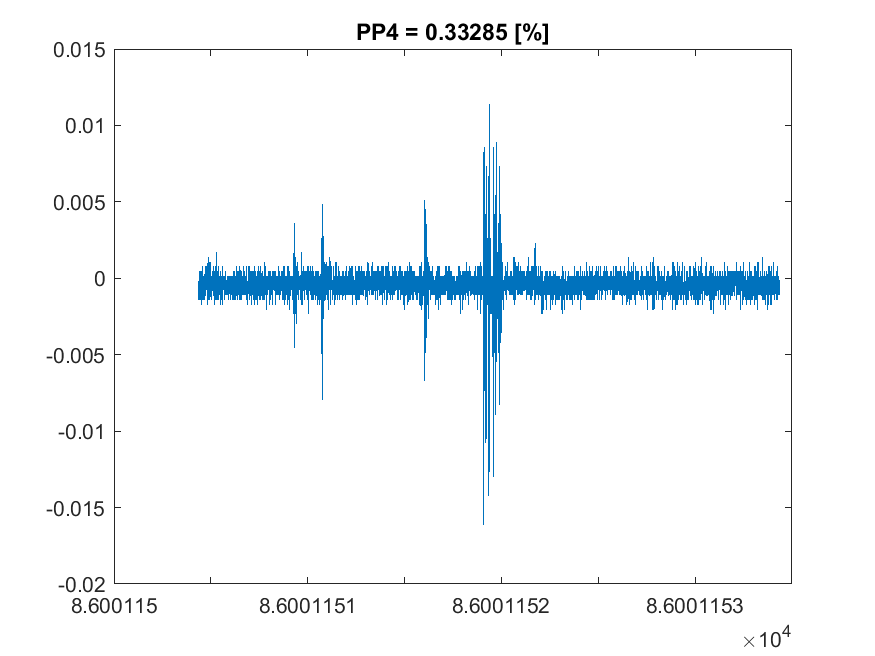
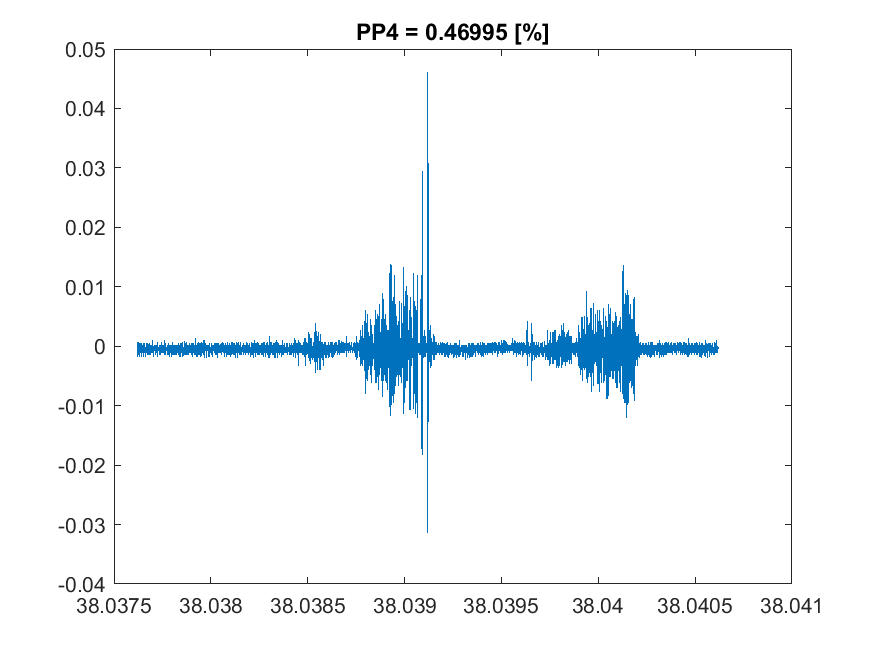
**Nema potencijalnih outlieri iznad 1.036 % i ispod -0.50514 % PP3:**

**PP4:**

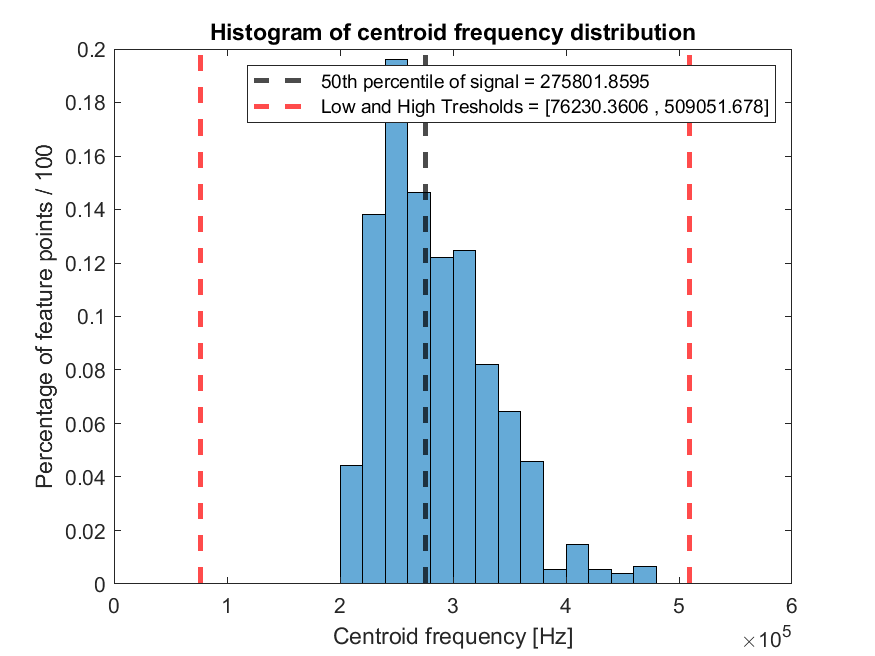
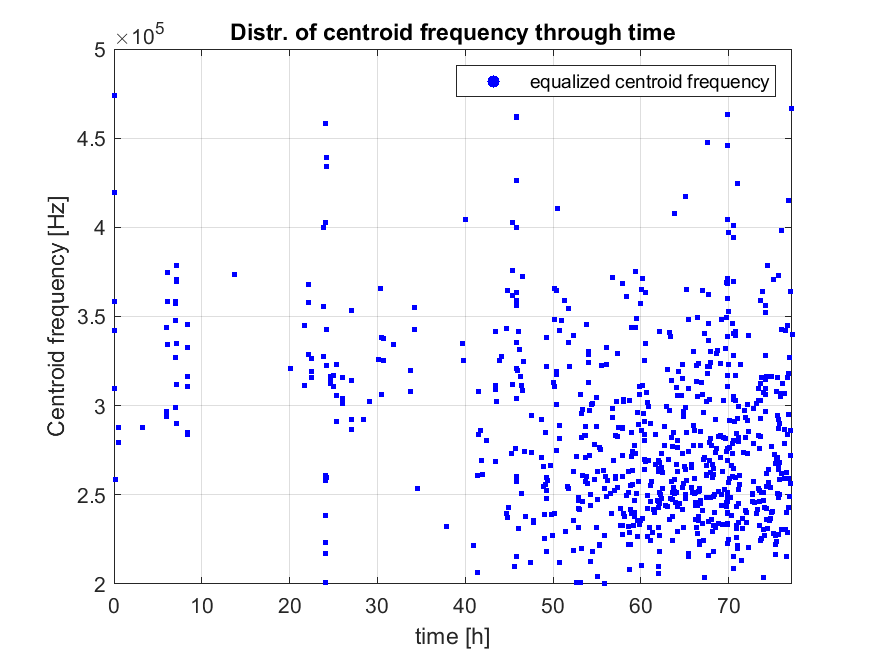


**Potencijalni outlieri iznad 0.32093 % i ispod -0.17753 % PP4:**

**Vidimo da ima ispravnih i neispravnih podataka pa emisije nisu pravi outlieri.**

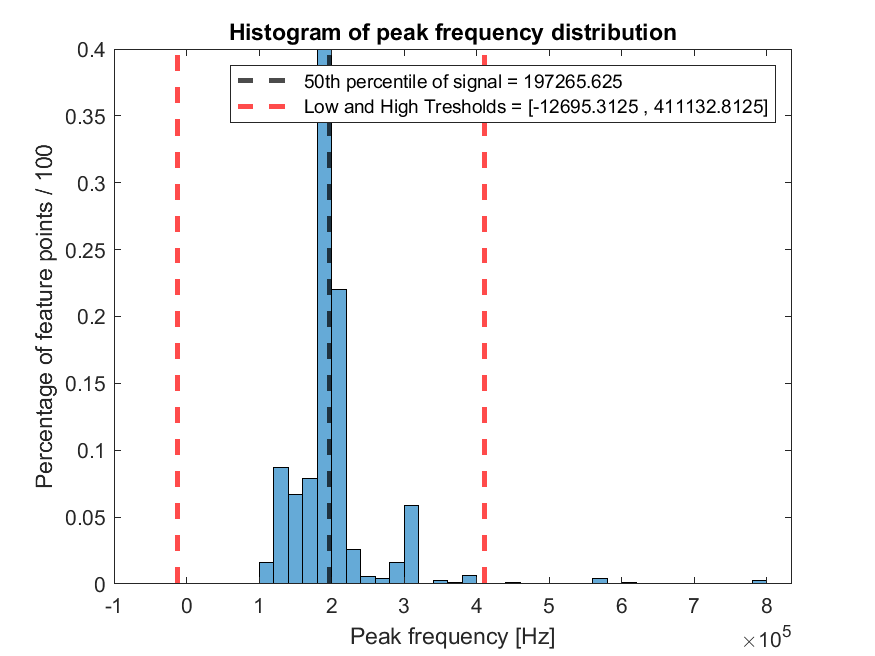
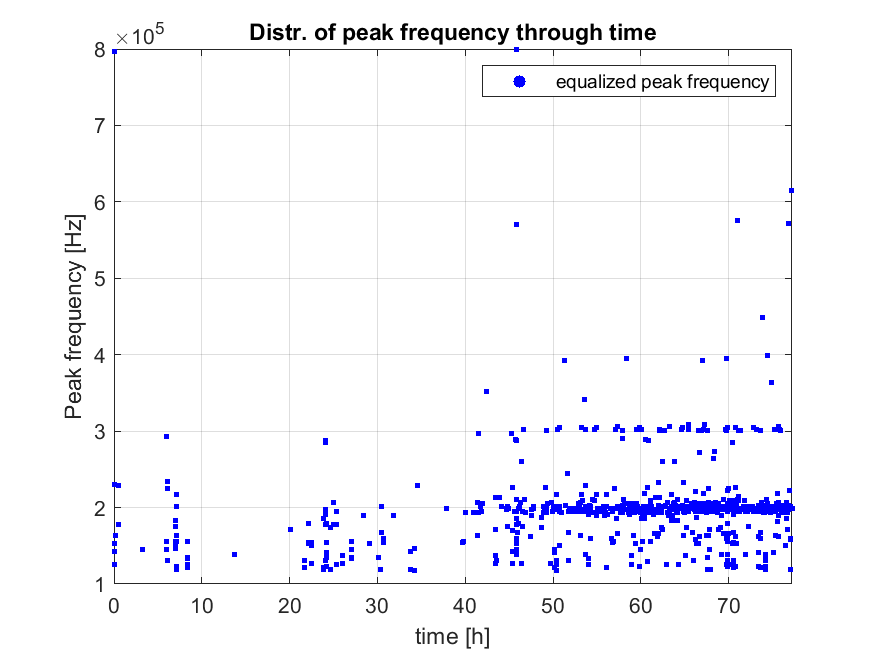


**CENTROID\_FREQUENCY:**



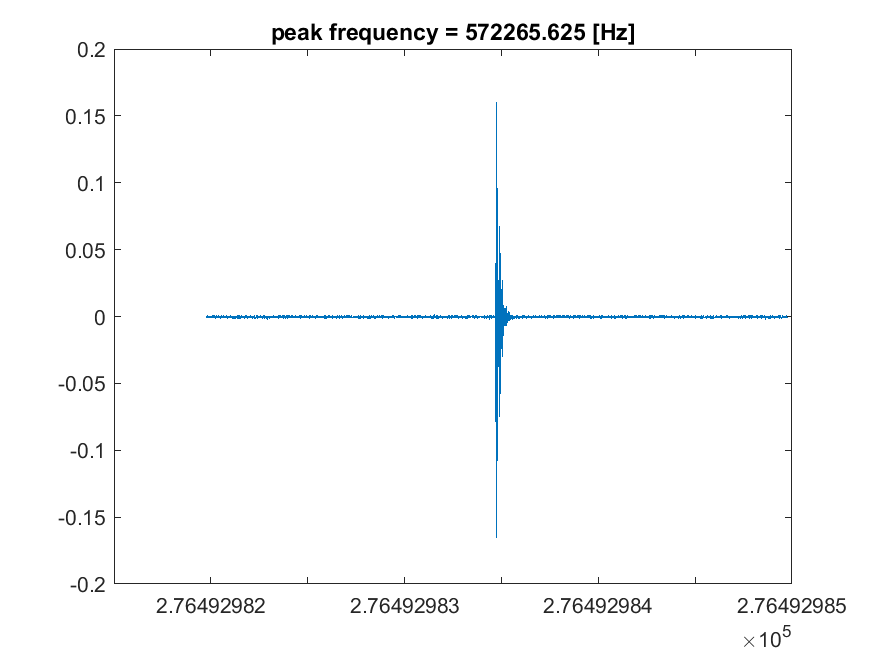
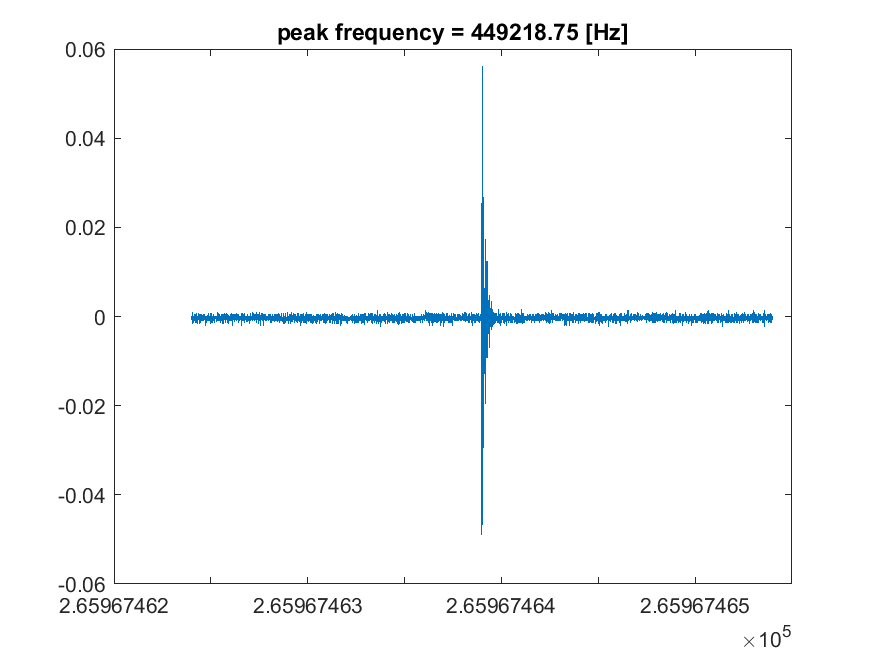
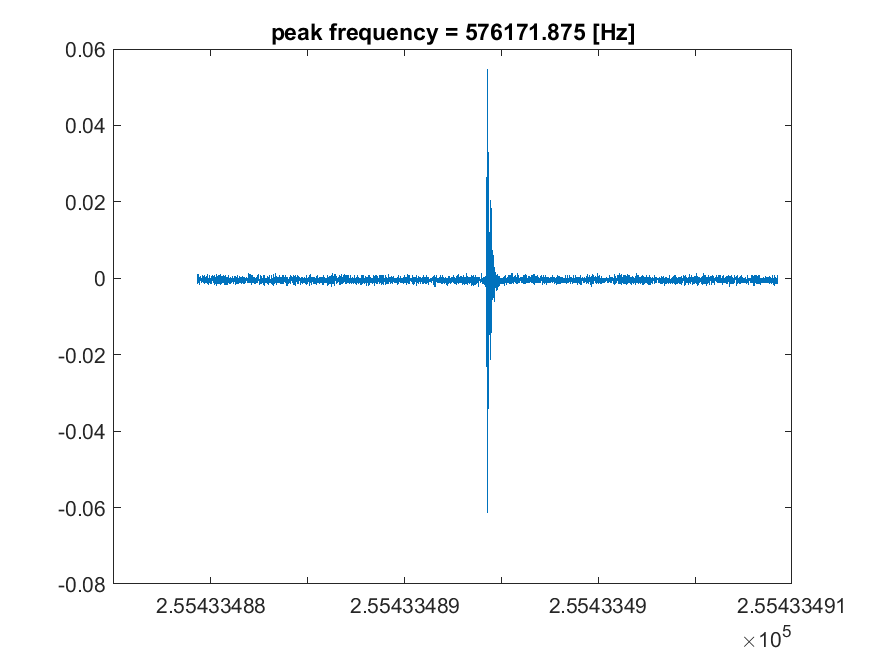
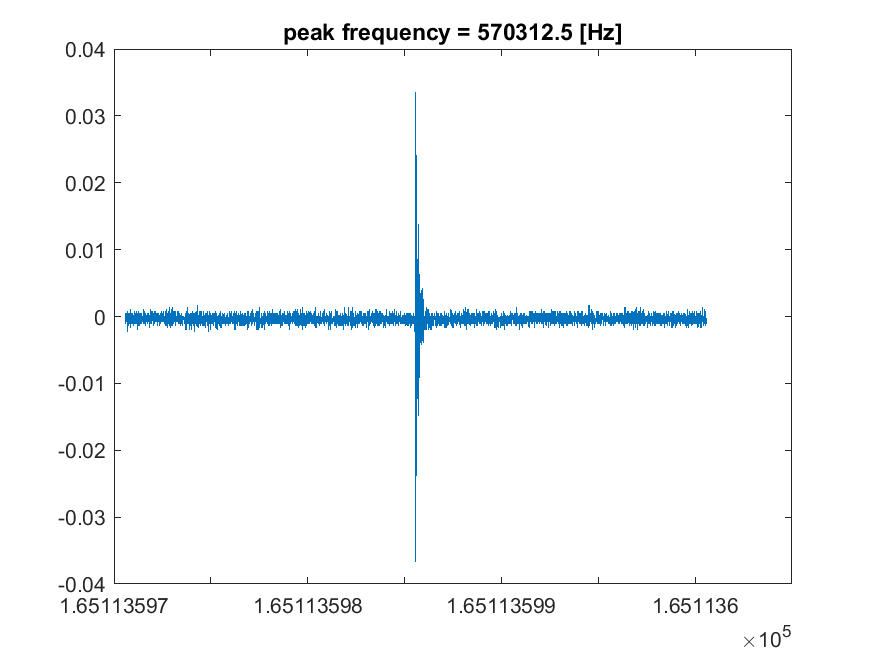
**Nema potencijalnih outlieri iznad 509051.678 Hz i ispod 76230.3606 Hz CENTROID FREQUENCY:**

**PEAK\_FREQUENCY:**

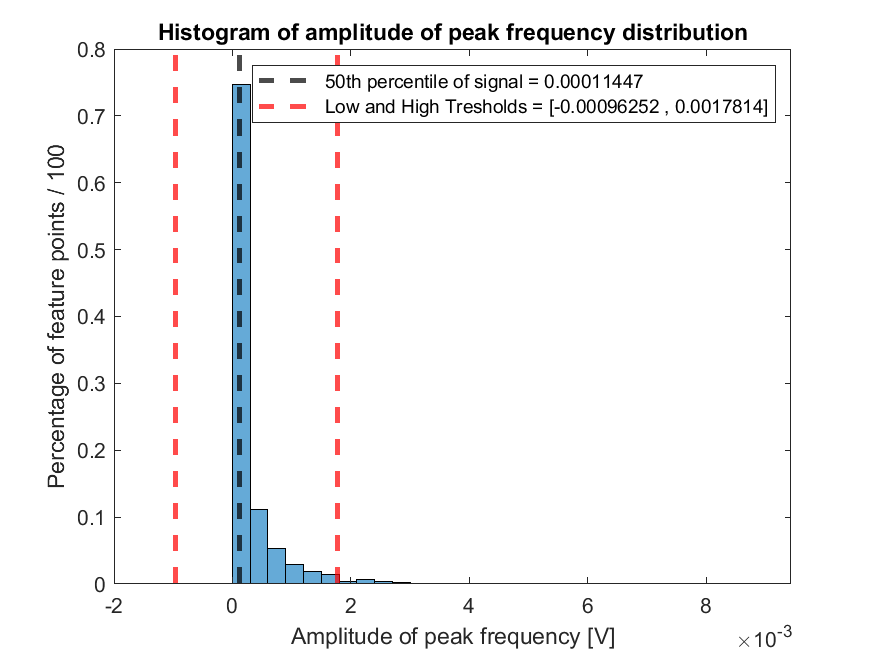
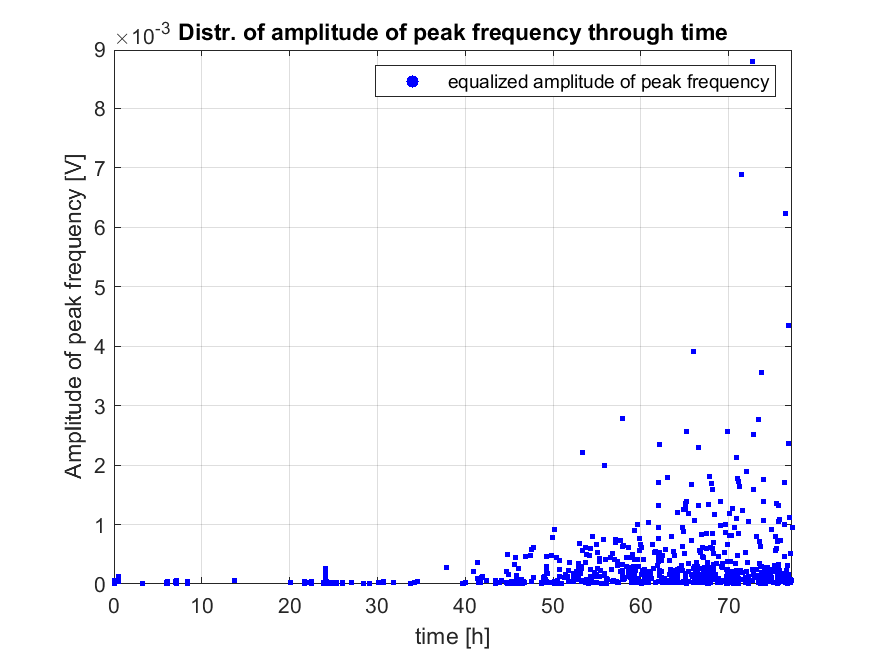


**Potencijalni outlieri iznad 411132.8125 Hz i ispod -12695.3125 Hz PEAK FREQUENCY:**

**Postoje validne emisije pa nisu pravi outlieri.**

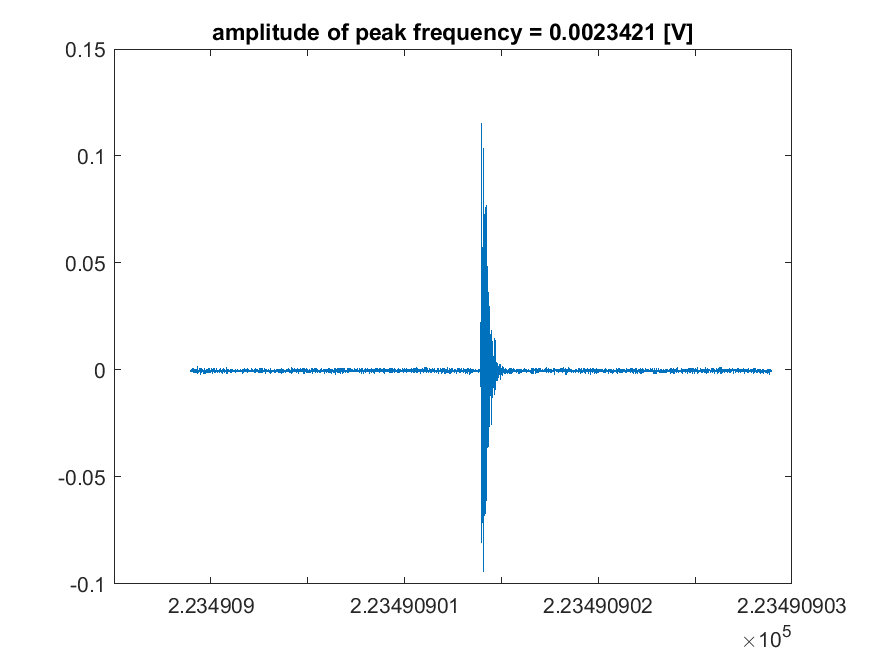
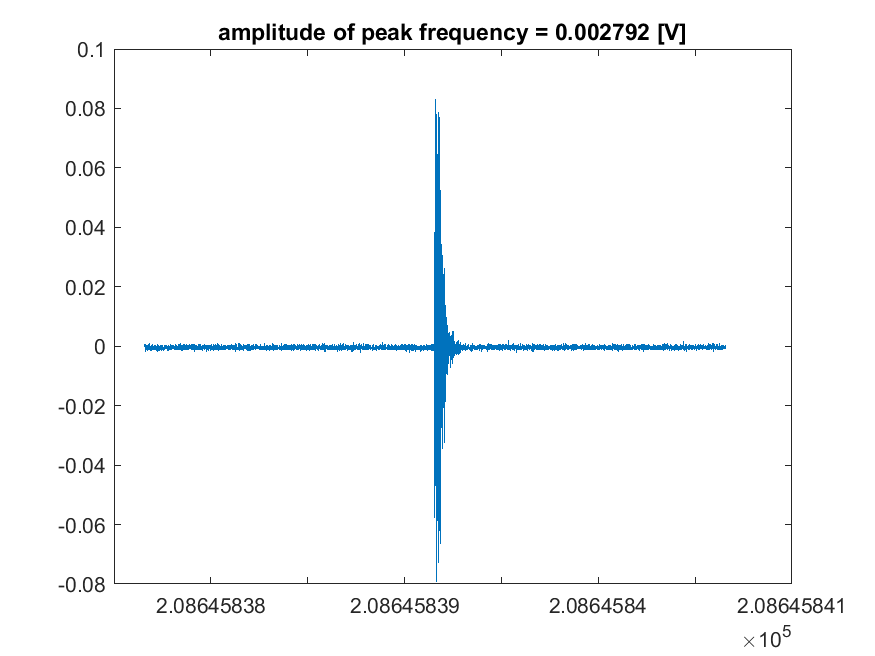
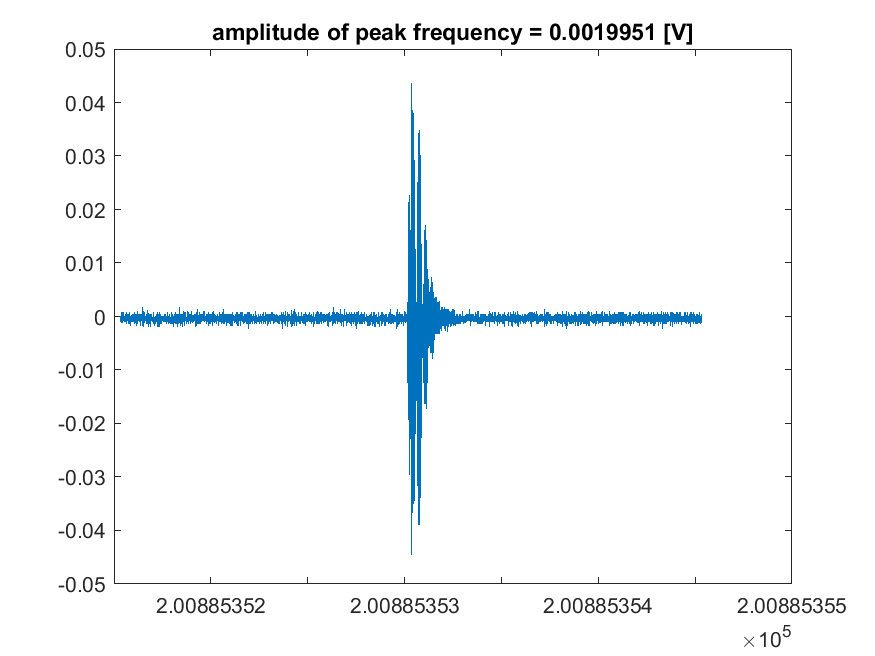
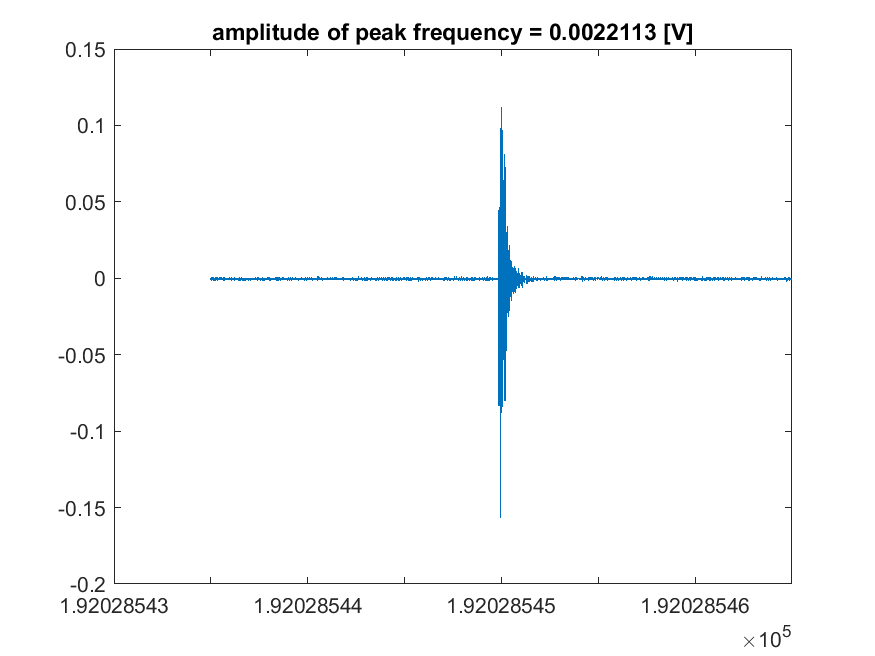


**AMPLITUDE\_OF\_PEAK\_FREQUENCY:**

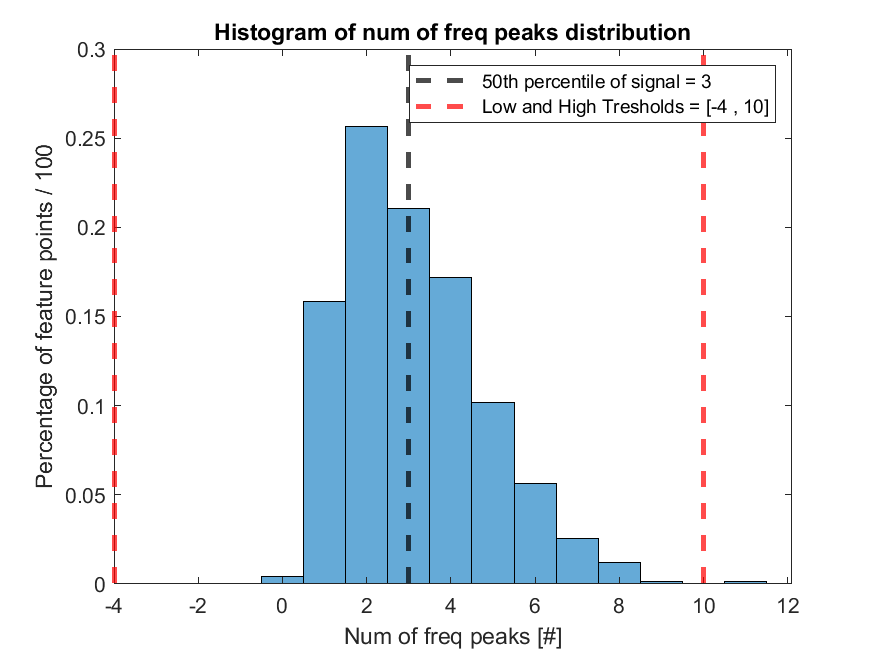
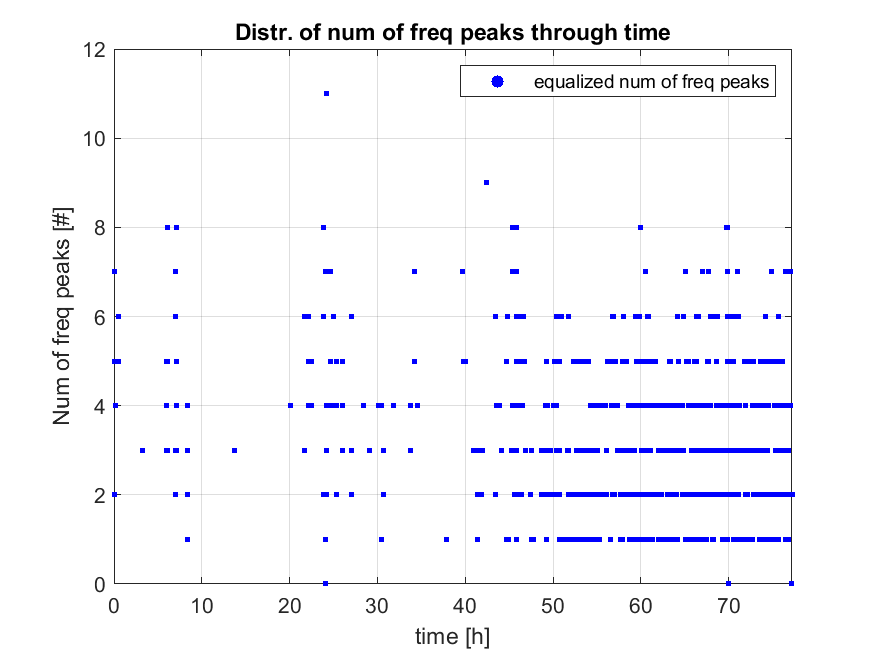


**Potencijalni outlieri iznad 0.0017814 V i ispod -0.00096252 V AMPLITUDE OF PEAK FREQUENCY:**

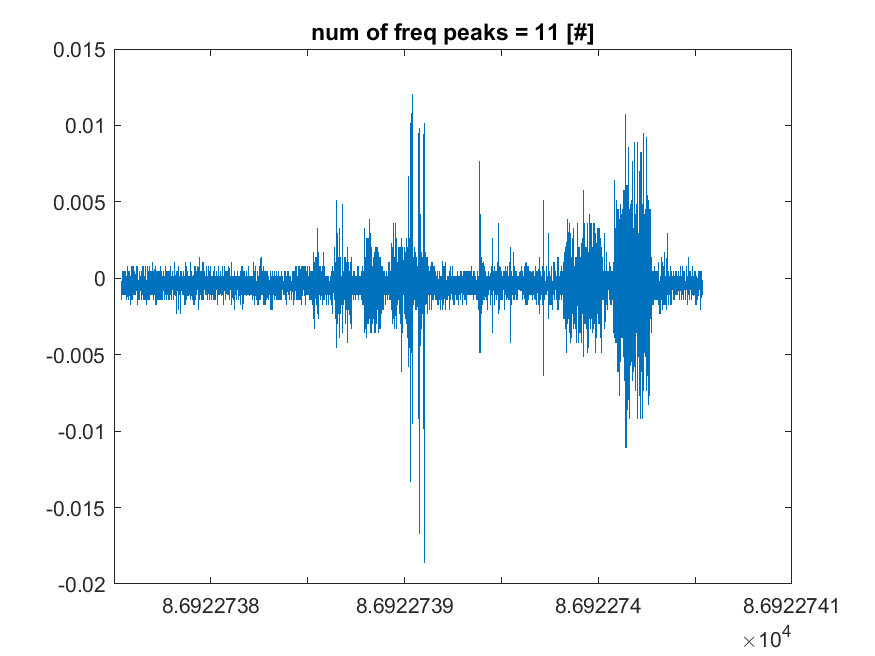
**Postoje validne emisije pa nisu pravi outlieri.**



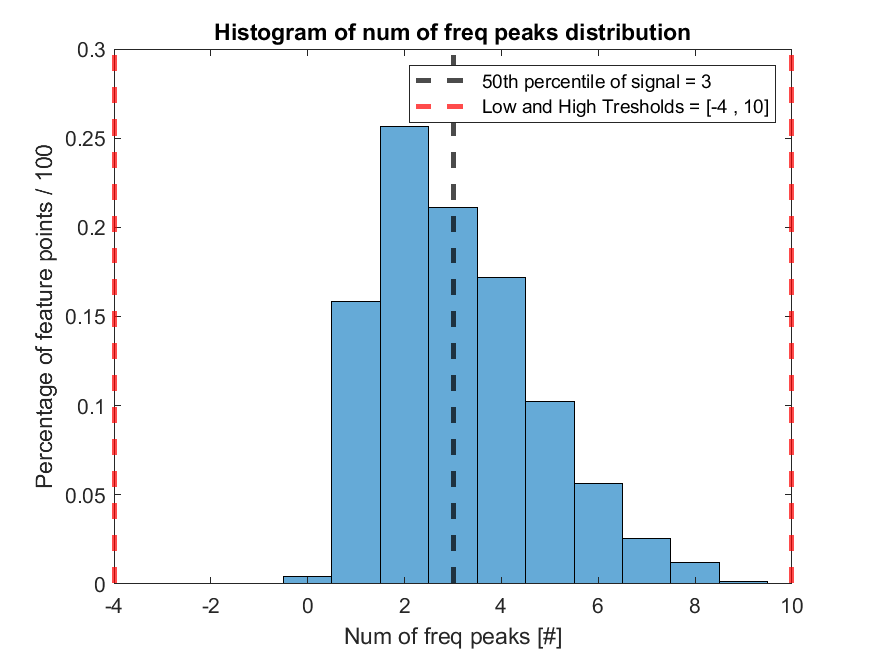
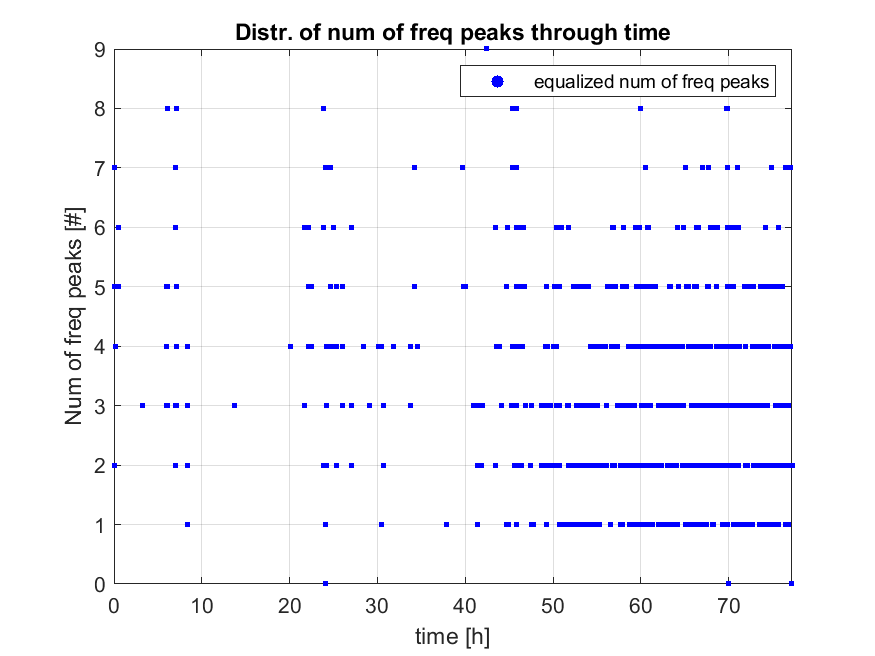
**NUM\_OF\_FREQ PEAKS:**



**Potencijalni outlieri iznad 10 # i ispod -4 # NUM OF FREQ PEAKS:**



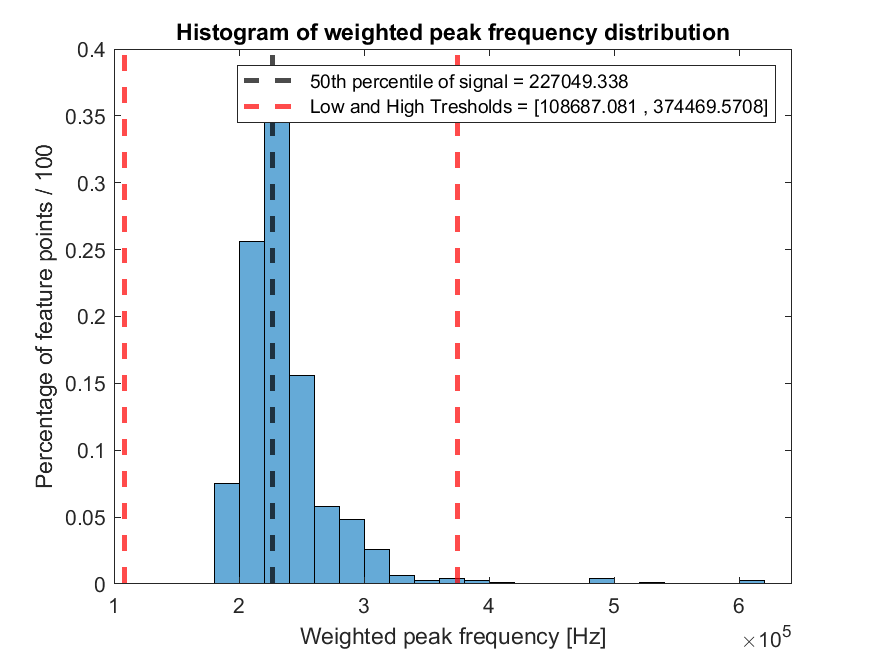
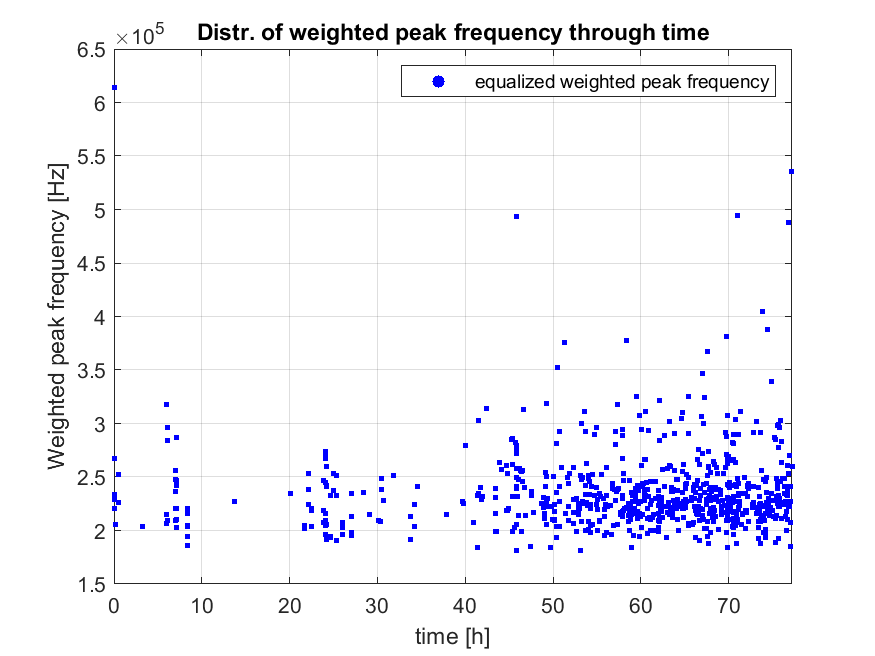
**Nakon izbacivanja outlieri iznad 10 # i ispod -4 # NUM OF FREQ PEAKS dobivamo slijedeću distribuciju:**



**NUM OF FREQ PEAKS outliers are emissions with value more than 10 [#]**

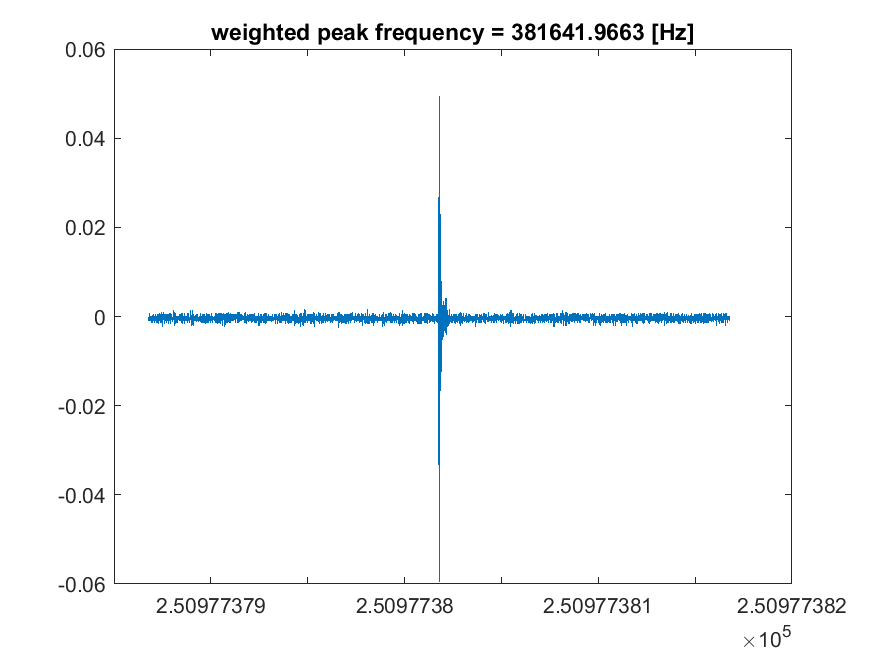
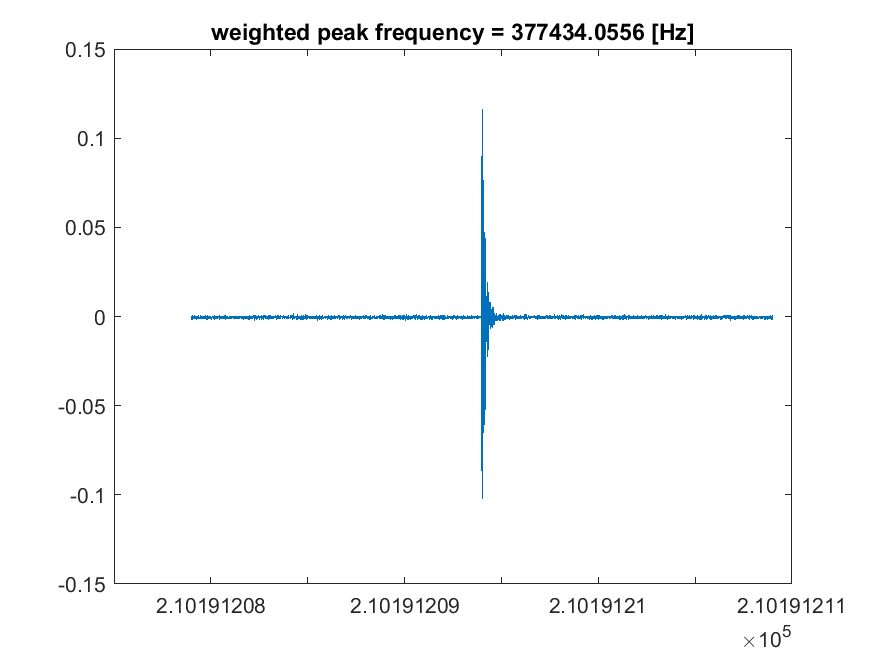
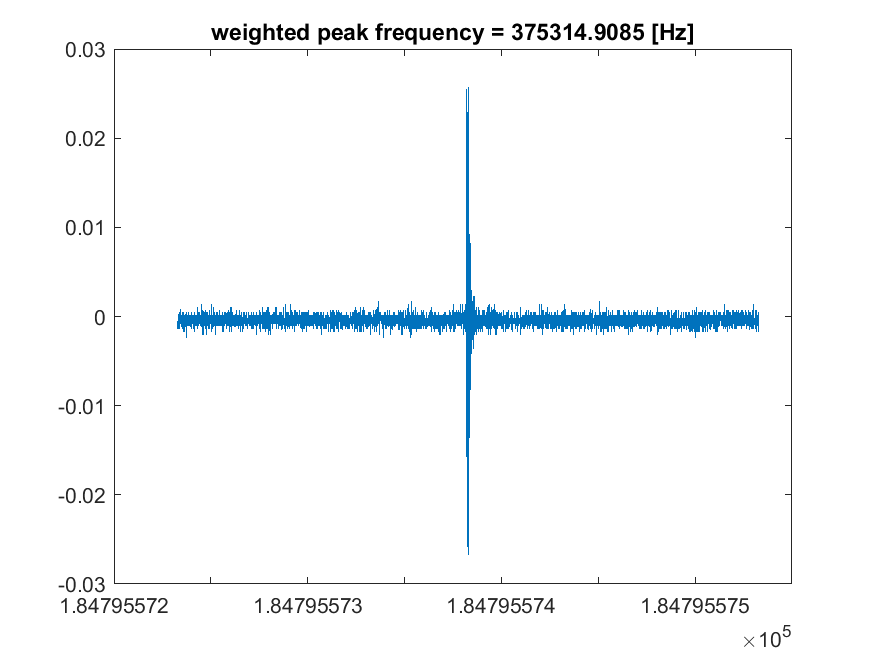
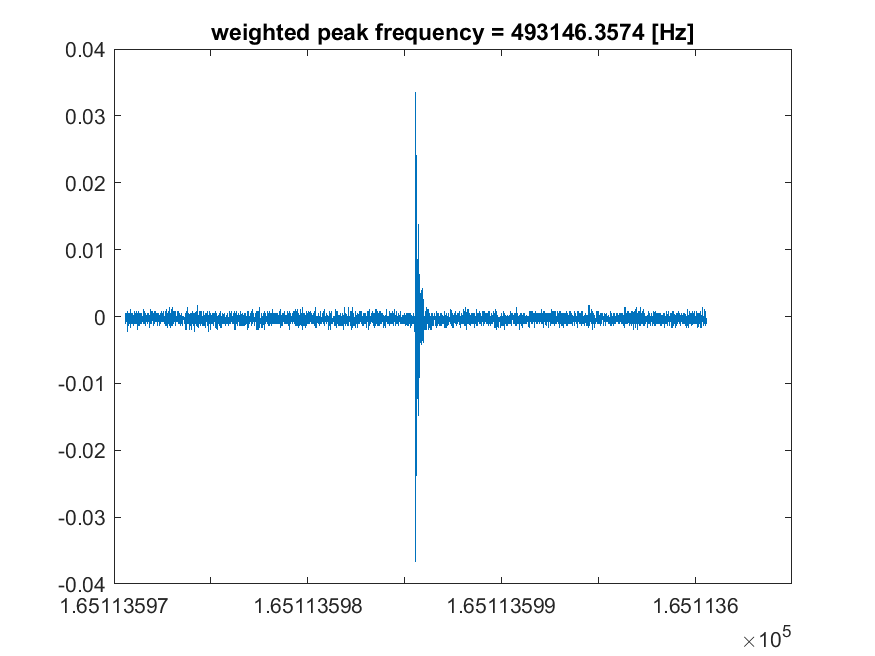
**Percentage of outliers for NUM OF FREQ PEAKS: 1/745 = 0.13423 %**

**WEIGHTED\_PEAK\_FREQUENCY:**

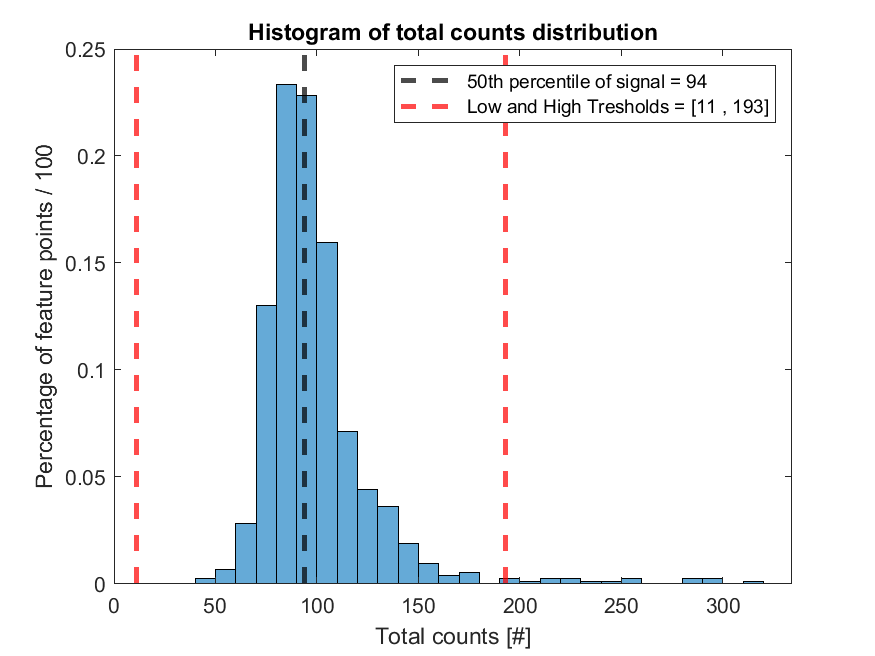
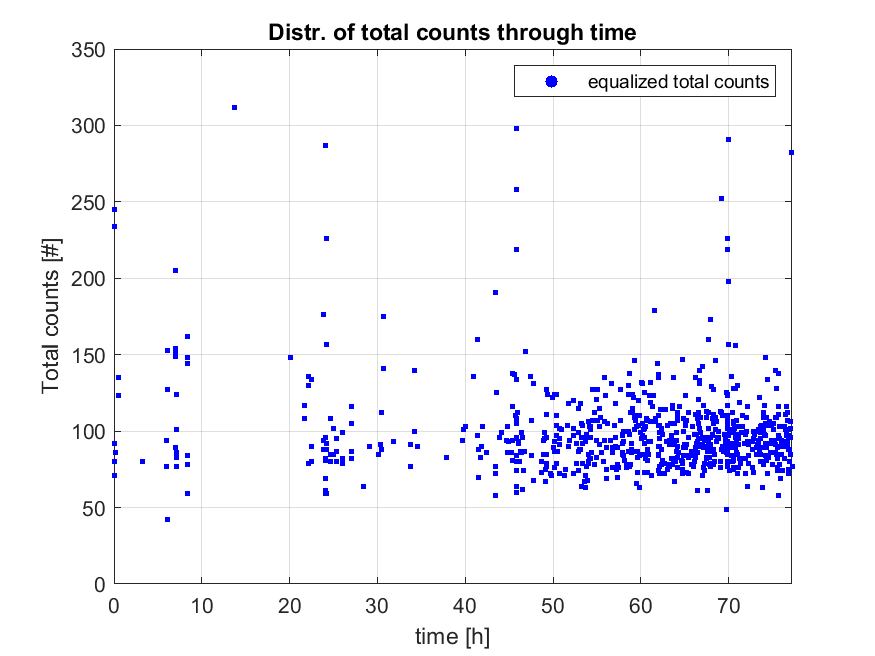


**Potencijalni outlieri iznad 374469.5708 Hz i ispod 108687.081 Hz WEIGHTED PEAK FREQUENCY:**

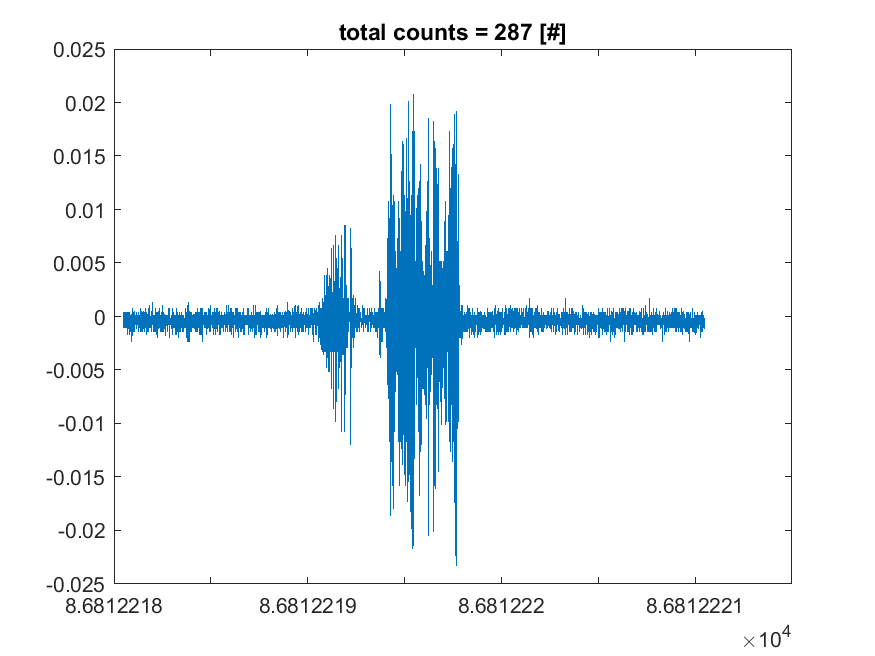
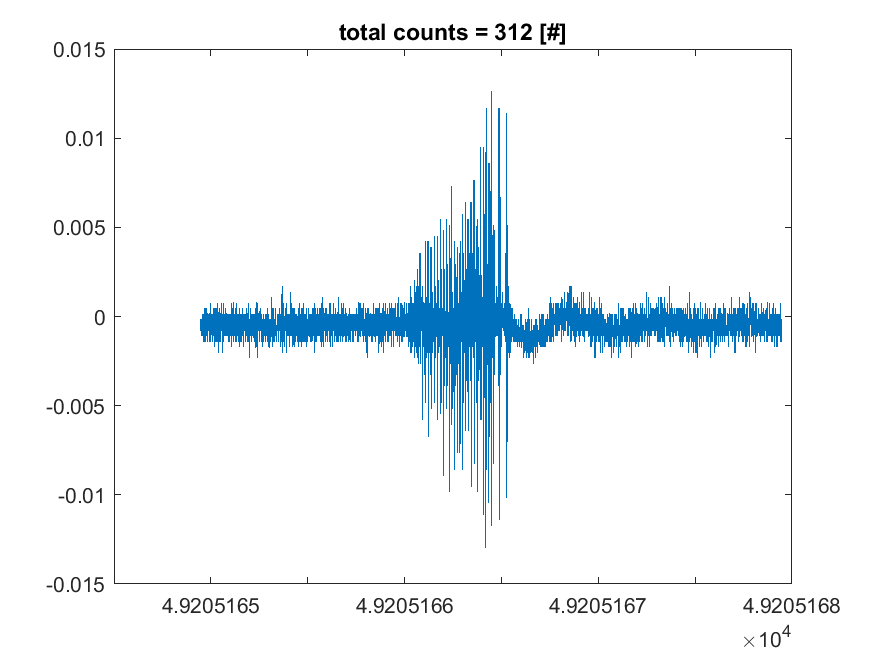
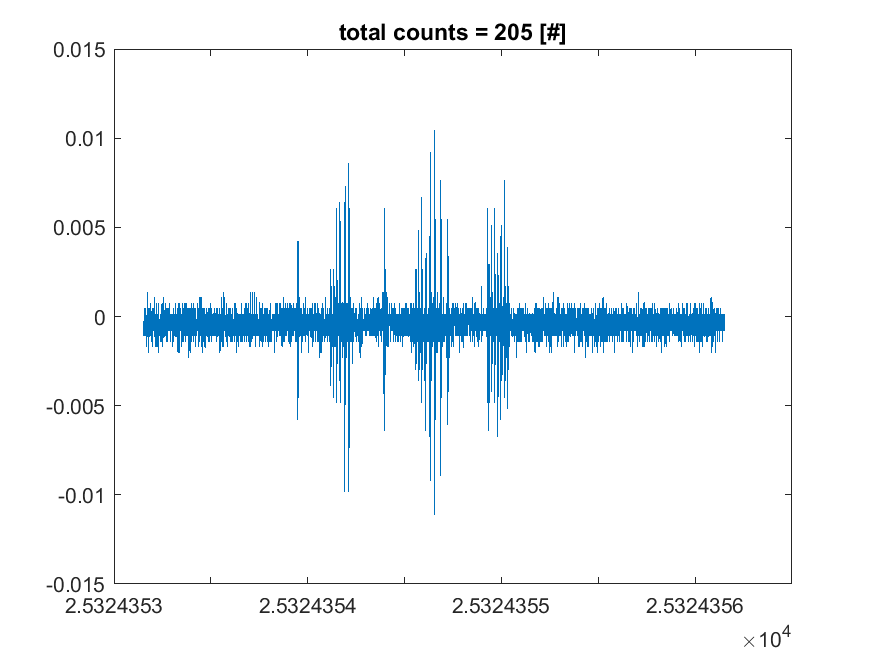
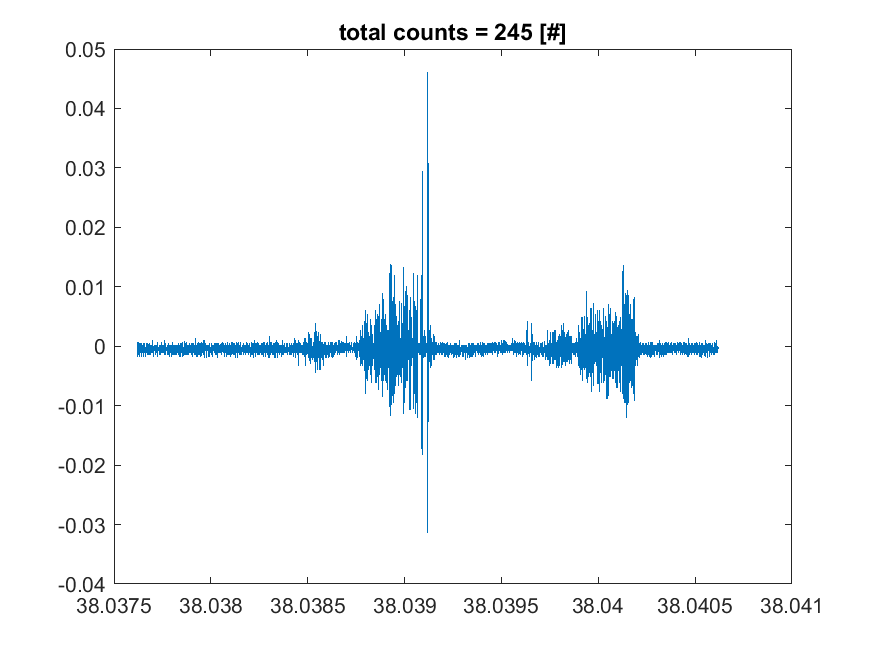
**Postoje validne emisije pa nisu pravi outlieri.**



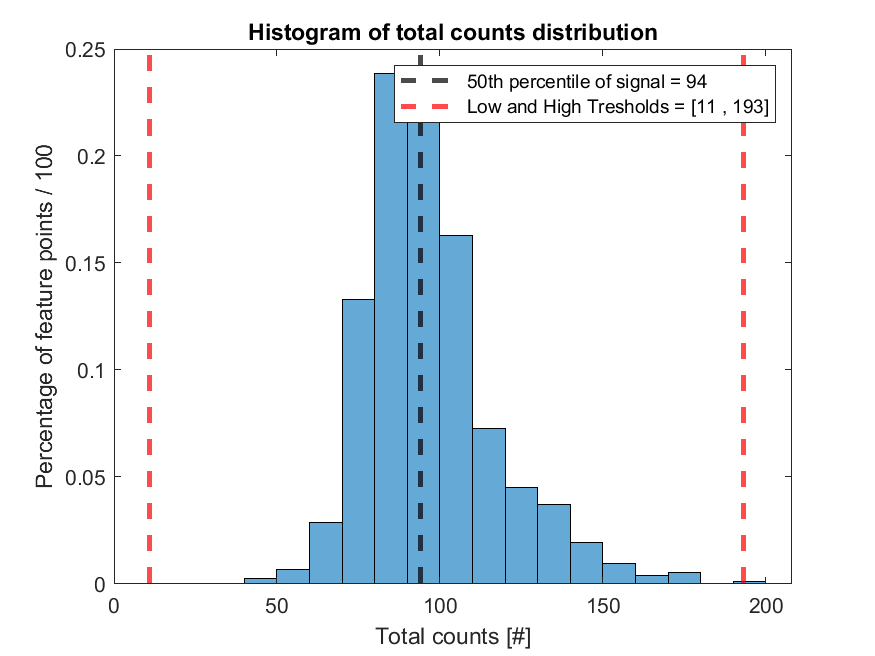
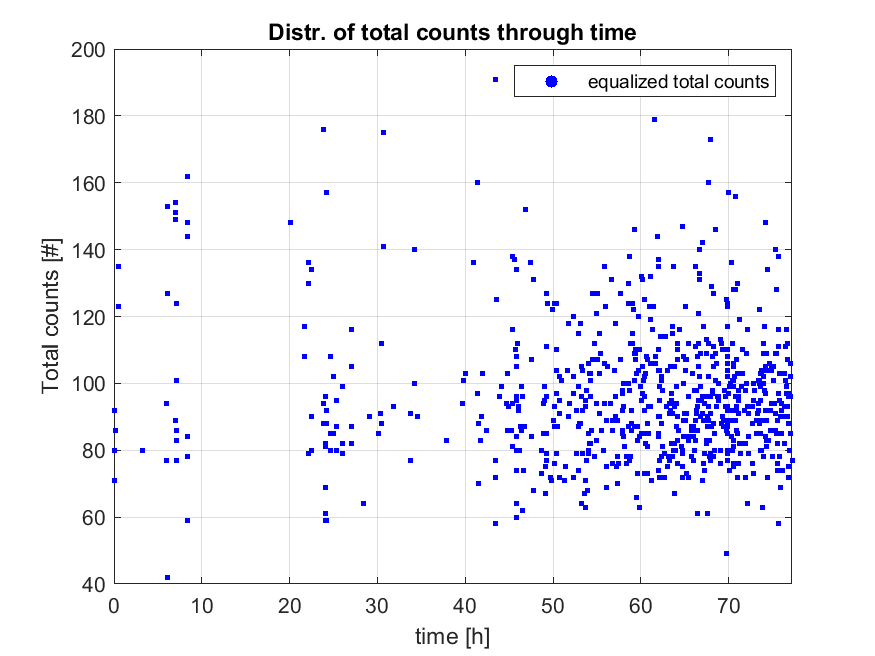
**TOTAL\_COUNTS:**



**Potencijalni outlieri iznad 193 # i ispod 11 # TOTAL COUNTS:**



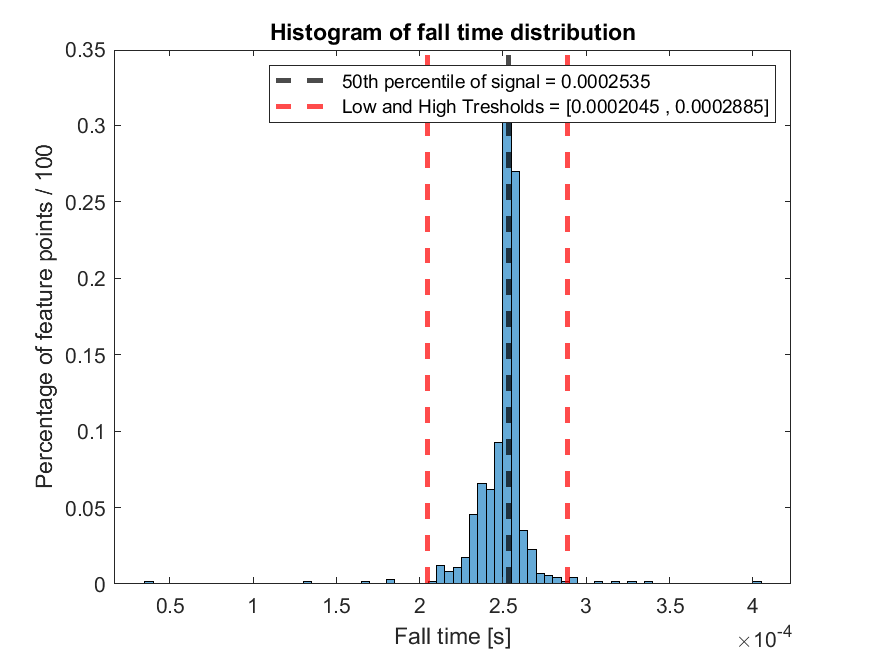
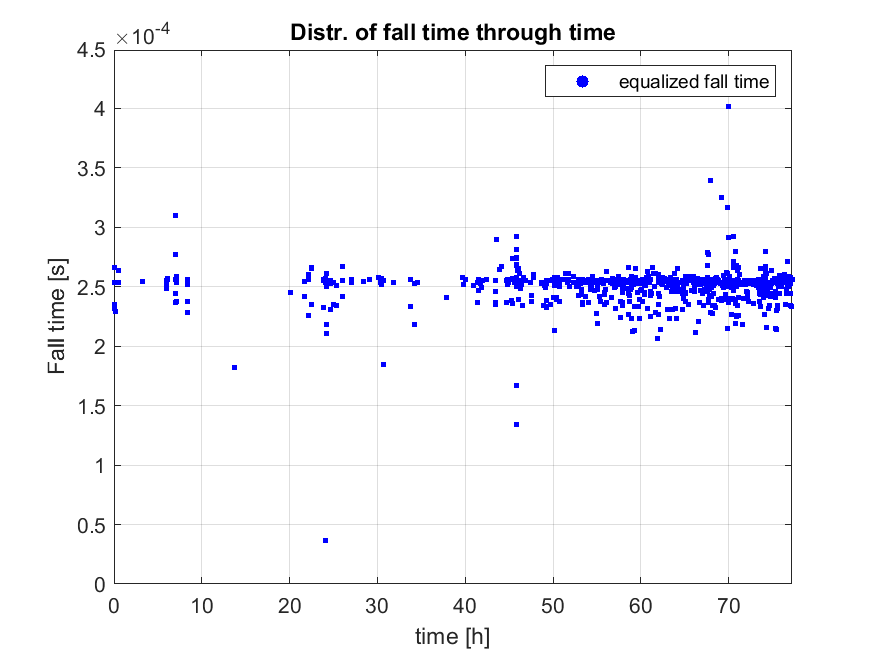
**Nakon izbacivanja outlieri iznad 193 # i ispod 11 # TOTAL COUNTS dobivamo slijedeću distribuciju:**



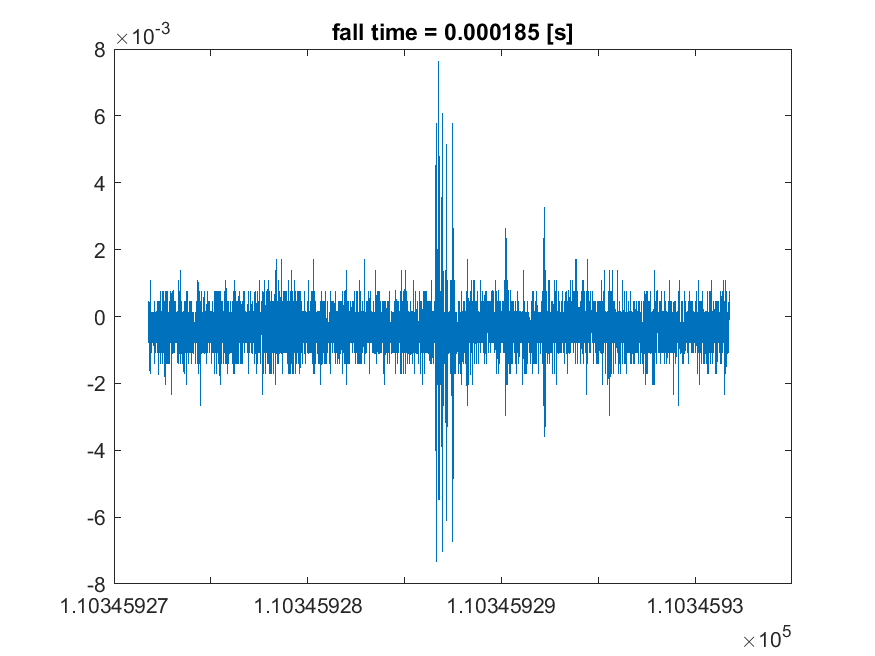
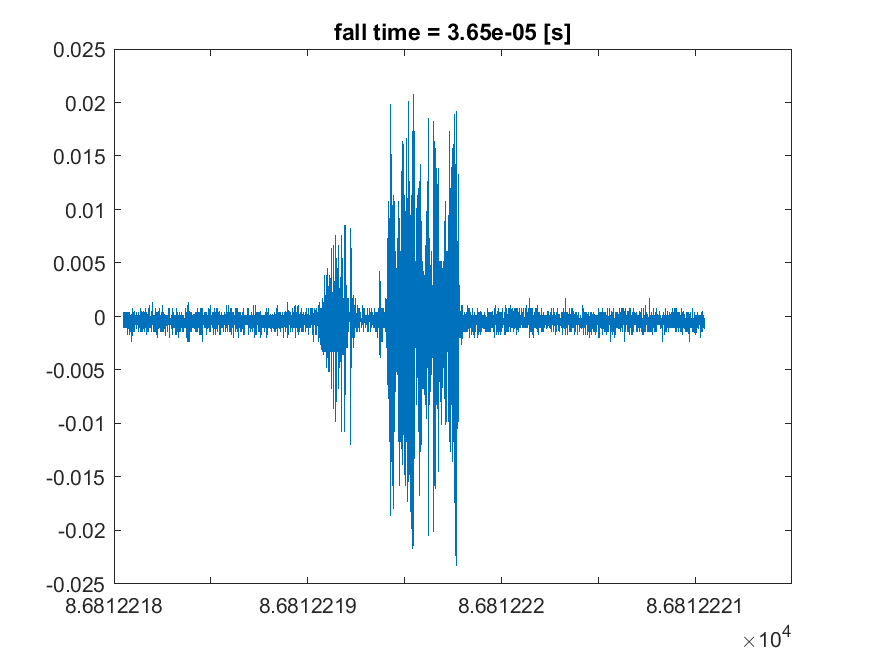
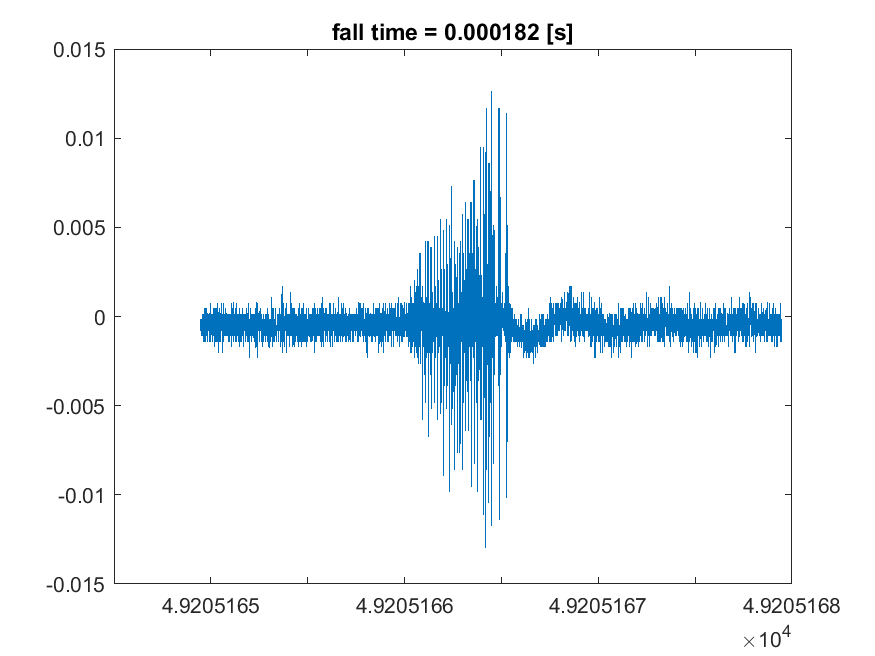
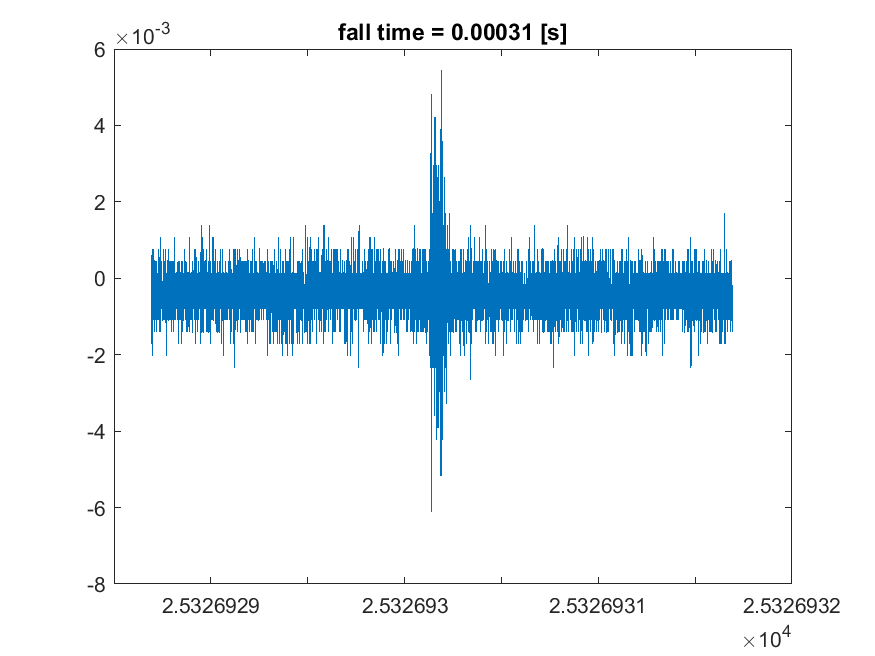
**TOTAL COUNTS outliers are emissions with value more than 193 [#]**

**Percentage of outliers for TOTAL COUNTS: 15/745 = 2.0134 %**

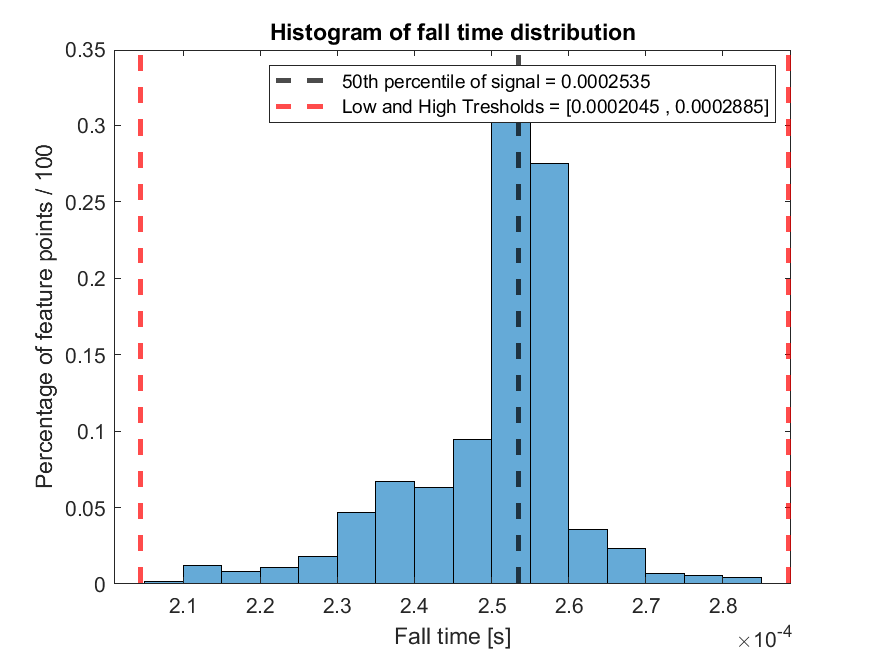
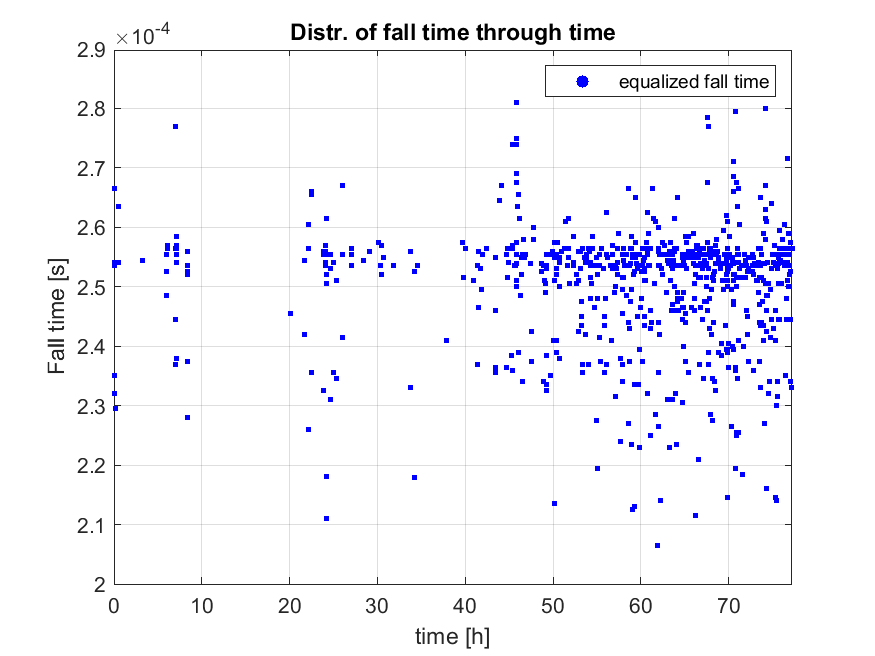
**FALL\_TIME:**



**Potencijalni outlieri iznad 0.0002885 s i ispod 0.0002045 s FALL TIME:**



**Nakon izbacivanja outlieri iznad 0.0002885 s i ispod 0.0002045 s FALL TIME dobivamo slijedeću distribuciju:**



**FALL TIME outliers are emissions with value more than 0.0002885 [s]**

**Percentage of outliers for FALL TIME: 14/745 = 1.8792 %**

**Vidimo kako za određene značajke je moguće automatskom metodom pronaći prave outliere. Za većinu značajki se ponavljaju outlieri ukupno izbacivanje za značajke koje su prepoznale prave outliere je ukupno izbacilo 40 emisija iz skupa podataka.**

**Percentage of total outliers in dataset: 40/745 = 5.3691 %**

**Normalizacija značajki:**

**Normalizacija je potrebna kako bi OPTICS algoritam svaku značajku gledao s jednakom važnosti.**

**Zbog toga trebamo skalirati sve distribucije značajki na jednaki raspon vrijednosti pomoću minimum-maksimum skaliranja. Gdje maksimalna vrijednost značajke postane jedan, dok minimalna vrijednost značajke postane nula.**

**Minimalna i maksimalna vrijednost značajki za određene značajke već su određene mjernom jedinicom značajke.**

**Npr. za parcijalne snage u frekvencijskoj domeni koje su znamo da mogu poprimiti vrijednosti između 0% i 100% tj. 0 i 1 pa ih ni ne trebamo skalirati.**

**Dok za maksimalne amplitude u vremenskoj domeni ne znamo maksimalnu granicu, osim onu određenu našim skupom podataka. Međutim u slučaju drugog skupa podataka s drukčijim granicama potrebno je odrediti fiksne granice čijim prelaskom vrijednost će se skalirati u 0 ili 1. Zbog toga potrebno je provjeriti granice u svim postojećim skupovima podataka i na temelju toga donijeti odluku.**

**Zbog ovakvog načina skaliranja izrazito je važno prvo izbaciti očite outliera iz skupa podataka, kao što je pokazano u koracima prije, kako minimalne i maksimalne vrijednosti ne bi pokvarile distribuciju na temelju koje će OPTICS algoritam clusterirati. Međutim IQR metoda prikladna je samo za manji dataset ,dok u ostalim slučajevima nije dala najbolje rezultate pa će se outlieri uzeti u obzir kod određivanja granica.**

## Granice u svim skupovima podataka:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Značajka/Skup podataka** | **Exp 6** | **Exp 7** | **Exp 8** | **Exp 9** |
| **rise time [s]** | [0.00e+00 , 4.16e-04] | [0.00e+00 , 2.38e-04] | [2.00e-06 , 1.19e-04] | [7.50e-06 , 1.88e-04] |
| **counts to [#]** | [1.00e+00 , 3.01e+02] | [1.00e+00 , 2.08e+02] | [3.00e+00 , 1.23e+02] | [2.00e+00 , 1.86e+02] |
| **counts from [#]** | [5.00e+00 , 3.13e+02] | [7.00e+00 , 2.45e+02] | [1.00e+01 , 1.40e+02] | [1.10e+01 , 1.59e+02] |
| **duration [s]** | [2.56e-04 , 5.12e-04] | [2.56e-04 , 5.10e-04] | [2.56e-04 , 3.57e-04] | [2.56e-04 , 4.70e-04] |
| **peak amplitude [V]** | [1.07e-04 , 2.42e-01] | [9.71e-05 , 3.07e+00] | [1.55e-04 , 3.97e-02] | [2.78e-04 , 1.92e-01] |
| **average frequency [Hz]** | [1.15e+05 , 9.32e+05] | [1.47e+05 , 8.54e+05] | [1.42e+05 , 5.13e+05] | [1.51e+05 , 6.56e+05] |
| **rms [V]** | [1.55e-05 , 5.55e-02] | [1.68e-05 , 8.05e-01] | [2.17e-05 , 5.74e-03] | [3.45e-05 , 4.01e-02] |
| **asl [dB]** | [-3.60e-05 , -1.26e-05] | [-3.53e-05 , -3.91e-06] | [-3.52e-05 , -2.10e-05] | [-3.32e-05 , -1.78e-05] |
| **reverbation frequency [Hz]** | [3.67e+04 , 1.04e+06] | [3.08e+04 , 7.95e+05] | [4.31e+04 , 4.62e+05] | [4.65e+04 , 5.47e+05] |
| **initiation frequency [Hz]** | [7.69e+04 , Inf] | [7.69e+04 , Inf] | [1.13e+05 , 2.40e+06] | [7.84e+04 , 2.00e+06] |
| **signal strength [Vs]** | [1.35e-09 , 8.35e-06] | [1.49e-09 , 1.92e-04] | [1.77e-09 , 7.13e-07] | [2.52e-09 , 4.52e-06] |
| **absolute energy [aJ]** | [6.44e-18 , 9.14e-11] | [7.53e-18 , 2.36e-08] | [1.21e-17 , 9.37e-13] | [3.15e-17 , 4.73e-11] |
| **PP1 [%]** | [1.55e-06 , 2.39e-02] | [7.86e-07 , 1.43e-02] | [2.84e-06 , 1.07e-02] | [5.03e-07 , 4.69e-03] |
| **PP2 [%]** | [2.94e-02 , 9.99e-01] | [1.70e-02 , 9.96e-01] | [6.95e-02 , 9.98e-01] | [1.38e-02 , 9.97e-01] |
| **PP3 [%]** | [3.87e-04 , 9.48e-01] | [3.22e-03 , 9.14e-01] | [1.89e-03 , 8.98e-01] | [2.36e-03 , 9.26e-01] |
| **PP4 [%]** | [4.74e-07 , 9.17e-01] | [1.43e-04 , 9.57e-01] | [4.56e-05 , 7.34e-01] | [2.11e-05 , 7.54e-01] |
| **centroid frequency [Hz]** | [1.58e+05 , 5.31e+05] | [1.89e+05 , 5.35e+05] | [1.88e+05 , 4.82e+05] | [1.78e+05 , 4.98e+05] |
| **peak frequency [Hz]** | [1.17e+05 , 8.01e+05] | [1.17e+05 , 8.03e+05] | [1.19e+05 , 6.19e+05] | [1.17e+05 , 5.72e+05] |
| **amplitude of peak frequency [V]** | [2.28e-06 , 1.70e-02] | [2.53e-06 , 3.30e-01] | [3.00e-06 , 1.89e-03] | [5.22e-06 , 1.21e-02] |
| **num of freq peaks [#]** | [0.00e+00 , 1.30e+01] | [0.00e+00 , 1.10e+01] | [1.00e+00 , 1.30e+01] | [0.00e+00 , 1.00e+01] |
| **weighted peak frequency [Hz]** | [1.44e+05 , 6.52e+05] | [1.63e+05 , 6.46e+05] | [1.64e+05 , 5.37e+05] | [1.70e+05 , 5.14e+05] |
| **total counts [#]** | [3.00e+01 , 3.41e+02] | [3.90e+01 , 2.84e+02] | [3.70e+01 , 1.83e+02] | [4.00e+01 , 2.57e+02] |
| **fall time [s]** | [1.75e-05 , 4.57e-04] | [7.50e-05 , 4.65e-04] | [2.11e-04 , 3.18e-04] | [1.77e-04 , 3.48e-04] |

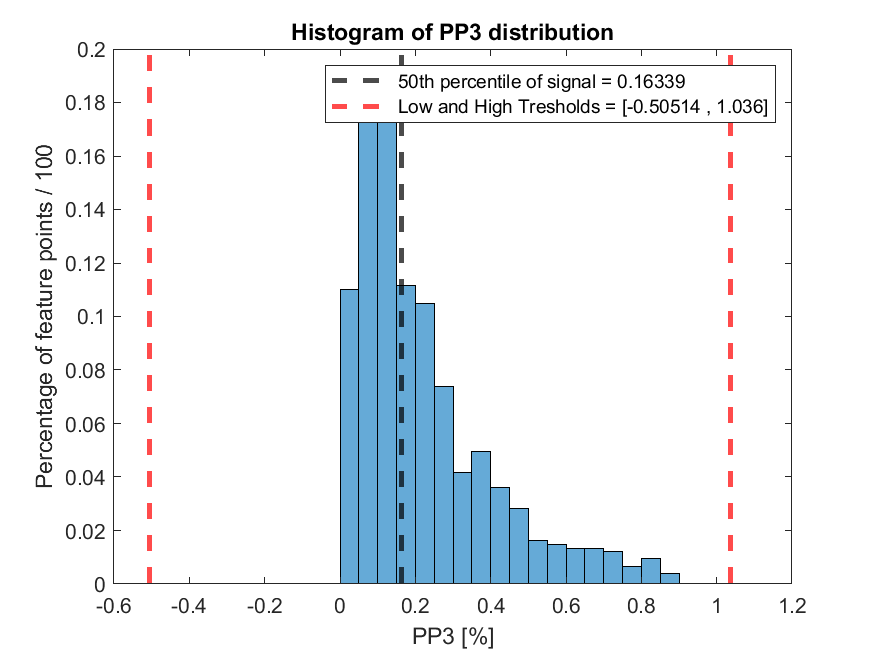
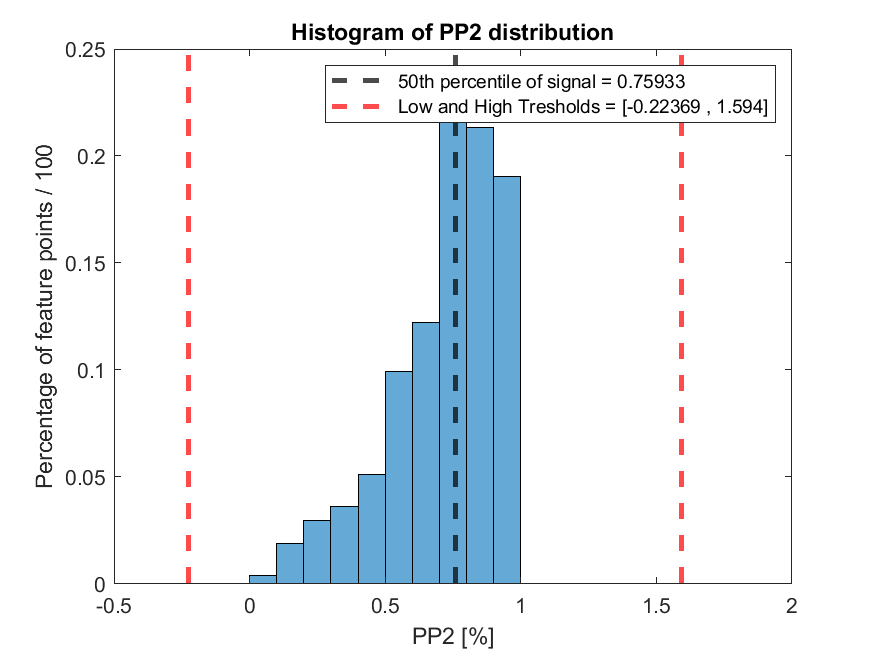
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Značajka/Skup podataka** | **Exp 10 part 1** | **Exp 10 part 2** | **Exp 12** | **Exp 13** |
| **rise time [s]** | [4.50e-06 , 7.30e-05] | [6.00e-06 , 2.01e-04] | [5.00e-07 , 3.66e-04] | [2.50e-06 , 3.77e-04] |
| **counts to [#]** | [3.00e+00 , 9.40e+01] | [2.00e+00 , 1.91e+02] | [2.00e+00 , 2.69e+02] | [4.00e+00 , 2.78e+02] |
| **counts from [#]** | [9.00e+00 , 1.15e+02] | [8.00e+00 , 2.30e+02] | [6.00e+00 , 2.22e+02] | [9.00e+00 , 1.95e+02] |
| **duration [s]** | [2.59e-04 , 3.11e-04] | [2.64e-04 , 4.90e-04] | [2.56e-04 , 5.12e-04] | [2.56e-04 , 5.12e-04] |
| **peak amplitude [V]** | [2.17e-04 , 2.17e-02] | [2.91e-04 , 1.80e-01] | [2.56e-04 , 3.05e-01] | [2.40e-04 , 1.98e-01] |
| **average frequency [Hz]** | [2.12e+05 , 5.25e+05] | [1.88e+05 , 7.75e+05] | [1.29e+05 , 9.03e+05] | [1.60e+05 , 7.06e+05] |
| **rms [V]** | [2.95e-05 , 4.65e-03] | [5.17e-05 , 3.79e-02] | [3.56e-05 , 4.74e-02] | [2.91e-05 , 3.00e-02] |
| **asl [dB]** | [-3.54e-05 , -2.14e-05] | [-3.27e-05 , -1.75e-05] | [-3.40e-05 , -1.53e-05] | [-3.42e-05 , -1.74e-05] |
| **reverbation frequency [Hz]** | [4.12e+04 , 4.53e+05] | [3.69e+04 , 8.02e+05] | [2.68e+04 , 9.27e+05] | [6.13e+04 , 6.27e+05] |
| **initiation frequency [Hz]** | [1.15e+05 , 2.00e+06] | [7.55e+04 , 1.80e+06] | [7.69e+04 , 4.00e+06] | [1.60e+05 , 2.40e+06] |
| **signal strength [Vs]** | [2.66e-09 , 6.07e-07] | [4.22e-09 , 4.61e-06] | [2.63e-09 , 8.53e-06] | [2.00e-09 , 3.28e-06] |
| **absolute energy [aJ]** | [2.41e-17 , 6.26e-13] | [7.39e-17 , 4.24e-11] | [3.30e-17 , 7.61e-11] | [2.23e-17 , 2.57e-11] |
| **PP1 [%]** | [8.97e-07 , 8.49e-03] | [3.97e-07 , 4.54e-03] | [1.52e-07 , 6.97e-03] | [1.14e-06 , 6.32e-03] |
| **PP2 [%]** | [5.44e-02 , 9.97e-01] | [3.93e-02 , 9.98e-01] | [1.30e-02 , 9.99e-01] | [5.91e-02 , 9.96e-01] |
| **PP3 [%]** | [2.29e-03 , 9.15e-01] | [1.57e-03 , 8.70e-01] | [6.26e-04 , 9.33e-01] | [3.68e-03 , 8.86e-01] |
| **PP4 [%]** | [5.82e-05 , 8.88e-01] | [6.10e-05 , 8.34e-01] | [8.17e-06 , 9.11e-01] | [5.23e-05 , 6.11e-01] |
| **centroid frequency [Hz]** | [2.06e+05 , 4.56e+05] | [1.96e+05 , 5.08e+05] | [1.69e+05 , 5.28e+05] | [2.00e+05 , 4.74e+05] |
| **peak frequency [Hz]** | [1.17e+05 , 5.70e+05] | [1.17e+05 , 6.29e+05] | [1.15e+05 , 8.03e+05] | [1.17e+05 , 7.99e+05] |
| **amplitude of peak frequency [V]** | [6.10e-06 , 1.78e-03] | [7.91e-06 , 1.28e-02] | [6.05e-06 , 1.66e-02] | [4.32e-06 , 8.78e-03] |
| **num of freq peaks [#]** | [1.00e+00 , 9.00e+00] | [0.00e+00 , 1.00e+01] | [0.00e+00 , 1.70e+01] | [0.00e+00 , 1.10e+01] |
| **weighted peak frequency [Hz]** | [1.85e+05 , 5.09e+05] | [1.78e+05 , 5.56e+05] | [1.49e+05 , 6.32e+05] | [1.81e+05 , 6.14e+05] |
| **total counts [#]** | [5.80e+01 , 1.50e+02] | [5.30e+01 , 2.62e+02] | [3.30e+01 , 3.12e+02] | [4.20e+01 , 3.12e+02] |
| **fall time [s]** | [2.10e-04 , 2.77e-04] | [1.94e-04 , 3.75e-04] | [1.31e-04 , 4.09e-04] | [3.65e-05 , 4.02e-04] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Značajka/Skup podataka** | **Exp 14 part 1** | **Exp 14 part 2** | **Ekstremni [min,max]** | **Srednji [min,max]** |
| **rise time [s]** | [2.00e-06 , 3.46e-04] | [2.00e-06 , 3.24e-04] | **[-3.60e-05 , 4.16e-04]** | **[3.30e+04 , 2.41e-04]** |
| **counts to [#]** | [3.00e+00 , 2.45e+02] | [3.00e+00 , 2.70e+02] | **[0.00e+00 , 3.01e+02]** | **[1.38e-02 , 1.97e+02]** |
| **counts from [#]** | [1.20e+01 , 1.92e+02] | [5.00e+00 , 2.14e+02] | **[0.00e+00 , 3.13e+02]** | **[4.72e-02 , 1.84e+02]** |
| **duration [s]** | [2.56e-04 , 4.89e-04] | [2.56e-04 , 5.11e-04] | **[0.00e+00 , 5.12e-04]** | **[1.48e-06 , 4.25e-04]** |
| **peak amplitude [V]** | [2.11e-04 , 5.06e-02] | [2.97e-04 , 1.03e-01] | **[0.00e+00 , 3.07e+00]** | **[1.24e-06 , 4.00e-01]** |
| **average frequency [Hz]** | [1.48e+05 , 9.56e+05] | [1.40e+05 , 8.78e+05] | **[0.00e+00 , 9.56e+05]** | **[8.82e+02 , 7.00e+05]** |
| **rms [V]** | [2.07e-05 , 5.80e-03] | [4.81e-05 , 1.15e-02] | **[0.00e+00 , 8.05e-01]** | **[1.74e-07 , 9.49e-02]** |
| **asl [dB]** | [-3.41e-05 , -1.89e-05] | [-3.45e-05 , -1.86e-05] | **[-3.60e-05 , 0.00e+00]** | **[-1.98e-07 , -1.50e-05]** |
| **reverbation frequency [Hz]** | [5.03e+04 , 9.18e+05] | [2.38e+04 , 9.95e+05] | **[0.00e+00 , 1.04e+06]** | **[2.29e+02 , 6.88e+05]** |
| **initiation frequency [Hz]** | [1.15e+05 , 2.18e+06] | [1.15e+05 , 2.50e+06] | **[0.00e+00 , Inf]** | **[5.78e+02 , Inf]** |
| **signal strength [Vs]** | [1.67e-09 , 4.70e-07] | [3.86e-09 , 1.00e-06] | **[0.00e+00 , 1.92e-04]** | **[1.39e-11 , 2.03e-05]** |
| **absolute energy [aJ]** | [1.13e-17 , 9.85e-13] | [6.00e-17 , 3.76e-12] | **[0.00e+00 , 2.36e-08]** | **[1.62e-19 , 2.17e-09]** |
| **PP1 [%]** | [1.21e-06 , 9.34e-03] | [6.11e-07 , 8.50e-03] | **[0.00e+00 , 2.39e-02]** | **[5.81e-09 , 8.89e-03]** |
| **PP2 [%]** | [2.83e-02 , 9.90e-01] | [4.60e-02 , 9.96e-01] | **[0.00e+00 , 9.99e-01]** | **[2.13e-04 , 9.06e-01]** |
| **PP3 [%]** | [1.00e-02 , 9.58e-01] | [3.77e-03 , 9.09e-01] | **[0.00e+00 , 9.58e-01]** | **[1.72e-05 , 8.32e-01]** |
| **PP4 [%]** | [3.67e-04 , 8.33e-01] | [5.59e-06 , 9.07e-01] | **[0.00e+00 , 9.57e-01]** | **[4.39e-07 , 7.59e-01]** |
| **centroid frequency [Hz]** | [2.15e+05 , 5.04e+05] | [1.81e+05 , 5.19e+05] | **[0.00e+00 , 5.35e+05]** | **[1.08e+03 , 4.58e+05]** |
| **peak frequency [Hz]** | [1.19e+05 , 7.97e+05] | [1.17e+05 , 7.99e+05] | **[0.00e+00 , 8.03e+05]** | **[6.76e+02 , 6.54e+05]** |
| **amplitude of peak frequency [V]** | [2.83e-06 , 1.07e-03] | [7.35e-06 , 2.43e-03] | **[0.00e+00 , 3.30e-01]** | **[2.74e-08 , 3.68e-02]** |
| **num of freq peaks [#]** | [0.00e+00 , 1.30e+01] | [0.00e+00 , 1.20e+01] | **[0.00e+00 , 1.70e+01]** | **[1.15e-03 , 1.08e+01]** |
| **weighted peak frequency [Hz]** | [1.77e+05 , 6.06e+05] | [1.58e+05 , 6.29e+05] | **[0.00e+00 , 6.52e+05]** | **[9.60e+02 , 5.36e+05]** |
| **total counts [#]** | [4.20e+01 , 3.13e+02] | [3.70e+01 , 3.25e+02] | **[0.00e+00 , 3.41e+02]** | **[2.37e-01 , 2.49e+02]** |
| **fall time [s]** | [3.00e-05 , 4.29e-04] | [5.95e-05 , 3.37e-04] | **[0.00e+00 , 4.65e-04]** | **[6.57e-07 , 3.47e-04]** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Značajka/Skup podataka** | **Ekstremni [min,max]** | **Srednji [min,max]** | **Konačan odabir granica** |
| **rise time [s]** | **[-3.60e-05 , 4.16e-04]** | **[3.30e+04 , 2.41e-04]** | **[15 , 500 us]** |
| **counts to [#]** | **[0.00e+00 , 3.01e+02]** | **[1.38e-02 , 1.97e+02]** | **[0 , 350]** |
| **counts from [#]** | **[0.00e+00 , 3.13e+02]** | **[4.72e-02 , 1.84e+02]** | **[0 , 350]** |
| **duration [s]** | **[0.00e+00 , 5.12e-04]** | **[1.48e-06 , 4.25e-04]** | **[15 , 500 us]** |
| **peak amplitude [V]** | **[0.00e+00 , 3.07e+00]** | **[1.24e-06 , 4.00e-01]** | **[0, 3 V]** |
| **average frequency [Hz]** | **[0.00e+00 , 9.56e+05]** | **[8.82e+02 , 7.00e+05]** | **[0 , 1 MHz]** |
| **rms [V]** | **[0.00e+00 , 8.05e-01]** | **[1.74e-07 , 9.49e-02]** | **[2.07e-05 , 5.80e-03]** |
| **asl [dB]** | **[-3.60e-05 , 0.00e+00]** | **[-1.98e-07 , -1.50e-05]** | **[-3.41e-05 , -1.89e-05]** |
| **reverbation frequency [Hz]** | **[0.00e+00 , 1.04e+06]** | **[2.29e+02 , 6.88e+05]** | **[0 , 1 MHz]** |
| **initiation frequency [Hz]** | **[0.00e+00 , Inf]** | **[5.78e+02 , Inf]** | **[0 , 2.5 MHz]** |
| **signal strength [Vs]** | **[0.00e+00 , 1.92e-04]** | **[1.39e-11 , 2.03e-05]** | **[0 , 0.2 mVs]** |
| **absolute energy [aJ]** | **[0.00e+00 , 2.36e-08]** | **[1.62e-19 , 2.17e-09]** | **[0 , 0.25 pJ]** |
| **PP1 [%]** | **[0.00e+00 , 2.39e-02]** | **[5.81e-09 , 8.89e-03]** | **[0, 1]** |
| **PP2 [%]** | **[0.00e+00 , 9.99e-01]** | **[2.13e-04 , 9.06e-01]** | **[0, 1]** |
| **PP3 [%]** | **[0.00e+00 , 9.58e-01]** | **[1.72e-05 , 8.32e-01]** | **[0, 1]** |
| **PP4 [%]** | **[0.00e+00 , 9.57e-01]** | **[4.39e-07 , 7.59e-01]** | **[0, 1]** |
| **centroid frequency [Hz]** | **[0.00e+00 , 5.35e+05]** | **[1.08e+03 , 4.58e+05]** | **[0, 1MHz]** |
| **peak frequency [Hz]** | **[0.00e+00 , 8.03e+05]** | **[6.76e+02 , 6.54e+05]** | **[0, 1MHz]** |
| **amplitude of peak frequency [V]** | **[0.00e+00 , 3.30e-01]** | **[2.74e-08 , 3.68e-02]** | **[0, 0.3 V]** |
| **num of freq peaks [#]** | **[0.00e+00 , 1.70e+01]** | **[1.15e-03 , 1.08e+01]** | **[0, 10]** |
| **weighted peak frequency [Hz]** | **[0.00e+00 , 6.52e+05]** | **[9.60e+02 , 5.36e+05]** | **[0 , 1 MHz]** |
| **total counts [#]** | **[0.00e+00 , 3.41e+02]** | **[2.37e-01 , 2.49e+02]** | **[0 , 350]** |
| **fall time [s]** | **[0.00e+00 , 4.65e-04]** | **[6.57e-07 , 3.47e-04]** | **[15 , 500 us]** |

**Zašto ne koristiti neku drugu vrstu normalizacije kao z-score [**[Z-scores (daylight.com)](https://www.daylight.com/meetings/emug97/Bradshaw/Significant_Similarity/Z-scores.html)**] gdje s vrijednosti značajke oduzmi sa srednjom vrijednosti i podijele sa standardnom devijacijom svake značajke. U tome slučaju normalizirani podatci imaju srednju vrijednost 0 te standardnu varijaciju 1. Time će se sve značajke grupirati oko nule kao što vidimo na slikama ispod da PP2 i PP3 koji bi trebali imati suprotne distribucije sada imaju većinu vrijednosti oko nule. Vidimo da kod clusteriranja će teško biti odvojiti na temelju gustoće ovako normalizirane značajke.**

**Prije z score normalizacije:**



**Poslije z score normalizacije:**

