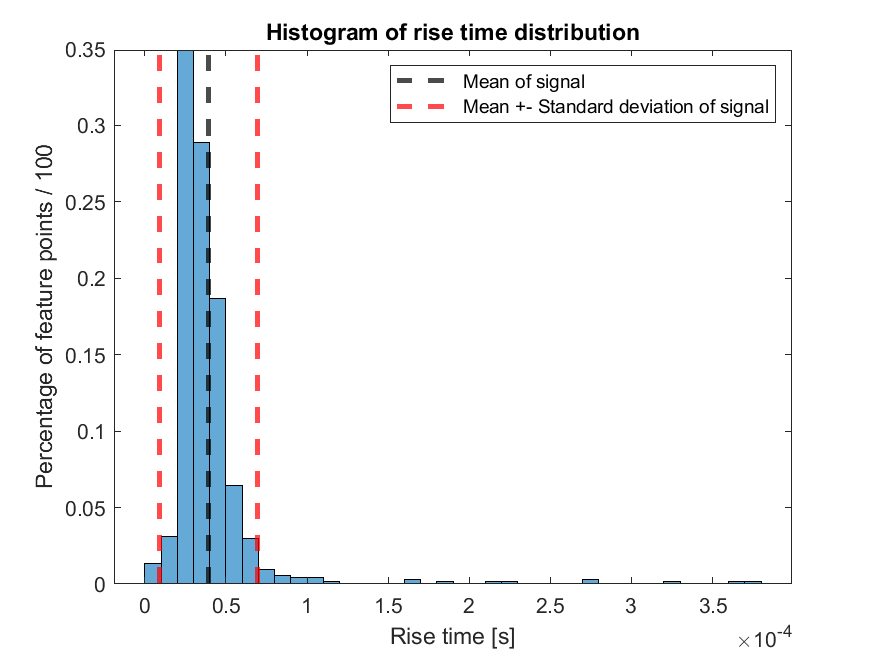
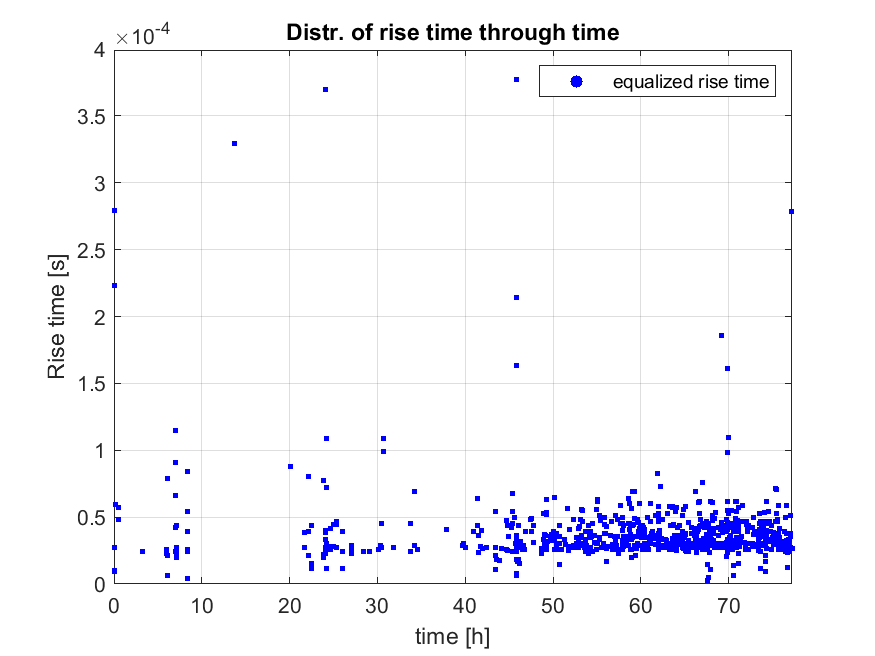
# Feature distribution and analysis for small\_dataset

**RISE\_TIME:**

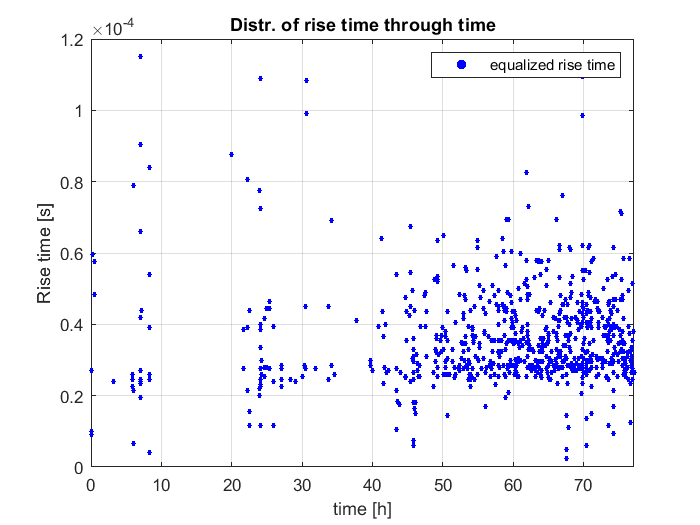


**Potencijalni outlieri iznad 0.12 ms:**

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



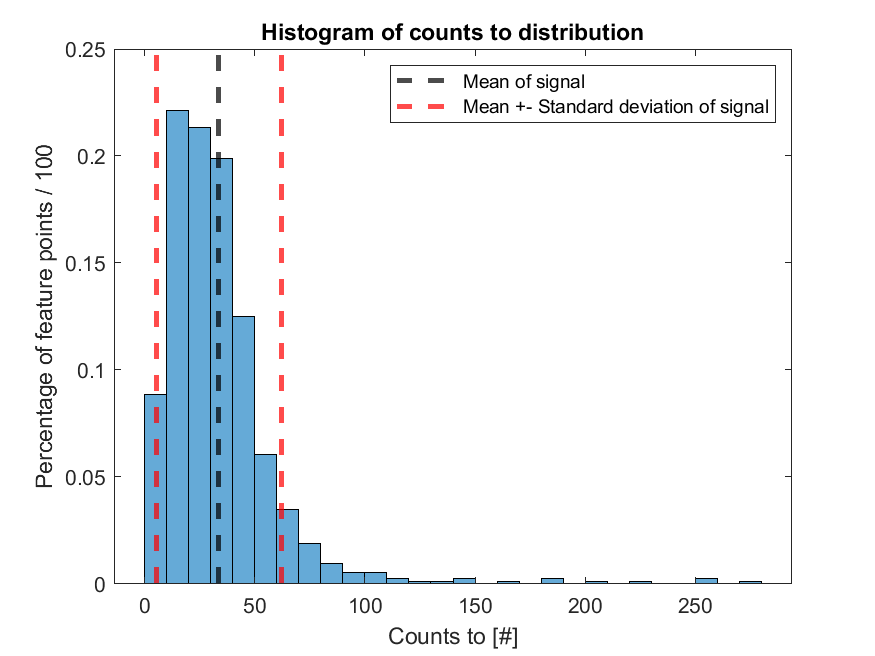
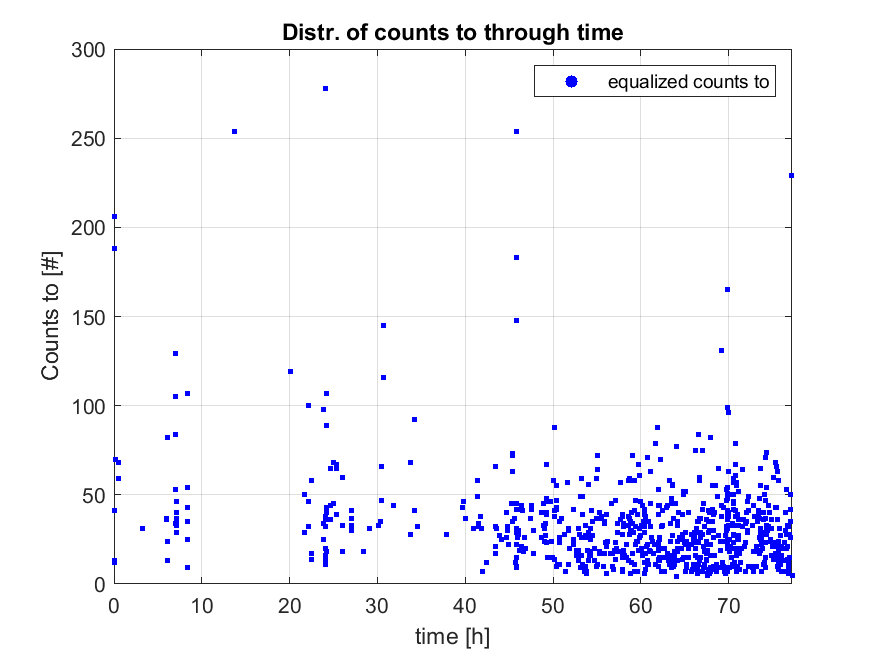
**Nakon izbacivanja outliera iznad 0.12 ms RISE TIME dobivamo slijedeću distribuciju:**



**RISE TIME are emissions with value more than 0.00012 [s]**

**Percentage of oultiers for RISE TIME: 10/745 = 1.3423 %**

**COUNTS\_TO:**



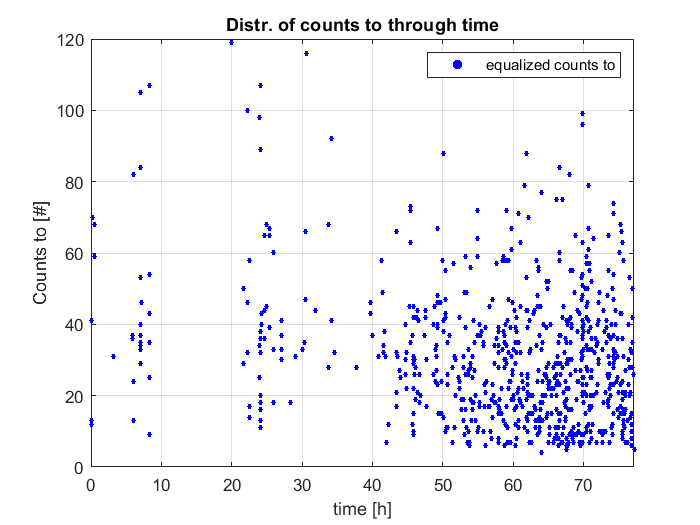
**Potencijalni outlieri iznad 120 vrhova: - isti razlog i emisije kao kod RISE\_TIME-a**



**Nove emisije zbog prevelikog šuma iznad 2 mV:**



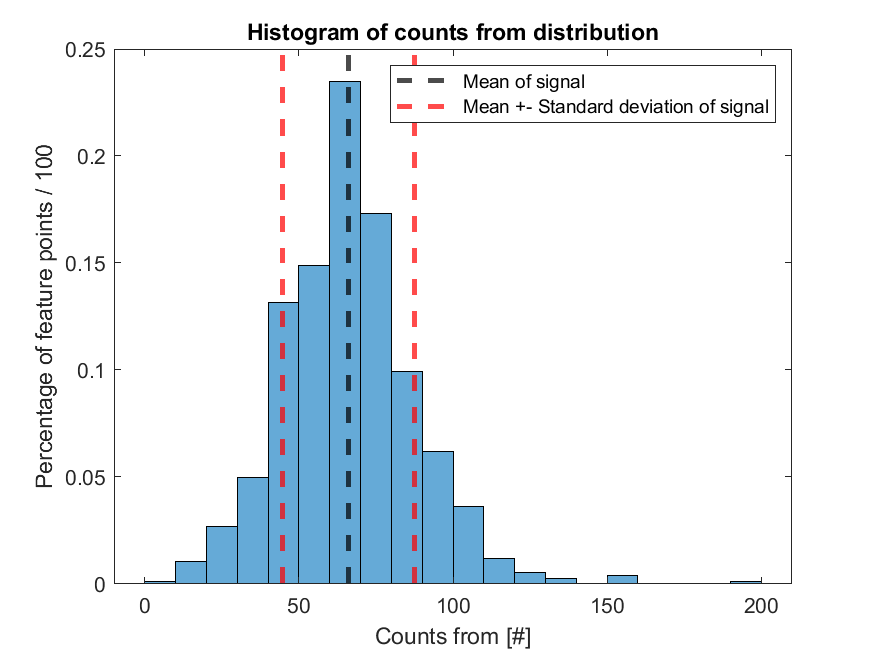
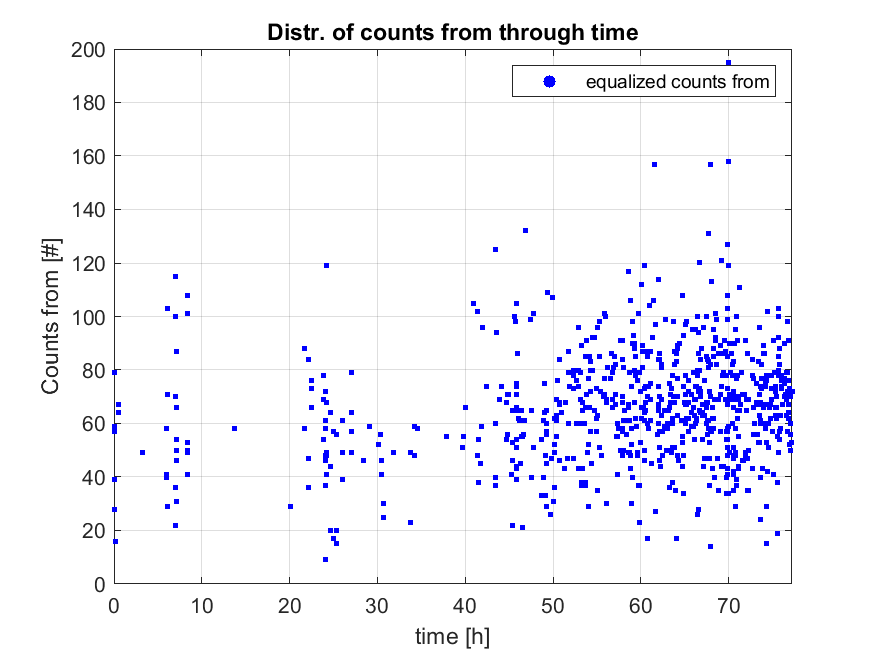
**Nakon izbacivanja outliera iznad 120 COUNTS TO dobivamo slijedeću distribuciju:**



**COUNTS TO are emissions with value more than 120 [#]**

**Percentage of oultiers for COUNTS TO: 12/745 = 1.6107 %**

**COUNTS\_FROM:**



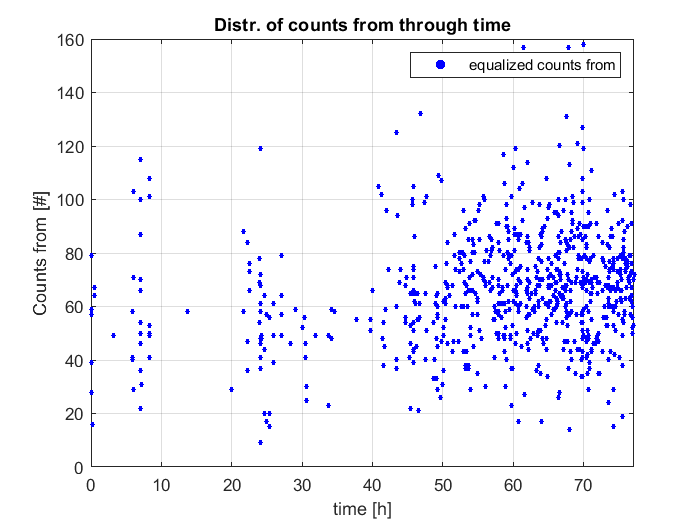
**Potencijalni outlieri iznad 140 vrhova:**



**Međutim ispod 160 vrhova imamo kvalitetnu emisiju pa postavljamo prag za izbacivanje outliersa na 160.**



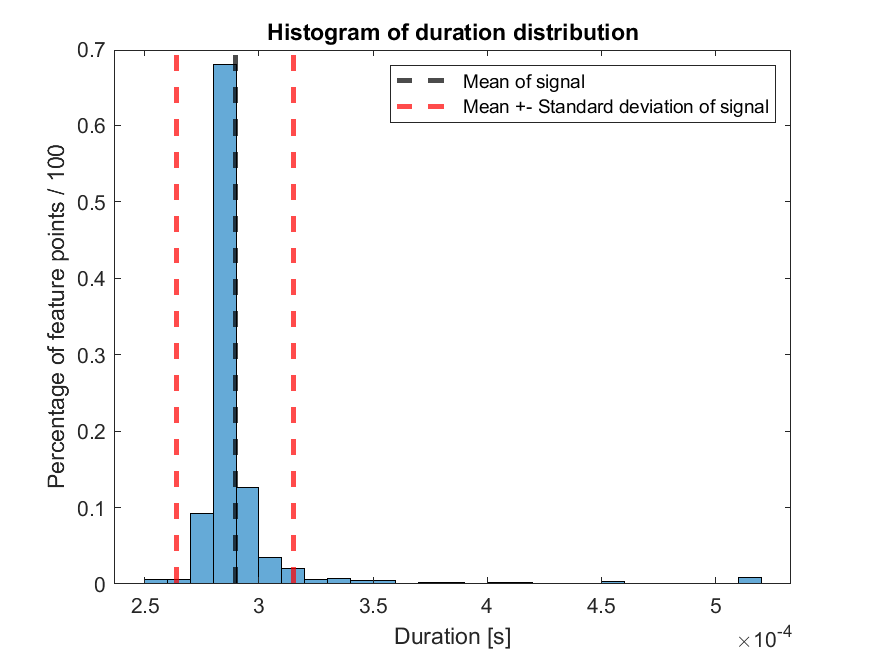
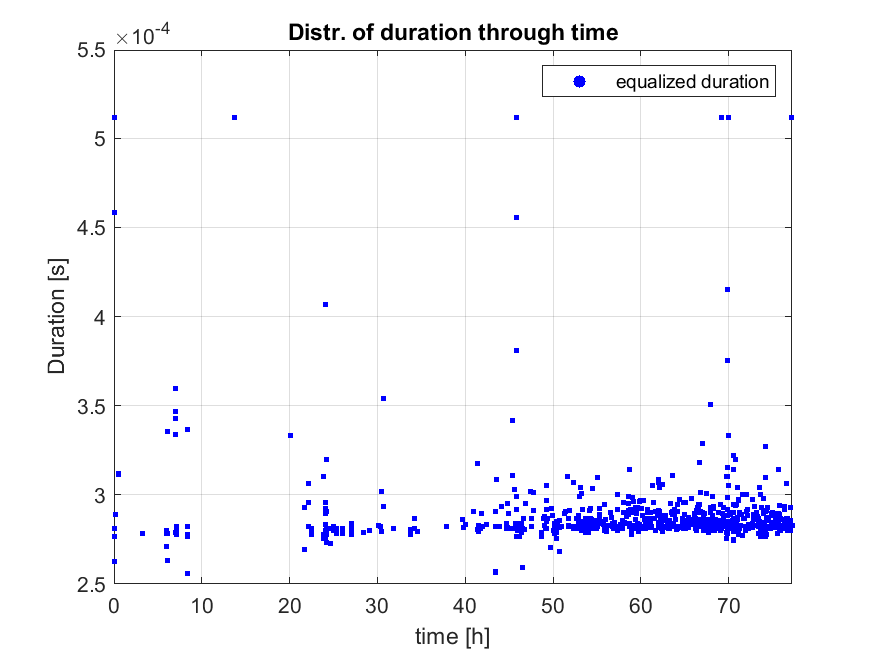
**Nakon izbacivanja outliersa distribucija:**



**COUNTS FROM are emissions with value more than 160 [#]**

**Percentage of oultiers for COUNTS FROM: 1/745 = 0.13423 %**

**DURATION:**



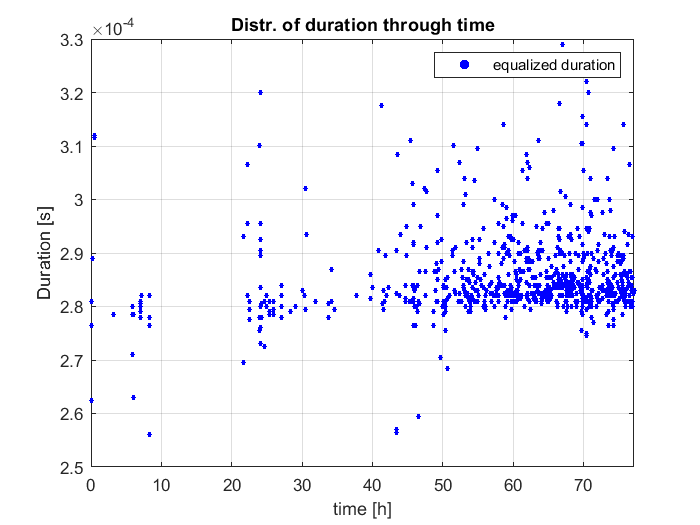
**Potencijalni outliersi iznad 0.33 ms:**



**Potencijalno dobre emisije ispod 0.4 ms , no imaju visoki šum pa ćemo ih izbaciti:**



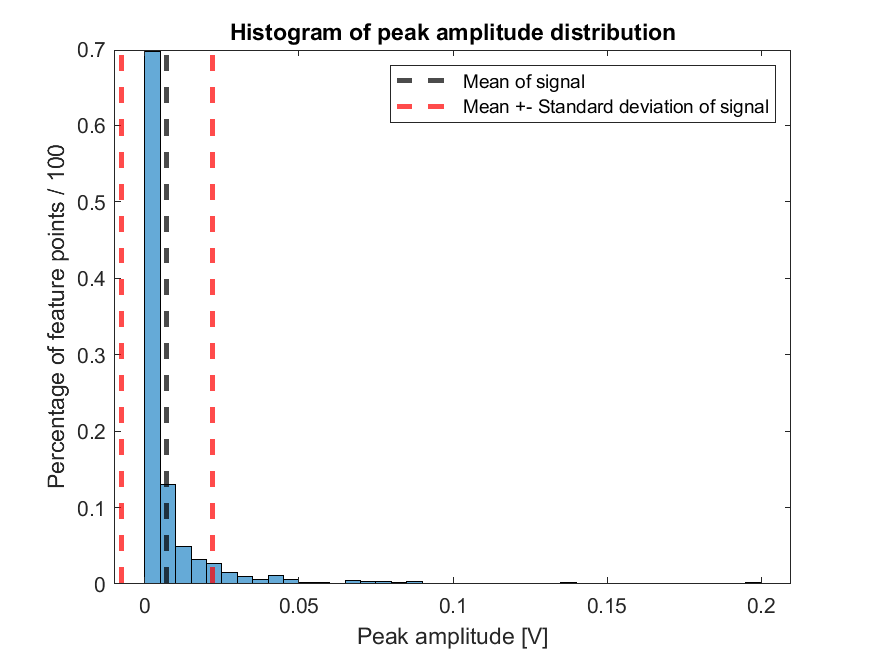
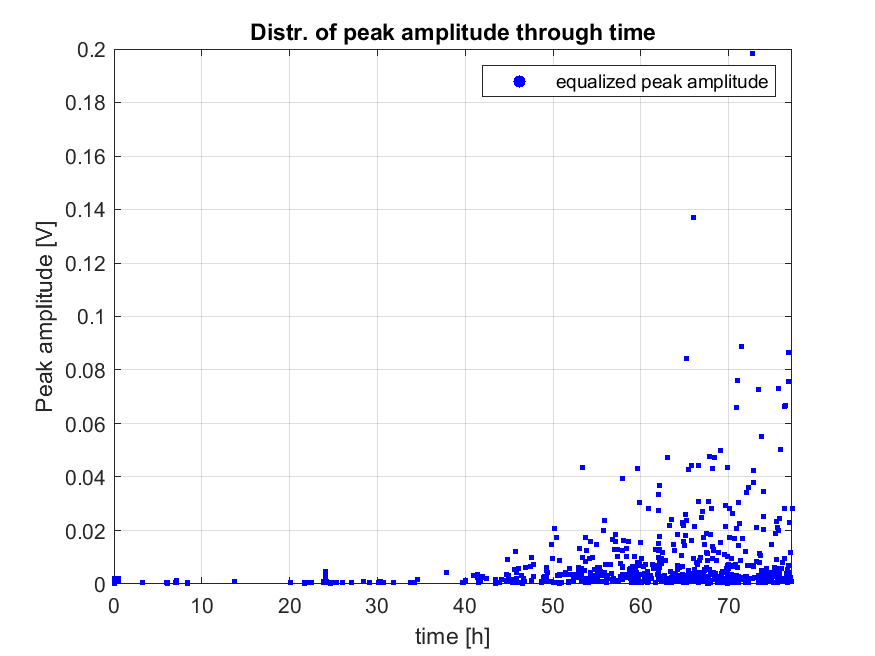
**Nakon izbacivanja outliersa distribucija:**



**DURATION are emissions with value more than 0.00033 [s]**

**Percentage of oultiers for DURATION: 23/745 = 3.0872 %**

**PEAK\_AMPLITUDE:**



**Potencijalni outlieri iznad 0.04-0.05 V:**

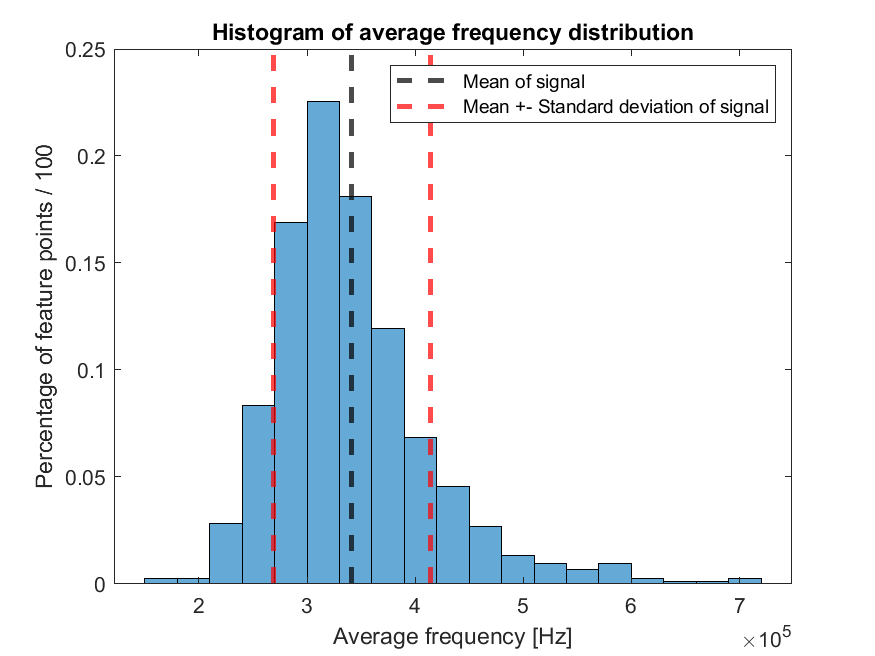
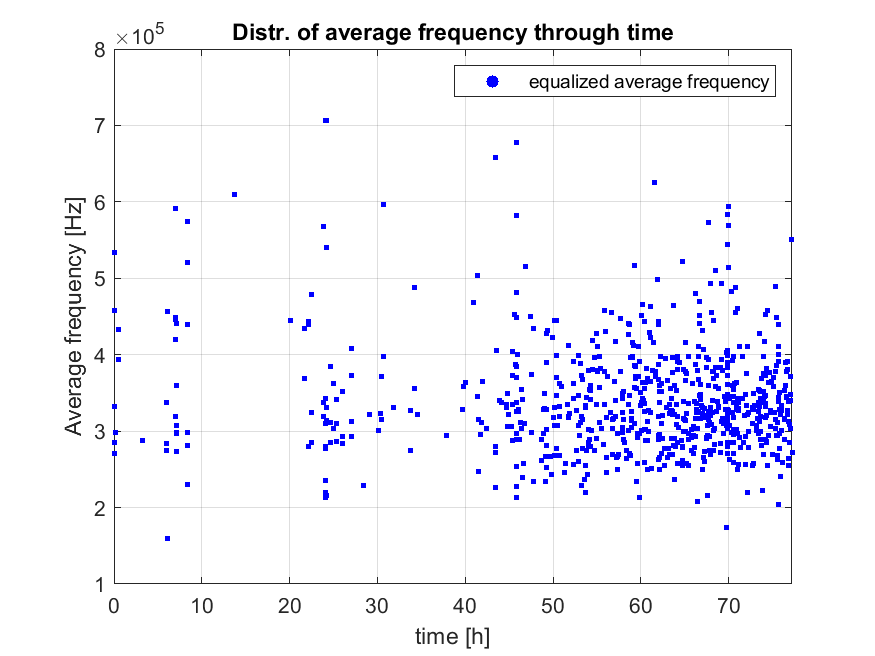
**Dobivamo emisije ispravnog oblika s malo šuma pa PEAK AMPLITUDE nema pravih outliera.**



**Navedeni signali skoro isključivo frekvencije peak amplitude u amplitudnom spektru na 200 kHz te drugi peak oko 300 kHz. Rijetko je obrnuto.**



**AVERAGE\_FREQUENCY:**

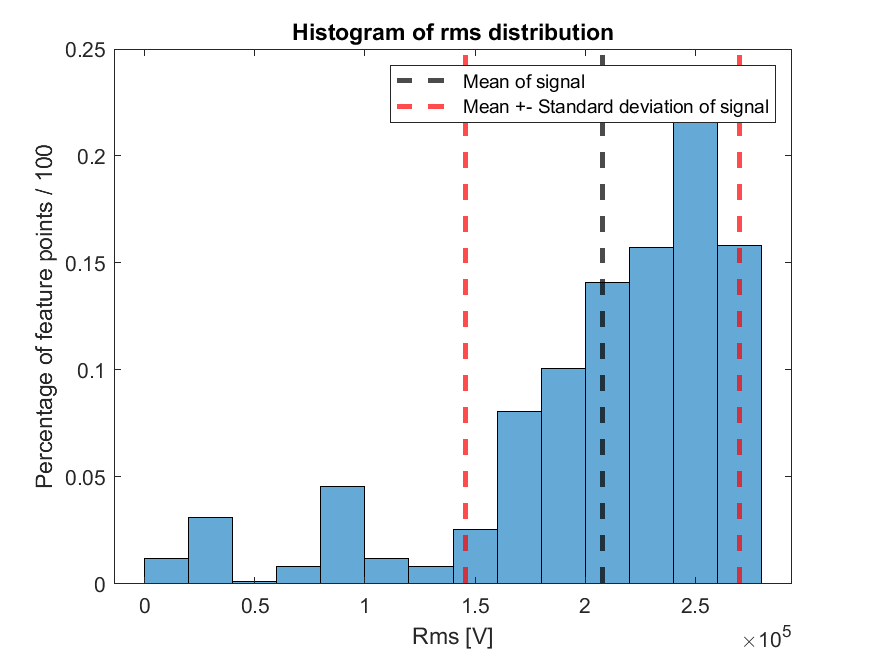
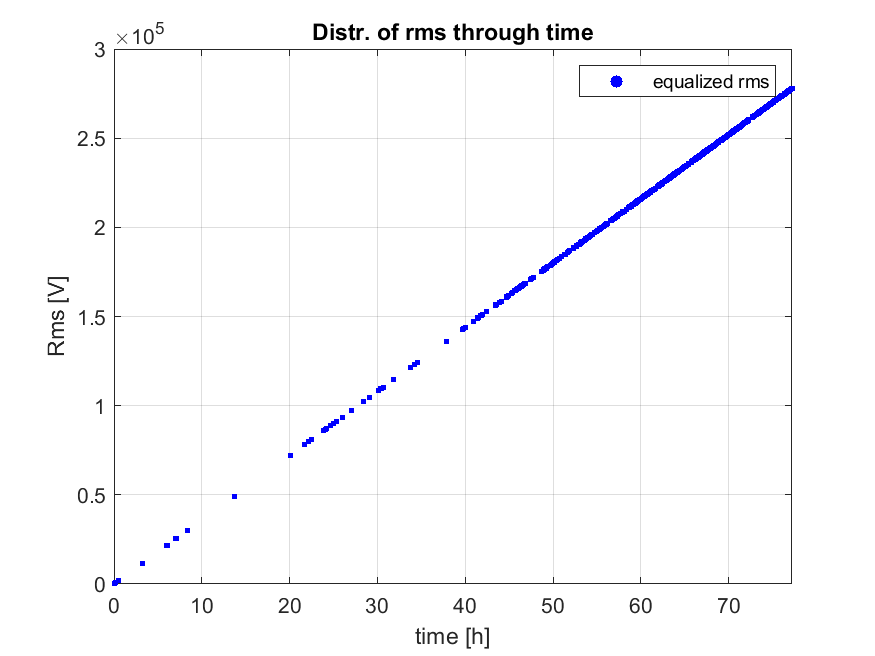


**Potencijalni outlieri iznad 600 kHz:**

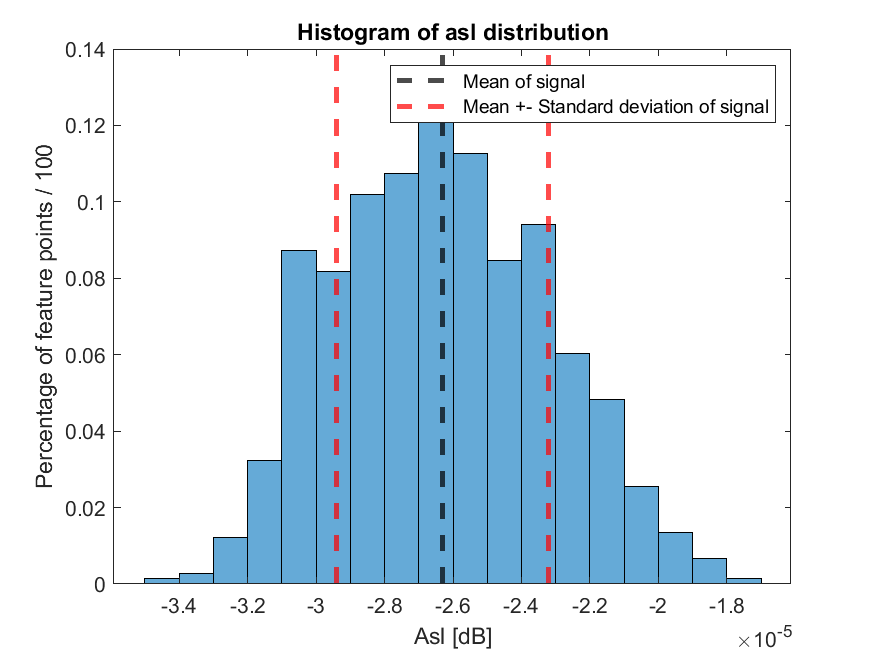
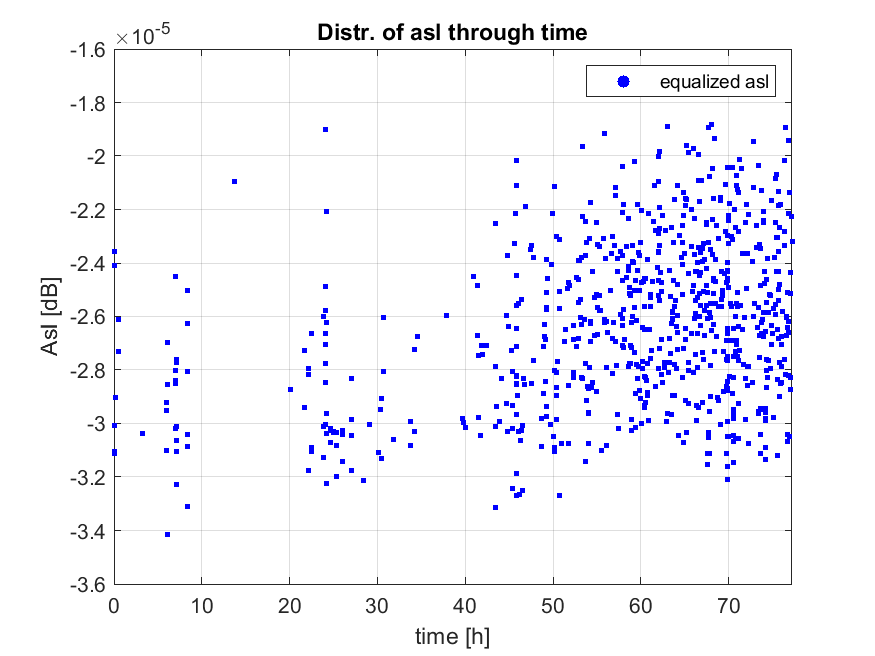
**Dobivamo mješovito ispravne i krive emisije pa nesmijemo izbaciti emisije, iako neke od emisija su već izbačene kao outlieri u prijašnjim koracima.**



**RMS:**

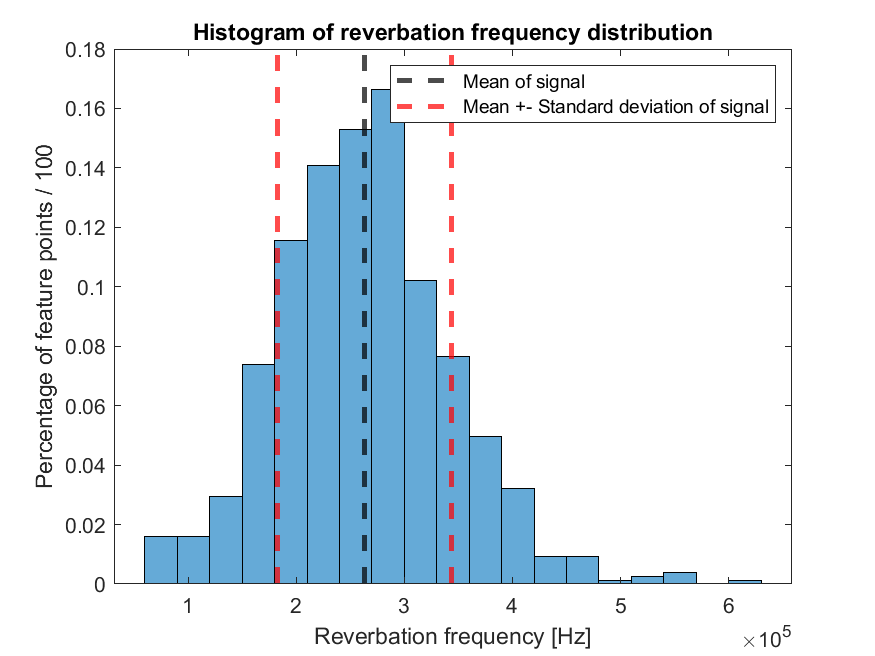
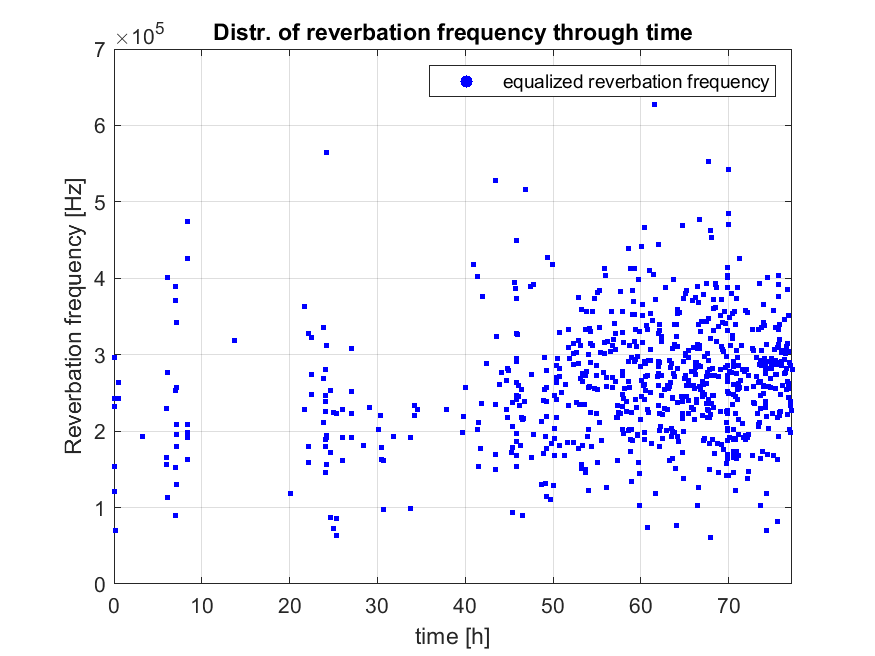


**ASL:**



**Ima kvalitetnu distribuciju bez uočljivih outliera.**

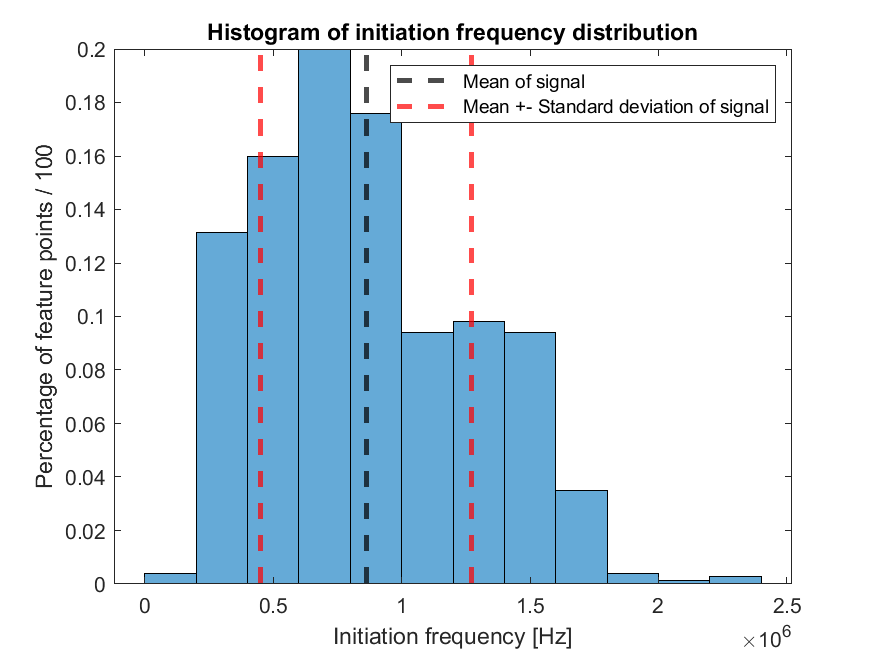
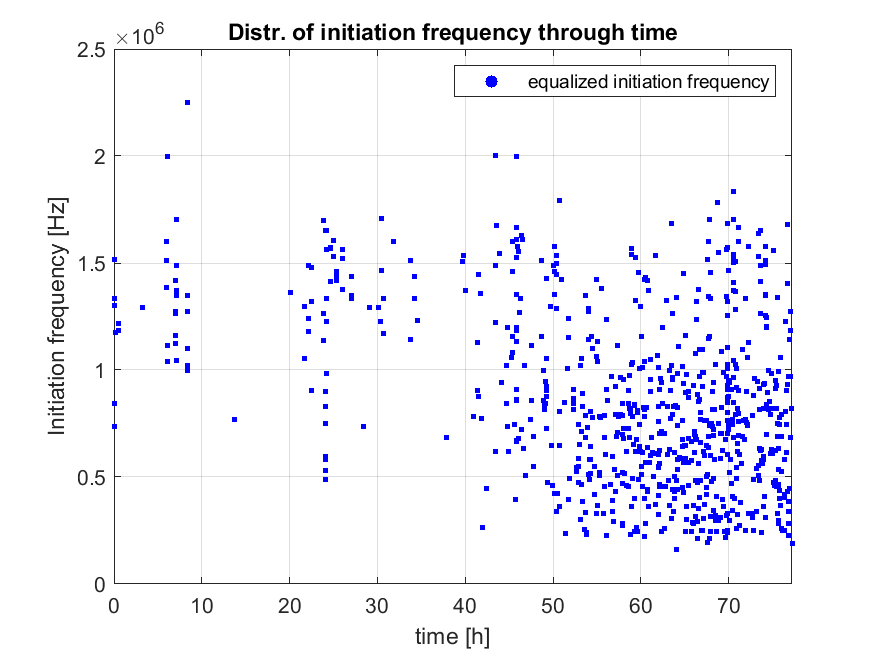
**REVERBATION\_FREQUENCY:**



**Mogući outlieri iznad 500 kHz međutim sadrže i ispravne i neispravne emisije pa nisu pravi outlieri.**



**INITIAL\_FREQUENCY:**

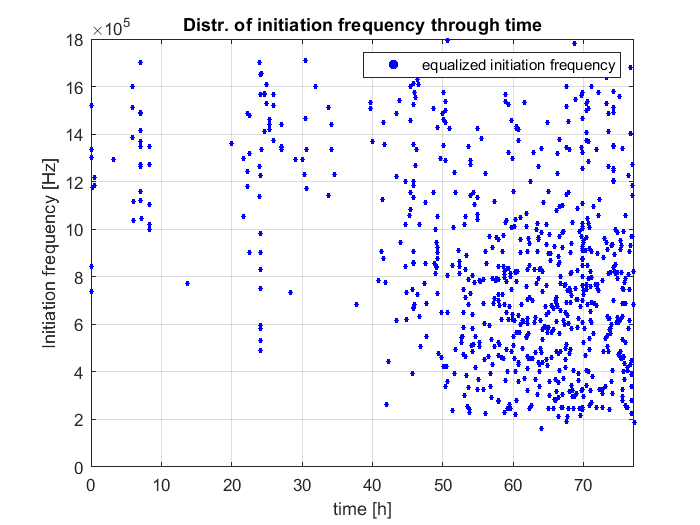


**Mogući outlieri iznad 1.8 MHz:**

**Pronašli smo jako šumovite signale koje je potrebno izbaciti.**



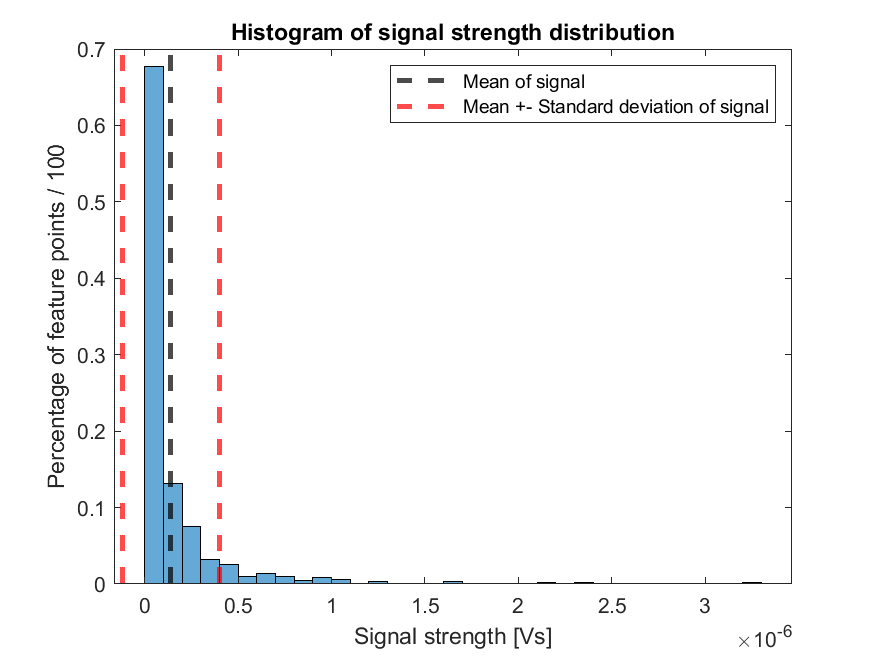
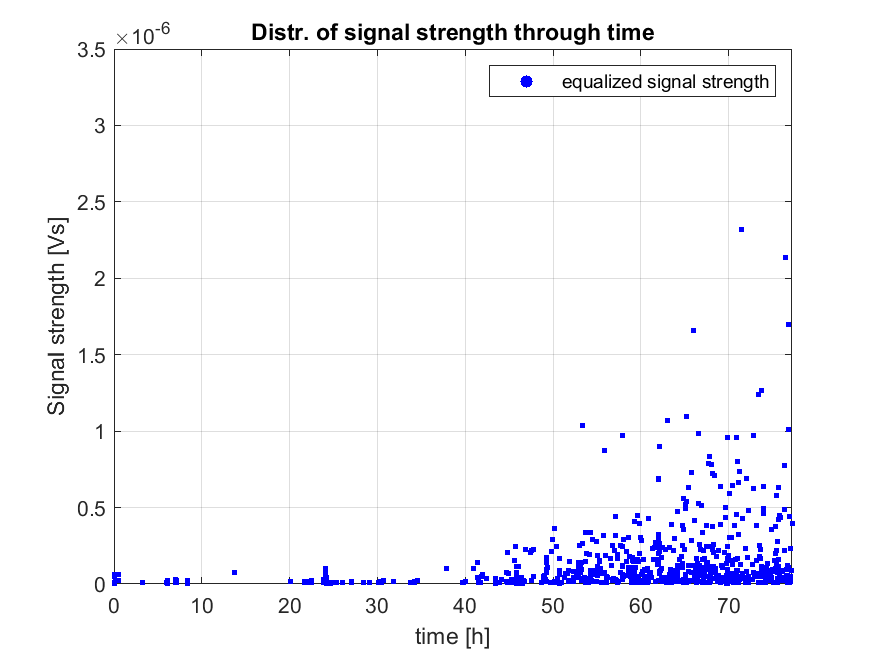
**Nakon izbacivanja outliera distribucija je:**



**INITIATION FREQUENCY are emissions with value more than 1800000 [Hz]**

**Percentage of oultiers for INITIATION FREQUENCY: 6/745 = 0.80537 %**

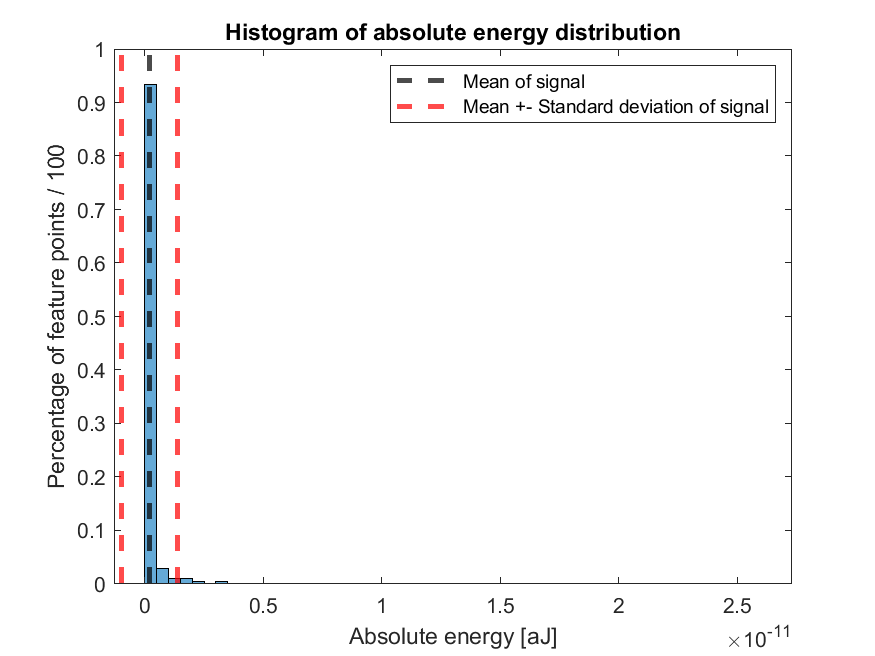
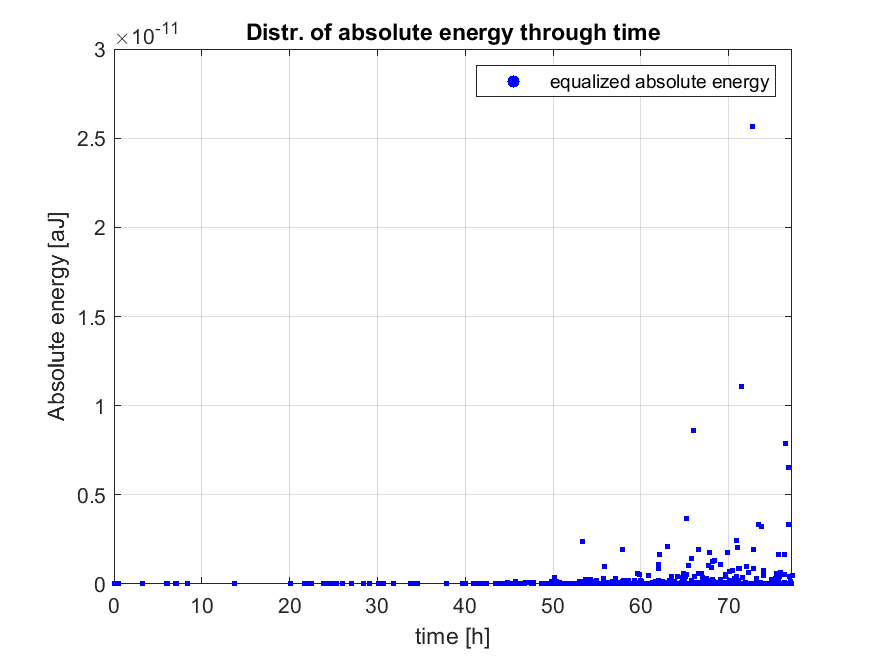
**SIGNAL\_STRENGTH:**



**Mogući outlieri iznad 1 uVs međutim isti pravilni signali kao kod PEAK AMPLITUDE.**



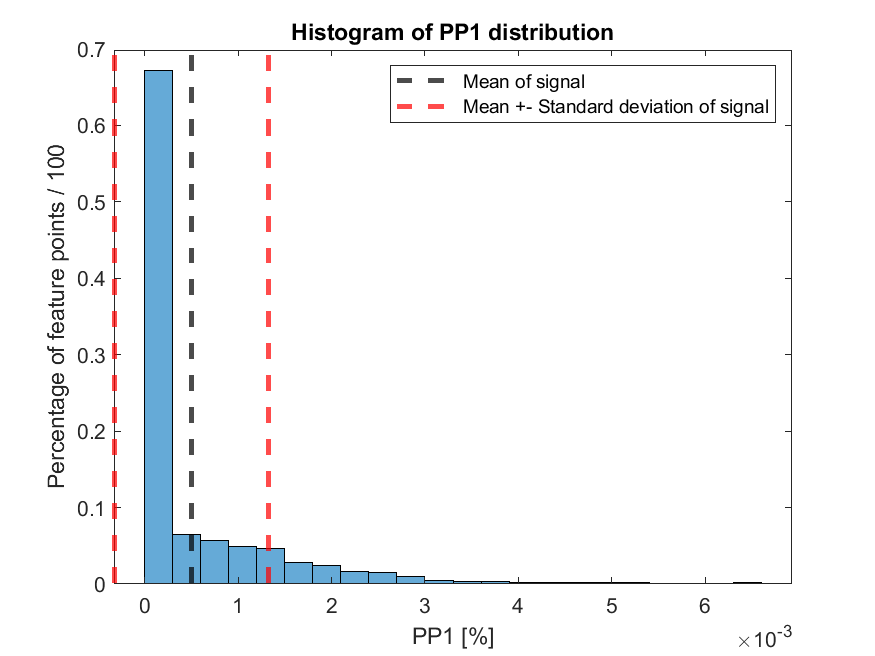
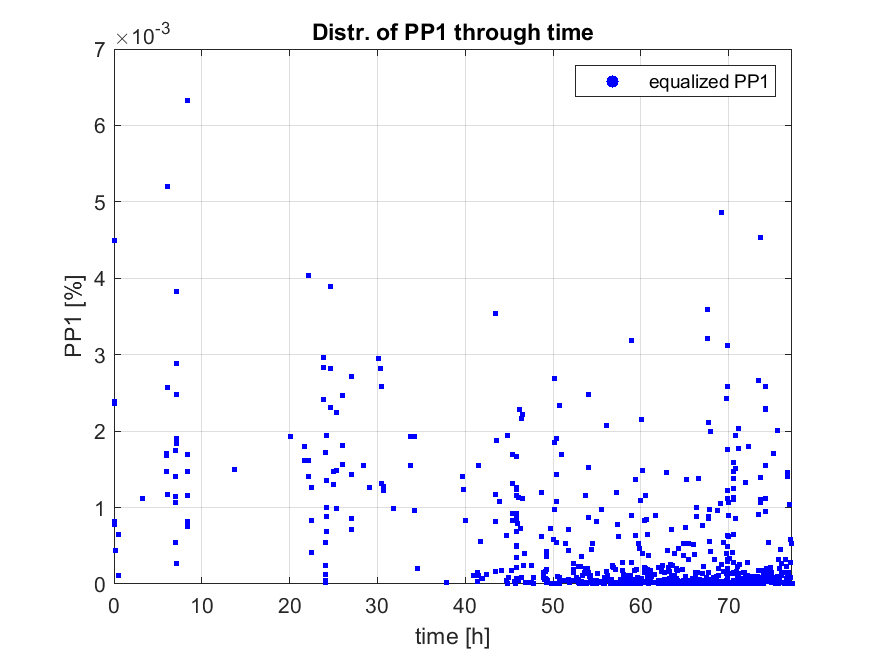
**ABSOLUTE\_ENERGY:**



**Mogući outlieri iznad 2.5 pJ međutim isti pravilni signali kao kod PEAK AMPLITUDE.**



**PP1:**

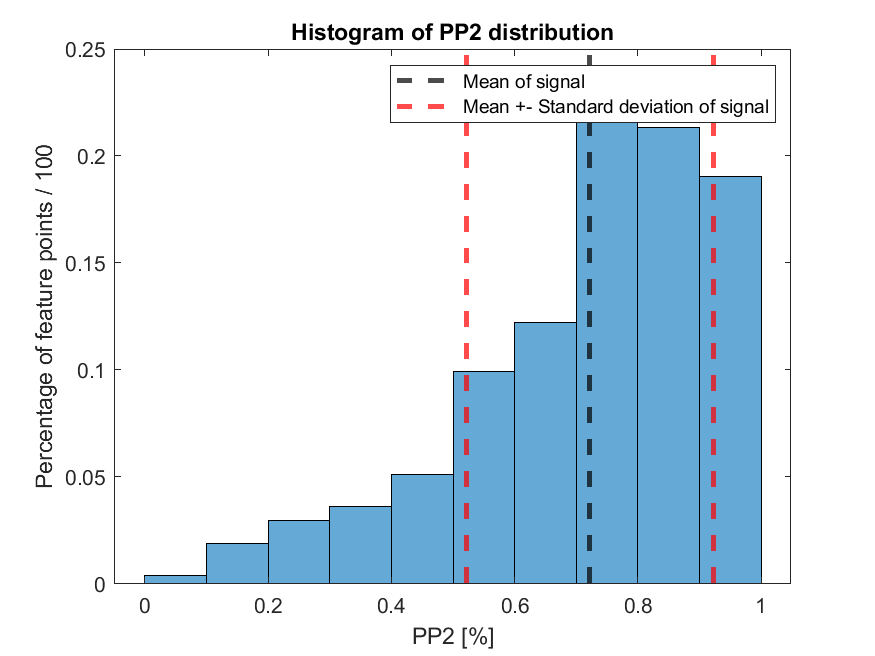
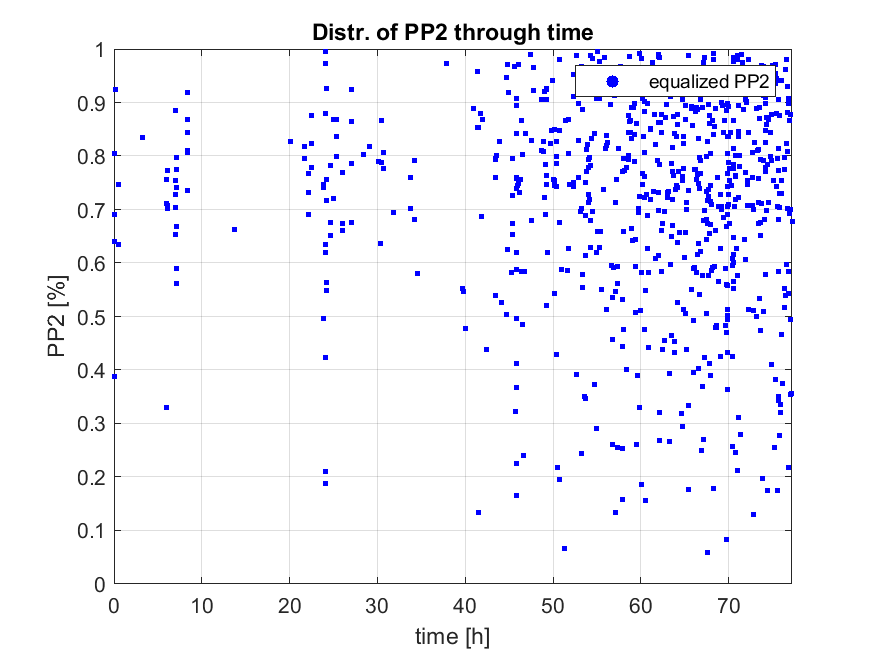


**Vrijednosti PP1 su praktički nula s obzirom na skalu od 0 do 100% zbog ekvilizacije frekvencija ispod 120 kHZ s -35 dB. Emisije koje imaju visoke vrijednosti PP1 su jako šumoviti signali u vremenskoj domeni. Izbacivanjem šumovitih emisija Kako ova značajka ima jako mali utjecaj zbog konstantne distribucije vrijednosti možemo je zanemariti iz selekcije značajki.**

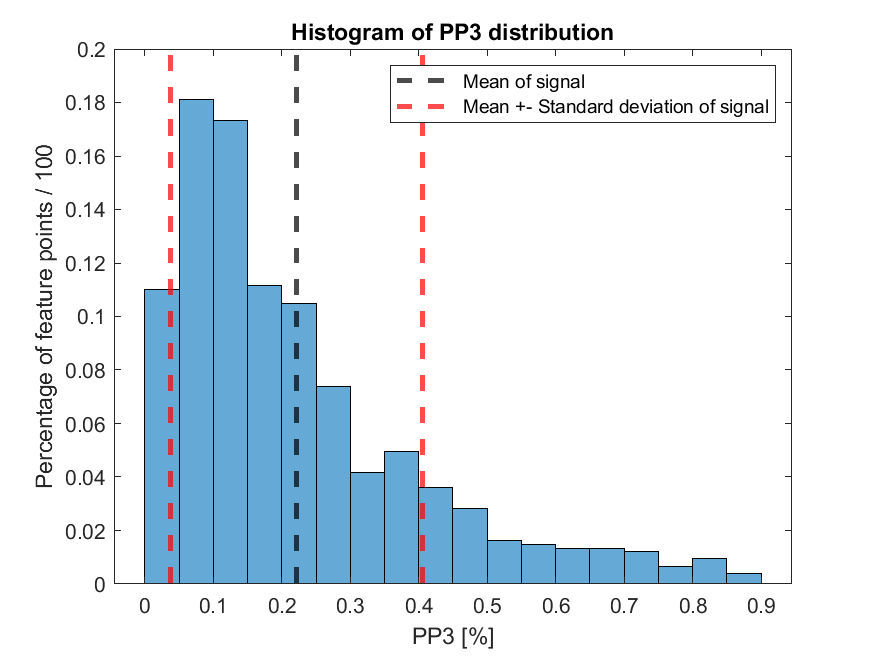
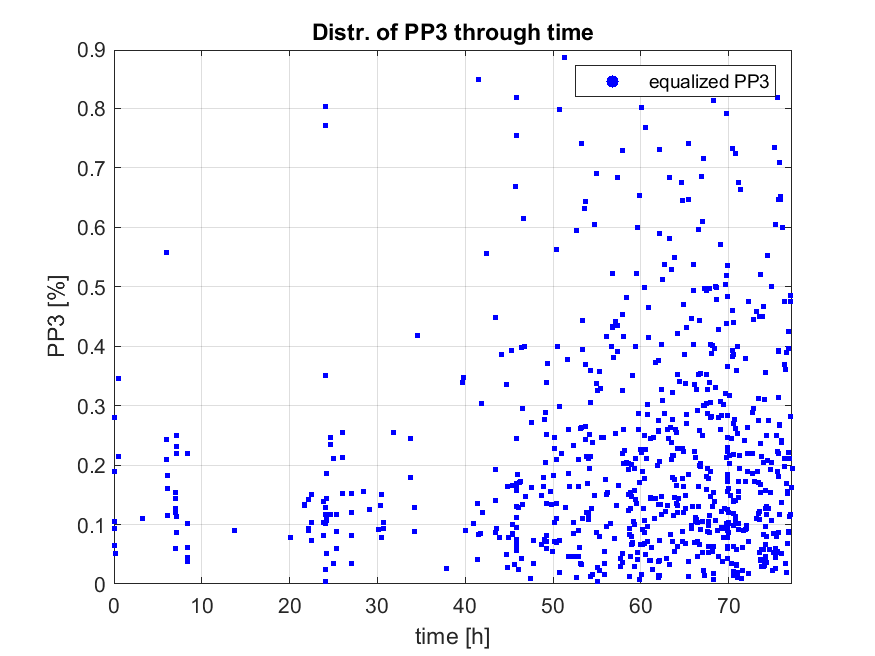
**Najveća vrijednost je 0.00632% prikazana na slici ispod u vremenskom i amplitudnom spektru.**



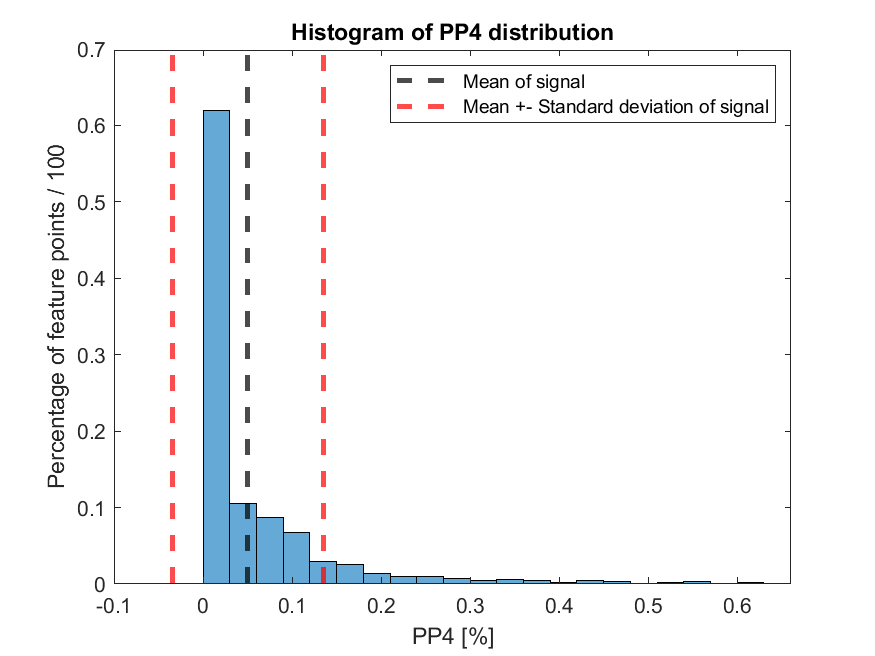
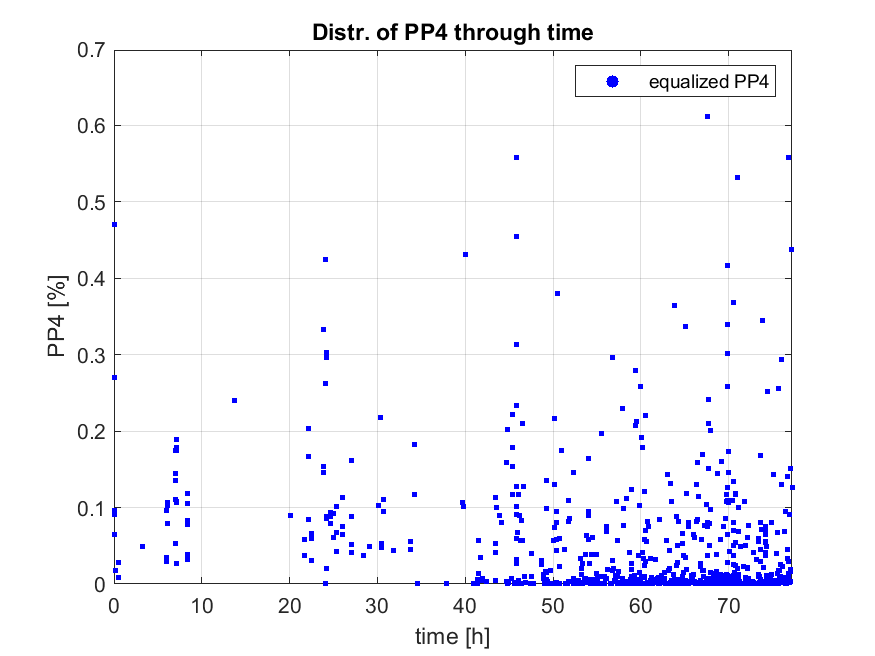
**PP2:**



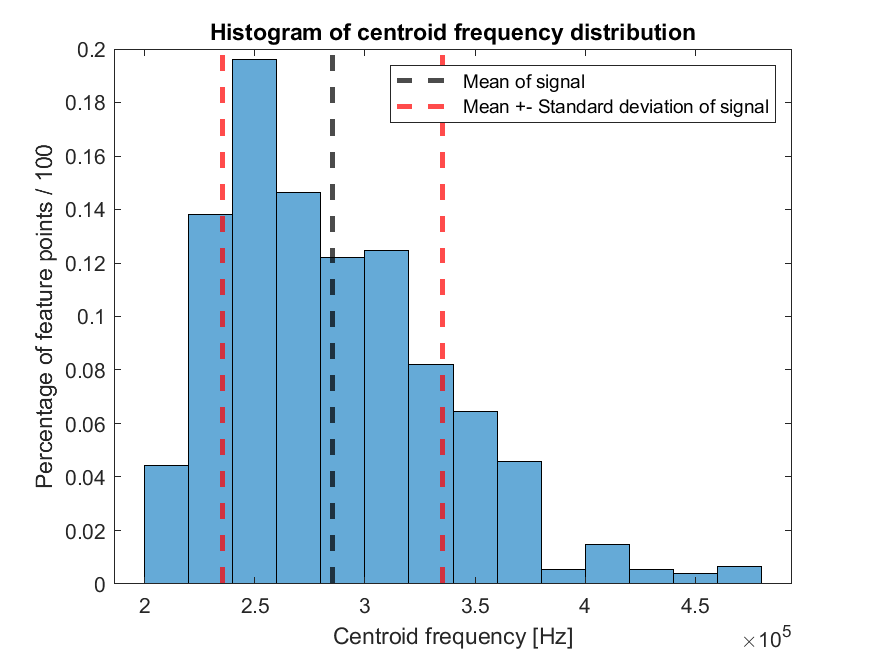
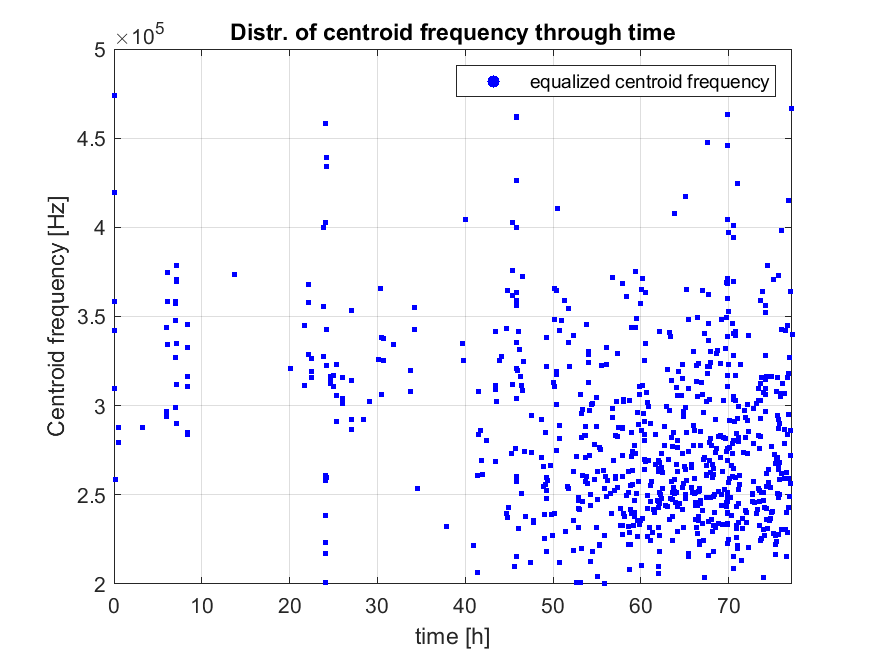
**PP3:**



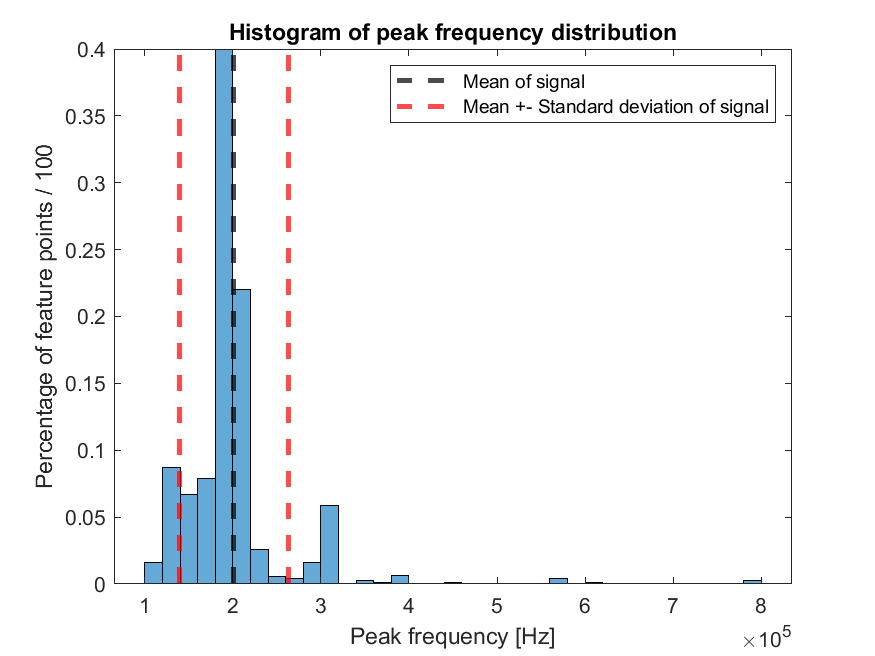
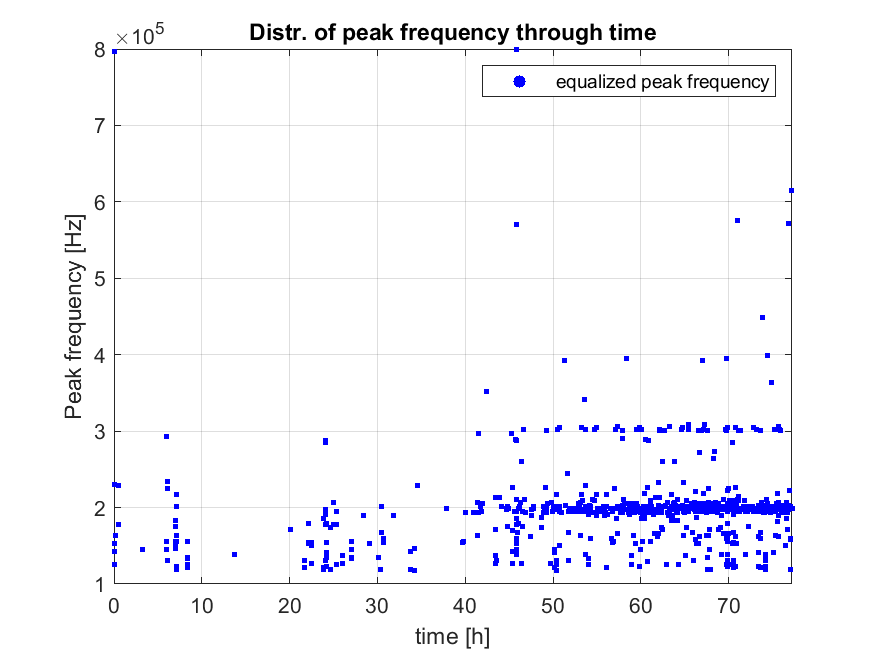
**PP4:**



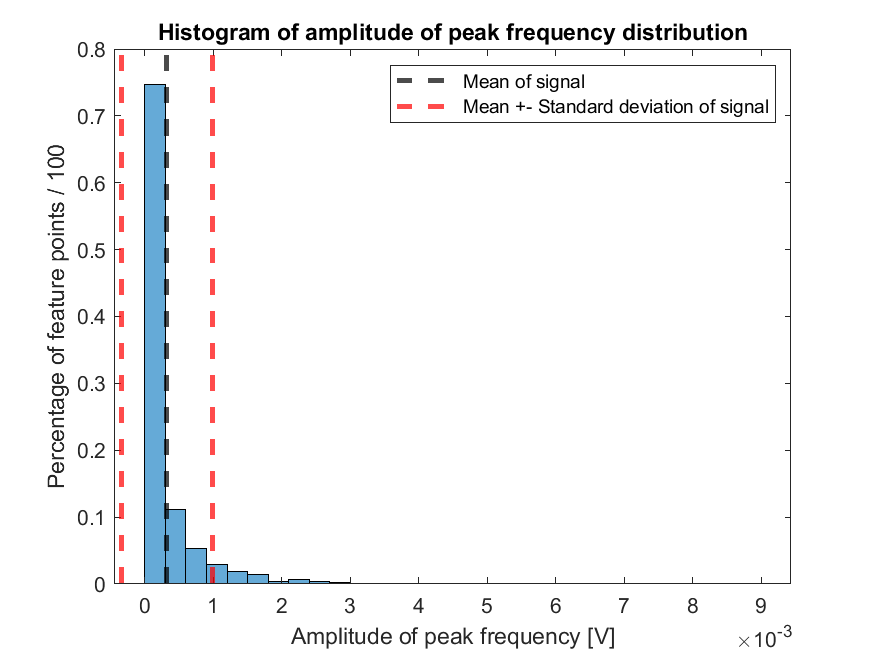
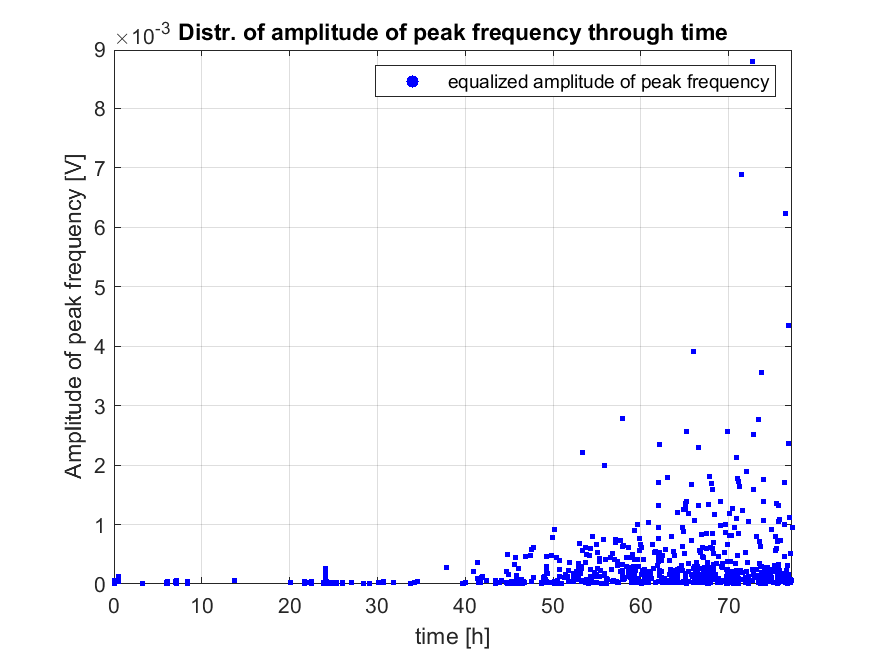
**CENTROID\_FREQUENCY:**



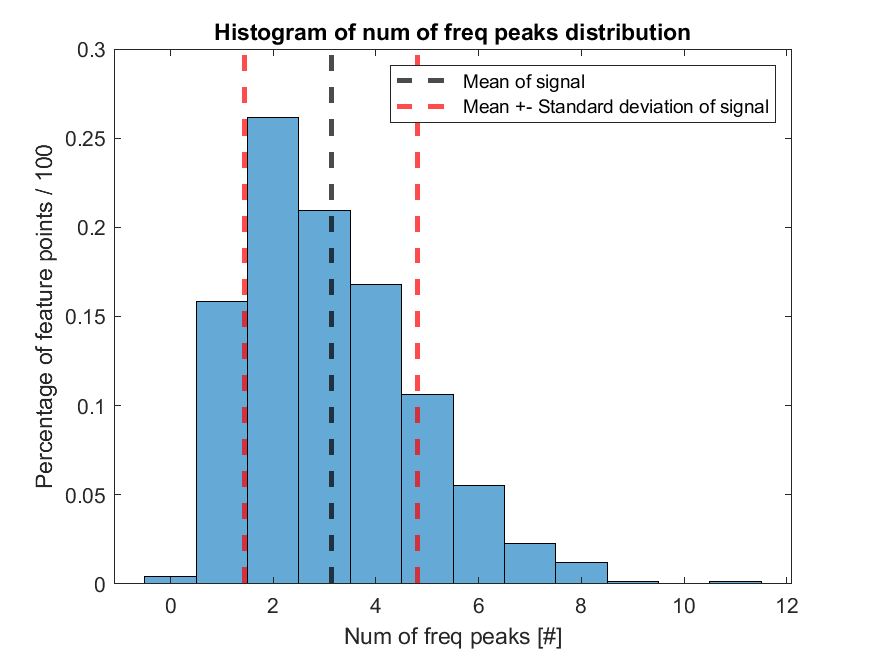
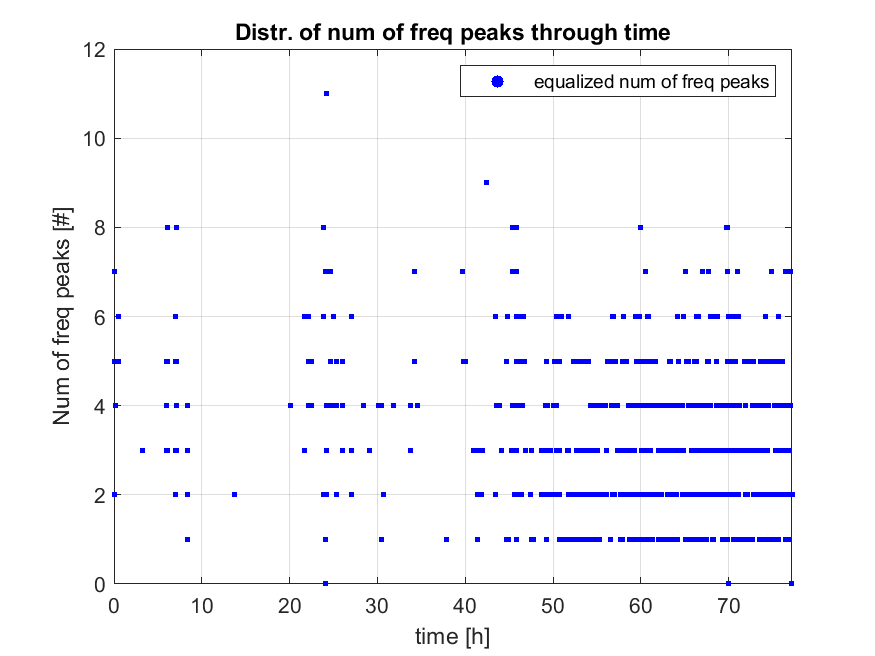
**PEAK\_FREQUENCY:**



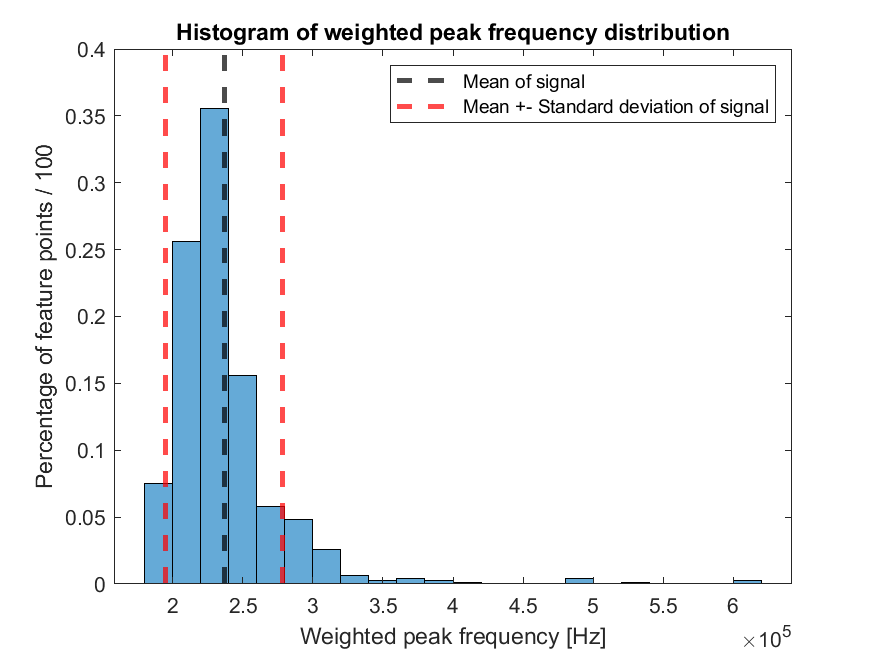
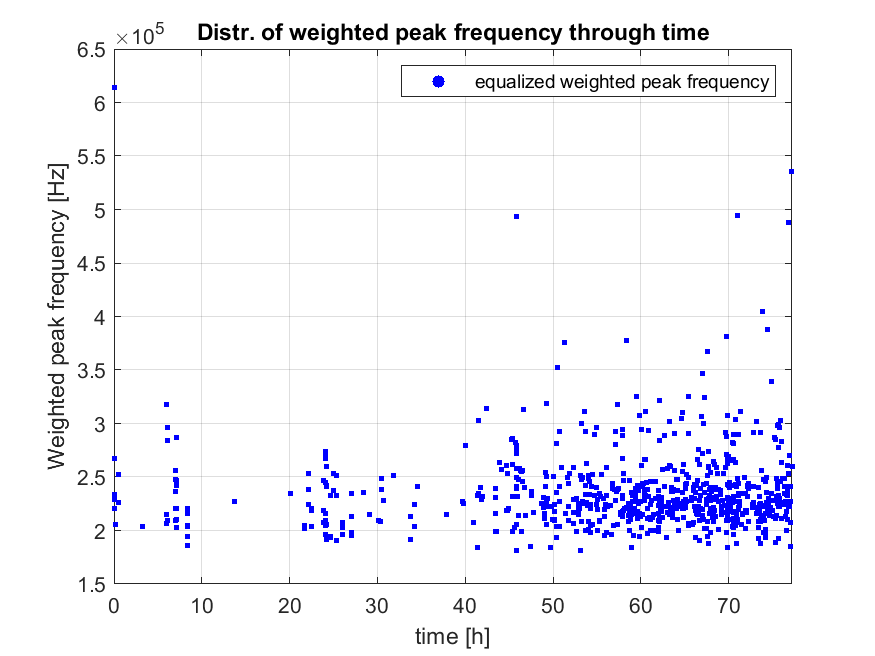
**AMPLITUDE\_OF\_PEAK\_FREQUENCY:**



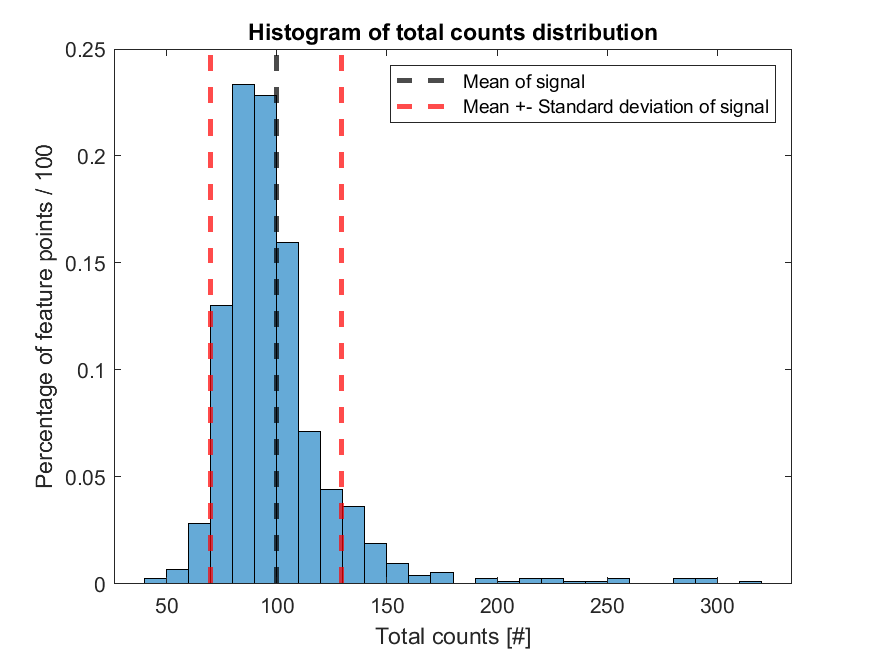
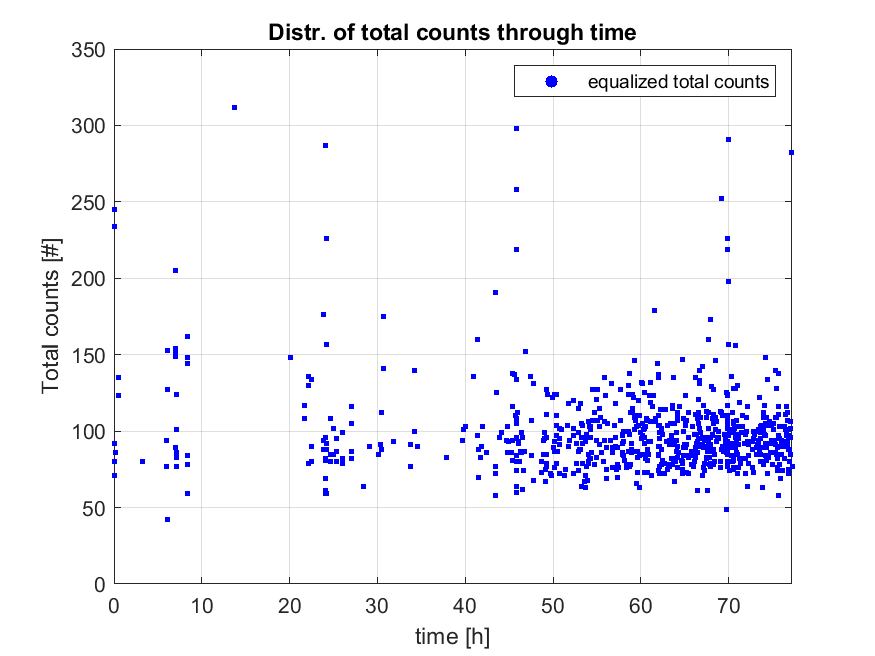
**NUM\_OF\_FREQ PEAKS:**



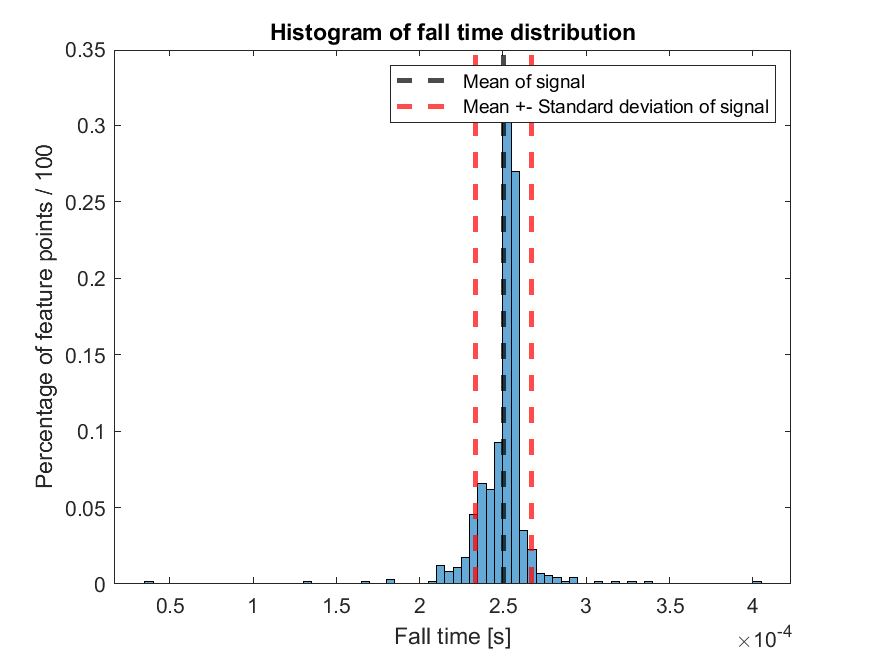
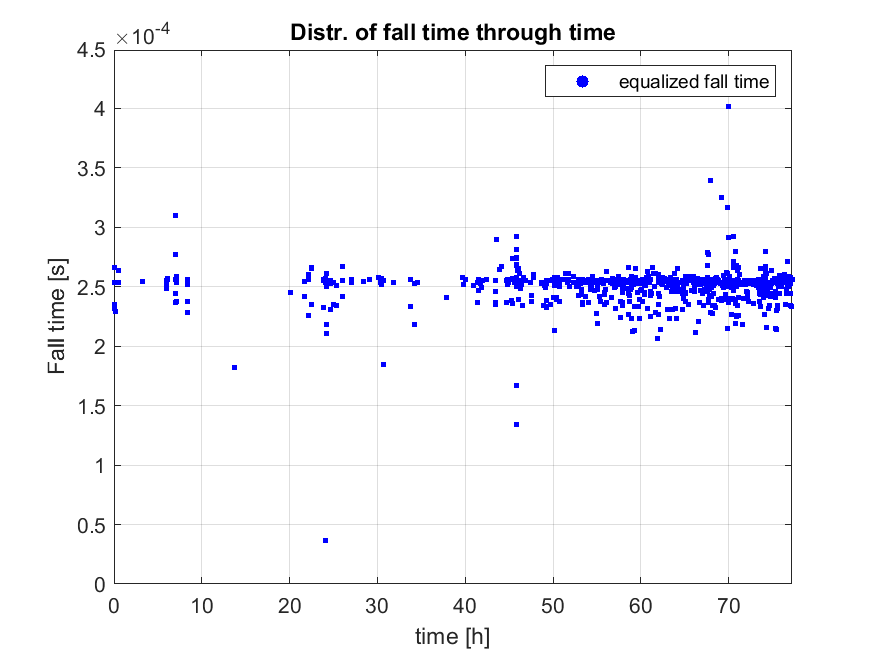
**WEIGHTED\_PEAK\_FREQUENCY:**



**TOTAL\_COUNTS:**



**FALL\_TIME:**



**Normalizacija značajki:**

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