Notes on using genetic algorithm for the user equilibrium problem

PyGAD library

https://pygad.readthedocs.io/en/latest/

Triangle Area Similarity – Sector Area Similarity (TS-SS)

- Alternative function for evaluating differences between computed flows and the real flows.
- TS-SS computes the similarity between two vectors from diverse perspective and generates the similarity value from two vectors not only from the angle and Euclidean distance between them, but also the difference between their magnitudes.
- Test runs with the same hyperparameters show that using TS-SS similarity yields slightly better results than using L2 norm.
- https://www.analyticsvidhya.com/blog/2021/06/nlp-answer-retrieval-from-document-using-ts-ss-similarity-python/
- https://www.computer.org/csdl/proceedings-article/bigdataservice/2016/2251a142/12OmNweBUID

TS-SS

```
# Try the TS-SS scoring function

tmp_v0 = np.random.rand(10)

tmp_v1 = np.random.rand(10)

print(ts_ss(tmp_v0, tmp_v1))

>> 0.0005569253043530737

print(ts_ss(tmp_v0, tmp_v0))  # Identical vectors

>> 0.0

print(ts_ss(tmp_v0, -5*tmp_v0))  # Vectors of different directions & very
different magnitudes

>> 6.407215524980118
```

Fitness function

- For both similarity functions (norm errors & TS-SS similarity), bigger scores indicate bigger differences -- we need to minimize this.
- The fitness function priorities higher fitness values. So we need to invert the similarity scores.

$$\text{fitness score} = \frac{1}{2^{\text{similarity score}}}$$

$$\text{similarity score} = 0 \text{ (identical vectors)} \iff \text{fitness score} = 1$$

Search for a good set of parameters

Focus on three parameters for the genetic algorithm:

- Number of generations: 50, 150, 500, 700, 850
- Solutions per generation: 5, 10, 16, 24, 32
- Mutation percentage: 1, 5, 10

Each parameter is tested against a number of values while the other two parameters are kept the same.

Criteria to focus on:

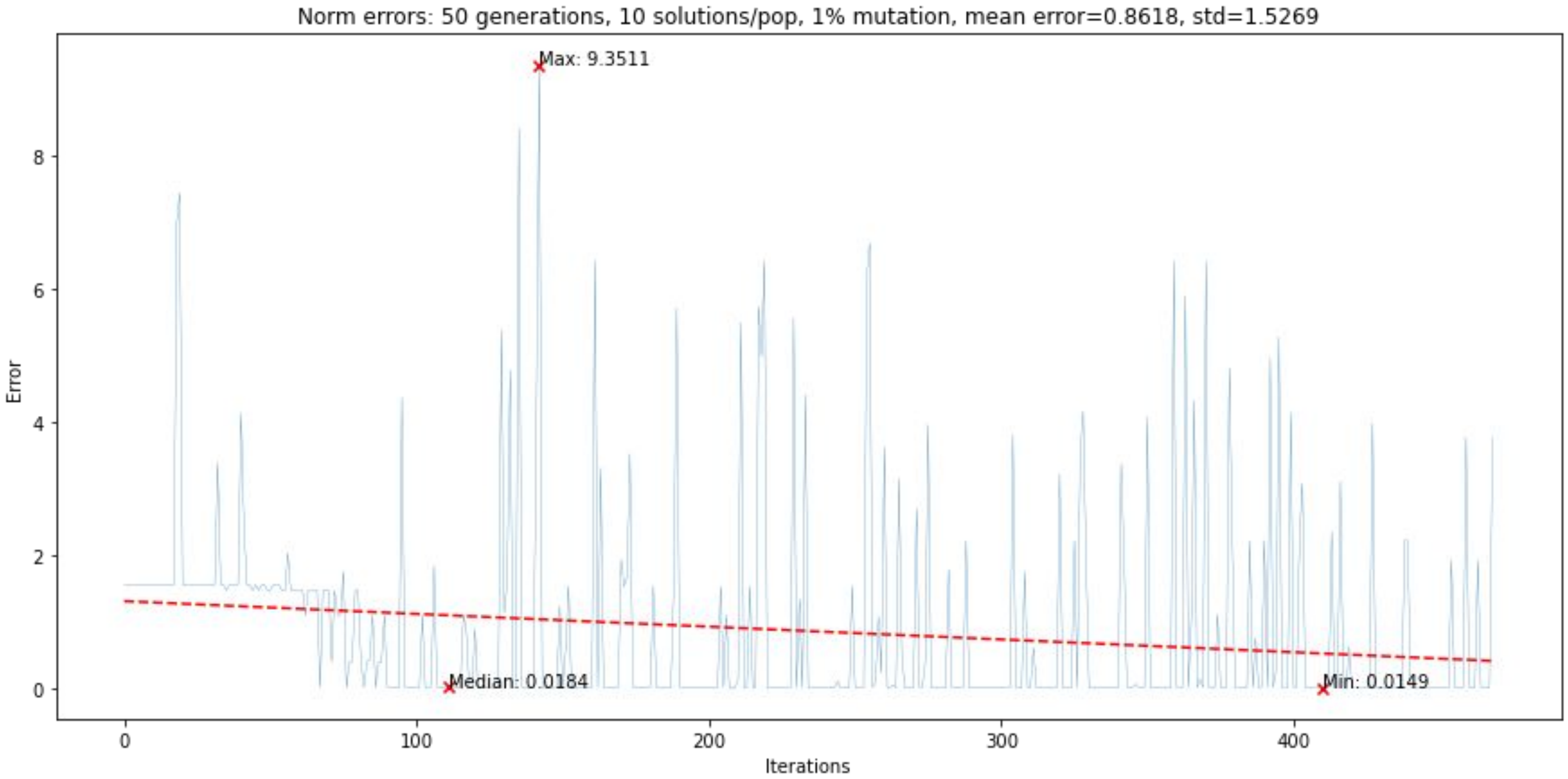
- Min, mean, median, standard deviation of errors (L2 norm errors)

Number of generations

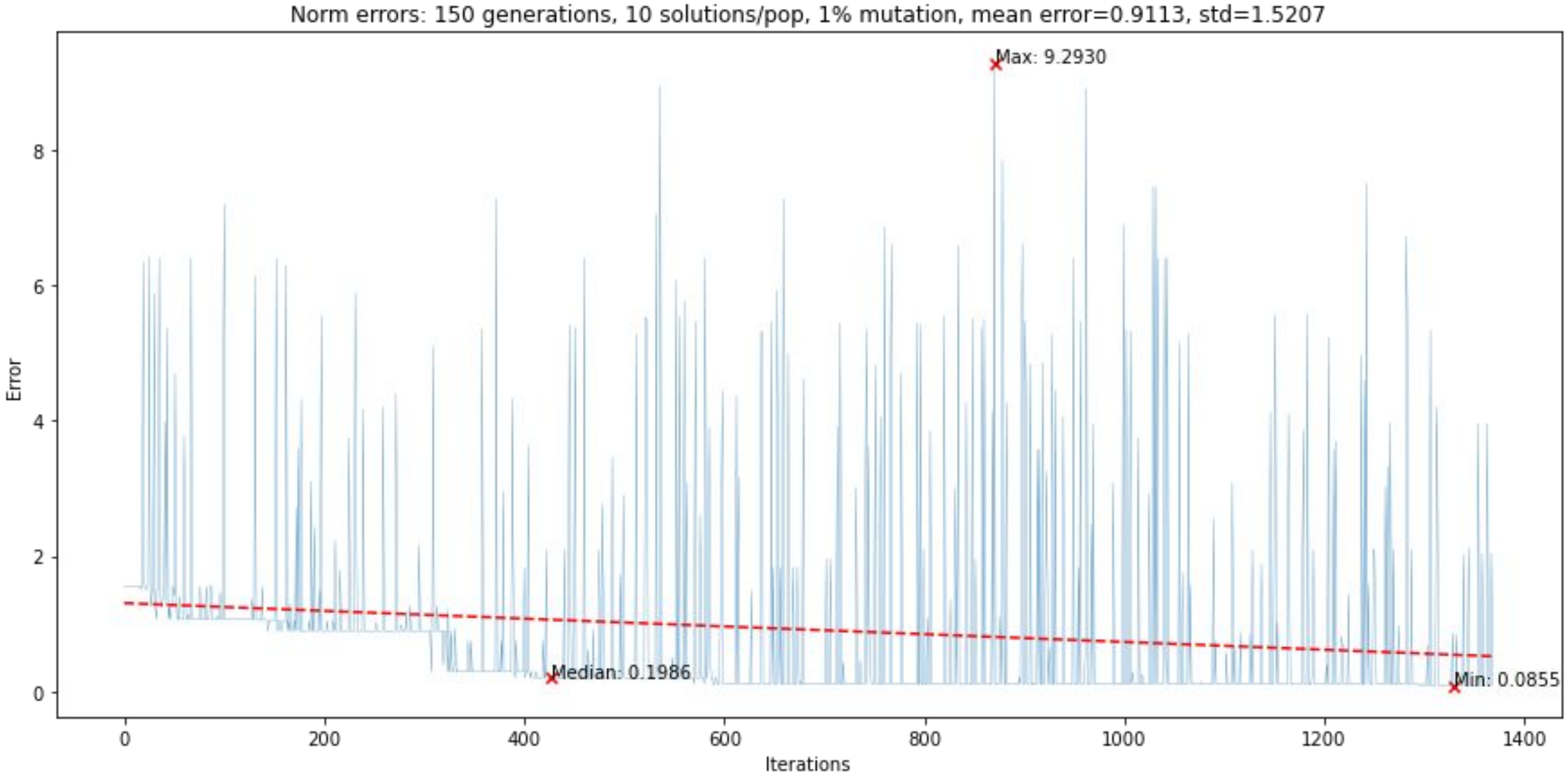
- Values to be tested: 50, 150, 500, 700, 850
- Number of solutions per generation: 10
- Mutation percentage: 1%

Values	Min	Mean	Median	STD
50	0.0149	0.8618	**0.0184**	1.5269
150	0.0855	0.9113	0.1986	1.5207
500	**0.0135**	0.7020	0.0229	1.5555
700	0.0237	**0.6232**	0.0288	**1.3782**
850	0.0202	0.6417	0.0364	1.5027

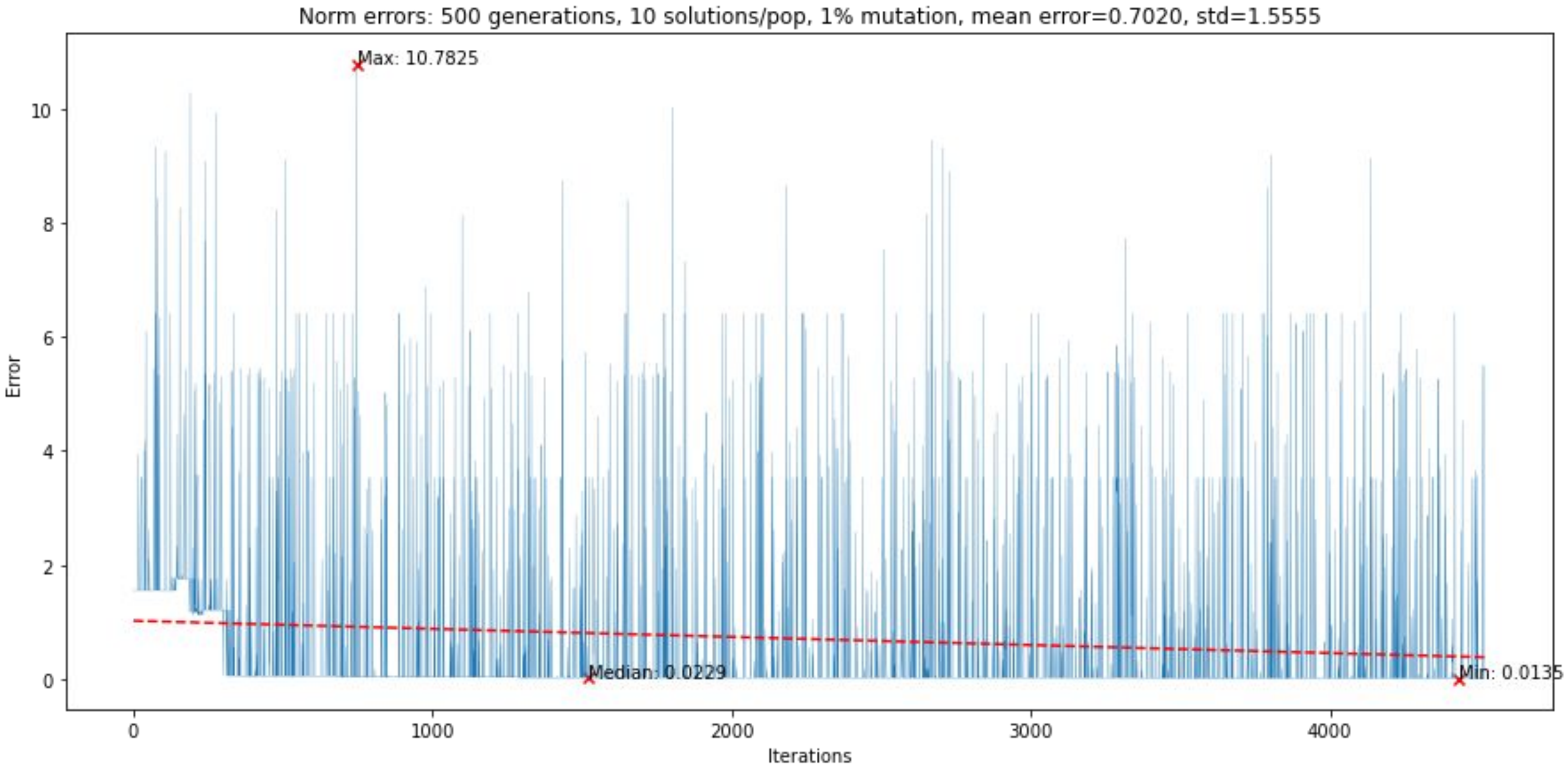
Number of generations - 50



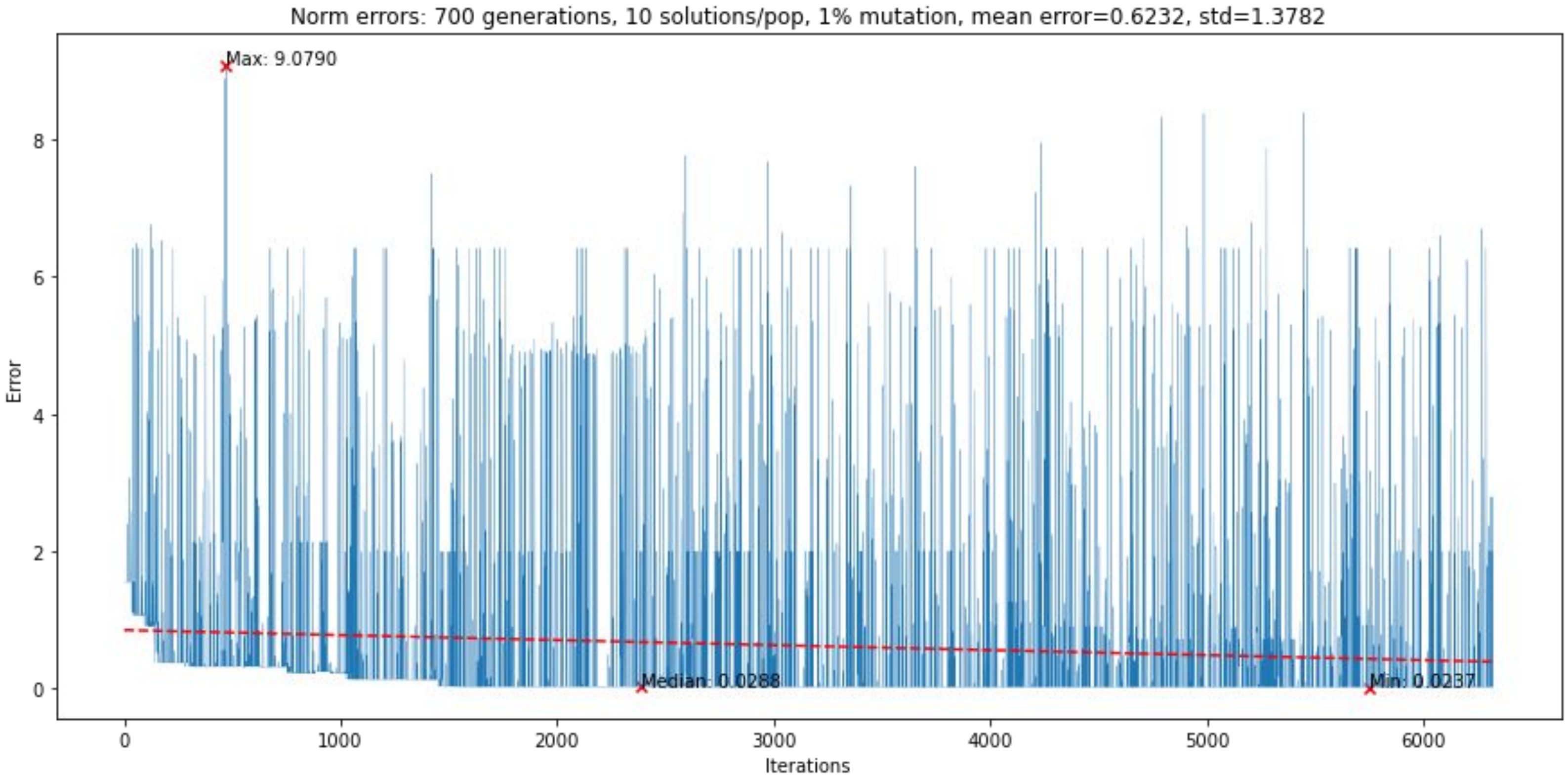
Number of generations - 150



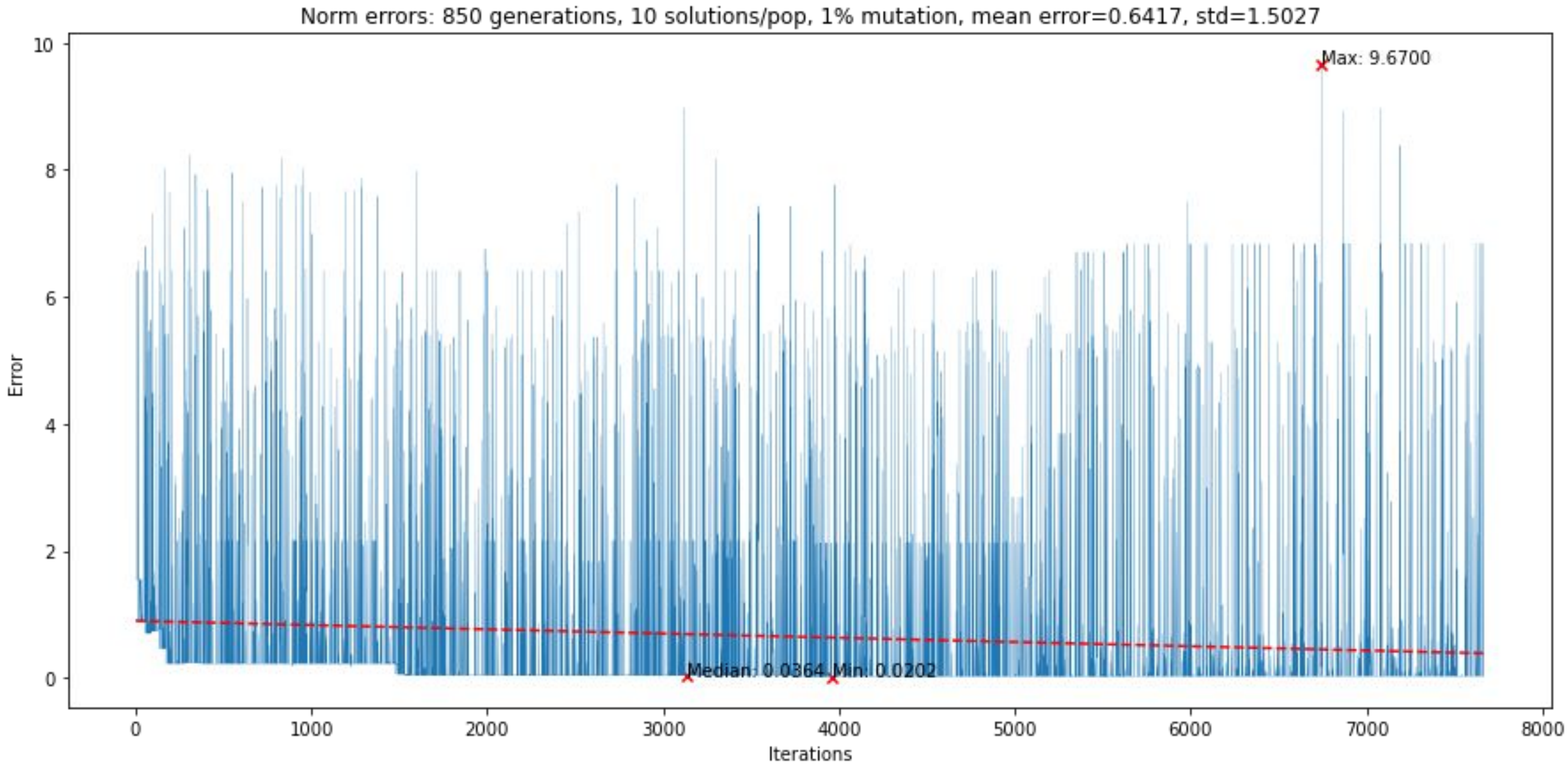
Number of generations - 500



Number of generations - 700



Number of generations - 850

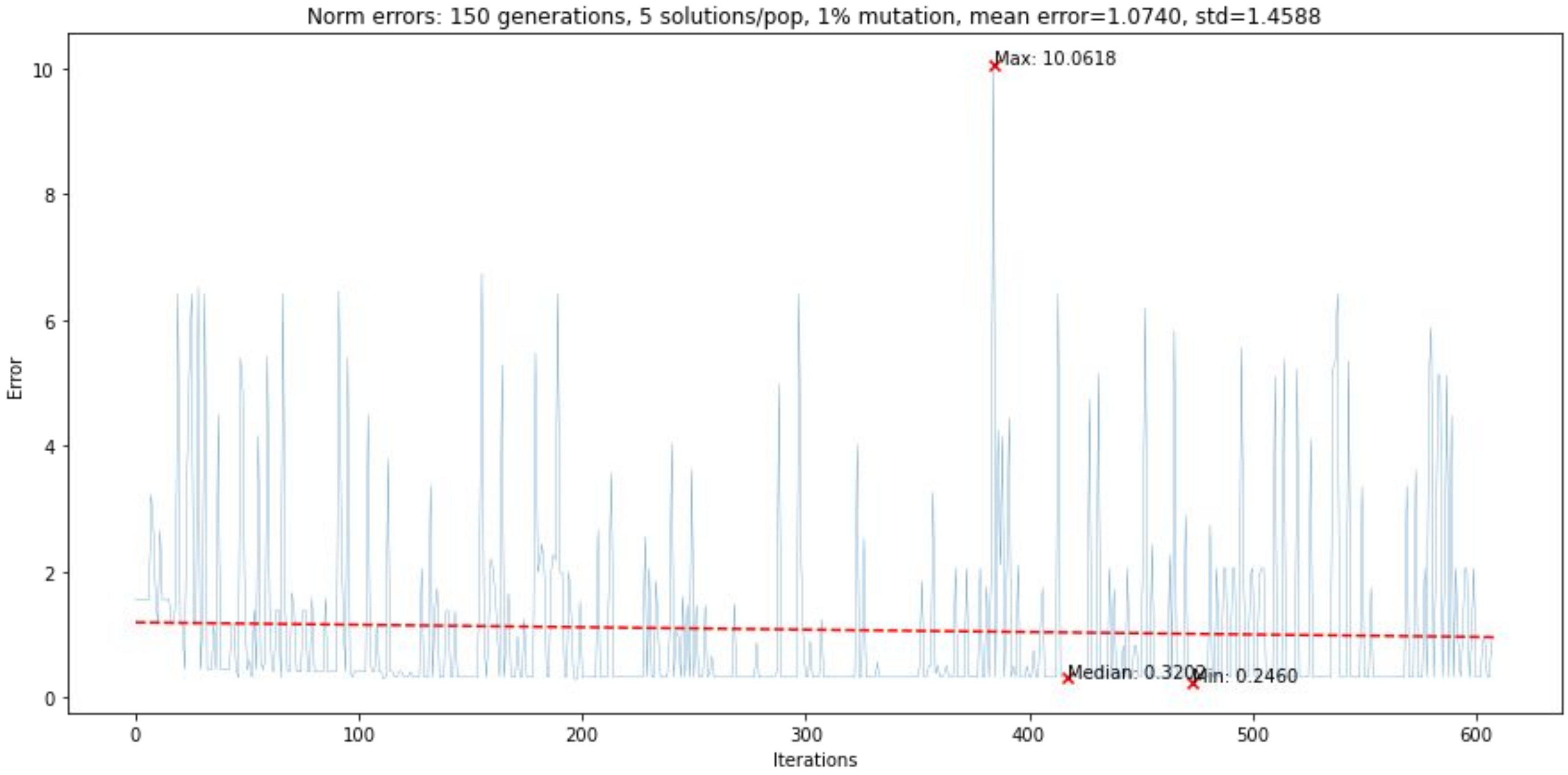


Number of solutions per generation

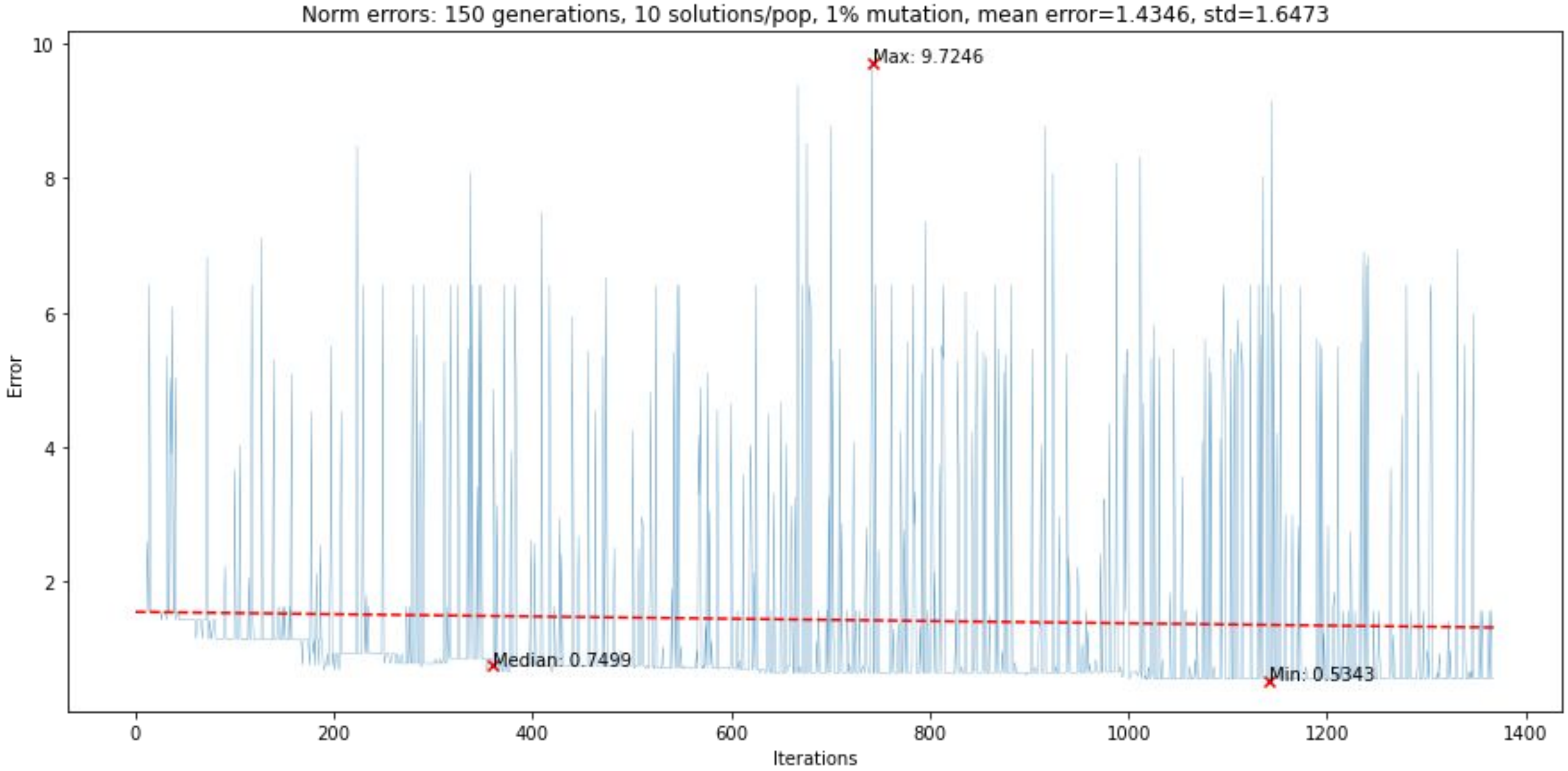
- Values to be tested: 5, 10, 16, 24, 32
- Number of generations: 150
- Mutation percentage: 1%

Values	Min	Mean	Median	STD
5	0.2460	1.0740	0.3202	1.4588
10	0.5343	1.4346	0.7499	1.6473
16	**0.0153**	**0.7173**	0.2060	1.4378
24	0.0382	0.7477	0.1492	1.4391
32	0.0411	0.7325	**0.0713**	**1.3706**

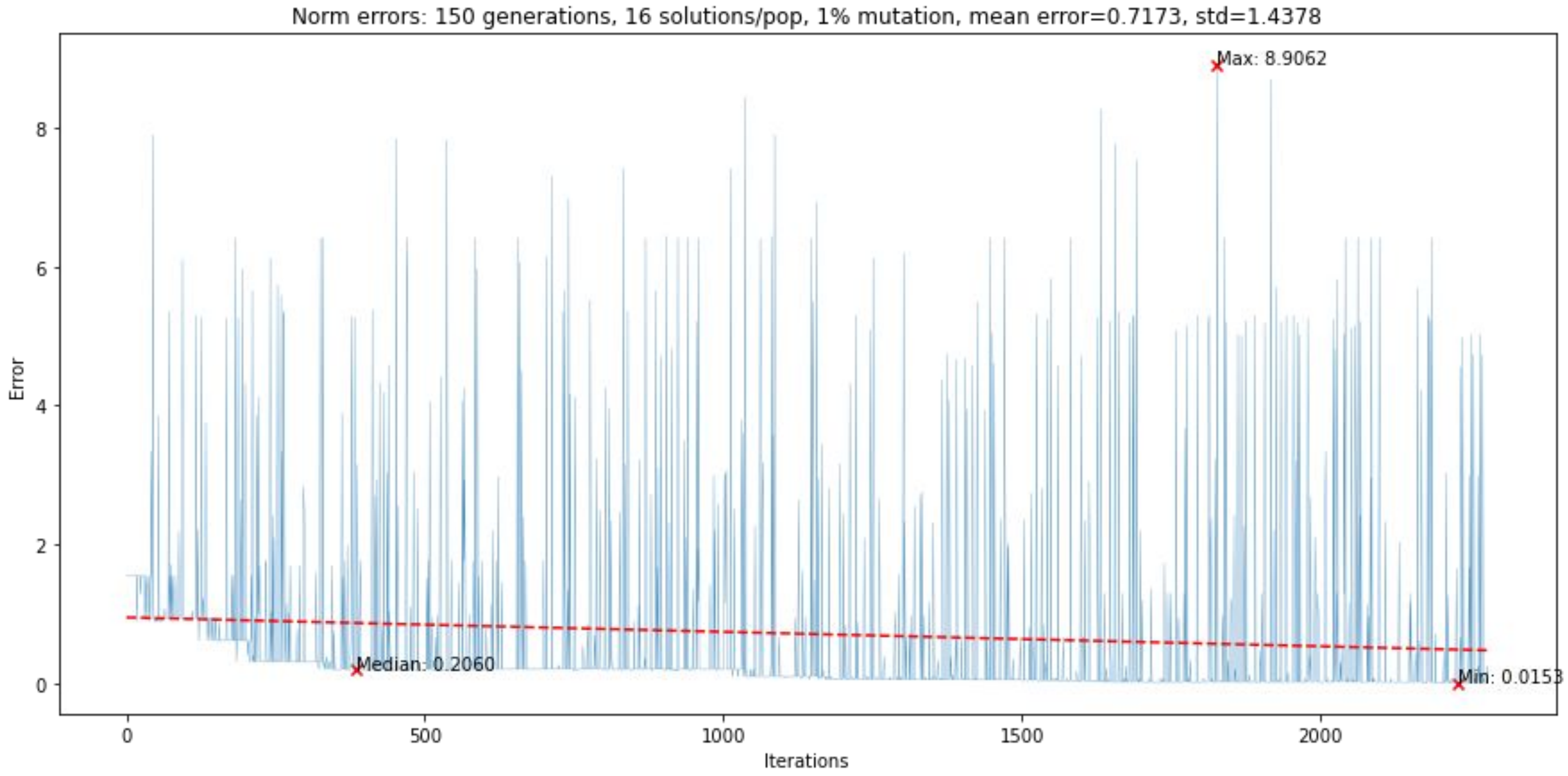
Number of solutions per generation - 5



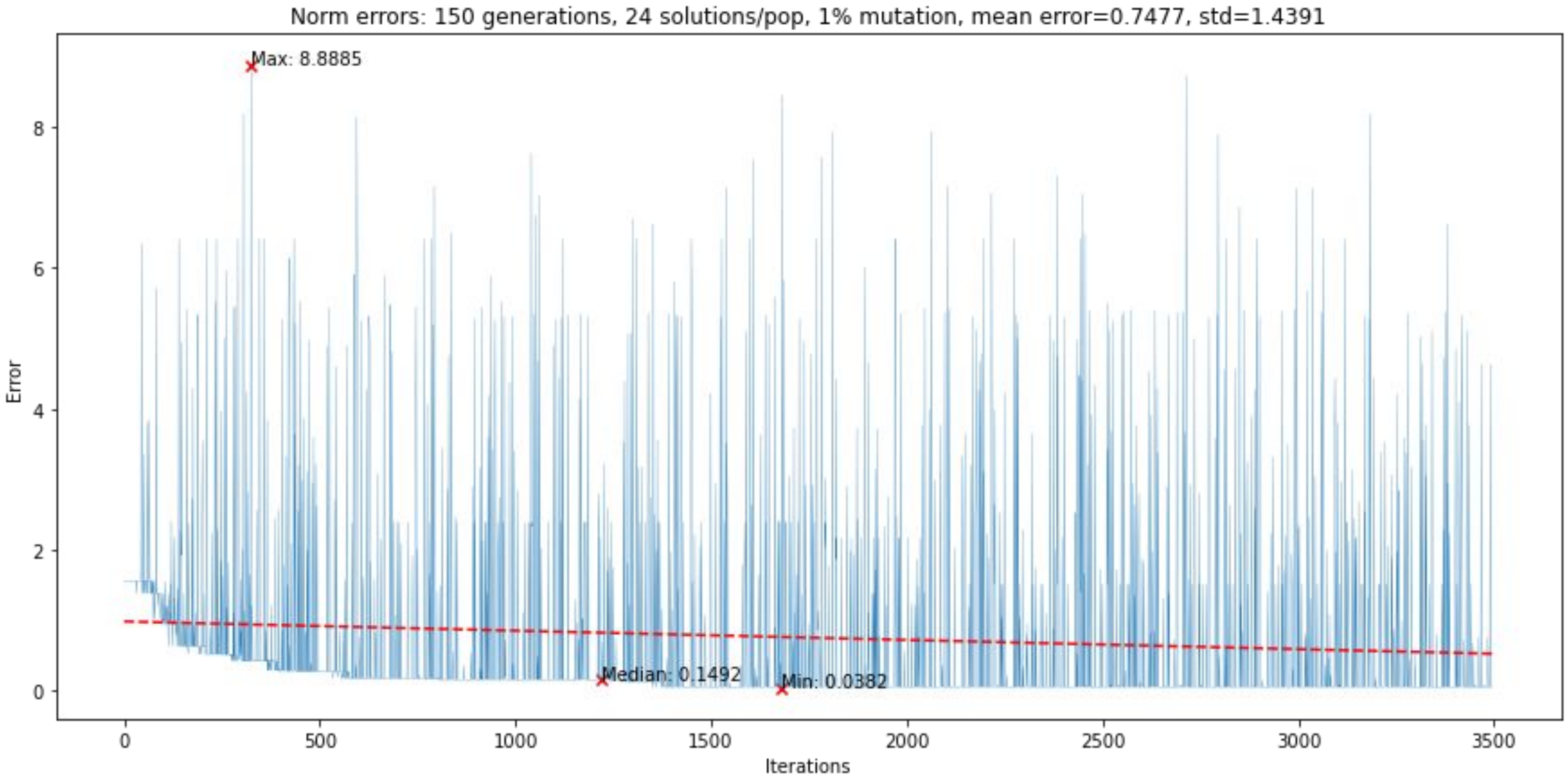
Number of solutions per generation - 10



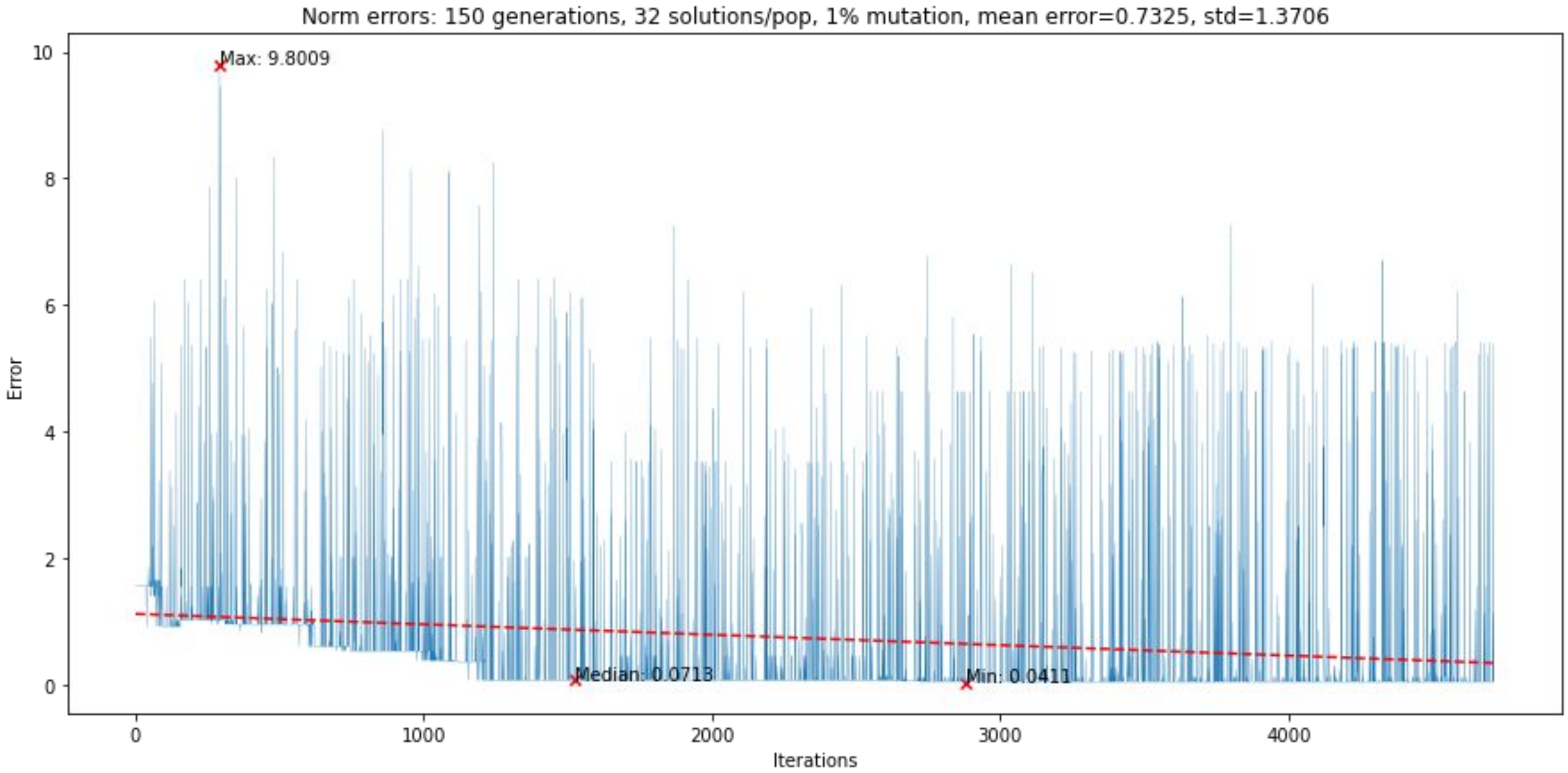
Number of solutions per generation - 16



Number of solutions per generation - 24



Number of solutions per generation - 32

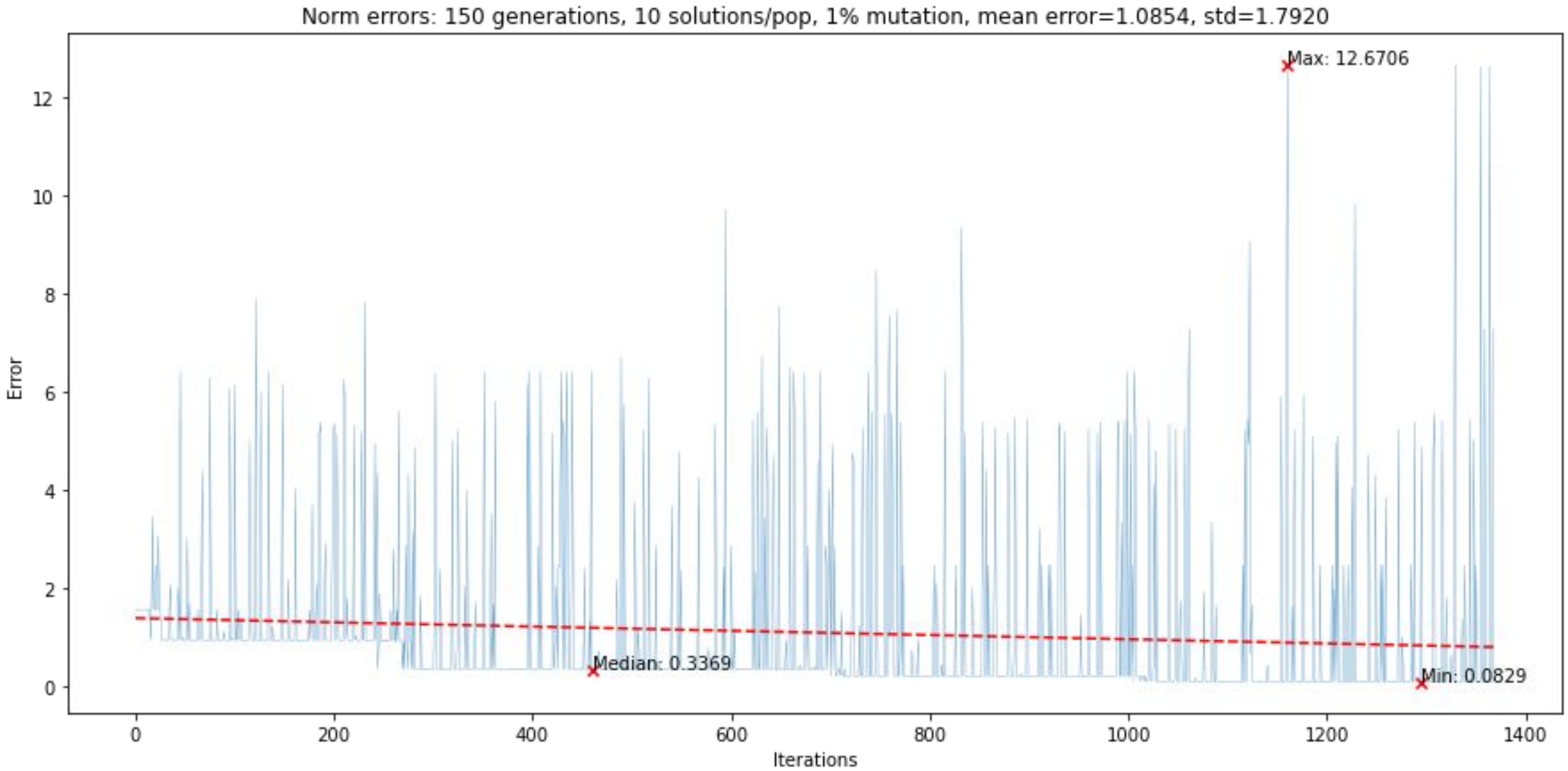


Mutation percentage

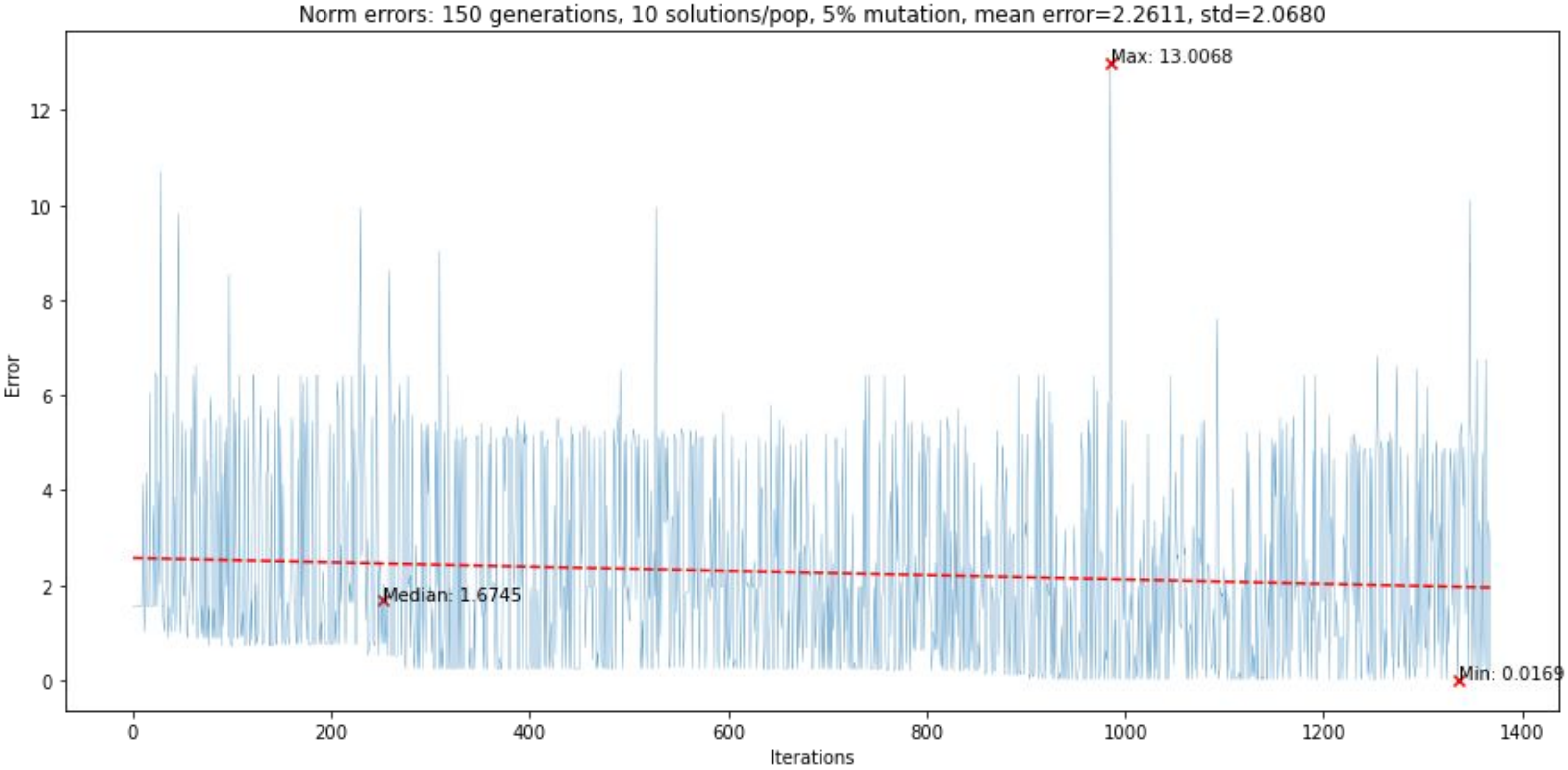
- Values to be tested: 1, 5, 10
- Number of generations: 150
- Number of solutions per generation: 10

Values	Min	Mean	Median	STD
1	0.0829	**1.0854**	**0.3369**	**1.7920**
5	**0.0169**	2.2611	1.6745	2.0680
10	0.0702	4.4281	4.3151	2.9707

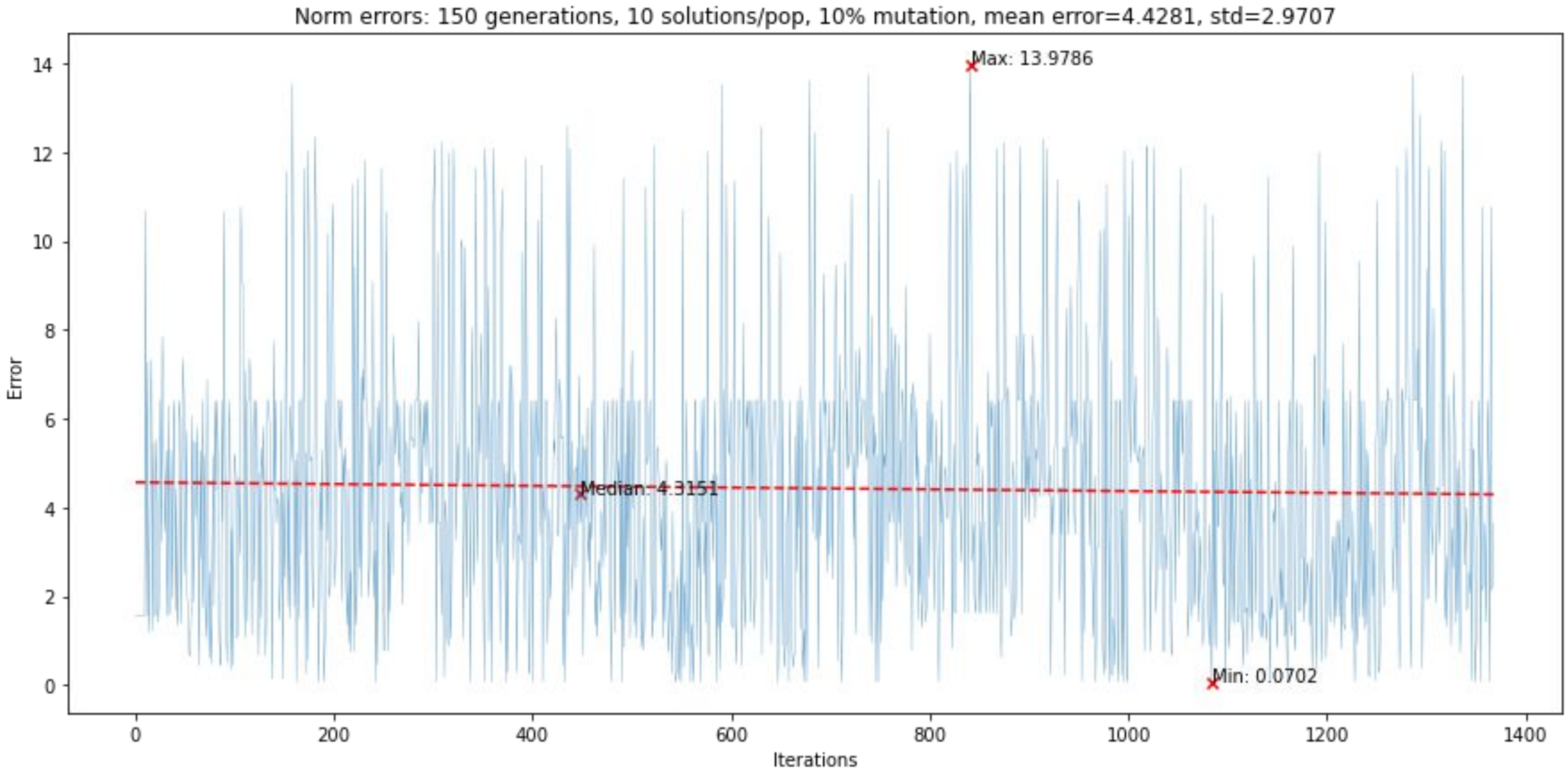
Mutation percentage - 1



Mutation percentage - 5



Mutation percentage - 10


Norm errors: 150 generations, 10 solutions/pop, 10% mutation, mean error=4.4281, std=2.9707
Max: 13.9786
Median: 4.3151
Min: 0.0702
Error
Iterations