Assignment Report 2

Name – Saankhya Subrata Mondal, Sr. No. 17990

DUCKWORTH – LEWIS METHOD

IMPLEMENTATION SUMMARY –

For data analysis, the first innings data for only those matches which weren't reduced due to rain were considered. The total number of data points totalled to 59,201. Firstly, data cleaning has been performed. It was observed that the entries of the 'Total.Runs' column which stored the cumulative sum of total runs scored were wrong. The wrong values in this column are replaced with correct values. After pre-processing, a final pandas dataframe was obtained which had the following columns - 'Match', 'Runs.Remaining', 'Wickets.in.Hand', 'Overs.Remaining', 'Innings.Total.Runs' and 59,201 rows.

For both the parts, the values of Z0 were initialized with mean of maximum runs obtained at that given wicket. The value of L was initialized to 15 and all the values of b were initialized to 0. Mean squared error was calculated over all the data points. A total of 10 functions were optimized in the first part and one function was optimized in the second part. L-BFGS-B technique was used for optimization in both the part.

RESULTS –

Part 1 – Run production function is given by

$$Z(u,w) = Z0(w)[1 - \exp\{-b(w)u\}]$$

w	1	2	3	4	5	6	7	8	9	10
Z0	14.53	31.47	56.96	93.39	118.75	151.23	183.98	226.57	259.33	304.39
b	0.726	0.268	0.192	0.107	0.085	0.071	0.057	0.047	0.0404	0.0338
MSE	115.5	222.0	399.2	550.0	721.2	921.2	1181.4	1507.1	2069.6	6120.1

Part 2 – Run production function is given by

$$Z(u,w) = Z0(w)[1 - \exp\{-Lu/Z0(w)\}]$$

$$L = 10.387$$

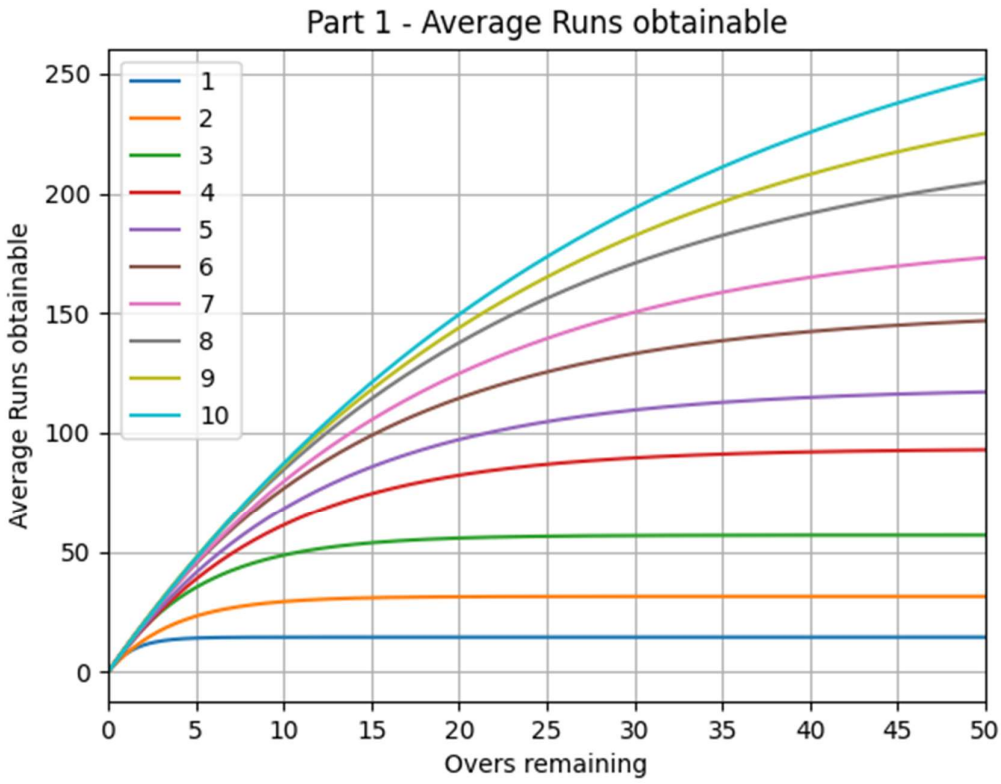
$$MSE = 5008.69$$

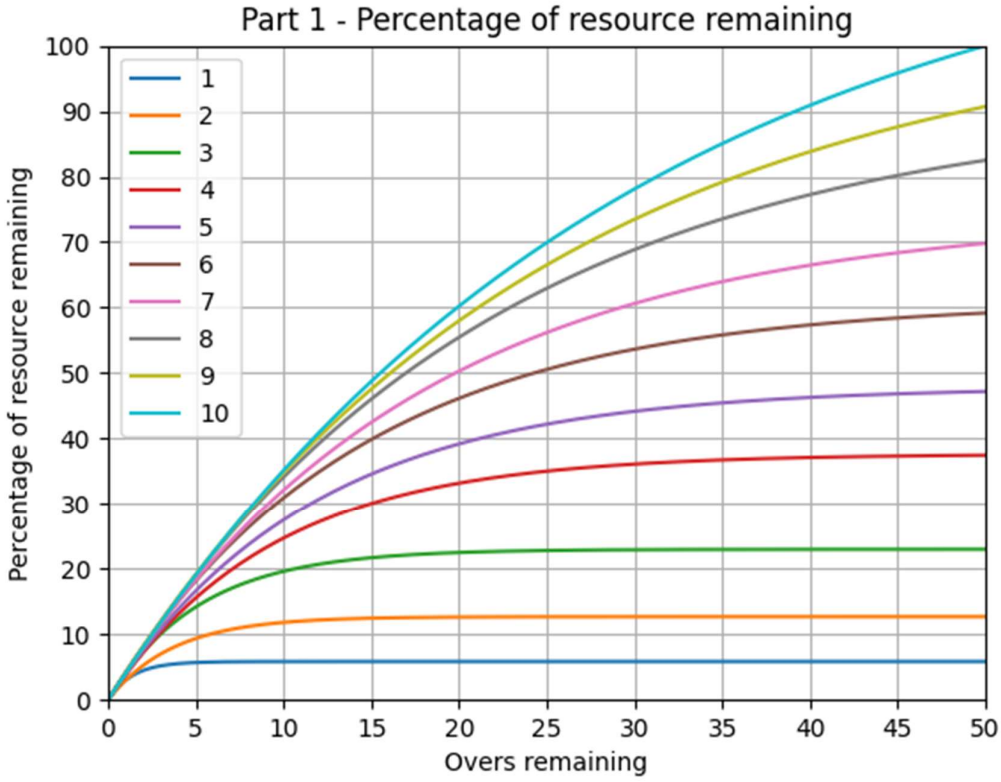
w	1	2	3	4	5	6	7	8	9	10
Z0	13.58	29.35	58.08	92.69	117.89	153.65	184.02	229.78	261.80	300.14

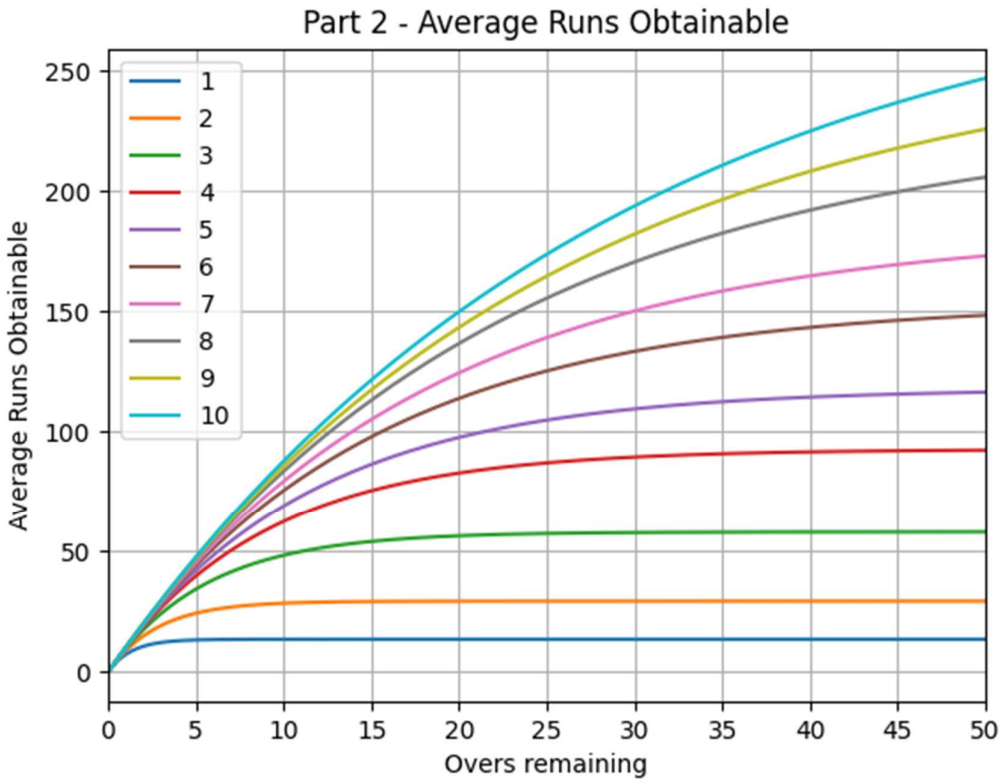
Slopes comparison – b is the slope in first part and Z0/L is the slope in second part

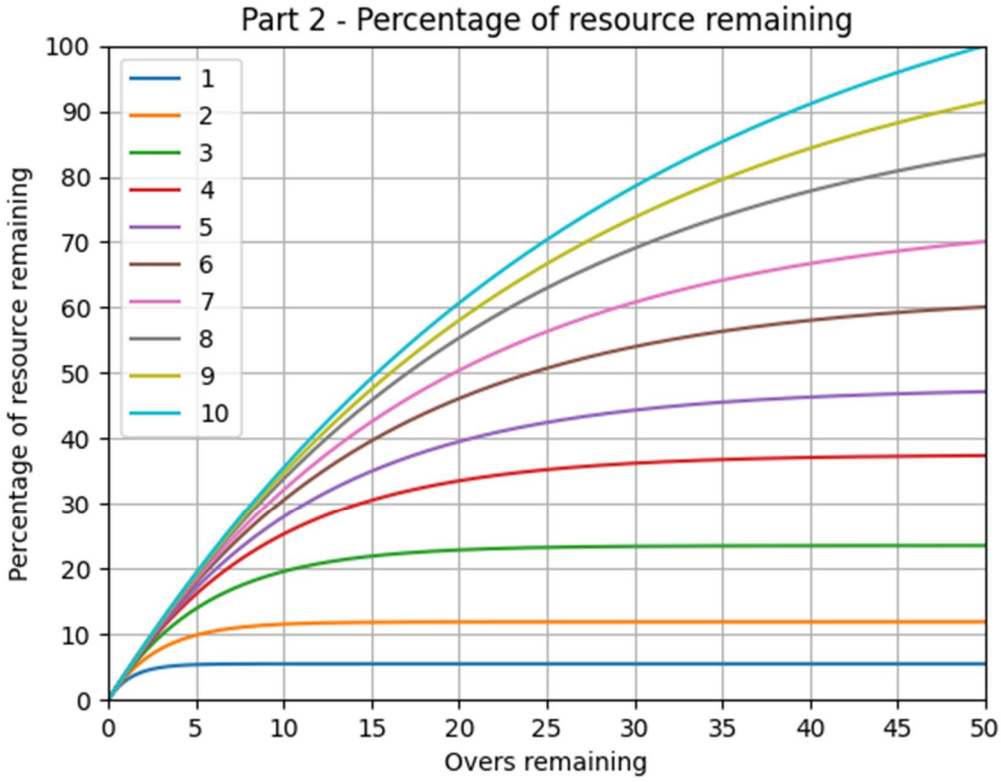
$$MSE = 0.0008997$$

w	1	2	3	4	5	6	7	8	9	10
Z0/L	0.764	0.354	0.179	0.112	0.088	0.067	0.056	0.045	0.040	0.0345
b	0.726	0.268	0.192	0.107	0.085	0.071	0.057	0.047	0.0404	0.0338


Part 1 - Average Runs obtainable
Average Runs obtainable
Overs remaining
0
50
100
150
200
250
0
5
10
15
20
25
30
35
40
45
50
1
2
3
4
5
6
7
8
9
10


Part 1 - Percentage of resource remaining
Percentage of resource remaining
Overs remaining
0
10
20
30
40
50
60
70
80
90
100
0
5
10
15
20
25
30
35
40
45
50
1
2
3
4
5
6
7
8
9
10


Part 2 - Average Runs Obtainable
Average Runs Obtainable
Overs remaining
0
50
100
150
200
250
0
5
10
15
20
25
30
35
40
45
50
1
2
3
4
5
6
7
8
9
10


Part 2 - Percentage of resource remaining
Percentage of resource remaining
Overs remaining
0
10
20
30
40
50
60
70
80
90
100
0
5
10
15
20
25
30
35
40
45
50
1
2
3
4
5
6
7
8
9
10