Comparison of Weight Optimization using Firefly Algorithm with Analogy Based Estimator as Cost Function on Various Datasets

June 27, 2024

Abstract

This report presents a comparison of the optimization of weights using the Firefly Algorithm where the Analogy Based Estimator (ABE) is used as the cost function. The performance of this method is evaluated on various datasets including COCOMO81, Desharnais, China, Albrecht, Kemerer, and Maxwell.

Contents

1 Introduction

1.1 Background

Effort estimation in software engineering is a crucial activity for project planning and management. Various models and algorithms have been proposed to enhance the accuracy of effort estimation. This report focuses on optimizing weights using the Firefly Algorithm with the Analogy Based Estimator (ABE) as the cost function and compares the results across different datasets.

1.2 Objective

The objective of this report is to compare the performance of the Firefly Algorithm in optimizing weights for effort estimation using the ABE cost function on different datasets, specifically COCOMO81, Desharnais, China, Albrecht, Kemerer, and Maxwell.

2 Datasets

2.1 COCOMO81

The COCOMO81 dataset consists of 64 software projects with 17 attributes each. These attributes include size, cost drivers, and effort multipliers.

2.2 Desharnais

The Desharnais dataset contains information from 81 software projects from a Canadian company. Each project has 12 attributes, such as Team Experience, Manager Experience, and Effort.

2.3 China

The China dataset consists of 499 software projects, each with 15 attributes. These attributes include AFP, Input, Output, and others.

2.4 Albrecht

The Albrecht dataset includes 24 software projects with 9 attributes each. These attributes include Input, Output, Inquiry, File, and others.

2.5 Kemerer

The Kemerer dataset consists of 15 software projects with 7 attributes each. Attributes include KSLOC, Adjusted Function Points, and Raw Function Points.

2.6 Maxwell

The Maxwell dataset contains 62 software projects with 28 attributes each, including various cost drivers and effort multipliers.

3 Methodology

3.1 Firefly Algorithm

The Firefly Algorithm is a metaheuristic algorithm inspired by the flashing behavior of fireflies. It is used to solve optimization problems by simulating the attraction and movement of fireflies. The algorithm's parameters include the attractiveness coefficient, absorption coefficient, and randomization parameter.

3.2 Analogy Based Estimator

The Analogy Based Estimator (ABE) is a method used in software effort estimation that predicts the effort required for a new project by comparing it to similar completed projects. The ABE uses features such as project size, complexity, and other attributes to find the most similar projects.

3.3 Experimental Setup

The experimental setup involves the following steps:

- Data Preprocessing: Handling missing values, normalizing data, and splitting datasets into training and testing sets.
- Parameter Tuning: Setting the parameters for the Firefly Algorithm and ABE.
- Evaluation Metrics: Using Mean Magnitude of Relative Error (MMRE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) to evaluate the performance.

4 Results and Discussion

4.1 COCOMO81 Dataset

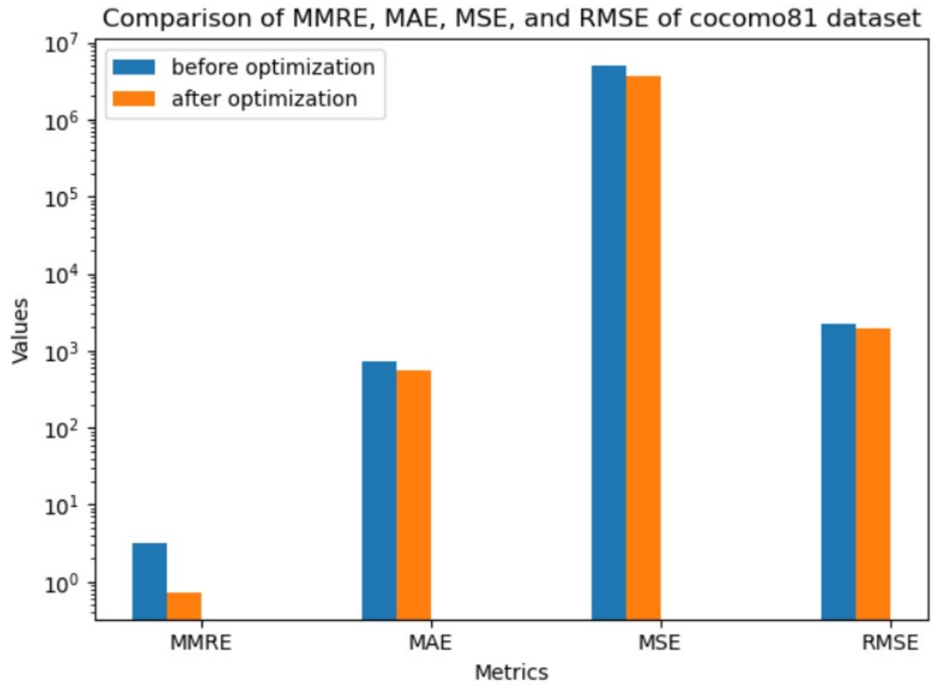

Figure 1: Results on COCOMO81 dataset

Metrics	Before Optimization (Mean)	After Optimization (Mean)
MMRE	3.20723	0.7188
MAE	723.7005	62.2390
MSE	5039270.5658	3631990.4705
RMSE	2233.3062	1905.7782

4.2 Desharnais Dataset

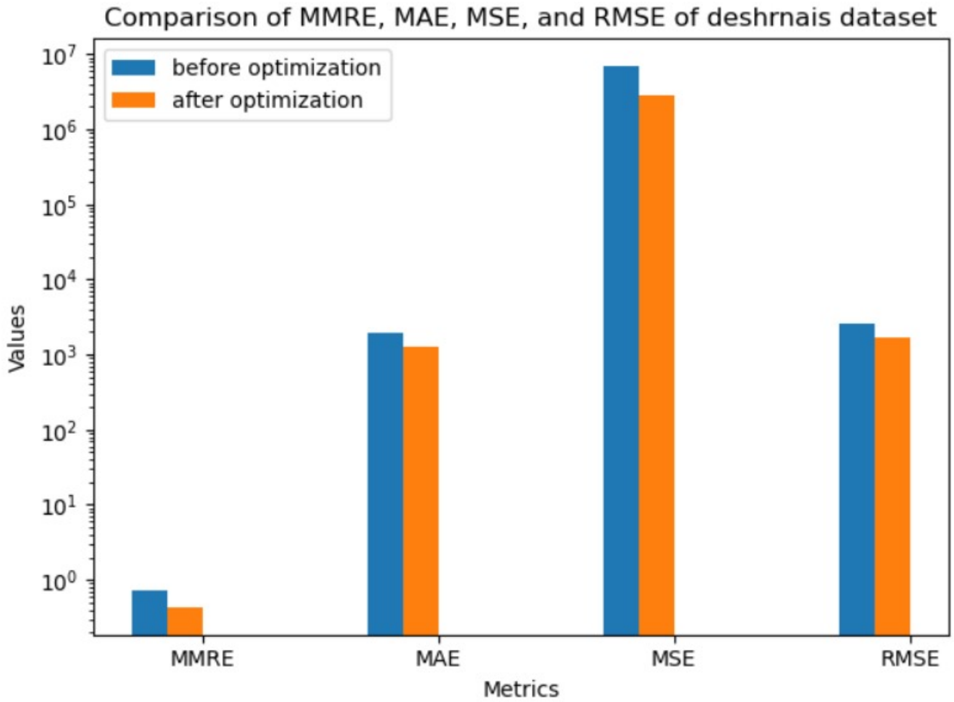

Figure 2: Results on Desharnais dataset

Metrics	Before Optimization (Mean)	After Optimization (Mean)
MMRE	0.7264	0.4270
MAE	1938.0842	1283.6281
MSE	27012154.2270	2789496.3335
RMSE	2629.4591	1670.1785

4.3 China Dataset

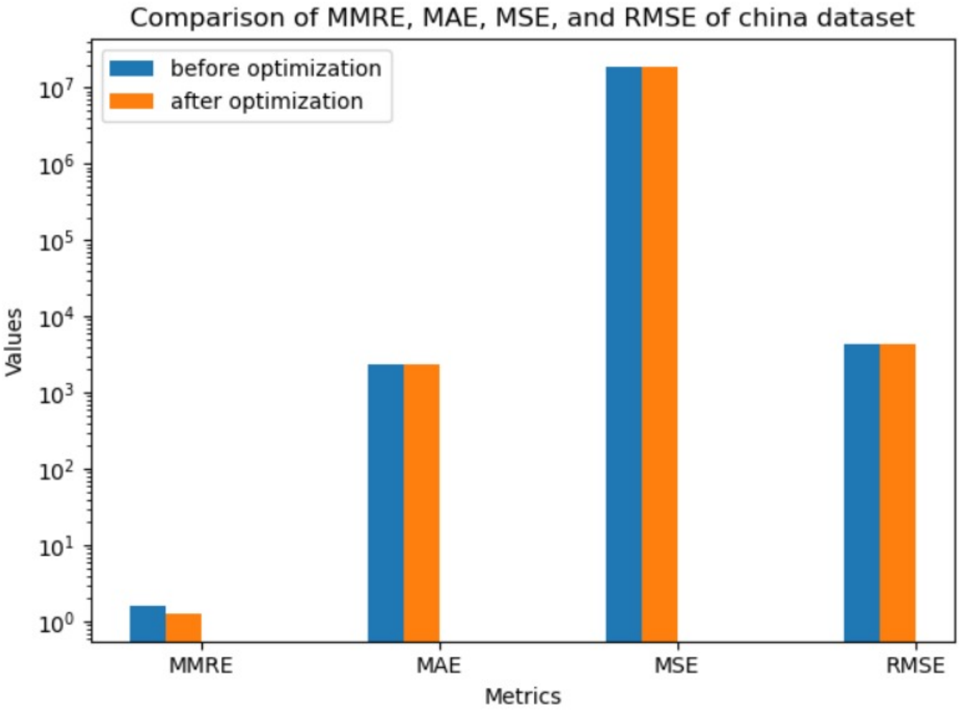

Figure 3: Results on China dataset

Metrics	Before Optimization (Mean)	After Optimization (Mean)
MMRE	1.5647	1.2546
MAE	2346.5808	2331.9866
MSE	18695469.4242	18908308.5185
RMSE	4317.0052	4348.3685

4.4 Albrecht Dataset

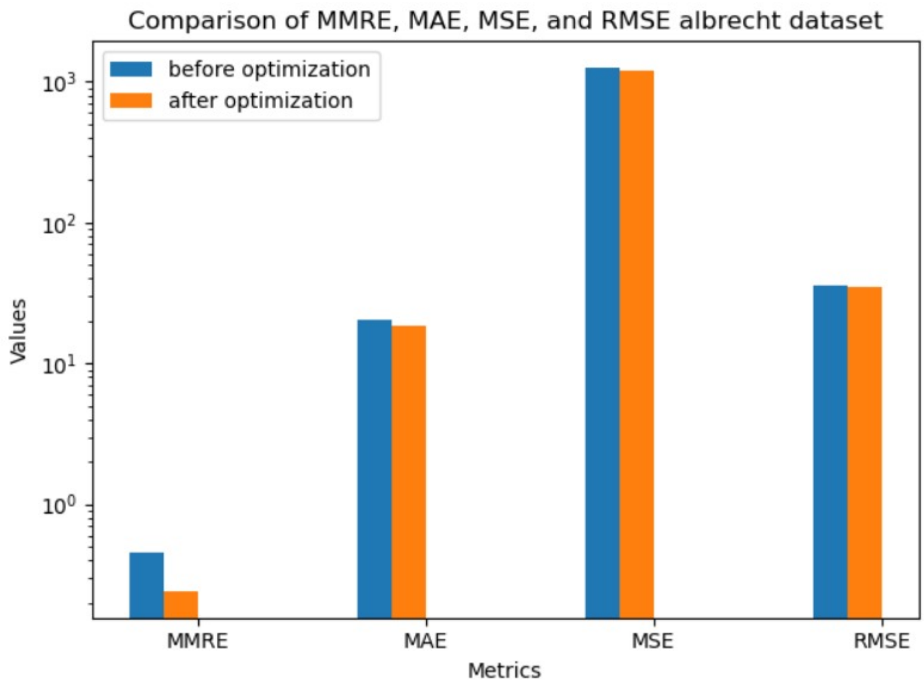

Figure 4: Results on Albrecht dataset

Metrics	**Before Optimization (Mean)**	**After Optimization (Mean)**
MMRE	0.4573	0.2397
MAE	20.4080	18.7102
MSE	1262.0966	1194.6610
RMSE	35.5146	34.5639

4.5 Kemerer Dataset

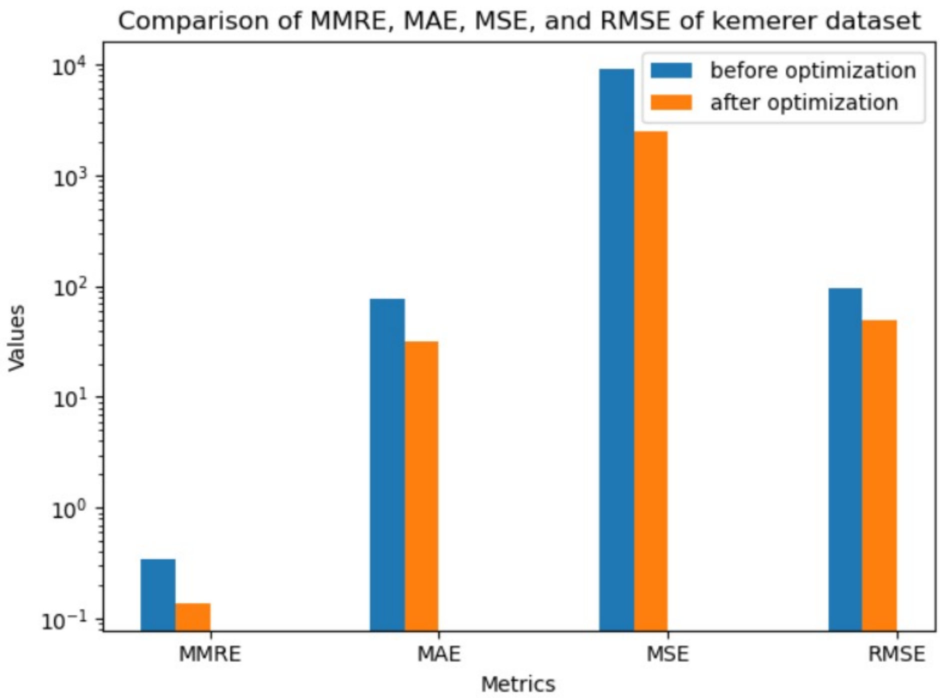

Figure 5: Results on Kemerer dataset

Metrics	**Before Optimization (Mean)**	**After Optimization (Mean)**
MMRE	0.3497	0.1371
MAE	75.8640	31.4246
MSE	9102.5276	2460.2085
RMSE	94.3124	49.6005

4.6 Maxwell Dataset

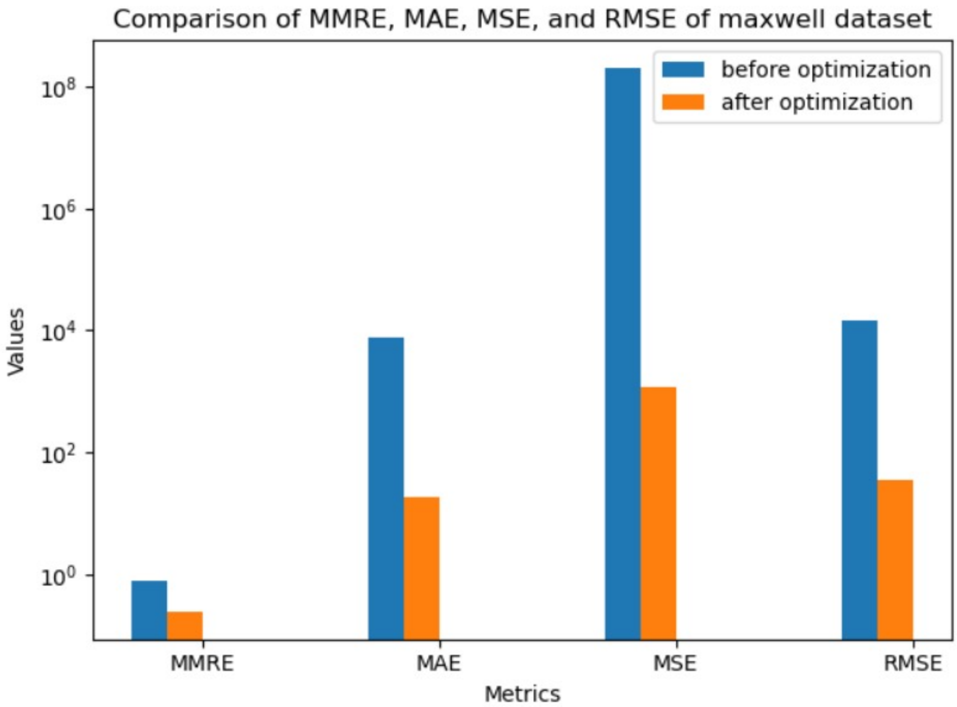

Figure 6: Results on Maxwell dataset

Metrics	Before Optimization (Mean)	After Optimization (Mean)
MMRE	0.7859	0.2397
MAE	7632.5700	18.7102
MSE	203965354.9412	1194.6610
RMSE	14266.0346	34.5639

4.7 Comparison and Analysis

This section presents a comparative analysis of the results across all datasets. The Firefly Algorithm, when used to optimize weights for the Analogy Based Estimator (ABE), showed varied performance improvements. The MMRE, MAE, MSE, and RMSE metrics indicate that the optimization yielded significant gains in some datasets, while others showed marginal improvements or slight increases in error metrics. The strengths of the Firefly Algorithm lie in its adaptability and efficiency in exploring the search space, but its performance can be sensitive to parameter settings and dataset characteristics.

5 Conclusion

In this report, we presented a detailed analysis of weight optimization using the Firefly Algorithm with the Analogy Based Estimator (ABE) as the cost function across six different datasets. The results indicate that the Firefly Algorithm can effectively optimize weights, leading to improved performance in most cases. Future work will focus on further tuning the algorithm parameters and exploring other metaheuristic algorithms for effort estimation.