The Effect of Number in Population and Differential Weight Parameter Values on Parameterised Response Differential Evolution Trading Strategy Performance

Karolina Kadzielawa
Department of Computer Science
University of Bristol
Bristol, England
zo19826@bristol.ac.uk

***Abstract*—Parameterised Response Differential Evolution (PRDE) is an adaptive trading strategy. Traders who employ this method each maintain a population of trading strategies, and use differential evolution to improve those strategies over time [1]. This report aims to explore the impact of *k* and *F* on PRDE's performance, where *k* denotes the number in population - the number of private candidate solutions held by each trader, and *F* is the the differential weight - a coefficient determining the extent to which the difference between two candidate solution vectors is mixed in with a base vector. Using BSE, a simulation limit-order-book financial exchange, we test a number of parameter configurations under a set of different market conditions including homogeneous and heterogeneous trader populations and static and dynamic price equilibria. Having conducted rigorous statistical analysis of the the experiment results we are not able to conclusively determine a superior value for either *k* or *F*.**

***Index Terms*—Automated Trading, Financial Markets, Adaptive Trader-Agents, Optimization, Differential Evolution**

I. Introduction

Parameterized-Response Differential Evolution traders maintain and update a population of candidate solutions using differential evolution - a population-based metaheuristic search algorithm. The differential evolution variant implemented in PRDE is specifically DE/rand/1 - the simplest form of the algorithm. BSE's implementation of differential evolution in PRDE works as follows: first the candidate solutions are shuffled, then a new candidate solution is created by mixing the second solution with the difference between the third and fourth solutions, multiplied by a coefficient known as the differential weight. If the new candidate solution outperforms the current first solution it is chosen to be the new first solution and the cycle repeats.

There are two parameters in this version of the algorithm which ought to be optimized for peak trader performance: the number of candidate solutions, often referred to as *NP* – for "Number in Population" and denoted with k, and the differential weight F. The standard differential evolution algorithm also makes use of *CR* - the crossover probability.This parameter however remains undefined in PRDE as the candidate solution being optimised is represented by a single number instead of a string or vector of real numbers. While in the original paper introducing PRDE by Cliff [1] they have initially been set to 4 and 0.8 respectively, there may exist a combination of the two that yields superior performance of the algorithm. A number of studies have explored the choice of the parameter values, often concluding with a general range or starting point recommendation (Storn [2] suggested that $k = 5D$–$10D$ and $F = 0.5$ to be sufficient for obtaining an optimal solution, where D denotes the dimension of the problem, while Gämperle et al. [3] recommended $k = 3D - 8D$, $F = 0.6$ for example). However there appears to be no research on how those values could affect PRDE performance specifically.

II. Experiment Setup

A. Bristol Stock Exchange

The experiments are conducted using a simulation of a limit-order-book financial exchange, the Bristol Stock Exchange (BSE) [4]. BSE, while a fairly minimal simulation of a real world financial market, has been used as a platform for research before [5], [6]. BSE allows to modify the supply and demand schedules a swell as the trader populations, which allows for simulating a continuous double auction with a sufficient level of control and realism.

In order to conveniently conduct the experiment some minor changes had to be made to the BSE code. In the original code the value of F is hard coded. It was made to be provided as a parameter together with the bounds on randomly-assigned strategy-value and k to the PRDE trader instead.

B. Experiment conditions

The experiment consists of a number of market simulations in four distinct sets of conditions: a static market homogeneously populated with PRDE traders, a static market with a mixed population including 5 PRDE traders, a trending market with the same population composition as the static heterogeneous one, and a market with a rapidly changing equilibrium, again with the same mixed population of traders. In all the experiment setups the total number of traders equals to 60 and is symmetrical, similarly to the original study on PRDE performance [1]. In all the experiments the PRDE traders' k and F parameter values change throughout the experiment, but at any given time the all the PRDE traders in the simulation share the same parameter values. The k ranges from 4 to 1024 in powers of 2. The choice of the upper bound is informed by Alić et al. [7] finding $NP = 400$ to yield the best performance. Piotrowski [8] advises against setting Differential Evolution population size lower than 50 individuals even for low dimensional problems as it may prevent convergence. We still choose to consider lower values of k since convergence is not the desired outcome in this context [1] and PRDE uses nonconvergent DE specifically. The F lies between 0 and 2 in 0.2 intervals.

Each combination is tested 10 times to account for the stochastic nature of the simulation. Each market session simulation lasts five minutes.

The static homogeneously populated market experiment is mostly conducted for validation and reference, for this reason the environment is simpler than in the subsequent experiments. The sessions implement a periodic order schedule (`'timemode'`) and fixed spacing in the price distribution (`'stepmode'`) for reduced randomness and noise. The supply and demand ranges both lie within $[100, 200]$.

The heterogeneously populated market sessions are meant to simulate PRDE performance in a real world trading exchange. The sessions implement a Poisson distribution order schedule (`'timemode"'`) like a real market [9], but still fixed spacing in the price distribution (`'stepmode'`). The trader population consists of BSE's implementation of 6 trading algorithm types, 5 of each. Those types are: PRDE, Zero-Intelligence Plus (ZIP) [10], Zero-Intelligence Constrained (ZIC) [11], Giveaway (GVWY), Sniper (SNPR) [12] and Shaver (SHVR) [13].

Additionally, in order to examine PRDE performance in a trending market the a similar environment is used, with the only difference being the supply and demand ranges changing over time. The supply and demand range can be represented as $[0.1 \times t + 100, 0.1 \times t + 200]$.

Finally, the experiment is conducted in a market with suddenly changing supply and demand curves, simulating market shocks and allowing to investigate k and F's impact on PRDE's adaptability. The supply and demand schedules simulate rapid changes in the equilibrium, where the price ranges from $[0.1 \times t + 100, 0.1 \times t + 200]$ in the first and last 2 minutes, and jumps to $[300, 400]|$ in the 1 minute in between.

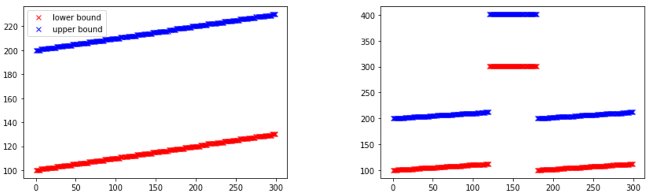

Fig. 1: Supply and demand range by simulation time for the continuously changing market on the left and for a volatile market on the right

III. Experiment Results

The following section contains the processed data gathered from the experiments described in the previous section, as well as the analysis of those results. The main metric by which the performance was assessed was average total profit per PRDE trader. The choice of the metric is motivated mostly by simplicity. We do not have to worry about any differences in expected profit between the market scenarios, as we only compare the strategies with different parameter values within the same market scenario. If we were to do a comparison between trading strategies in different environments a metric like allocative efficiency would have been more informative.

The box-and-whisker plots included in this section have been plotted for both a second parameter fixed to one value as well as for average profits across all trials. Those fixed values are $k = 4$ and $F = 0.8$, same as in the original study introducing PRDE [1]. The aim of including those plots is to visualise the isolated impact of only one of the parameters. Visual analysis of the same type of plots for all the other k and F values, which are not included in this section, has informed the analysis and conclusions as well.

A. Static Homogeneously Populated Market

The best performing combination of parameter values in these experiment conditions is $k = 64$ and $F = 0.6$. Those values are close to the common practice of keeping a fixed 50-100 population size and starting out with the differential weight close to 0.5. Table I shows the 5 highest mean profit values from all 10 trials.

k	F	profit
64	0.6	411.15
128	2.0	406.08
32	0.4	400.06
32	1.1	399.70
8	0.6	394.52

TABLE I: The parameter values yielding the highest average profit per trader

The next best performing values tell a similar story, with most of the values relatively close to those recommendations as well.

Upon initial visual inspection, the average profit per trader appears not to vary to any significant degree with changes in the value of k. As seen on the box-and-whisker plots in Fig.2 the inter-quartile ranges for all of the values of k overlap, meaning no clear "winner" can be determined. That remains to be the case even when we plot box-and-whisker plots for all values of F separately, like we did for $F = 0.8$ in the left half of Fig.2. Therefore there appears not to be a superior choice of k even in one of the narrowed conditions of a particular fixed F value.

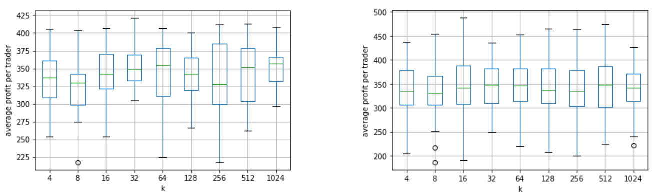

Fig. 2: Average profit per trader for different k values where $F = 0.8$ on the left and over all F values on the right

A similar conclusion can be drawn from visualising the same for F. While $k = 4$ and $F = 1.1$ visually appears to yield superior results as seen on Fig.3, its inter-quartile range still overlaps with that of $F = 1.7$. What is more this tendency does not hold for other k values.

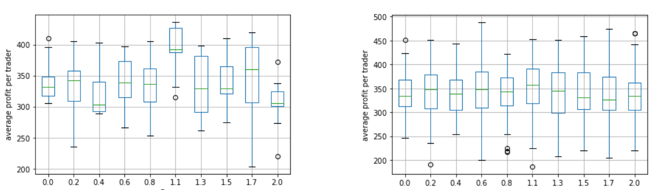

Fig. 3: Average profit per trader for different F values where$k = 4$ on the left and over all k values on the right

B. Static Heterogeneously Populated Market

Next we evaluated the performance of PRDE traders in a market populated with multiple types of traders and a static price equilibrium. The highest profit was accumulated by traders with $k = 4$ and $F = 1.5$. As seen in Table II, 1.5 is a common value of F among the best performing strategies. Therefore the best performing parameter combinations in a heterogeneous market seem to have little in common with those in a homogeneous one.

Upon further analysis however, the presence of other types of traders does not appear to have let traders with any particular choice of parameter values conclusively outperform others. Both the box-and-whisker plots illustrating the relation of profit to k (Fig.4) and F (Fig.5) appear similar to the homogeneous market outcomes (Fig.2, Fig.3) in the sense that while for fixed values of the other parameter some values might yield slightly higher profits than others, when all the values are included those differences disappear.

k	F	profit
4	1.5	547.33
64	1.5	486.11
128	0.8	474.02
128	1.5	464.81
256	1.5	458.39

TABLE II: The parameter values yielding the highest average profit per trader

What needs to be kept in mind when comparing the homogeneous market results to the heterogeneous ones is that the aforementioned plots are not exactly equivalent. For the homogeneous market the average profit per trader is taken from 60 traders - the whole market, making it more akin to a representation of overall economic efficiency of the market. Meanwhile the average profit per trader in a market where only $1/6$ of the population is comprised of PRDE traders tells us more about how the parameters affect their ability to compete with other traders, leaving out the information about overall economic efficiency of the market. We choose to focus on the latter, therefore the dynamic equilibrium tests are conducted in a heterogeneous environment as well instead of a homogeneous one.

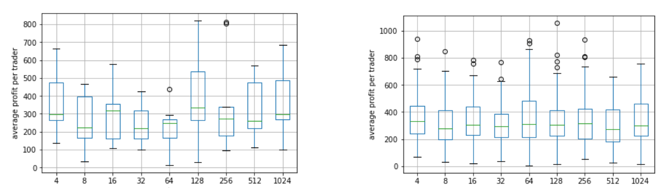

Fig. 4: Average profit per trader for different k values where $F = 0.8$ on the left and over all F values on the right

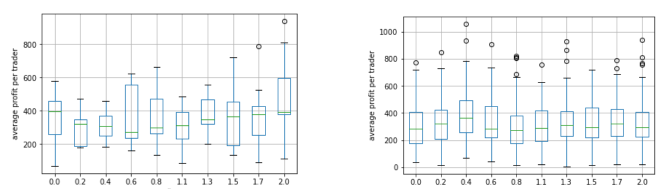

Fig. 5: Average profit per trader for different F values where $k = 4$ on the left and over all k values on the right

C. Trending Heterogeneously Populated Market

k	F	profit
256	0.0	536.77
8	1.5	478.89
1024	1.5	433.36
8	0.6	428.41
1024	0.6	427.48

TABLE III: The parameter values yielding the highest average profit per trader

Table III containing the best performing parameter values in a trending heterogeneously populated market shows little consistency with the results in a static market and is topped by $k = 256$ and $F = 0.0$. The plots in Fig.6 and Fig.7 however are nearly identical to Fig.4 and Fig.5, meaning that again no value combination appears to give traders' a marked advantage.

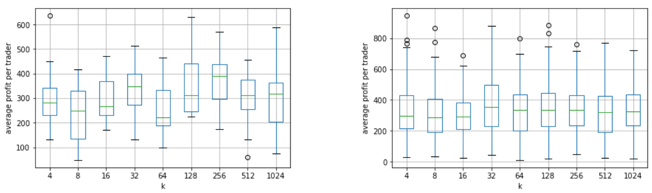

Fig. 6: Average profit per trader for different k values where $F = 0.8$ on the left and over all F values on the right

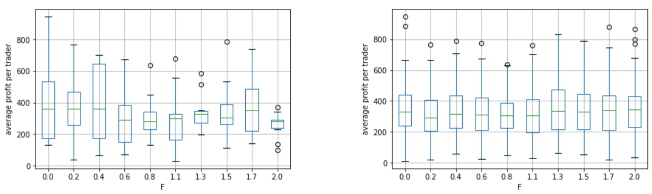

Fig. 7: Average profit per trader for different F values where $k = 4$ on the left and over all k values on the right

D. Volatile Heterogeneously Populated Market

The last set of tests was meant to simulate market shocks and involved running the simulation with a rapidly changing price equilibrium. The results seem just as if not even more inconclusive, as the average profit across all the k and F values evaluated appears even more even than in the previously implemented sets of conditions.

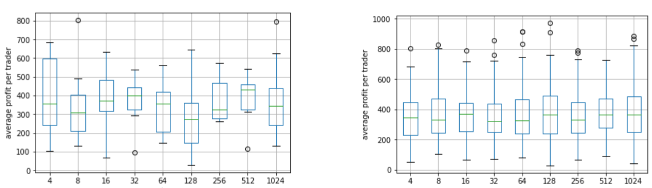

Fig. 8: Average profit per trader for different k values where$F = 0.8$ on the left and over all F values on the right

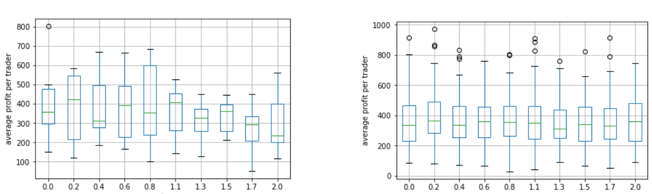

Fig. 9: Average profit per trader for different F values where $k = 4$ on the left and over all k values on the right

The best performing parameter values in these conditions were $k = 1024$ and $F = 1.3$. Again, very far off from the top value combinations in all the other tested scenarios. The lack of consistency is not completely unexpected. The environment is different for each experiment after all. However paired with how narrow the gap between the profits made by traders with different k and F value combinations is, as well as the presence of a degree of randomness in the simulation and the fact that we are taking averages from only 10 trials for each combination it does not seem unlikely that those differences are a result of pure chance.

k	F	profit
1024	1.3	541.72
32	0.2	509.57
16	0.0	501.32
256	1.7	487.45
256	0.6	484.02

TABLE IV: The parameter values yielding the highest average profit per trader

IV. Further analysis of the experiment results

For all of the market scenarios the best performing parameter combinations were different and their advantage did not appear to be statistically significant. In order to rigorously evaluate that statement it is paramount that we conducted further statistical analysis of the results.

First we test whether the data from each experiment setup conforms to the normal distribution using Shapiro-Wilk. For the static homogeneous market test the profits of 83 value combinations proved to approximately follow the normal distribution, 7 did not. In the static heterogeneous market the numbers were 69 for normal and 21 for not normal. In the trending market experiment the values were 75 and 15 and in the volatile environment the ratio was 78 to 12. While the vast majority of average profit per trader values for different parameter value combinations adhere to the normal distribution, we cannot assume normality for all the data. Therefore for hypothesis testing we decide to opt for Kruskal-Wallis - a rank-based nonparametric test. The hypoteses for the test are:

- *H0*: population medians are equal.
- *H1*: population medians are not equal.

If the $p - value > 0.05$ we cannot reject the null hypothesis. In that case none of the parameter combinations tested in the experiment could be considered to yield superior PRDE performance. Table V shows the p-values for all market scenarios obtained by grouping the profits by k and F.

Looking at Table V it is clear that neither of the parameters on its own has a statistically significant impact on the profits obtained by the traders. Since all the p-values are significantly higher than 0.05, all the groups with different k and F parameter values have equal population medians.

market conditions	by k	by F
Static, homogeneous	0.73	0.43
Static, heterogeneous	0.68	0.30
Trending, heterogeneous	0.47	0.38
Volatile, heterogeneous	0.73	0.74

TABLE V: p-values for all experiment setups obtained by grouping the profits by k and F

V. Conclusions

In the original paper introducing PRDE [1] the values of k and F are fixed. In this report we have concluded this was unlikely to have a significant impact on the traders' performance, as we failed to find superior values for neither the differential weight nor the number in population in any of the market conditions tested.

A. Limitations and further work

While our analysis has shown a a lack of statistically significant advantage of any of the parameter combinations, there is a number of steps that could be taken to explore the problem in more depth and possibly come to a different conclusion.

One of the most obvious shortcomings of the study is a possibly insufficient dataset size. Many studies on algorithmic trader-agent performance repeat their experiments 100 [14], 2500 [15] times in order to achieve stronger results. 10 could have very likely been insufficient to obtain meaningful results. What is more, the simulation time is very short due to the limited computation power and time available. Some experiments simulate as much as multiple days of market activity [16] and had we done it as well our results might have been more insightful.

Furthermore, the range of conditions under which the traders' performance was tested was very limited. It could be that certain k and F parameter values are advantageous only in a particular environment or situation. One of the examples of extending the experiment would be adding more market shocks to the rapidly changing price equilibrium scenario. Most trading strategies can adapt to a single shock so in order to evaluate the algorithm more thoroughly it might be beneficial to test it in a more challenging environment, as suggested by Snashall and Cliff [17].

Finally, this study does not implement the methods suggested by Cliff [1], namely populating the market with PRDE traders with heterogeneous distributions of k and F or replacing the DE/rand/1 by other instances of DE such as DE/best/1, JADE or SHADE.

Taking all this into account this exploration could serve as a starting point for further work. It shows that the impact of the choice of k and F values is not immediately apparent and might require either more resources and time or more sophisticated simulation and analysis techniques to be meaningfully explored.

References

[1] D. Cliff, ”Metapopulation Differential Co-Evolution of Trading Strategies in a Model Financial Market”, 2022

[2] R. Storn, ”On the usage of differential evolution for function optimization,” *Proceedings of North American Fuzzy Information Processing*, 1996, pp. 519-523, doi: 10.1109/NAFIPS.1996.534789.

[3] R. Gaemperle, S. D. Mueller, and P. Koumoutsakos, “A Parameter Study for Differential Evolution,” in *Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation*, 2002, pp. 293–298.

[4] D. Cliff, *Bristol Stock Exchange: open-source financial exchange simulator*. `https://github.com/davecliff/BristolStockExchange`, 2012–2022

[5] H. Hanifan,, J. Cartlidge. ”Fools Rush In: Competitive Effects of Reaction Time in Automated Trading.”, *arXiv preprint arXiv:1912.02775*, 2019.

[6] D. Cliff, M. Rollins. “Methods Matter: A Trading Agent with No Intelligence Routinely Outperforms AI-Based Traders.” *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020.

[7] A. Alić, K. Berkovič, B. Bošković, J. Brest, ”Population Size in Differential Evolution.”, *Communications in Computer and Information Science* vol 1092, 2020.

[8] A. Piotrowski, “Review of Differential Evolution population size.” *Swarm Evol. Comput. 32:* 1-24, 2017

[9] M.B. Garman, ”Market Microstructure”, *Journal of Financial Economics*, 3, 257-275, 1976.

[10] D. Cliff, J. Bruten, ”Minimal-Intelligence Agents for Bargaining Behaviors in Market Based Environments” *Technical Report HPL-97-91*, 1997

[11] D.K. Gode, S. Sunder,”Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality,” *Journal of Political Economy, University of Chicago Press*, vol. 101(1), pages 119-137, 1993.

[12] J. Rust, J. H. Miller, R. Palmer, “Behavior of Trading Automata in a Computerized Double Auction Market” in *The Double Auction Market:Institutions, Theories, and Evidence*. pp.155—198, 1993.

[13] D. Cliff, ”BSEguide1.2e”, Accessed: 2022. [Online]. Available: `https://raw.githubusercontent.com/davecliff/BristolStockExchange/master/BSEguide1.2e.pdf`

[14] B. Liu, ”Learning the Trading Algorithm in Simulated Markets with Non-stationary Continuum Bandits”, 2022.

[15] W. Walsh, R. Das, G. Tesauro, J. Kephart. ”Analyzing complex strategic interactions in multi-agent systems.”, *AAAI-03 Workshop on Game Theoretic andDecision Theoretic Agents*, pages 109–118, 2002.

[16] P. Vytelingum,D. Cliff, N.R. Jennings, ”Strategic bidding in continuous double auctions”. *Artificial Intelligence*, 172 (14), 1700-1729, 2008

[17] D. Snashall, D. Cliff, ”Adaptive-aggressive traders don't dominate”. *ICAART 2019. LNCS (LNAI)*, vol. 11978, pp. 246–269, 2019