Erroneous Truncation Selection – A Breeder's Decision Making Perspective

Hans-Michael Voigt[1][*] and Heinz Mühlenbein[2]

[1] GFaI, Rudower Chaussee 5, 12484 Berlin, Germany,
[2] GMD – Forschungszentrum Informationstechnik, 53754 Sankt Augustin, Germany

Abstract. Based on experiences from livestock breeding we introduce erroneous truncation selection for the Breeder Genetic Algorithm (BGA). The decision behavior of the breeder is given by a simple model. It is shown that there is no benefit to the BGA by using erroneous selection though the variance of the parent population is increased by increasing the decision error variance.

1 Introduction

Selection is one of the most fundamental operators in Evolutionary Algorithms, i.e. it is inherent to Genetic Algorithms, Evolution Strategies, Evolutionary Programming, and Genetic Programming. Generally, for all these algorithms a directed selection scheme is used which favors individuals having smaller – for minimization – or larger – for maximization – fitness values. With respect to Evolutionary Algorithms we will use the notion of fitness, trait, and character synonymously to characterize one measurable variable or function. Corresponding to the great variety of algorithms we also have a great variety of selection operators. This ranges from proportionate selection [5], tournament selection [6, 2], (μ, λ)–selection [12, 13, 1], Boltzmann selection [9], linear and exponential ranking selection [4, 17] to truncation selection [10, 3, 8].

Perhaps the most widespread used selection schemes in modern Evolutionary Algorithms are tournament selection and truncation or (μ, λ) selection.

With this paper we consider selection from a breeder's perspective. The ultimate selection scheme in livestock breeding is truncation selection. *The outward results of selection to a population* [8] is given by the *response to selection*

$$R(t) = \overline{f(t+1)} - \overline{f(t)}. \tag{1}$$

The amount of selection is measured by the selection differential

$$S(t) = \overline{f_s(t)} - \overline{f(t)}, \tag{2}$$

where $\overline{f_s(t)}$ is the average fitness of the selected parents. The equation for the response to selection relates R and S by

$$R(t) = b(t) \cdot S(t). \tag{3}$$

---

[*] HMV is also with the Technical University of Berlin

$b(t)$ is called the *realized heritability*. For many fitness functions and selection schemes the selection differential can be expressed as a function of the phenotypic standard deviation σ. For truncation selection (selecting the $T \cdot N$ best individuals) one obtains

$$S(t)/\sigma(t) = I. \tag{4}$$

I is called the *selection intensity*. With these notions the famous equation for the response to selection is obtained [3]:

$$R(t) = I \cdot b(t) \cdot \sigma(t). \tag{5}$$

Usually the response to selection equation is considered for a population having a normal fitness distribution. But it holds also for other unimodal fitness distributions [14, 16]. It should be noted that the response to selection equation does not explain *how selection changes the genetic composition of a population* [8].

Truncation selection in real breeding is hard to realize because of the precise truncation threshold which implies a perfect measurement of the trait under consideration. Therefore, breeders are well aware of possible errors in selection: *The kind of selection pictured in Fig. 1 (left) corresponds to that actually practiced for important traits in stock breeding where many different traits must be considered.* **Some animals which are mediocre or even inferior in the characteristic pictured are saved** *because they are unusually desirable in several other characteristics or* **because the breeder is careless or confused** [8].
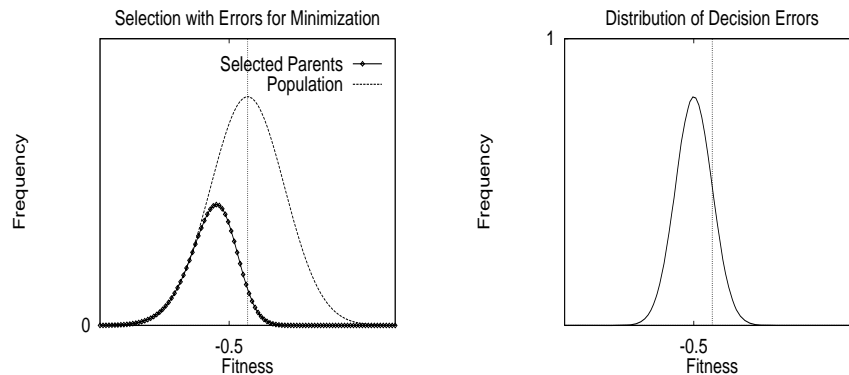

Fig. 1. Parent population selected from a population with erroneous truncation selection (left) and distribution of decision errors made by a breeder around the truncation threshold $f_T = -0.5$ (right)

In this paper we consider only one trait which is measured, and therefore selected with errors. The case of considering more then one trait is closely con-

nected with multiobjective and/or subjective decision problems and will be analyzed in a forthcoming paper. A usual way breeders cope with such problems is the application of selection indices. Furthermore, we assume that selection is done without replacement.

The questions to be raised are: How does selection errors influence the behavior of an Evolutionary Algorithm? How can the decision behavior of a breeder be modeled? Is the response to selection equation sufficient for predicting the convergence of an Evolutionary Algorithm?

To give some first answers to these questions we proceed in the following way: Based on a very simple but comprehensive model for erroneous breeder decisions we present a general method for computing the selection intensity and the standard deviation of a parent population with respect to decision errors. This will be compared for selected fitness functions with empirical observations from simulation. Based on these results first conclusions concerning decision errors in truncation selection are presented.

2 Truncation Selection with Errors

The analysis of truncation selection with errors has to be based on a decision model for the breeder. As a first attempt we use the following model:

Definition 1 (Breeder decision model). *The decision behavior of a breeder is given by a probability density function (pdf) $p_e(f)$ with the expectation $E(f) = f_T$ and the variance $V(f) = \sigma_e^2$. f_T is the truncation threshold and σ_e^2 the variance of the decision or selection error.*

This model implies that decision errors of the breeder are most probable at the truncation threshold. That is a rational behavior because at this point it is most difficult to decide whether an individual will be saved as a parent or not. The shrewdness of the breeder, or the precision of measurement of the considered trait, is given by the selection error variance. The smaller this variance is the less is the decision error. For $\lim \sigma_e \rightarrow 0$ we have a one point distribution at the truncation threshold f_T which corresponds to the usual truncation selection. An example for such a decision model is shown in Fig. 1 (right) for a truncation threshold $f_T = -0.5$.

Obviously, the selection scheme will be influenced by the variance of the selection error. Having a population with a fitness probability density function $p(f)$ with expectation $E(f) = \overline{f}$ and variance $V(f) = \sigma^2$ we will analyze the dependency of the selection intensity $I = \overline{f}_p$ and the standard deviation of the selected parents σ_p on the decision error standard deviation σ_e.

We consider the standardized and normalized probability density function of the population fitness distribution which is given by $p(z)$ with
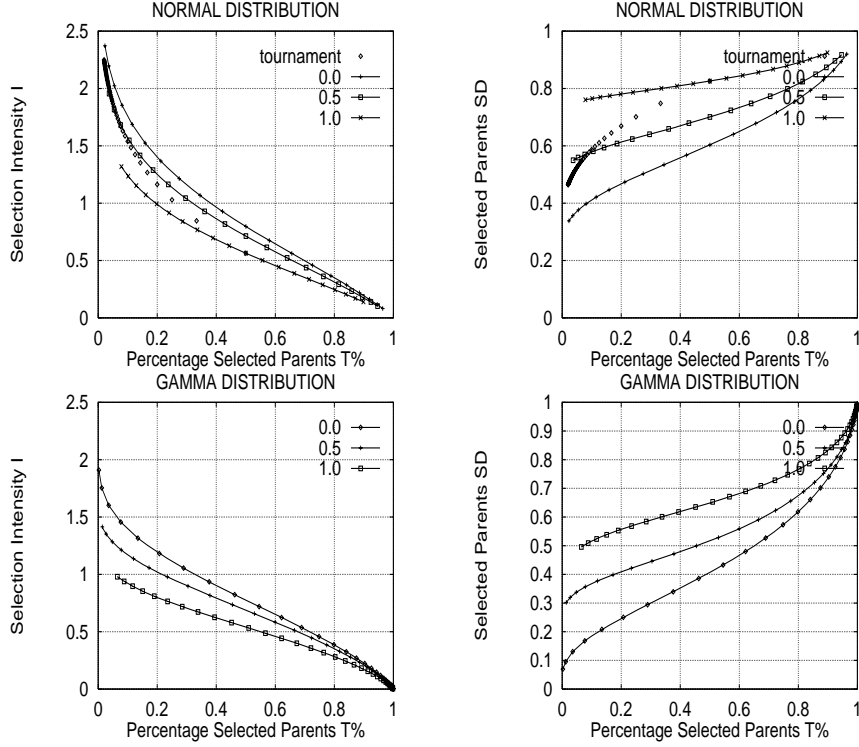
$$z = \frac{f - \overline{f}}{\sigma} \qquad (6)$$

Fig. 2. Selection intensity $|I|$ (left) and standard deviation σ_p (right) of the selected parents for different values of the selection error standard deviation $\sigma_e = 0.0, 0.5, 1.0$ and different values of the proportion of selected parents T; Normal distribution (upper row) and Gamma distribution with $p = 5$ (lower row). For the normal fitness distribution there are also shown $|I|$ and σ_p for tournament selection with tournament sizes from 2 to 50 ($\diamond$) (upper row)

with $E(z) = \overline{z} = 0$ and $\sqrt{V(z)} = \sigma_z = 1$. This probability density function will be used subsequently. The decision error probability density function will be standardized in the same way, i.e. we have $p_e(z)$ with $E(z) = z_T = (f_T - \overline{f})/\sigma$ and $\sqrt{V(z)} = \sigma_e(f)/\sigma(f) = \sigma_e = k \cdot \sigma_z$. The cumulative density function (cdf) of the standardized decision error probability density function is given by

$$P_e(z) = \int_{-\infty}^{z} p_e(t)dt \tag{7}$$

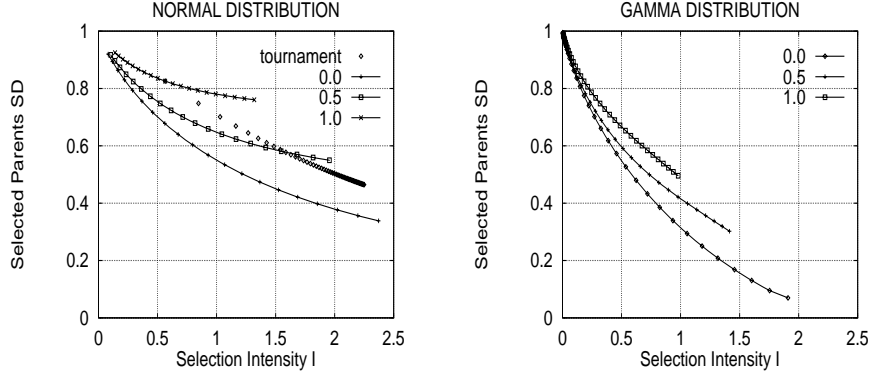such that the proportion of selected parents T with the smallest fitness values are given by

Fig. 3. Standard deviation σ_p vs. selection intensity $|I|$ for different values of the selection error standard deviation $\sigma_e = 0.0, 0.5, 1.0$, Normal distribution (left) and Gamma distribution with $p = 5$ (right)

$$T = \int_{-\infty}^{+\infty} (1 - P_e(z)) \cdot p(z) dz. \tag{8}$$

Finally, the probability density function of the selected parents has the form

$$p_p(z) = \frac{1}{T}(1 - P_e(z)) \cdot p(z). \tag{9}$$

Now we are able to compute the selection intensity I

$$E_p(z) = I = \int_{-\infty}^{+\infty} z \cdot p_p(z) dz \tag{10}$$

and the standardized fitness variance of the parents

$$V_p(z) = \sigma_p^2 = \int_{-\infty}^{+\infty} (z - I)^2 \cdot p_p(z) dz. \tag{11}$$

For a quantitative characterization of erroneous truncation selection we have to assume certain probability density functions for $p(z)$ and $p_e(z)$. A reasonable first assumption is a normal distribution for the breeder's decision errors, i.e.

$$p_e(z) = \frac{1}{\sqrt{2\pi}\sigma_e} e^{-\frac{(z - z_T)^2}{2\sigma_e^2}}. \tag{12}$$

For the fitness distribution we assume a standardized and normalized normal distribution $p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ because up to now the quantitative analysis of selection schemes in quantitative genetics and Evolutionary algorithms is based

on this assumption. Furthermore, we consider a standardized and normalized Gamma distribution $p(z) = \frac{\sqrt{p}}{\Gamma(p)}(z\sqrt{p} + p)^{p-1} \quad e^{-(z\sqrt{p} + p)}$ encountered already in the analysis of gene pool and fuzzy recombination schemes [14, 15].

For some values of the selection error standard deviation $\sigma_e = 0, 0, 0.5, 1.0$ the selection intensity I and the standard deviation σ_p of the selected parents dependent on the proportion of the selected parents T are shown in Fig. 2. The Figures in the upper row refers to a normal fitness distribution, the lower row to a Gamma distribution with $p = 5$. It it obvious that a smaller σ_e implies a larger selection intensity I and a smaller σ_p. In general, the selection intensity and the parents standard deviation for the Gamma distribution are lower then that for the normal distribution because the Gamma distribution is skewed to the right. It is interesting that I and σ_p for tournament selection can be generated by a suitable choice of σ_e. Fig. 3 depicts the dependence of σ_p on I for different values of σ_e.

The response to selection equation does not contain any information about the standard deviation σ_p of the fitness values of the selected parents. This gives rise to the question: What is the influence of σ_p on the response $R = \overline{f(t+1)} - \overline{f(t)}$? We will try to find first answers based on simulation experiments.

3   Empirical Observations

To study the effect of selection errors we consider the following optimization problem

$$\min_{x \in G} f(x) = f^* \quad \text{with} \quad G = R_+^n \tag{13}$$

with $R_+^N = \{x : x_i \geq 0, i = 1, ..., n\}$ for the following two functions

$$f_{\text{PLANE}}(x) = \sum_{i=1}^{n} x_i \quad \text{and} \quad f_{\text{SPHERE}}(x) = \sum_{i=1}^{n} x_i^2. \tag{14}$$

For both functions the optimal value is $f^* = 0$.

The simulations are done using discrete gene pool recombination (DGPR) [14, 11] with $\sigma_e = 0.0$, i.e. for pure truncation selection, and $\sigma_e = 1.0$, i.e. truncation selection with a rather high selection error. The proportion of selected parents $T = 0.2$ is chosen to be equal for all runs. DGPR was applied to avoid other stochastic influences which might be introduced using e.g. fuzzy gene pool recombination [15, 14].

Fig. 4 shows the results for the PLANE function. I and σ_p are dropping from high values at generation 1 to almost constant values after 3 ($\sigma_e = 0$) and 5 ($\sigma_e = 1$) generations, respectively. This is due to a non-stationary population fitness distribution which changes from an approximate symmetric distribution to a skewed to the right distribution. This change is shown for $\sigma_e = 1$ in Fig. 5. Therefore, in the beginning we have values for I and σ_p corresponding to that of the normal distribution from Fig. 2. These values are changing approximately to
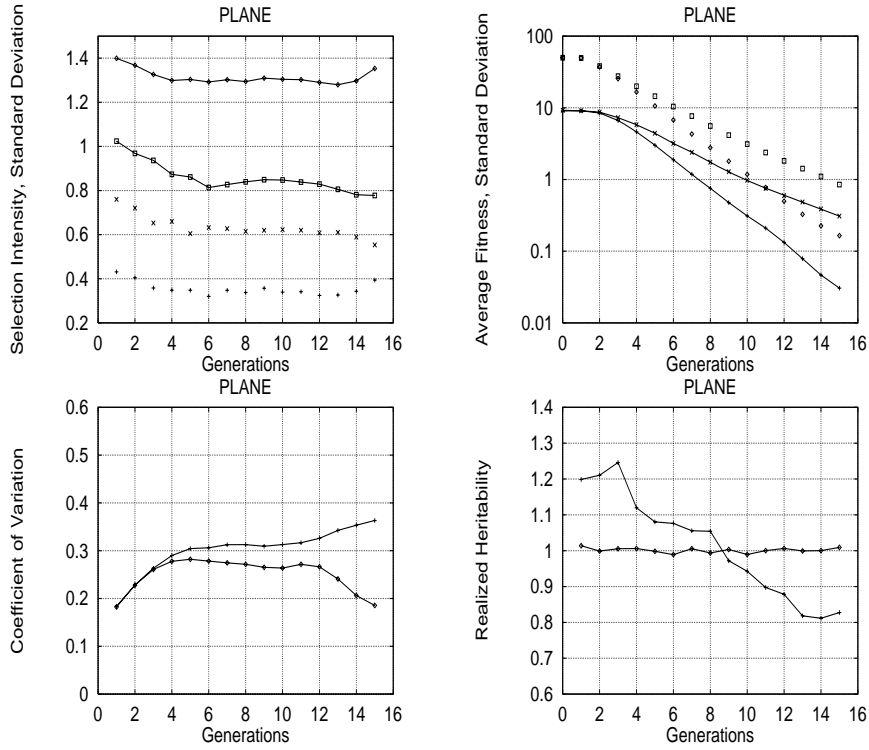
Fig. 4. **PLANE – Upper, left**: Selection intensity $|I|$ $(\sigma_e = 0 : \text{–}\diamond\text{–})$, $(\sigma_e = 1 : \text{–}\square\text{–})$ and standard deviation σ_p of the selected parents $(\sigma_e = 0 : +)$, $(\sigma_e = 1 : \times)$, **upper, right**: average fitness $\overline{f}$ $(\sigma_e = 0 : \diamond)$, $(\sigma_e = 1 : \square)$ and standard deviation σ $(\sigma_e = 0 : \text{–+–})$, $(\sigma_e = 1 : \text{–}\times\text{–})$ of the population, **lower, left**: coefficient of variation CV $(\sigma_e = 0 : \text{–}\diamond\text{–})$, $(\sigma_e = 1 : \text{–+–})$, **lower, right**: realized heritability b $(\sigma_e = 0 : \text{–}\diamond\text{–})$, $(\sigma_e = 1 : \text{–+–})$

those of the Gamma distribution from Fig. 2. The deviation of the values of I and σ_p for $\sigma_e = 0$ is less then that for $\sigma_e = 1$ because the fitness distribution for the former case is less skewed then that for the latter case. This would support the idea that using a low selection error standard deviation leads to a low skewness of the fitness distribution of the population and vice versa.

To confirm this assumption we made a simulation with $T = 0.475$ and $\sigma_e = 0$. This gives approximately the same selection intensity as for the erroneous selection and a somewhat smaller $\sigma_p = 0.49$. The population fitness distribution is approximately symmetric for all considered generations. This leads to the same progress rate as for the erroneous selection though the standard deviation of the fitness of the selected parents σ_p is smaller.

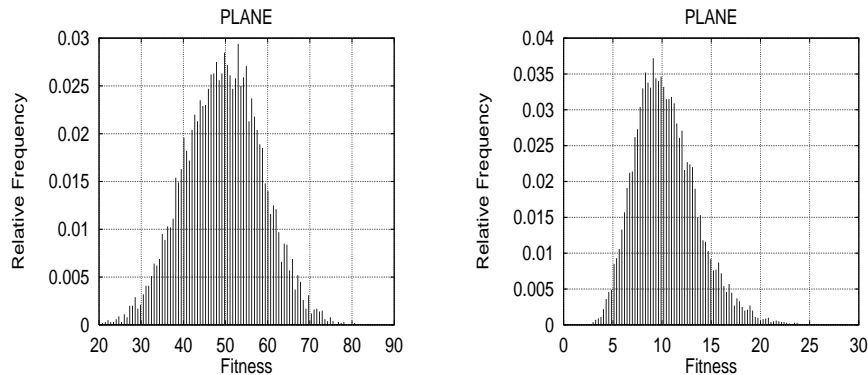The most striking property is the dramatic decline of the realized heritability

Fig. 5. **PLANE** Fitness distribution of the population at generations 1 and 6 for $\sigma_e = 1$

for erroneous selection over generations. Though we have a high fitness variability of the selected parents this will not ensure a high heritability. But this means that high fitness variability does not necessarily lead to high genetic variability.

We made much more simulations. In all cases of truncation selection with more or less severe selection errors we found a more or less rapid decline of the heritability. For pure truncation selection without errors we have an average $b = 1 = $ const. such that the lower fitness variability of the parents is compensated. Finally, the coefficient of variation $CV(t) = \sigma(t)/\overline{f(t)}$ is approximately constant and equal for both σ_e after some generations.

The results for the SPHERE function are similar to those of the PLANE function but the considered effects are stronger. They are shown in Fig. 6. The main difference is the definitely higher coefficient of variation for erroneous selection then that for pure truncation selection.

4 Conclusions

Summarizing the empirical observations leads to the conclusion that having pure and erroneous truncation selection with $I(\sigma_e = 0) = I(\sigma_e > 0)$, which means for an equal fitness distribution of the population that $\sigma_p(\sigma_e = 0) < \sigma_p(\sigma_e > 0)$, that pure truncation selection should be preferred because of its sustained higher heritability. Obviously, the selection error has a decisive influence on the skewness of the fitness distribution of the population which may contribute to the declining heritability. Since we are selecting without replacement erroneous truncation selection causes always the loss of individuals with good fitness values. The same may hold for tournament selection without replacement.

The advantage of pure truncation selection was already mentioned in [8]: Pure truncation selection *is the extreme kind of selection which might be practiced in a laboratory* (or computer) *experiment on selection for one trait alone, disregarding*
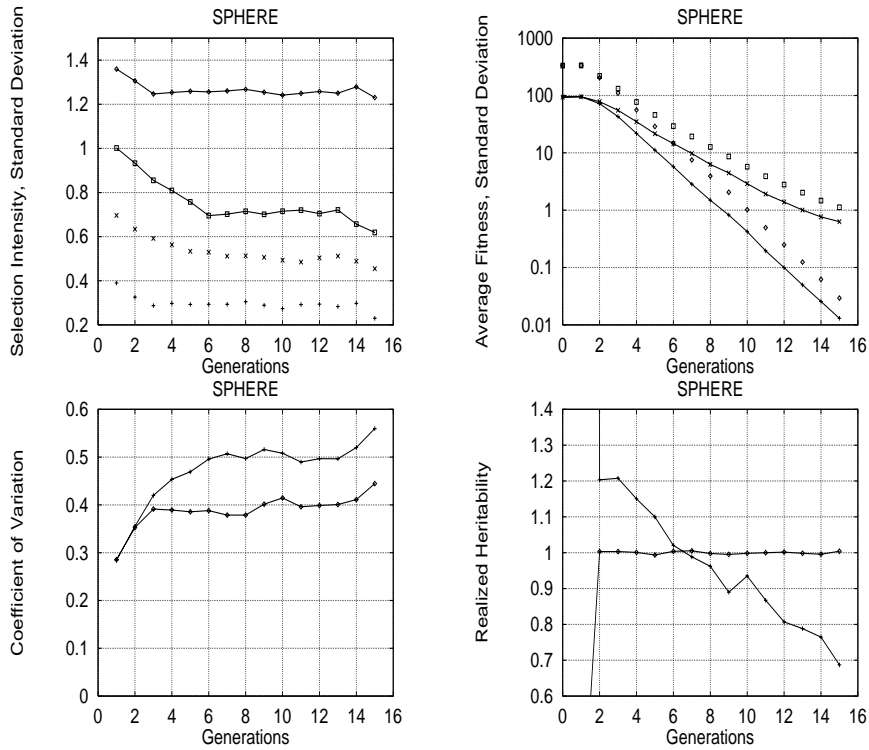
Fig. 6. **SPHERE – Upper, left**: Selection intensity $|I|$ ($\sigma_e = 0 : \diamondsuit$), ($\sigma_e = 1 : \square$) and standard deviation σ_p of the selected parents ($\sigma_e = 0 : +$), ($\sigma_e = 1 : \times$), **upper, right**: average fitness $\overline{f}$ ($\sigma_e = 0 : \diamondsuit$), ($\sigma_e = 1 : \square$) and standard deviation σ ($\sigma_e = 0 : +$), ($\sigma_e = 1 : \asterisk$) of the population, **lower, left**: coefficient of variation CV ($\sigma_e = 0 : \diamondsuit$), ($\sigma_e = 1 : +$), **lower, right**; realized heritability b ($\sigma_e = 0 : \diamondsuit$), ($\sigma_e = 1 : +$)

all others. It *is, of course more effective if the percentage saved is the same as in* erroneous truncation selection.

From the experiments it can be seen that the best choice, at least for simple fitness functions as observed here, is the use of pure, error free truncation selection. With respect to the response to selection there is no benefit from a high parents standard deviation caused by erroneous selection as conjectured in [2]. A further analysis is necessary for erroneous truncation selection with replacement which is usually outside the scope of breeder decisions.

Finally, it should be noted that a further analysis is necessary to confirm these conclusions, e.g. for multi-modal fitness functions. But for these functions it is an open question whether the response to selection equation can be applied.

References

1. T. Bäck "Generalized Convergence Models for Tournament– and (μ, λ)–Selection". In: L. J. Eshelman (Ed.) "Proc. 6th Int. Conf. on Genetic Algorithms", 2–8, Morgan Kaufmann Publ. 1995
2. T. Blickle & L. Thiele "A Mathematical Analysis of Tournament Selection". In: L. J. Eshelman (Ed.) "Proc. 6th Int. Conf. on Genetic Algorithms", 9–16, Morgan Kaufmann Publ. 1995
3. D. S. Falconer "Introduction to Quantitative Genetics" Longman 1981
4. J. J. Grefenstette & J. E. Baker "How Genetic Algorithms work: A Critical Look at Implicit Parallelism. In: J. D. Schaffer (Ed.) "Proc. 3rd Int. Conf. on Genetic Algorithms", 20–27, Morgan Kaufmann Publ. 1989
5. D. E. Goldberg "Genetic Algorithms in Search, Optimization and Machine Learning" Addison–Wesley Publ. Comp. 1989
6. D. E. Goldberg, B. Korb & K. Deb "Messy Genetic Algorithms: Motivation, Analysis, and First Results" Complex Systems 3:493–530, 1990
7. D. E. Goldberg & K. Deb "A Comparative Analysis of Selection Schemes Used in Genetic Algorithms". In: G. J. E. Rawlins (Ed.) "Foundations of Genetic Algorithms", 69–93, Morgan Kaufmann Publ. 1991
8. J. L. Lush "Animal Breeding Plans" Iowa State University Press 1945
9. M. de la Maza & B. Tidor "An Analysis of Selection Procedures with Particular Attention Paid to Proportional and Boltzmann Selection". In: S. Forrest (Ed.) "Proc. 5th Int. Conf. on Genetic Algorithms", 124–131, Morgan Kaufmann Publ. 1993
10. H. Mühlenbein & D. Schlierkamp–Voosen "Predictive Models for the Breeder Genetic Algorithm I. Continuous Parameter Optimization". Evolutionary Computation 1:335–360, 1993
11. H. Mühlenbein & H.-M. Voigt "Gene Pool Recombination in Genetic Algorithms" In: J. P. Kelly & I. H. Osman (Eds.) "Proc. International Metaheuristic Conference", Kluwer Academic Publ. 1996
12. I. Rechenberg "Evolutionsstrategie '94" Frommann–Holzboog 1994
13. H.-P. Schwefel "Evolution and Optimum Seeking" Wiley 1995
14. H.-M. Voigt & H. Mühlenbein "Gene Pool Recombination and Utilization of Covariances for the Breeder Genetic Algorithm" In: "Proc. IEEE 2nd Int. Conf. on Evolutionary Computation", 172–177, IEEE Press 1995
15. H.-M. Voigt, H. Mühlenbein & D. Cvetcovic "Fuzzy Recombination for the Breeder Genetic Algorithm" In: L. J. Eshelman (Ed.) "Proc. 6th Int. Conf. on Genetic Algorithms",104–112, Morgan Kaufmann Publ. 1995
16. H.-M. Voigt & H. Mühlenbein "The Response to Selection Equation for Skew Fitness Distributions" To appear in: "Proc. IEEE 3rd Int. Conf. on Evolutionary Computation", IEEE Press 1996
17. L. D. Whitley "The GENITOR Algorithm and Selection Pressure: Why Rank–Based Allocation of Reproductive Trials is Best" In: D. E. Schaffer (Ed.) "Proc. 3rd Int. Conf. on Genetic Algorithms" 116–121, Morgan Kaufmann Publ. 1989