Original article

Prediction of genetic gain from quadratic optimisation with constrained rates of inbreeding

Beatriz VILLANUEVA[a]*, Santiago AVENDAÑO[a],
John A. WOOLLIAMS[b]

[a] Scottish Agricultural College, West Mains Road, Edinburgh, EH9 3JG, Scotland, UK
[b] Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, UK

Abstract – There are selection methods available that allow the optimisation of genetic contributions of selection candidates for maximising the rate of genetic gain while restricting the rate of inbreeding. These methods imply selection on quadratic indices as the selection merit of a particular individual is a quadratic function of its estimated breeding value. This study provides deterministic predictions of genetic gain from selection on quadratic indices for a given set of resources (the number of candidates), heritability, and target rate of inbreeding. The rate of gain was obtained as a function of the accuracy of the Mendelian sampling term at the time of convergence of long-term contributions of selected candidates and the theoretical ideal rate of gain for a given rate of inbreeding after an exact allocation of long-term contributions to Mendelian sampling terms. The expected benefits from quadratic indices over traditional linear indices (*i.e.* truncation selection), both using BLUP breeding values, were quantified. The results clearly indicate higher gains from quadratic optimisation than from truncation selection. With constant rate of inbreeding and number of candidates, the benefits were generally largest for intermediate heritabilities but evident over the entire range. The advantage of quadratic indices was not highly sensitive to the rate of inbreeding for the constraints considered.

prediction of genetic gain / quadratic indices / control of inbreeding / genetic contributions

1. INTRODUCTION

Quadratic optimisation [7, 8, 10, 11] provides a solution to the problem of optimising selection decisions in breeding schemes for maximising genetic gain (ΔG) with constrained rates of inbreeding (ΔF). The selection merit of individual candidates is a quadratic function of their estimated breeding values and therefore selection is based on quadratic indices. Stochastic simulation studies have shown that the extra response from quadratic indices over traditional linear indices (*i.e.* truncation selection) at the same ΔF is large, ranging from

---

* Corresponding author: beatriz.villanueva@sac.ac.uk

20% to 60% (*e.g.* [10]). Also, substantial increases in ΔG are expected when applying quadratic indices to real livestock populations [1, 9].

Although the benefits in terms of gain from quadratic optimisation are clear from empirical studies, there is not yet a framework available for obtaining deterministic predictions of ΔG when using this approach. Grundy *et al.* [7] showed that the ideal optimal solution for a given constraint on ΔF could be obtained after an exact linear allocation of long-term genetic contributions of selected candidates (r) to their Mendelian sampling terms (a). Using stochastic simulations, Avendaño *et al.* [2] empirically confirmed that quadratic optimisation allocates contributions of selected candidates according to the best information on their Mendelian sampling terms and not on their breeding values. This provided the link between the optimisation of breeding schemes using quadratic indices and the maximisation of the covariance between r and a implicit in the definition of genetic gain of Woolliams and Thompson [18] (*i.e.* $E\,[\Delta G] = \sum r_i a_i$).

Grundy *et al.* [7] also provided a deterministic prediction for the ideal rate of genetic gain (ΔG_{ideal}) but the empirical gains obtained after applying quadratic optimisation were consistently lower than ΔG_{ideal}. The lack of knowledge about the true Mendelian sampling term at selection time was one of the reasons of the inability of quadratic optimisation to attain the ideal optimal solution [20]. Recently, Avendaño *et al.* [3] have provided deterministic predictions of the accuracy of Mendelian sampling terms both at selection time and at convergence of long-term contributions that can prove to be useful for filling the gap in the prediction of gain from selection on quadratic indices.

The aim of this study was to develop a deterministic framework for predicting the rate of genetic gain under quadratic optimisation when BLUP is used to estimate breeding values. The approach combined predictions of the accuracy of the estimated Mendelian sampling term at the time of convergence of long-term contributions [3] with predictions of the ideal rate of genetic gain [7]. Predictions of gain from quadratic optimisation were then compared to those from traditional truncation selection and the expected benefits from using quadratic indices were quantified.

2. METHODS

2.1. Ideal and upper bound genetic gain under a constrained rate of inbreeding

Assuming that Mendelian sampling terms are normally distributed with standard deviation equal to one, Grundy *et al.* [7] showed that the ideal

theoretical rate of genetic gain (ΔG_{ideal}) can be defined in terms of the standardised truncation point (x) and the selection intensity (i) as

$$\Delta G_{ideal} = (i - x)^{-1} \text{ or equivalently } i/k \tag{1}$$

where $k = i(i - x)$ and i and x are the solution for

$$(4T\Delta F)^{-1} = p(i - x)^2(1 + x^2 - ix)^{-1}, \tag{2}$$

being T the total number of candidates per generation ($T/2$ males and $T/2$ females), ΔF the desired rate of inbreeding, and p the proportion of selected candidates. Values for p, and hence x and i, can be found using the Newton-Rapson method for defined values of $T\Delta F$ (App. A).

The variance of the Mendelian sampling terms is not reduced by selection through gametic phase disequilibrium (*i.e.* the Bulmer effect). Hence, ΔG_{ideal} from expression (1) can be conveniently expressed in terms of the trait additive genetic variance in the base population ($\sigma^2_{A_0}$), resources available (T) and the target ΔF as

$$\Delta G_{ideal} = \frac{i\sqrt{(0.5)\sigma^2_{A_0}}}{k}. \tag{3}$$

This assumes that breeding values and Mendelian sampling terms are known without error (*i.e.* the initial heritability, h^2_0, is equal to one and the environmental variance equals to zero). It also assumes a constant Mendelian variance such that losses from inbreeding are restored by mutational variance. In other circumstances this will provide a medium term approximation.

It is valid to ask how the expression for ΔG_{ideal} which implies an exact allocation of long-term contributions to Mendelian sampling terms compares to conventional formulae for predicting the rate of gain under standard truncation selection under the same circumstances of known breeding values and Mendelian sampling terms. The equilibrium rate of gain after accounting for gametic phase disequilibrium generated by selection [6] for mass truncation selection can be approximated as

$$\Delta G_{m\_tru} = \frac{i\sigma^2_{A_0}\sigma^{-1}_P}{(1 + kh^2)} \tag{4}$$

where σ^2_P and h^2 are respectively the phenotypic variance and heritability, both at equilibrium. With the assumption of $h^2_0 = h^2 = 1$ (*i.e.* breeding values are assumed known without error) then $\sigma_P = \sigma_A$, $\sigma_{A_0}/\sigma_A = \sqrt{1 + k}$ and the equation (4) can be re-written as

$$\Delta G_{m\_tru} = \frac{i\sigma_{A_0}}{\sqrt{1 + k}}. \tag{5}$$

Equation (5) shows that the form of ΔG_{ideal} in equation (3) is not unexpected and that the expressions for the rate of genetic gain after linear allocation of Mendelian sampling terms and long-term contributions for constrained ΔF, and that for truncation selection are closely related. Importantly, it arises that for the same σ_{A_0}, $\Delta G_{m\_tru} < \Delta G_{ideal}$. The validity of this inequality is shown in Figure 1, where both expressions for the rate of gain are compared at the same x and k (*i.e.* not at the same ΔF). Figure 1 shows that, assuming an exact allocation of long-term contributions to Mendelian sampling terms, and known Mendelian sampling terms for each selected candidate, ΔG_{ideal} is always greater than ΔG_{m_tru} when compared at the same truncation point. As expected, the difference between ΔG_{ideal} and ΔG_{m_tru} was reduced as x and k increased, approaching zero at the limit ($k = 1$). Typical values for x in practical livestock breeding programmes could range from -1.0 to 3.5, giving a corresponding range for k of 0.34 to 0.95.

As pointed out by Grundy *et al.* [7] and Woolliams *et al.* [20], one of the factors for which the ΔG_{ideal} is unattainable is that the true Mendelian sampling terms are unknown, so contributions of selected candidates are allocated according to initial estimates at the time of selection. Grundy *et al.* [7] suggested an intuitive expression for predicting the upper bound genetic gain obtained under quadratic optimisation (ΔG_{quad}) that accounts for this lack of knowledge of the true Mendelian sampling terms:

$$\Delta G_{quad} = \rho_{conv}\, \Delta G_{ideal}, \tag{6}$$

where ρ_{conv} is the accuracy of the Mendelian sampling term at convergence of long-term contributions.

2.2. Predicting the effective number of parents at the time of selection

The population structure at the time of selection is required for obtaining the predictions for ρ_{conv} from the extended pseudo-BLUP index of Avendaño *et al.* [3]. In their study, the 'effective number of parents' of equal contributions [12] at selection time was calculated empirically (*i.e.* from simulations) as $N_c = \left[\sum c^2\right]^{-1}$, where the sum of squares of contributions was taken over all selected candidates resulting from the optimisation. The population structure at selection time was derived as $n_s = n_d = \frac{1}{2}N_c$ and $n_o = T/n_d$, where n_s, n_d and n_o are respectively the number of sires, the number of dams and the number of offspring per dam. A mating ratio equal to one was therefore assumed.

Since the primary objective here was to obtain a completely deterministic prediction for ΔG_{quad}, an approach was derived for predicting N_c.
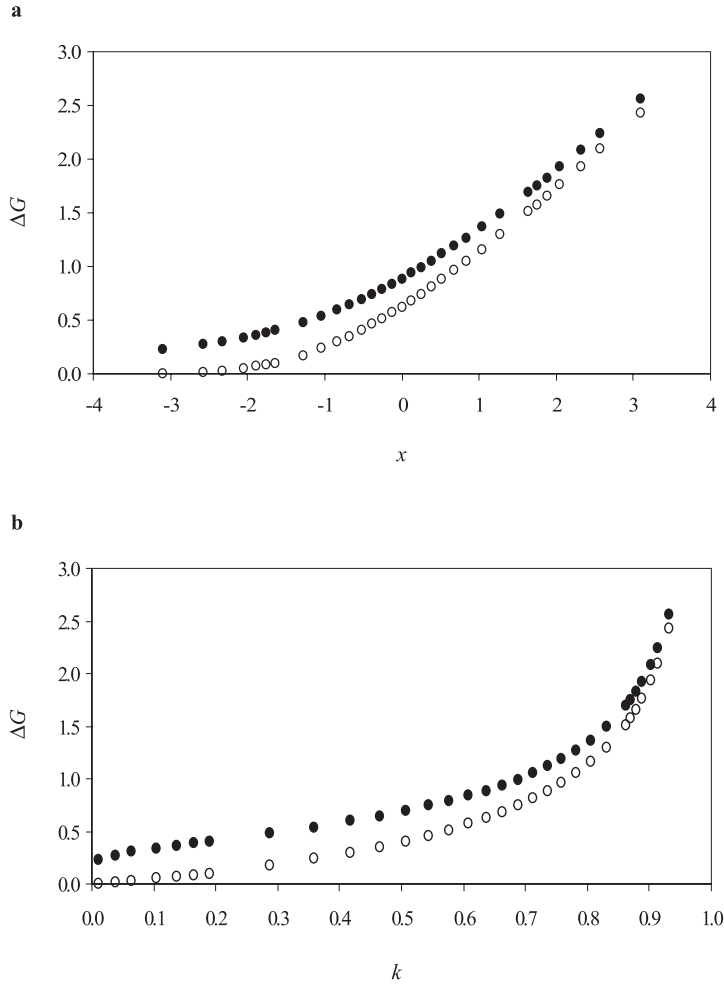
a

b

Figure 1. Rate of genetic gain from mass truncation selection (○) and theoretical ideal rate of gain from optimised selection (●) after an exact allocation of long-term contributions to Mendelian sampling terms for a range of deviations of the truncation point from the mean (x, in a) and coefficients of variance reduction (k, in b).

The approach was based on predictions for the ratio N_c/N_r, where N_r is the 'effective number of ancestors'. The ratio N_c/N_r measures the change in the equivalent number of parents from initial selection that are still represented when their contributions converge. The effective number of ancestors can be derived directly from the inbreeding constraint as $N_r = [4\Delta F]^{-1}$ given that $E(\Delta F) = \frac{1}{4} \sum r_i^2$ [22]. Therefore, N_c could be obtained from a prediction of

N_c/N_r and the ΔF constraint. The prediction approach for the ratio N_c/N_r is described in Appendix B.

2.3. Predicting the accuracy of the Mendelian sampling term

Predictions for the accuracy of the Mendelian sampling term needed to predict ΔG_{quad} from equation (6) were obtained using the approach of Avendaño *et al.* [3]. Briefly, a selection index theory framework was used where the original pseudo-BLUP index of Wray and Hill [21], which assumed a nested mating structure with m dams mated to each sire, was extended to accommodate three extra sources of information related to the Mendelian sampling term. Here a mating ratio $m = 1$ was assumed. The three sources were the estimated Mendelian sampling term of the dam, the mean estimated Mendelian sampling term of all dams mated to the sire, and the estimated Mendelian sampling term of the sire. This extension allowed the accuracy of the Mendelian sampling term to be predicted for selected candidates just after being selected (ρ_{sel}), and after becoming sires (ρ_{sir}) or dams (ρ_{dam}) with recorded offspring. The latter was considered to be the accuracy at the convergence of long-term contributions. With a mating ratio equal to one, $\rho_{sir} = \rho_{dam} = \rho_{conv}$.

2.4. Simulations

The rate of gain ΔG_{quad} predicted from equation (6) was compared to the empirical responses obtained from stochastic simulations (ΔG_{sim}) over several generations. An additive infinitesimal model [6] was assumed for the trait under selection. The true breeding values of base population individuals were obtained from a normal distribution with mean zero and variance equal to the initial heritability (h_0^2) so the phenotypic variance was assumed equal to one. In subsequent generations, the true breeding value of the progeny was obtained as half the sum of the true breeding values of their parents plus a random Mendelian sampling term. The Mendelian sampling term of an individual was sampled from a normal distribution with mean zero and variance $\frac{1}{2}h_0^2$ (*i.e.* the effects of inbreeding on the genetic variance were ignored). The phenotypic value for any individual at any generation was obtained by adding an environmental component sampled from a normal distribution with mean zero and variance $1-h_0^2$ to the true breeding value. Populations with discrete generations were evaluated over 10 generations of selection. In the base generation ($t = 0$), T individuals ($T/2$ males and $T/2$ females) with family full-sib structure were

generated. The first generation of selection was obtained from the mating of animals selected at $t = 0$.

Selection decisions were optimised for each generation by using the optimisation algorithm described by Meuwissen [10] which gives the maximum genetic gain while constraining the inbreeding rate to a specific value. A BLUP animal model was used to estimate breeding values. The constant rate of inbreeding in the long-term was achieved by setting the constraint on the average coancestry of selected candidates ($\mathbf{c}^{\mathbf{T}}\mathbf{A}\mathbf{c}/2$) to $1 - (1-\Delta F)^t$, where $\mathbf{A}$ is the average relationship matrix, $\mathbf{c}$ is a vector of projected use of candidates (*i.e.* mating proportions), ΔF is the desired rate and t is the generation number [7]. Mating among selected candidates was at random. Populations of $T = 100$, 200 or 300 candidates per generation were simulated. A broad range of trait h_0^2 (0.01, 0.25, 0.50, 0.75 and 0.99) was studied. ΔF was constrained to 0.05, 0.025, 0.0125 and 0.01 per generation. The parameter values go beyond those commonly encountered in practice which correspond to the low to intermediate values for h_0^2, and the lower values of ΔF. One hundred replicates were performed for each scenario.

2.5. Comparison of deterministically predicted genetic gains from optimised selection and truncation selection at the same rate of inbreeding

The rate of gain (ΔG_{quad}) predicted from equation (6) was also compared to the predicted rate of gain obtained from truncation selection (ΔG_{tru}) at the same ΔF for a range of scenarios all using BLUP genetic evaluation. The predictions of genetic gain and rate of inbreeding for truncation selection schemes were obtained using the Sel*A*ction software [4, 5, 13]. A mating ratio equal to one was assumed for all schemes under truncation selection (*i.e.*, $n_s = n_d = T/2$ and $n_o = T/n_d$ across generations).

Comparisons between truncation and optimised selection at the same ΔF were carried out in two ways. Firstly, ΔG_{quad} and ΔG_{tru} were compared for an extensive range of ΔF at two fixed levels of h_0^2 (0.10 and 0.35). Comparisons across ΔF were carried out with $T = 100$ or 300 candidates. In order to obtain an extensive range of ΔF (from 0.1% up to 4.0%), the population structure was varied by changing the proportion of selected individuals at regular intervals. For each population structure, Sel*A*ction [13] was used to predict ΔG_{tru} and the corresponding ΔF. For each level of ΔF predicted under truncation selection, N_c was obtained from T, ΔF and h_0^2 using the approach described above. Then ΔG_{quad} under optimised selection was calculated after predicting

ρ_{conv} and ΔG_{ideal} from the modified pseudo-BLUP index [3] and equation (3), respectively.

Secondly, ΔG_{quad} and ΔG_{tru} were compared for an extensive range of h_0^2 and two fixed levels of $\Delta F(0.01$ and $0.025)$. Comparisons across h_0^2 (with h_0^2 ranging from 0.05 to 0.85) were carried out with $T = 100$, 300 and 1000. For a given combination of T and h_0^2, the population structure that gave the desired ΔF under truncation selection was obtained using Sel*Action*. A maximum deviation of 0.05 with respect to the target ΔF was allowed. Then, as before the corresponding N_c was obtained from T, ΔF and h_0^2, and ΔG_{quad} was calculated from ρ_{conv} and ΔG_{ideal}.

3. RESULTS

3.1. Deterministic prediction of the rate of genetic gain under quadratic optimisation

Table I shows a comparison between ΔG_{obs} (simulations) and ΔG_{quad} (predictions) for the broad range of h_0^2 and ΔF constraints studied. The % error ranged from –0.7% ($h_0^2 = 0.5$ and $\Delta F = 0.0250$) to 15.0% ($h^2 = 0.99$ and $\Delta F = 0.0100$). Over-prediction was the norm for the most relaxed ΔF constraint (*i.e.* $\Delta F = 0.0500$) and for $h_0^2 = 0.99$. However, under-prediction was typically observed for the most stringent ΔF constraints with $h_0^2 < 0.99$. It should be noted that accurate (% error less than 10%) predictions were obtained for the range of ΔF constraints most likely to be applied in practice (*e.g.* from 0.0100 to 0.0250) in selection programmes aimed to improve typical traits of economic importance (*i.e.* with h_0^2 levels of up to 0.50).

In Figure 2, predictions of ΔG_{quad} (using predicted N_c) are compared with the rates of gain from simulations for two scheme sizes ($T = 100$ and $T = 300$), a broad range of h_0^2 and two ΔF constraints ($\Delta F = 0.0100$ and $\Delta F = 0.0250$). Standard errors for simulated rates of gain were always less than 0.02 and ranged from 0.001 to 0.014. Predictions for ΔG_{quad} were very similar to those using empirical N_c for both scheme sizes (not shown). Excluding the extreme $h_0^2 = 0.99$, Figure 2a shows a good agreement between ΔG_{quad} and ΔG_{obs} for the smaller scheme ($T = 100$) at both ΔF constraints. For $\Delta F = 0.0100$, ΔG_{quad} under-predicted ΔG_{obs} on average by –5.9%, but the differences between predicted and observed values were non-significant ($p > 0.05$) for h_0^2 ranging from 0.01 to 0.7. Similarly, for $\Delta F = 0.0250$, the difference between ΔG_{quad} and ΔG_{obs} was non-significant ($p > 0.05$) for $h_0^2 \leq 0.6$.

Table I. Rate of gain obtained from simulations (ΔG_{obs}), theoretical ideal upper limit (ΔG_{ideal}), predicted rate of gain from quadratic optimisation (ΔG_{quad}), and predicted accuracy of the Mendelian sampling term for sires (ρ_{conv}) at the time of convergence of the long-term contributions for a range of h_0^2 and ΔF constraints and $T = 100$.

h_0^2	ΔF	$\Delta G_{obs}^{1,2}$	ΔG_{ideal}	ρ_{conv}	ΔG_{quad}	% error[3]
0.01	0.0500	0.021	0.153	0.148	0.023	7.8
	0.0250	0.015	0.132	0.123	0.016	8.0
	0.0125	0.011	0.108	0.100	0.011	−3.5
	0.0100	0.010	0.099	0.096	0.009	−4.4
0.25	0.0500	0.374	0.766	0.558	0.427	14.3
	0.0250	0.316	0.660	0.495	0.327	3.4
	0.0125	0.250	0.538	0.435	0.234	−6.5
	0.0100	0.227	0.494	0.419	0.207	−9.0
0.50	0.0500	0.668	1.083	0.673	0.729	9.1
	0.0250	0.587	0.933	0.625	0.583	−0.7
	0.0125	0.463	0.761	0.573	0.436	−5.9
	0.0100	0.420	0.700	0.560	0.392	−6.7
0.75	0.0500	0.983	1.327	0.749	0.994	1.2
	0.0250	0.849	1.143	0.714	0.816	−3.9
	0.0125	0.677	0.932	0.683	0.636	−6.0
	0.0100	0.612	0.856	0.679	0.581	−5.0
0.99	0.0500	1.323	1.525	0.947	1.444	9.2
	0.0250	1.132	1.313	0.952	1.251	10.5
	0.0125	0.908	1.071	0.961	1.029	13.3
	0.0100	0.825	0.984	0.964	0.948	15.0

[1] Obtained at $t = 3$.
[2] Standard errors over replicates ranged from 0.003 to 0.001 for $h_0^2 = 0.01$, from 0.010 to 0.006 for $h_0^2 = 0.25$, from 0.0174 to 0.007 for $h_0^2 = 0.50$, from 0.018 to 0.009 for $h_0^2 = 0.75$ and from 0.016 to 0.009 for $h_0^2 = 0.99$. The higher and lower bound of each range corresponds to $\Delta F = 0.05$ and $\Delta F = 0.01$, respectively.
[3] % error = [($\Delta G_{quad} - \Delta G_{obs}$)/ ΔG_{obs}] 100.

For the largest scheme ($T = 300$), ΔG_{quad} still provided good predictions of gain for the most stringent constraint. For $\Delta F = 0.0100$, the difference between ΔG_{quad} and ΔG_{obs} was non-significant ($p > 0.05$) for $h_0^2 < 0.9$ (Fig. 2b). For $h_0^2 \leq 0.7$, ΔG_{quad} over-predicted ΔG_{obs} by about 3.9%, and for $h_0^2 \geq 0.6$ ΔG_{quad} under-predicted ΔG_{obs} by about −3.7%. Predictions were less accurate for
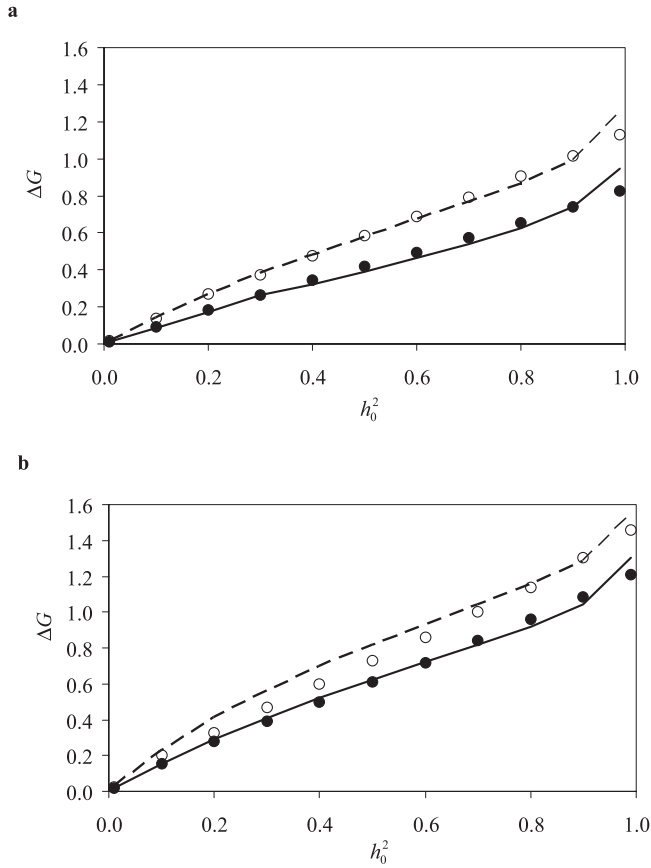
a

b

Figure 2. Observed (symbols) and predicted (lines) rate of genetic gain for a range of base heritabilities (h_0^2) and two population sizes ($T = 100$, in a and $T = 300$, in b). ——: ΔG_{quad} under $\Delta F = 0.010$; - - -: ΔG_{quad} under $\Delta F = 0.025$; •: ΔG_{obs} under $\Delta F = 0.010$; ○: ΔG_{obs} under $\Delta F = 0.025$.

$\Delta F = 0.0250$ and ΔG_{quad} over-predicted ΔG_{obs} by on average 11.8% in the whole range of h_0^2 up to 0.8. The difference between ΔG_{quad} and ΔG_{obs} was significant ($p < 0.05$) for h_0^2 ranging from 0.1 to 0.6.

3.2. Comparison of predicted rates of genetic gain from truncation and quadratic optimisation at the same rate of inbreeding

The rates of gain under truncation and optimised selection for different ΔF levels are presented in Figure 3. At a given ΔF, the gain from optimised selection (ΔG_{quad}) was always higher than that from truncation selection (ΔG_{tru})
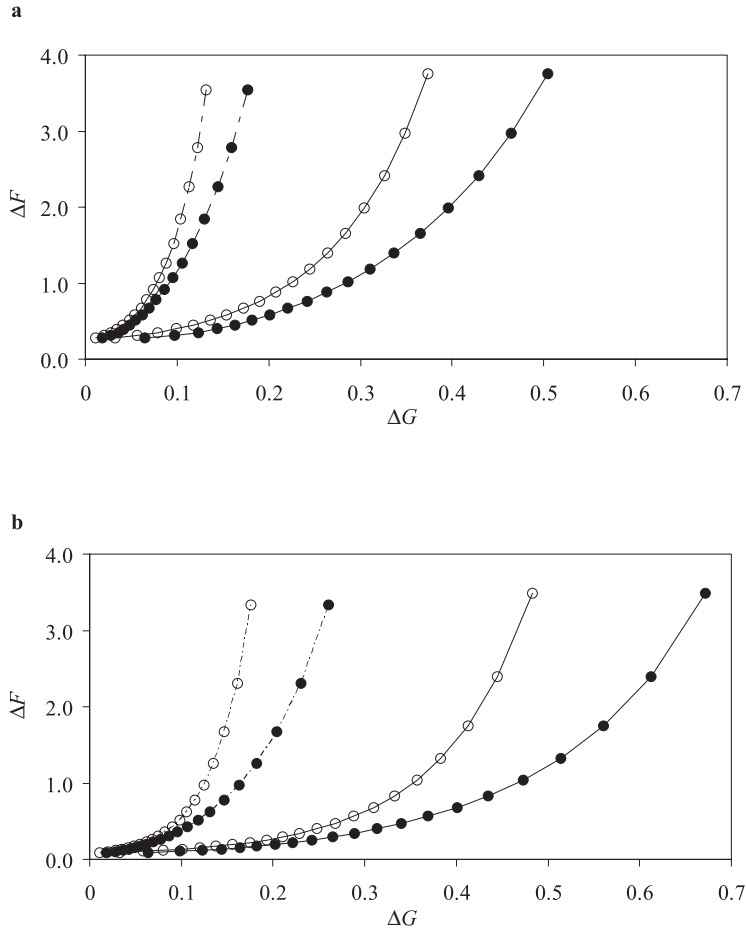
a

b

Figure 3. Relationship of predicted rates of gain (ΔG) and rates of inbreeding (ΔF, in %) for schemes under BLUP truncation (○) and BLUP quadratic optimisation (●) selection for different base heritabilities (- - -: $h_0^2 = 0.10$; ——: $h_0^2 = 0.35$) and population sizes ($T = 100$, in a and $T = 300$, in b).

for both heritabilities and scheme sizes. The advantage of ΔG_{quad} over ΔG_{tru} increased as ΔF increased (*i.e.* as the constraint in quadratic optimisation becomes less stringent) and was the greatest for the highest h_0^2 and the largest breeding scheme. For instance, for $\Delta F = 0.0100$ and $h_0^2 = 0.35$, the advantage of ΔG_{quad} over ΔG_{tru} was 26.5% (0.286 *versus* 0.226) for $T = 100$ and 32.5% (0.473 *versus* 0.357) for $T = 300$. For $\Delta F = 0.0100$ and $h_0^2 = 0.10$ the advantage of ΔG_{quad} over ΔG_{tru} was 17.3% (0.095 *versus* 0.081) for $T = 100$ and 30.4% (0.163 *versus* 0.125) for $T = 300$. The same profile was found when

a

b

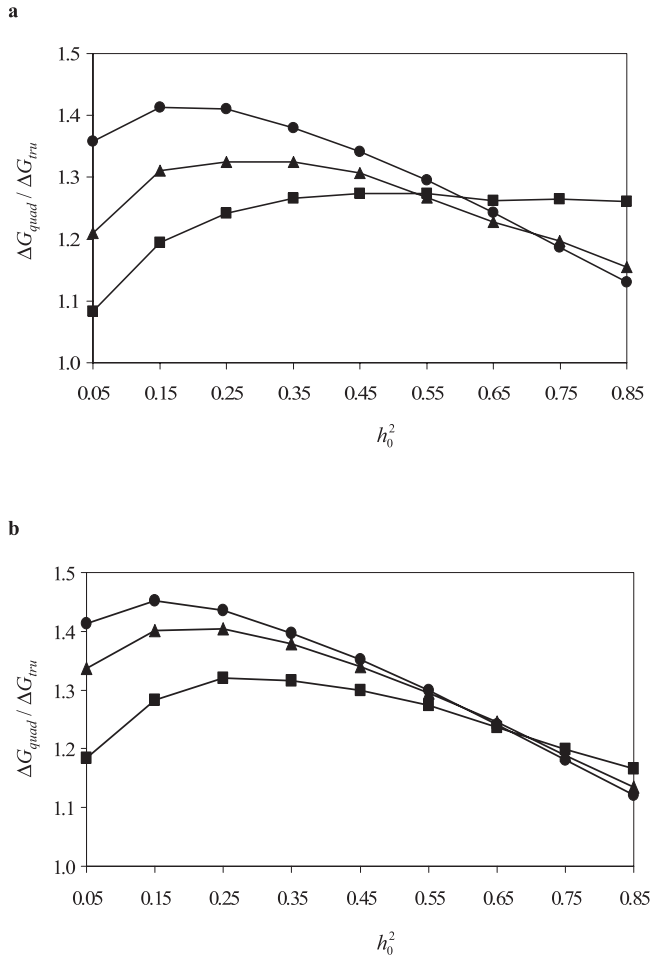**Figure 4.** Ratio of predicted rate of gain under BLUP quadratic optimisation (ΔG_{quad}) over predicted rate of gain under BLUP truncation selection (ΔG_{tru}) for a range of base heritabilities (h_0^2), populations sizes (■: $T = 100$; ▲: $T = 300$; ● : $T = 1000$) and two inbreeding constraints ΔF ($\Delta F = 0.0100$, in a, and $\Delta F = 0.0250$, in b).

T was increased up to 1000 candidates (results not shown). In this scenario, the superiority of ΔG_{quad} over ΔG_{tru} was about 40% for $\Delta F = 0.0100$.

Figure 4 shows the change of $\Delta G_{quad}/\Delta G_{tru}$ with h_0^2 for three population sizes and two fixed values of ΔF (0.0100 and 0.0250). Two situations can be identified according to ranges of h_0^2 below or above 0.55. For h_0^2 up to 0.55, $\Delta G_{quad}/\Delta G_{tru}$ was greater for $\Delta F = 0.0250$ than for $\Delta F = 0.0100$ and the maximum ratio increased with T. For $\Delta F = 0.0100$, the maximum $\Delta G_{quad}/\Delta G_{tru}$ was 1.41 for $T = 1000$, 1.32 for $T = 300$ and 1.27 for $T = 100$, and it

occurred at h_0^2 = 0.15, 0.25 and 0.45, respectively. For ΔF = 0.0250, the maximum $\Delta G_{quad}/\Delta G_{tru}$ was 1.45 for T = 1000, 1.40 for T = 300 and 1.32 for T = 100, and it occurred at h_0^2 = 0.15, 0.25 and 0.25, respectively. For h_0^2 above 0.55 and ΔF = 0.0100, $\Delta G_{quad}/\Delta G_{tru}$ was fairly constant up to h_0^2 = 0.85 for T = 100, whereas that for T = 300 and T = 1000 the ratio decreased down to 1.15 and 1.13, respectively. For ΔF = 0.0250, $\Delta G_{quad}/\Delta G_{tru}$ decreased for the three population sizes reaching minimum values of 1.16, 1.13 and 1.12 (for T = 100, 300 and 1000, respectively) at h_0^2 = 0.85. Thus the effect of the size of the scheme on the superiority of quadratic optimisation over truncation selection decreased as h_0^2 increased (with the exception of very small schemes at tight ΔF constraints).

4. DISCUSSION

This study presents deterministic predictions of the maximum potential rate of genetic gain that can be obtained in schemes using selection on quadratic indices where the rate of inbreeding is restricted to pre-defined levels. The two key components required for the prediction are the following: (i) a prediction of the ideal rate of gain after a perfect allocation of long-term contributions to Mendelian sampling terms for a given ΔF, and (ii) a prediction of the ultimate accuracy of the Mendelian sampling term at the time of convergence of the long-term contributions of selected candidates. The first component (*i.e.* prediction of ΔG_{ideal}) was available from Grundy *et al.* [7] and the second component (*i.e.* prediction of the accuracy) from Avendaño *et al.* [3]. The approach provided accurate predictions of ΔG_{quad} for realistic h_0^2 and target ΔF values in breeding schemes of contrasting size. However, ΔG_{quad} was less satisfactory and over-predicted gain with the extreme values of ΔF (5%) or h_0^2 (0.99). The deterministic method provides a solution to the gap in prediction tools for the design of breeding schemes under constrained rates of inbreeding and gives the necessary accompanying tool to the current available operational dynamic selection algorithms.

Deterministic methods for the strategic optimisation of breeding schemes, that is, the maximisation of genetic gain for a pre-defined ΔF were previously developed for schemes under mass [15,16], and index selection [15]. However, these methods dealt with truncation selection, and were unlinked to dynamic selection algorithms based upon quadratic indices [7, 10]. For truncation selection, Villanueva *et al.* [17] and Bijma *et al.* [5] used predictions of ΔG and/or ΔF based on lifetime contributions (u_i) of selected candidates. The basic framework for predicting u_i was laid down by Woolliams *et al.* [19] and is

based on $u_i = E(r_i|s_i)$ where s_i is the selective advantage of the individual i. However, although this framework is relevant for truncation selection it may not be appropriate for quadratic indices [19]. Under truncation selection, selective advantages are defined as functions of the breeding value of the candidates whereas under quadratic optimisation the selective advantage is related to the Mendelian sampling term [2]. The main limitation for extending the prediction approach based on u_i to constrained optimisation under BLUP selection is the difficulty of predicting the selection intensity for the group of selected candidates for which the average coancestry has been restricted [20].

Here, the predicted accuracy of the Mendelian sampling term was the key parameter for obtaining predictions of genetic gain under constrained ΔF. The lack of knowledge of the Mendelian sampling term (the selective advantage in quadratic indices) and the availability of only initial estimates at the time of selection is one of the central factors that prevent attaining the ideal optimal solution of an exact allocation of long-term contributions to Mendelian sampling terms [7, 20]. Therefore, by using predictions of the ultimate accuracy, the over-prediction represented by ΔG_{ideal} is directly adjusted. The inability to set the contributions of selected candidates to their desired values (*i.e.* independently from the contributions of future descendants) is the second factor that drives the observed outcome of the constrained optimisation away from the ideal solution [7, 20]. Not accounting for this effect when adjusting ΔG_{ideal} does not seem to represent a strong limitation since predictions ignoring it were reasonably accurate for a broad range of parameters.

Although the deterministic framework presented is self-contained and can be used by the specification of solely T, h_0^2 and ΔF, it should be recognised that it has an element of empirical nature and so is limited. This arises from the regression approach for predicting the ratio N_c/N_r that implied the empirical relationship between the optimisation outcome at selection time ($\sum c^2$) and at convergence ($\sum r^2$) for different h_0^2, ΔF restrictions and population sizes. Apart from the method of inferring equilibrium structure, further sources of error may arise from the use of standard selection intensities that assume infinite population sizes, particularly in the smaller schemes where indices among relatives may be highly correlated, and from the use of linear indices in the method to calculate ρ_{conv} for quadratic optimisation. The relatively poor performance of the prediction for the extreme heritability $h_0^2 = 0.99$ is not unexpected since a very high accuracy at the start of the optimisation process implies that the contributions assigned initially will differ greatly among individuals and there will be no scope to change them over time *e.g.* to accommodate desired contributions from their offspring. In contrast, with lower values of h_0^2, the assigned

initial contributions will be less variable and these can be adjusted more easily, if necessary, in subsequent generations as information accumulates.

Predictions for ΔG_{quad} ignore the effects of inbreeding on the genetic variance and thus refer to either (i) a near-asymptotic response to selection in the medium term after accounting for the Bulmer effect, or (ii) where losses from inbreeding are balanced by mutational variance entering the population. Villanueva and Woolliams [16] accounted for the effect of inbreeding on genetic variance and found for sib-indices that the optimum schemes for maximising gain at different generations (*e.g.* $t = 5$ or $t = 20$) under restricted inbreeding were similar. In essence, by applying a restriction on the accumulation of inbreeding, the loss of variance is also restricted. In any case, comparisons of predicted and simulated gains were free of bias since the rate of gain obtained from simulations also ignored reductions of genetic variance due to inbreeding.

A novel result of this study was the deterministic comparison of the rate of gain for schemes under BLUP optimised and truncation selection at the same ΔF. The results in Figure 4 suggest that, for realistic target ΔF values in commercial breeding populations, the maximum advantage of quadratic optimisation (ΔG_{quad}) over truncation selection (ΔG_{tru}) occurs at h_0^2 ranging from 0.15 to 0.35. In addition, the results indicate that for any level of ΔF, ΔG is higher for optimised than for truncation selection and that for h_0^2 up to 0.55 the superiority of ΔG_{quad} over ΔG_{tru} increases as the population size increases. Benchmark maximum values found for the ratio $\Delta G_{quad}/\Delta G_{tru}$ were 1.40 and 1.45 for the largest scheme (*i.e.* $T = 1,000$), and 1.27 and 1.32 for the smallest scheme (*i.e.* $T = 100$) for $\Delta F = 0.0100$ and $\Delta F = 0.0250$, respectively. The predicted superiority of quadratic optimisation over truncation selection agreed with empirical evidence from real livestock populations of Aberdeen Angus (beef cattle) and Meatlinc (sheep) [1]. At the observed ΔF in these populations, $\Delta G_{quad} / \Delta G_{tru}$ was 1.30 and 1.17 for Aberdeen Angus (for $\Delta F = 0.0020$) and Meatlinc (for $\Delta F = 0.0100$), respectively.

A number of points should be addressed in relation to the predicted superiority of ΔG_{quad} over ΔG_{tru}. First, it might be counter intuitive that $\Delta G_{quad} / \Delta G_{tru}$ increased with the size of the scheme. At first sight it might be expected that the benefits would be the greatest when the constraint on ΔF is hardest to meet (*i.e.* when T is small). However, large resources enable larger family sizes to be available, consequently increasing the accuracy of the evaluation of all terms required, and also allowing for selection and use more closely aligned to the optimum indicated by the quadratic index. Second, it should be noted that values for $\Delta G_{quad} /\Delta G_{tru}$ might be overestimated for two reasons. The benefits

of the quadratic index over truncation selection shown in Figures 3 and 4 include the benefits of factorial mating (assumed with quadratic indices) over hierarchical mating with a mating ratio equal to one (assumed with truncation selection). Nevertheless, the effect of the family structure on $\Delta G_{quad} / \Delta G_{tru}$ is expected to become less important as the scheme size increases [14]. Also, following the approach of Villanueva and Woolliams [16], higher responses could be obtained from truncation selection if the weights given to the different sources of information in the pseudo-BLUP index were optimised for maximising ΔG_{tru} for a given constraint on ΔF. In any case, the results are indicative of the potential benefits of quadratic indices and the potential trends in the benefits for the key parameters involved.

As it stands today, the prediction framework can be readily applied in species with no restriction in reproductive rates (*e.g.* fish and trees) to improve traits with phenotypes available in both sexes. Extensions of this framework, particularly to accommodate reproductive limitations would be relevant to allow its application in livestock populations, particularly of cattle and sheep. Although a mating ratio of one has been assumed here, the current predictions of the Mendelian sampling accuracy allow any mating ratio and are thus not a limitation. In contrast, an expression for the ideal optimal ΔG when reproductive limitations exist needs to be developed.

These results are critical from the point of view of the strategic design of practical breeding schemes, since with this framework breeders will be able to set their risk preference (*i.e.* the target ΔF) and predict the benefit from changing from their current selection practices (namely, truncation selection) to optimised selection. In a subsequent stage, available operational tools [7,10] for making selection decisions on a day-to day basis can be applied. Both the design tool (*a priori*) and the operational (*a posteriori*) tool have the same underlying definition of genetic gain that is expressed in terms of long-term genetic contributions and Mendelian sampling terms [18]. Thus, a prediction framework for ΔG_{quad} is a necessary accompanying tool for the available operational dynamic selection algorithms.

ACKNOWLEDGEMENTS

REFERENCES

[1] Avendaño S., Villanueva B., Woolliams J.A., Expected increases in genetic merit from using optimised contributions in two livestock populations of beef cattle and sheep, J. Anim. Sci. 79 (2003) 2964–2975.

[2] Avendaño S., Woolliams J.A., Villanueva B., Mendelian sampling terms as a selective advantage in optimum breeding schemes with restrictions on the rate of inbreeding, Genet. Res. Camb. 83 (2004) 55–64.

[3] Avendaño S., Woolliams J.A., Villanueva B., Prediction of accuracy of estimated Mendelian sampling terms, J. Anim. Breed. Genet. 122 (2005) 302–308.

[4] Bijma P., Rutten M.J.M., Sel*A*ction: Software for optimisation of breeding programs, in: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, 19–23 August 2002, Montpellier, CD-ROM communication No. 28-15.

[5] Bijma P., van Arendonk J.A.M., Woolliams J.A., Predicting rates of inbreeding for livestock improvement schemes, Anim. Sci. 79 (2001) 840–853.

[6] Bulmer M.G., The effect of selection on genetic variability, Amer. Nat. 105 (1971) 201–211.

[7] Grundy B., Villanueva B., Woolliams J.A., Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development, Genet. Res. Camb. 72 (1998) 159–168.

[8] Grundy B., Villanueva B., Woolliams J.A., Dynamic selection for maximising response with constrained inbreeding in schemes with overlapping generations, Anim. Sci. 70 (2000) 373–382.

[9] Kearney J.F., Wall E., Villanueva B., Coffey M.P., Inbreeding trends and application of optimised selection in the UK Holstein population, J. Dairy Sci. 87 (2004) 3503–3509.

[10] Meuwissen T.H.E., Maximising the response of selection with predefined rate of inbreeding, J. Anim. Sci. 75 (1997) 934–940.

[11] Meuwissen T.H.E., Sonesson A.K., Maximizing the response of selection with predefined rate of inbreeding: Overlapping generations, J. Anim. Sci. 76 (1998) 2575–2583.

[12] Roberston A., The interpretation of genotypic ratios in domestic animal populations, Anim. Prod. 7 (1965) 319–324.

[13] Rutten M.J.M., Bijma P., Woolliams J.A., van Arendonk J.A.M., Sel*A*ction: Software to predict selection response and rate of inbreeding in livestock breeding programs, J. Hered. 93 (2002) 456–458.

[14] Sonesson A.K., Meuwissen T.H.E., Mating schemes for optimum contribution selection with constrained rates of inbreeding, Genet. Sel. Evol. 32 (2000) 231–248.

[15] Villanueva B., Woolliams J.A., Optimisation of breeding programs under index selection and constrained inbreeding, Genet. Res. Camb. 69 (1997) 145–158.

[16] Villanueva B., Woolliams J.A., Gjerde B., Optimum designs for breeding programs under mass selection with an application in fish breeding, Anim. Sci. 63 (1996) 563–576.

[17] Villanueva B., Bijma P., Woolliams J.A., Optimal mass selection policies for schemes with overlapping generations and restricted inbreeding, Genet. Sel. Evol. 32 (2000) 339–355.

[18] Woolliams J.A., Thompson R., A theory of genetic contributions, in: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, 7–12 August 1994, vol. 19, Guelph, pp. 127–134.

[19] Woolliams J.A., Bijma P., Villanueva B., Expected genetic contributions and their impact on gene flow and genetic gain, Genetics 153 (1999) 1009–1020.

[20] Woolliams J.A., Pong-Wong R., Villanueva B., Strategic Optimisation of Short- and Long-Term Gain and Inbreeding in MAS and Non-MAS Schemes, in: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, 19–23 August 2002, Montpellier, CD-ROM communication No. 23–02.

[21] Wray N.R., Hill W.G., Asymptotic rates of response from index selection, Anim. Prod. 49 (1989) 217–227.

[22] Wray N.R., Thompson R., Prediction of rates of inbreeding in selected populations, Genet. Res. Camb. 72 (1990) 41–54.

APPENDIX A: IMPLEMENTATION OF THE NEWTON-RAPHSON METHOD FOR FINDING THE TRUNCATION POINT AND THE SELECTION INTENSITY FOR A GIVEN ΔF CONSTRAINT

In order to find the values of i and x that satisfy $(4T\Delta F)^{-1} = p(i - x)^2(1 - ix + x^2)^{-1}$ for given values of T and ΔF, a starting value of zero was given to x. Then, iterations were performed such as $x_t = x_{t-1} - [f(x_{t-1})/f'(x_{t-1})]$, where $f(x_{t-1}) = [p(i - x)^2/(1 - ix + x^2)] - (4T\Delta F)^{-1}$, $f'(x_{t-1}) = 2p(i - x)[i(i - x) - 1]/(1 - ix + x^2)^2$, and p and i are taken from the truncated normal distribution for the normal deviate x_{t-1}.

APPENDIX B: PREDICTION OF THE RATIO OF EFFECTIVE NUMBER OF PARENTS AT THE TIME OF SELECTION (N_c) TO EFFECTIVE NUMBER OF PARENTS AT THE TIME OF CONVERGENCE OF CONTRIBUTIONS (N_r)

The prediction of N_c/N_r was carried out in several steps. Firstly, the empirical ratios N_c/N_r observed from the simulation were plotted against $1 - h_0^2$ for
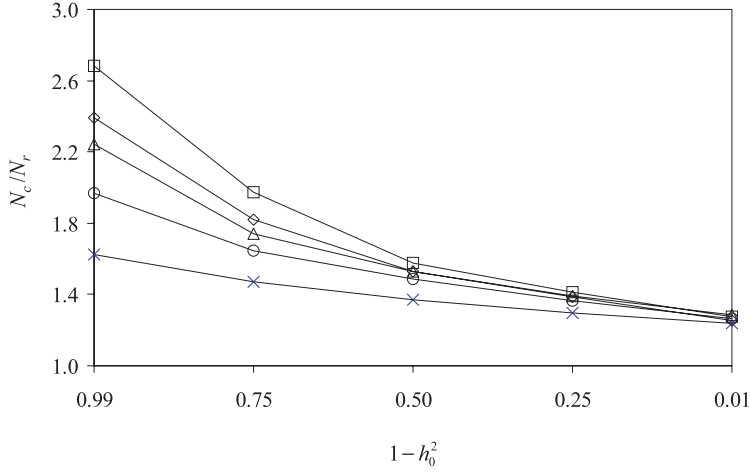
Figure B1. Relationship between the empirical ratio N_c/N_r and $(1 - h_0^2)$ for different levels of $T\Delta F$ resulting from $T = 100$, 200 and 300, and $\Delta F = 0.0250$, 0.0125 and 0.0100. x: $T\Delta F = 1.0$; ∘: $T\Delta F = 2.0$; △: $T\Delta F = 3.0$; ◇: $T\Delta F = 5.0$; □: $T\Delta F = 7.5$.

different values of $T\Delta F$ (Fig. B1). Secondly, the regression coefficients of the double natural logarithm (ln) of N_c/N_r on $1 - h_0^2$, namely $b_{T\Delta F}$, were estimated for several levels of $T\Delta F$ using the model

$$\ln[\ln(N_c/N_r)] = a_1 + b_{T\Delta F}(1 - h_0^2) \tag{B1}$$

where a_1 is the intercept. The double natural logarithm of N_c/N_r was taken to make the regression of N_c/N_r on $1 - h_0^2$ linear. Thirdly, the regression coefficients for each level of $T\Delta F$ were regressed on $\ln(T\Delta F)$ as

$$b_{T\Delta F} = a_2 + b_2 \ln(T\Delta F) \tag{B2}$$

where a_2 is the intercept. A full prediction model can be written by substituting (B2) in (B1):

$$\ln[\ln(N_c/N_r)] = a_1 + [a_2 + b_2 \ln(T\Delta F)](1 - h_0^2)$$

which resulted in

$$\ln[\ln(N_c/N_r)] = -1.459 + [0.755 + 0.367 \ln(T\Delta F)](1 - h_0^2)$$

where the intercept corresponds to the common intercept of the nested regression of $\ln[\ln(N_c/N_r)]$ on $1 - h_0^2$ for each level of $T\Delta F$. After taking antilogarithms, N_c/N_r can be expressed in terms of $1 - h_0^2$ and $T\Delta F$ as

$$\ln(N_c/N_r) = e^{a_1} T\Delta F^{b_2} e^{a_2(1-h_0^2)},$$

and predicted from

$$\ln(N_c/N_r) = 0.2325(T\Delta F)^{0.3671} e^{[0.7553(1-h_0^2)]}.$$

The correlation between the ratio N_c/N_r resulting from the simulation (empirical) and the predicted N_c/N_r was 0.92. This correlation was obtained using different values for h_0^2 and $T\Delta F$ resulting from $T = 100$, 200 and 300, and $\Delta F = 0.025$, 0.0125 and 0.010.