

Making the Web equal profit – surfing for genetics

Dorian J. Garrick

Department of Animal Sciences, Colorado State University, Fort Collins, CO 80523

Introduction

Genetic improvement is a straightforward result that follows the use of truly superior sires within the context of particular production, management and economic circumstances. But how does one identify truly superior sires? The rancher's answer to this question has varied over the last century, partly because of changes in production and economic circumstances but also because of changes in scientific knowledge, education and fashion. Nevertheless, the World Wide Web (www) has contributed very little to the rancher's ability to identify superior animals to increase their profit. This will all change with the next generation of ranchers for two reasons. One is about convenience; the other is about customized computation and decision support.

The convenience of the web

A hardcopy sire summary was a fantastic new resource when they were first released to the industry. However, a major problem with hardcopy is deciding the order in which the bulls will be presented to the reader. A telephone book is a useful resource to find a person's number, but not so useful to find the name of the person that corresponds to a particular phone number. So it is with the sire summary. It is a good resource to look up for details, by name, on one or a few bulls that you know to be of interest. For those interested in extreme animals for a particular trait, separate sections of trait leaders are useful.

However, most ranchers are interested in multiple trait improvement and the bulls that are trait leaders for one trait are seldom the bulls of most interest from a multi-trait perspective. Many breeders practice multiple trait selection through the use of independent culling levels. For example, they will not consider using a bull whose birthweight EPD exceeds a particular benchmark value. In high altitude conditions they may not consider using a bull with a PAP EPD that exceeds a particular threshold. Some breeders avoid animals that are at either high or low weight extremes. The web delivery of so-called database queries linked to an on-line sire summary provides a fantastic tool for filtering bulls according to any number of criteria and then sorting the subset of bulls that meet the filter. Commonly-used filters might be on the minimum or maximum EPD, the accuracy of the EPD, the breeder name or ranch

location. Most Breed Associations already provide on-line access to the results of their evaluations.

The true genetic merit of a bull does not change from the moment of conception through the time the bull produces its own performance records and then produces offspring after use as a sire. However, our estimate of the merit of the bull (EPD), typically does change over time, starting with the parent average value for each trait with that estimate being modified upwards or downwards according to the superiority or inferiority of the individual relative to its contemporaries and then again based on the superiority or inferiority of its offspring. Bulls that are widely used in many regions will be continually accruing new information which can improve the EPD estimate. However, for historical reasons, most national beef cattle genetic evaluations are only undertaken once or twice a year. This requires that a deadline be used to determine the information that is included in any particular analysis. Any information collected after that deadline will not be used to improve the accuracy of sire EPDs until the next national analysis, six months or a year later. In the future, continuous genetic evaluation will become a reality. Continuous national evaluation is already used in livestock evaluations in other countries. Continuous evaluation provides some challenges for hardcopy sire summaries as these may become outdated before they can even be widely distributed. Web delivery can overcome these problems and readily provide the most up-to-date information. However, there are many more compelling reasons to surf the web in search of more profitable sires than simply the convenience of electronically sorting through lists of bulls.

Customized Computation using the web

Making profitable selection decisions on a repeatable basis requires one to simultaneously quantify the consequences of using particular animals as parents across a portfolio of economically relevant traits. Typically, sires that are more popular enjoy a premium price such that identifying the most profitable option involves weighing up the benefits of a particular sire in comparison to the cost. One of the early tools for quantifying the relative impact of alternative sires was the EPD. From a simplistic viewpoint, the EPD would appear to provide the required information. For example, suppose a cow-calf rancher sells animals at

weaning. If a particular sire has a weaning weight EPD that is 20 lb above an alternative sire, it might be argued that this is all the information needed to determine the benefit of that sire. However, there are at least six reasons why such EPDs are not sufficient information to make good decisions without further analysis. These six reasons have their basis in statistics, genetics, systems biology, nutrition, economics and probability respectively. A consequence of these six issues is that arithmetic analysis of the entire portfolio of EPDs of a particular sire along with other genetic, production, management and economic circumstances are required to provide the utility that one might have traditionally (naively) expected from an EPD. This arithmetic analysis might be referred to as comprising sire selection by simulation (Bourdon, 1998). The web is ideally suited for providing such an analytical tool. These six reasons will now be explored in more detail.

1. Interpretation of threshold traits

Threshold analysis is commonly undertaken on categorical data, such as calving ease (Hoeschele et al., 1995). Philosophically it assumes that there is underlying continuous genetic and environmental variation with respect to the trait, but the phenotypic observation that is recorded by the rancher is limited to two or a few ordered categories. A threshold dictates which category is observed for given values of the underlying genetic and environmental factors and the fixed effects. Threshold analyses may concurrently account for correlated continuously observed variables. For example, the analysis of calving ease typically account for correlated information on birthweight as an indicator trait. Threshold analyses are also used for stayability (score 1 reflects a cow that was in the herd as a two-yr old and stayed in the herd up to or beyond six years of age or five calving opportunities or score 0 indicates the cow failed to stay that long) and for heifer pregnancy (score 1 reflects a heifer that was confirmed pregnant whereas score 0 reflects a heifer that failed to get pregnant).

The genetic merit computed from a threshold analysis is on an underlying scale that does not have conventional measurement units. Accordingly, estimates of merit from a threshold model are converted or transformed from the underlying scale back to the original observed scale of measurement in terms of a probability. A feature of this transformation is that it requires assuming some average circumstance or average incidence of the various observed categories. Put another way, this means that a sire that will provide a given shift in the underlying scale will be equivalent to different observed effects (or progeny

differences) in different production circumstances. This is best demonstrated by graphic example as in Figures 1 & 2 for a bull with an underlying EPD that improves heifer pregnancy by 0.38 units. In the herd depicted in Figure 1, this bull will improve heifer pregnancy by 8% (8 calves per 100 heifers) by lifting pregnancy rate from 80 to 88%. The same bull used in herd 2 will only lift the pregnancy rate by 4.5%, from 90 to 94.5%.

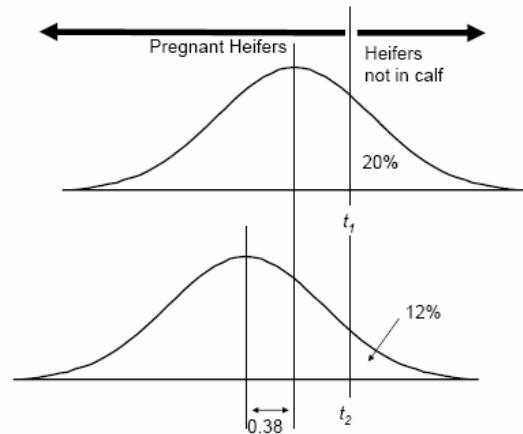


Figure 1¹. The effect of a sire that improves heifer pregnancy (in his daughters) by 0.38 underlying units in a herd with an average 20% heifers not in calf is to increase the heifer pregnancy rate from 80% (100-20) to 88% (100-12).

A consequence of the underlying nature of these traits is that the published EPD and the progeny difference you expect to observe in your particular circumstances will not be the same unless your average levels of performance happen to be exactly the same as the value that was used in the published transformation. An incidence of 50% is typically used in the published transformation. This is close to the average incidence for stayability (where a typical herd has about half the number of six-year old cows compared to two-year old cows), but is not realistic for heifer pregnancy (which is typically 75-85%) or calving ease.

¹ Normal distribution tables are required to show that the threshold t_1 is at 0.84σ in order for 20% heifers to exceed the threshold. For a phenotypic $sd=1.17\sigma$, then threshold t_2 is at $0.84+0.38/1.17=1.165\sigma$. Normal distribution tables can be used to show that the area to the right of that threshold will be 12%.

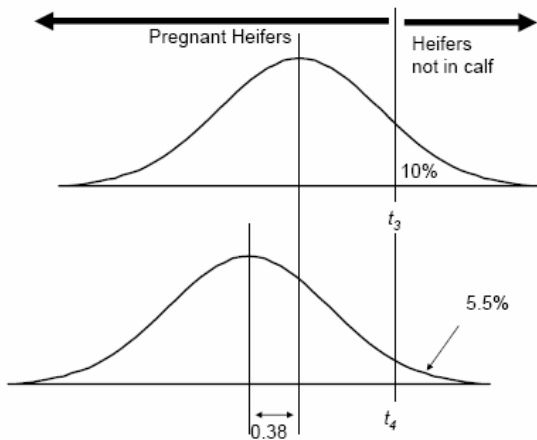


Figure 2². The effect of the same sire as in Figure 1 that improves heifer pregnancy rate by 0.38 underlying units in a herd with an average 10% heifers not in calf is to increase the pregnancy rate from 90% to 94.5%.

The value of all threshold trait EPDs (heifer pregnancy, stayability and calving ease) thus varies with the current average level of performance. The effect on the observed scale of a given shift in the underlying scale is greatest if the average is 50% and declines as the average increases or decreases from that value. Furthermore the effect of a given increase (eg +0.38) or the equivalent decrease (-0.38) on the underlying scale is not typically the same on the observed scale.

It is not appropriate or practical for a hardcopy sire summary to produce these threshold EPDs for more than one observed scale. Altering the transformation will have no influence on the ranking of the sires, but it may have enormous influence on the productive consequences that would result from using a particular bull in your own circumstances. The web provides the ability to undertake the calculations required for a custom transformation so that the consequences of using a particular bull can be more appropriately quantified for your circumstances.

2. Multibreed evaluation and crossbreeding

Theoretically, the breeding value (or EPD) is determined by the sum (or half the sum) of all the

² Threshold t_3 must be at 1.28σ in order for 10% individuals to exceed the threshold. For a phenotypic $sd = 1.17 \sigma$, then threshold t_4 is at $1.28 + 0.38 / 1.17 = 1.605 \sigma$ above the mean giving only 5.5% above that threshold.

average effects of the genes an individual carries. The effect of a gene is not generally expected to be the same in every population, due to dominance effects and to differences in gene frequencies between populations. In practice, various studies have indicated that EPDs perform reasonably well in terms of predicting differences in performance, in different populations, including in different breeds, with two exceptions. The first occurs when the environmental circumstances differ markedly in terms of nutritional, climatic or disease stress. The second exception occurs when the bulls comprise more than one breed or cross, or their mates represent more than one breed or cross. There are two possible explanations as to why EPDs may not perform in the context of such crossbreeding. The first is that the EPDs from each breed may be on a different base. This is currently the case with most evaluations where the EPD base is set independently of the base for other breeds. Using records from USDA's Clay Center, breed base adjustments for mainstream traits are published annually at BIF meetings (eg Van Vleck and Cundiff, 2004). The second explanation is due to the phenomenon of heterosis or hybrid vigor.

The National Beef Cattle Evaluation Consortium (NBCEC) is currently prototyping, at Cornell University, a multibreed analysis for many US beef breeds for growth traits. Similar analyses for carcass traits and for calving ease will soon be prototyped at the University of Georgia. These analyses will produce EPDs for all breeds that are comparable on the same base. It remains to be seen whether the resulting EPDs will be published on a common base or whether each Breed Association adjusts the results back to their usual base. It seems counterproductive to pool data across many breeds for joint analysis and then to report breeds on different bases. Multibreed dairy cattle analyses in New Zealand have been published on a common across-breed base for some years (for examples see www.aeu.org.nz).

In order to interpret the expected progeny performance in an across-breed setting, heterosis values must be taken into account. The heterosis values are expected to be different for each trait. Theoretically, the amount of heterosis is also a reflection of the genetic distance between breeds, as heterosis is simply recovering historical losses in performance that have resulted from inbreeding. Accordingly, the heterosis among breeds of similar type is expected to be more similar than among breeds of different backgrounds. Heterosis values can be predicted from multibreed evaluation systems. However, experience with the data structures that are represented within the context of US national beef cattle evaluations shows that Breed Association datasets are not particularly reliable for estimating

Table 1. Yearling Weight EPDs used for theoretical and actual calculation of expected performance from two Angus and two Simmental (Simm) sires mated to different cow breeds^a.

YWT	Offspring Performance by cow breed					
			Angus	Simm	Hereford	Angus-Simm
	EPD	EPD				
Breed/Bull	Within-breed	Across-breed				
Angus 1	+65	+65	Base	+h _{AS}	+h _{AH}	½h _{AS}
Angus 2	+80	+80	+15	+h _{AS} +15	+h _{AH} +15	½h _{AS} +15
Simm 3	+58	+80	+h _{AS} +15	+15	+h _{HS} +15	½h _{AS} +15
Simm 4	+68	+90	+h _{AS} +25	+25	+h _{HS} +25	½h _{AS} +25
Angus 1			850 lb	863	873	857
Angus 2			865	878	888	872
Simm 3			878	865	878	872
Simm 4			888	875	888	882

^a Angus 1 is breed average (Cundiff, 2004) for yearling weight. Angus 2 is 15 lb superior for yearling weight EPD. Simmental 3 is breed average for yearling weight. Simmental 4 is 10 lb above breed average. Across-breed EPDs are on an Angus base. The base adjustment for yearling weight in the Simmental breed is 22 lb (Van Vleck and Cundiff, 2004). Crossbred cows are ½Angus ½Simmental. Heterosis values for yearling weight F1’s are taken to be different between each pair of breeds and are h_{AS} = 13 lb for Angus-Simmental, h_{AH} = 23 lb for Angus-Hereford, and h_{HS} = 13 lb for Hereford-Simmental.

heterosis. The approach originally adopted by Cornell University (Pollak and Quaas, 1998) and now used more widely involves a Bayesian procedure that introduces prior knowledge on likely heterosis values from previous well-designed published studies.

The use of within-breed and across-breed EPDs to predict crossbred performance is best demonstrated by example considering yearling weight in offspring resulting from Angus or Simmental sires over Angus, Simmental, Hereford or crossbred cows. The upper portion of Table 1 demonstrates the nature of the calculations for predicted performance using across-breed EPDs whereas the lower portion presents possible numerical values of EPDs given assumed base adjustments and heterosis values.

The example in Table 1 demonstrates that the ranking of the four sires is sensitive to the nature of the cow breed. Using straightbred Angus cows, the ranking of the sires for yearling weight is 4>>3>>2>>1. Using straightbred Simmental cows the ranking depends upon the value of h_{AS} and is 2>4>>3>1. Over Hereford cows the ranking depends upon the relative magnitude

the heterosis values h_{AH} and h_{HS} and is 2=4>>3>1. Over these crossbred cows, the ranking is 4>>3=2>>1.

The consequences of these results are that any rancher intending to use bulls from more than one breed must currently deal with bulls listed in more than one sure summary. They must then know where to find and how to use the base EPD adjustments relevant to their circumstances. They must also know where to find and how to use the appropriate heterosis values. The situation is even more complex in multibreed circumstances for maternally-influenced traits where maternal and direct heterosis will have different coefficients when the dam and the offspring breed composition are not identical. All these problems can be overcome by web delivery of a single file of all EPDs from all breeds in a multibreed analysis. The arithmetic (shown in Table 1) to adjust for base (if required) and to account for heterosis can be readily achieved behind the scenes for the convenience of the user.

The shift to routine multibreed analysis will introduce challenges for Breed Associations in terms of data

Table 2. Influence of example EPDs on a number of economically relevant traits on sale weight at weaning from a 1,000 cow herd.

Bull ID	EPD ¹				Performance		
	WWD	STAY	HPG	CED	Weaning Wt	#Sold	Extra Wt sold per cow
Romeo	Average	Average	Average	Average	451 lb	744	Base ²
Sierra	+30 lb	Average	Average	Average	481 lb	744	22 lb
Tango	+30 lb	+8%	Average	Average	482 lb	759	30 lb
Uniform	+30 lb	+8%	+12%	Average	482 lb	765	33 lb
Victor	+30 lb	+8%	+12%	+11%	482 lb	766	34 lb

¹ EPDs are for weaning weight direct (WWD), Stayability (STAY), Heifer Pregnancy (HPG) and calving ease direct (CED).

² Base herd of 1,000 cows weans 938 calves at an average 451 lb and sells 358 lb calf per cow wintered.

deadlines. An obvious approach to overcome this aspect is to provide continuous genetic evaluation. Each Association would upload their pedigree and performance information at their convenience, knowing that the next analysis (perhaps monthly, weekly or daily) will use all that information.

3. Interactions between economically relevant traits

Many of the economically relevant traits interact, to the extent that the observed differences in actual performance are not identical to those that would be predicted from EPDs. This occurs for even the simplest of traits, such as weaning weight. It is best demonstrated by example (Table 2), considering five alternative bulls used to generate all the replacements in the context of a straight-bred self-replacing cow-herd.

Relative to a base herd scenario (using Romeo), the use of a bull such as Sierra with an increased weaning weight direct of 30 lb will increase the average weaning weight by 30 lb provided all other EPDs are unchanged. Some cows fail to rear a live calf to weaning and a proportion of the weaned heifers need to be retained as replacements. Accordingly, the additional sale weight expressed per cow is 22 lb. Tango is a bull that has increased stayability as well as increased weaning weight direct. Increased stayability impacts sale weights at weaning in two ways. First, there are fewer heifer replacements required so more female calves can be sold at weaning. Second, the cow herd has a smaller fraction of first calvers and a larger fraction of mature cows. This increases average weaning weight as mature cows wean larger calves than first calvers. The combined effect of increased stayability and increased weaning weight direct leads to an increase of 30 lb saleable weaning weight per

cow. Uniform is a bull with all the features of Tango in addition to an improved EPD for heifer pregnancy. For the same number of required pregnant rising two-year old replacement cows, Tango's daughters need fewer weanlings retained. This has the effect of increasing weaner sale numbers by a further 3 lb per cow compared to Tango, giving 33 lb more than Romeo. Victor is a bull with all the features of Uniform, in addition to improved calving ease direct. Calving ease and birthweight are traits that many producers emphasize. In this herd, with 22% first-calf heifers requiring assistance, the improved calving ease only results in an additional 1 lb sale weight per cow.

This example demonstrates that even a simple production system involving 1,000 cows with a straightforward goal based on weaning sale weight will be influenced by a portfolio of trait EPDs in addition to the obvious influence of weaning weight direct. Weaning weight direct EPDs alone are not a good indication of system performance, even for weaning weight.

4. Assessment of nutritional (input) implications

The nutritional or dry matter intake requirements of a cow herd and its replacements depend upon a number of factors. The major requirement for feed is in supporting the maintenance requirements of the mature cows. From the perspective of a typical mixed-age mature cow, this is influenced by its mature size, its condition score at maturity and its milk production potential. Added to these requirements, mature cows need feed to support gestation (varying with birth weight) and lactation (varying with milk production). The calves themselves require feed for maintenance

and growth up to sale age (eg weaning) and to meet the requirements for replacement heifers. The replacement rate in terms of number of cows at first calving will vary with the stayability of the herd. The number of heifer calves that need to be retained to provide sufficient replacements will vary further with heifer pregnancy rate. In some circumstances, any change in the feed requirements of the cow herd can be met by purchasing in feed at some given feed cost. In other extensive grazing cases, the primary source of feed for the cow herd, their calves and replacements, is provided by the amount of pasture produced. This is principally determined by the land area and the amount of precipitation.

In order to identify the impact that sires will have on profit, it is therefore necessary to predict the feed requirements that will result from their use. Given this information, the cost of additional purchased feed can be determined, or the required modification to the stocking rate can be quantified. Computing the nutritional requirements of a cow-calf herd according to its age structure and other aspects of performance should be straightforward for any well-educated Animal Science graduate, but nevertheless requires access to the relevant tables (NRC, 1996) and a considerable amount of arithmetic. Both this knowledge and the arithmetic can be readily provided via a web-based decision support tool.

For example, consider a livestock system using the sires that were introduced in Table 2. Suppose the ranch environment was capable of supporting 1,000 cows of merit reflective of Romeo. Replacing the herd with daughters of Sierra with 30 lb increases in weaning weight (but no change in birth, yearling or mature weights) would increase nutritional requirements to support the faster pre-weaning growth. This would require a reduction in stocking rate equivalent to 2 cows to give a total herd of 998 cows in order to consume the same amount of feed on an annual basis. Increasing herd stayability through the use of Tango will reduce the number of replacements required to be kept postweaning, allowing 999 cows to be calved. Increasing heifer pregnancy with Uniform will further reduce heifer retention allowing herd size to increase to 1001 cows for the same annual feed consumption. The use of sires that modify maintenance energy requirements (through altered mature size or milk production potential) will have much more dramatic influence on the number of cows that can be managed in order to achieve the same annual feed consumption than do the examples above. The high genetic correlation between weights at various ages results in most bulls with higher growth rates being associated with higher mature size and

maintenance energy. Failure to properly account for any such increases in nutritional requirements and resultant feeding costs will lead to overestimation of the value of improved growth and a tendency to overlook more profitable bulls with more moderate growth and improved stayability.

5. Assessment of financial implications – accounting for prices and costs

We have already demonstrated the need for considerable knowledge and arithmetic in order to properly quantify the levels of outputs and the number of inputs required for a particular production and management circumstance. In order to surf for profit, one must then combine the outputs according to their output values (which may vary with quality attributes that are modified by selection) and subtract the cost of inputs. Feed costs, costs that vary with the number of cows, veterinary and other labor costs may need to be taken into account. All this is easily achieved using the web. The example bulls used in Table 2 show that Sierra increases profit by about \$22 per cow (relative to Romeo) and Tango, Uniform and Victor increase profit by \$33, \$34 and \$35 respectively in the particular management, productive and economic circumstances modeled. These increases in net income are well worthwhile when one considers the number of cows a bull can breed over its lifetime.

6. Accounting for risk associated with the use of bulls with less than perfect accuracy

A few so-called proven bulls may have high accuracy EPDs for some traits, indicating that the current estimate of bull merit is unlikely to change much if additional, new information became available. Even proven bulls are likely to have some trait EPDs with reduced accuracy, such as traits that are measured late in life or after slaughter including stayability, heifer pregnancy, maintenance requirements or perhaps carcass merit. Most bulls will have only moderate accuracy EPDs as they will have been evaluated based on their parental and individual performance, without yet having the benefit of recorded offspring. Such bulls are equally likely to be better than their current estimate suggests, or worse than their current estimate. Limiting selection to proven bulls is therefore overlooking some of the young bulls that will turn out to be much better than can be currently assessed. The hardcopy sire summaries typically publish likely change tables, but it is no easy matter to simultaneously determine the impact of profit on inaccuracies in a whole portfolio of economically relevant trait EPDs. Sire selection by simulation can

achieve this end, by simulating a number of possible realizations of each bull and determining the distribution of likely profit that will result from using each bull. This can be delivered by web-based decision support tools.

Web-based decision support is not just another index

It could be argued that the same kind of models that might underlie a decision support tool can be used to construct relative economic values for each economically relevant trait. Such weights could then be used to combine the individual trait EPDs into a single index figure, to reflect profitability (see for example MacNeil, these proceedings). Some of the six factors addressed above can be accounted for in a selection index. These include the interpretation of threshold traits, some of the interactions between economically-relevant traits, the assessment of nutritional requirements and the financial implications. However, the index must assume average values for many characteristics and to the extent that your circumstances may vary from that average, the index may be sub optimal. Index selection will not account for the multibreed context unless an index is created for every mating strategy. Nor will they typically account for risk.

Finally, there is a philosophical distinction between index selection and decision support. Index selection essentially makes decisions for you with little clarification as to why particular animals get the rankings and index values they receive, other than what might be able to be determined by inspection of the index weights. There is nothing wrong with this if you believe in the index and there are many examples of the positive improvements that can be achieved from the use of index selection.

Web-based decision support is more than simply on-line customized indexes. It can also provide justification as to why particular animals get the values they get. In this context it supports your decision by providing you with relevant information as to the ramifications of selection with respect to your production, management and economic circumstances.

Summary

Web-based sire selection allows you to select sires with quantified prediction and to scrutinize your resulting whole system performance, along the same lines that EPDs and indexes had attempted to provide. It can account for the peculiarities of threshold trait

interpretation, the complexity of trait interactions, the knowledge of heterosis in crossbreeding contexts, and the arithmetic for predicting nutritional and economic ramifications.

The National Beef Cattle Evaluation Consortium (NBCEC) is developing such a decision support tool in concert with its other research activities regarding new EPDs for economically-relevant traits (such as maintenance energy) and multibreed evaluation. The prototype website is accessible at <http://ert.agsci.colostate.edu>.

Many current ranchers will never use web-based decision support. However, those early adopters of this technology have the opportunity to identify sires that they can be confident will increase their profit, rather than using sires that may simply lead to genetic change, without providing genetic improvement.

References

- Bourdon, R.M. 1998. Shortcomings of current genetic evaluation systems. *J. Anim. Sci.* 76:2308-2323.
- Cundiff, L.V. 2004. Mean EPDs reported by different breeds. Proceedings of the Beef Improvement Federation 36th Annual Research Symposium and Annual Meeting, Sioux Falls, SD. 36:62-63.
- Hoeschele, I., B. Tier and H.U. Graser. 1995. Multiple-trait genetic evaluation for one polychotomous trait and several continuous traits with missing data and unequal models. *J. Anim. Sci.* 73:1609.
- NRC. 1996. Nutrient requirements of beef cattle. (7th Edition). National Academy Press, Washington, DC.
- Pollak, E.J. and R.L. Quaas. 1998. Multibreed genetic evaluations in beef cattle. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Armidale, Australia. Vol 23.
- Van Vleck, L.D. and L.V. Cundiff. 2004. Across-breed EPD tables for the year 2004 adjusted to breed differences for birth year of 2002. Proceedings of the Beef Improvement Federation 36th Annual Research Symposium and Annual Meeting, Sioux Falls, SD. 36:46-61.