Data Filters....Risk Managing Software

By Larry Keenan, Mark Enns and Scott Speidel

During the 54 year history of the Red Angus Association, leadership has made a consistent effort to provide members and their bull customers with the most reliable, unbiased genetic predictions in the beef industry. Many critical decisions have led to the achievement of this goal. Two highlights are mandatory performance reporting through Total Herd Reporting; and formation of a Technical Committee to remove any opportunity for politics to influence the science of EPD calculations. However, quite possibly the most significant decision occurred a decade ago with the implementation of data filters to remove data that simply is not biologically probable in a contemporary group.

Data Filters

Developed and owned by Colorado State University Center for Genetic Evaluation of Livestock (CSU CGEL), the data filtering software is simply one component of the bi-annual RAAA National Cattle Evaluation (NCE). As with all of the software that encompasses the NCE, the filter software is housed at CSU CGEL so that there are no issues with the integrity of the software's use in the RAAA analysis. The RAAA maintains a "hands-off" policy with the filter technology and does not use this software to randomly investigate member integrity. Additionally, RAAA is not privy to the specifics of the software's operation, which avoids any breeders having unfair knowledge of the system.

How does this software work?

In contrast to popular belief, observations are not filtered simply due to a predefined ratio or standard deviation. Rather, they are based on complex statistical procedures that associate a probability with each observation. In all, four separate filters are used to identify unlikely observations for both birth and weaning weights. These filters assign probability values to observations within a contemporary group on the basis of whether or not each observation would be naturally occurring given the current underlying RAAA genetic variation. If an animal's observation is identified as highly unlikely, meaning they received a low probability of occurring naturally in their contemporary group, that observation is "filtered". In other words, that observation or piece of performance data is removed from the NCE. The animal itself remains in the evaluation but its performance for that particular trait (birth weight or weaning weight) is not used in the analysis. In some instances, an entire contemporary group can be excluded from the NCE.

The probability beyond which an animal has its own record removed from the evaluation is set by the RAAA technical committee. In order for an individual's observation to be removed from the analysis, its probability must be less than 1 in 10 thousand for birth weight and less than 1 in 100 thousand for weaning weight. For example, if the probability of an animal's weaning observation is 1 in 1 million, it will get filtered. If the probability of that animal's weaning observation is 1 in 50 thousand it will remain in the analysis. It is important to remember that no matter what level of tolerance is used in the filter, there is a very remote chance that the animal's observation is reflective of true genetic merit. If the animal is truly a great outlier, its other performance measures and the performance of its progeny will eventually confirm their level of performance and the EPD will reflect that superiority. So, while an individual's measurement may be filtered, such animals always have the potential to prove their superiority through their progeny's performance.

In the scenario of an animal getting a performance observation filtered out of the NCE, its EPDs will be calculated using information on relatives (the individual's sire, dam, and possibly progeny) and other related traits. For example, we know that a portion of the genes that have an impact on birth weight also have an impact on weaning weight. Therefore, if an animal gets their weaning weight filtered, their weaning weight EPD can be calculated based on the weaning weight genetics of the Sire and Dam plus birth weight information taken on the animal itself. As stated above, eventually performance of these animal's progeny will confirm whether the extreme measure truly represented the genetic merit of that individual.

Data Filters....Risk Managing Software

Remember, some filters may remove the performance of individual animals while others may remove the performance of entire contemporary groups. While removing performance from entire contemporary groups may seem extreme, it ensures that this performance data is of high quality. As an example, let's say birth weights on 20 calves in a single contemporary group were submitted and they ranged from 69 to 71 pounds. Even if all of the calves were by the same sire, this is not sufficient variation, and EPDs generated from such data would also not reflect true genetic difference in birth weight between animals in the contemporary group. In these instances, data on a trait from an entire contemporary group might be filtered to further ensure that EPDs reflect true genetic differences. In the grand scheme of things, the amount of data removed is quite small. The chart on the right shows the percentage of birth and weaning records removed for a given birth year.

It is important to realize that proper management grouping of your animals plays a key role in EPD calculations, which include the filter software. All animals given the same management code should have been provided equal opportunity to perform.
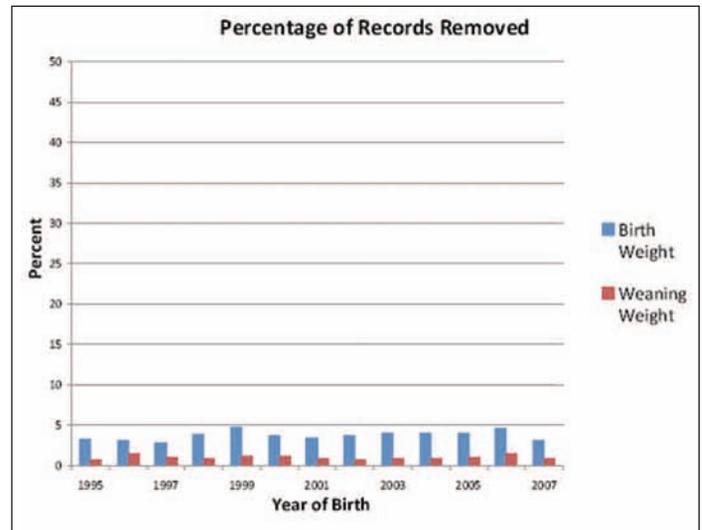
Animals that are incorrectly identified through producer supplied management codes could increase the likelihood of having an observation filtered from the National Cattle Evaluation.

Data filters play an important role in maintaining the reliability of the RAAA EPD calculations; ensuring Red Angus EPDs are a more meaningful prediction of that animal's true genetic merit. ■