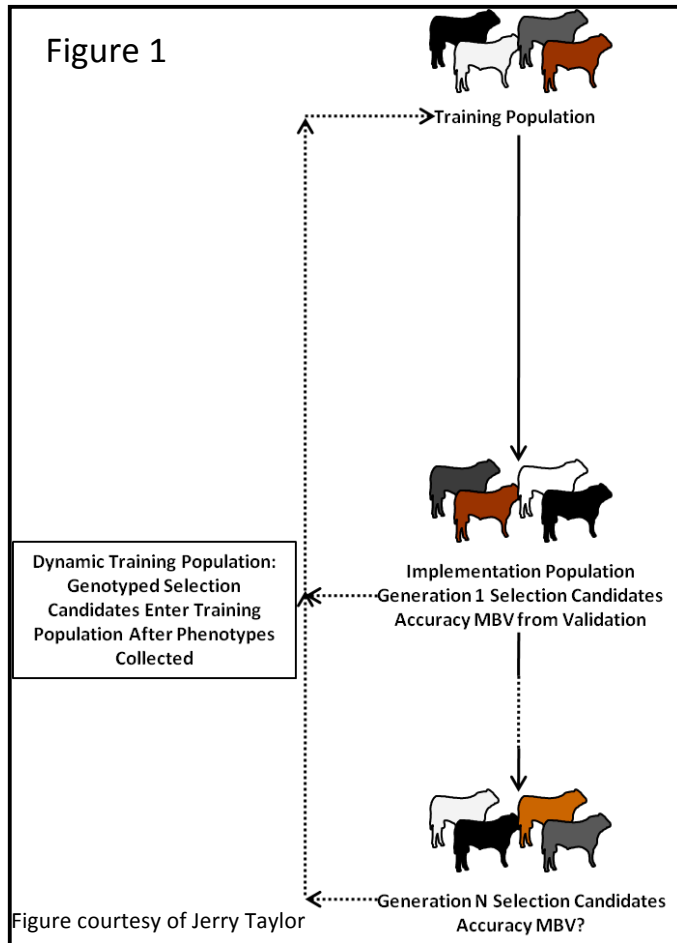


Decay in MBV Accuracy Over Time

Megan M. Rolf, Oklahoma State University

Generating molecular breeding values (MBVs) for beef cattle is a two-step process. In the first step, we use a training population (See Figure 1) to estimate marker effects. These animals are typically from a resource population and generally possess high density genotypes (50K+) along with unique and interesting phenotypes or high accuracy EPDs. The marker effect estimates obtained in this population of animals are applied in the implementation population.



When this process is performed by a testing company, these procedures are performed “in house” and the only data returned is a single value for each trait. When this analysis is performed independently, a breeding value for each trait is generated, but the data can be retained for further use. In both cases, this data is incorporated into the genetic evaluations at the breed associations and results in increased accuracy of EPDs for the genotyped selection candidates (bulls or replacement females).

This process operates recursively, and without the addition of new data, the prediction accuracies will decay over time. This decay in prediction accuracy necessitates “retraining” the data, where all animals that have previously been genotyped now move to the training population and marker effects are re-estimated. Ownership of the genotype data will dictate how this process is performed. The rate of decay seen for genomic predictions is not well established.

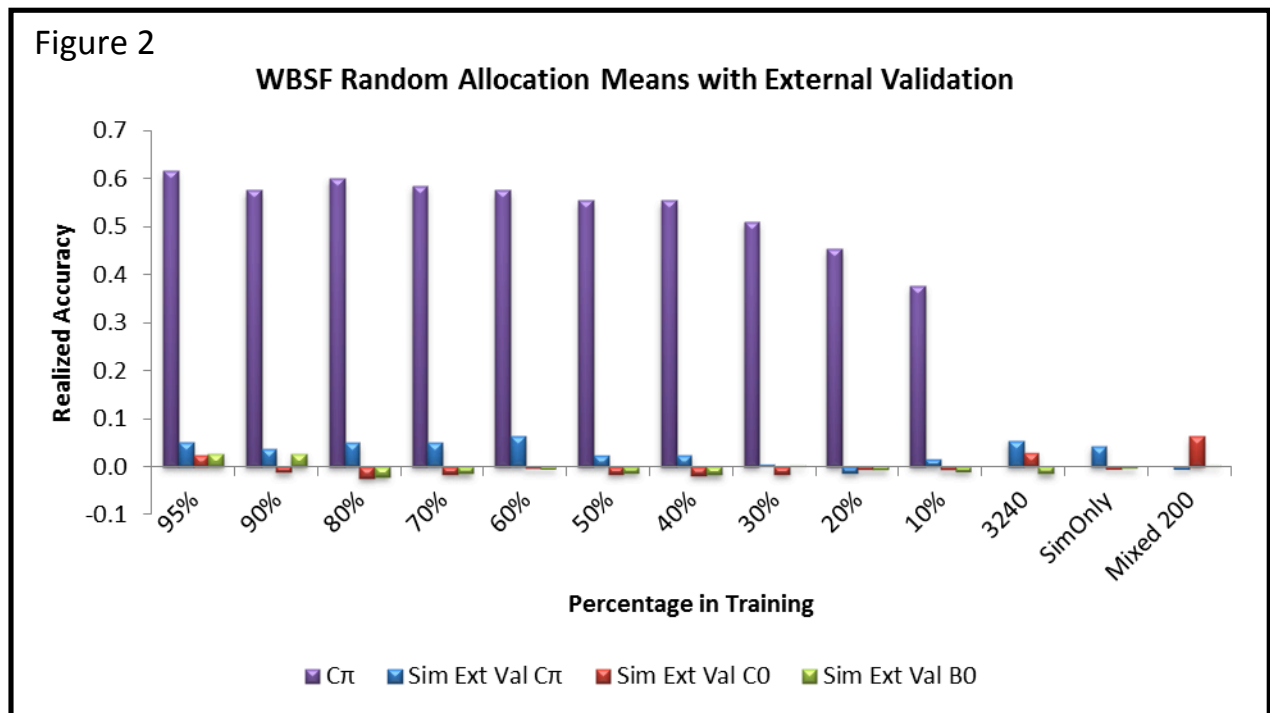
However, MBV accuracies in purebred populations generally depend on two different factors:

1. Accuracies of EPDs for animals used in training
2. Extent of relatedness between training and implementation populations

MBV accuracies do not appear to be directly influenced by the SNP density used in their estimation provided that a sufficient density is used. Using animals with higher initial accuracies in the training population coupled with high degrees of relationship between animals in the training and implementation populations will increase prediction accuracies. Further research is essential to

determine how often re-training needs to be performed in each population to prevent the associated decay in accuracy seen over time.

The University of Missouri has genotyped animals (n=3,240) from the NCBA Carcass Merit Project (CMP) on the 50K SNP chip. The animals in this population were comprised of five different beef cattle breeds (Angus, Charolais, Hereford, Limousin, and Simmental). Animals were sampled randomly from this population and were used as the training population to predict marker effects for Warner-Bratzler Shear Force (WBSF, steak tenderness). Anywhere from 10% to all of the animals were used in the training population in this study. The implementation population consisted of either CMP animals that were not used in training (purple bars) or 443 Simmental steers and heifers that also possessed WBSF phenotypes (blue, red, and green bars) but were separated by approximately 8-10 years (and an unknown number of generations) from the CMP population and showed little relationship to the original CMP animals. Results are shown in Figure 2 below.



When the implementation population consisted of CMP animals (purple bars), prediction accuracies were quite high (approximately 0.6). However, when the implementation population consisted of the independent external population of Simmentals stratified over time, the prediction accuracies were approximately 90% lower (0.6-0.05). Such a drastic reduction in prediction accuracy over a short period of time underscores the imperative to retrain prediction models in a timely manner to retain accuracy in the predictions as well as ensuring a broad sample of animals from a breed to ensure that all animals in the implementation population are related to animals within the training population.