The impact of utilizing previous generations of genotyped animals in genomic selection.

Jeremy T. Howard[1] Steve D. Kachman[2] Matt L. Spangler[1]

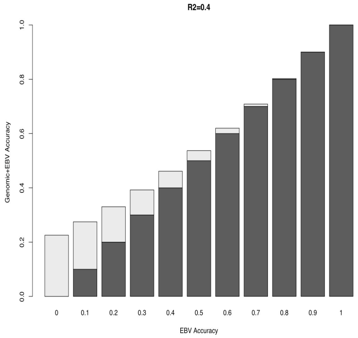
[1]University of Nebraska, Department of Animal Science
[2]University of Nebraska, Department of Statistics

s-jhowar20@unl.edu

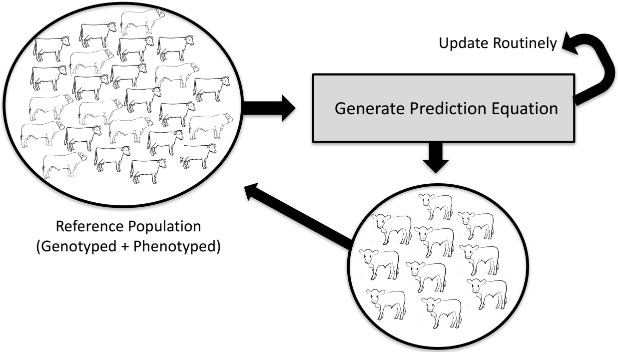
11th World Congress on Genetics Applied to Livestock Production
Auckland, New Zealand — February 11-16, 2018

Introduction

- The use of genetic markers to predict breeding values (GEBV) has become a routine practice across the majority of all major livestock species.
- As a result, a large number of animals (> 2 million Holstein) now have genotype information.
- As the number of genotyped animals increase:
  - Computing the inverse becomes an issue.
  - Computation time increases.



Introduction



Reference Population
(Genotyped + Phenotyped)

Selection Candidates
(Genotyped)

How much information are selection candidates getting from animals born ”n” generations ago?

Objectives

- Use simulation to determine the impact of truncating data on older animals on the long-term genetic gain when genomic selection is practiced.
  - Impact of replacement rate.
  - Impact of trait heritability.
  - Impact of breeding value prediction method.

Materials and Methods

```
########################################################################
#########################################################   ############
#####################################################   /~|   ##########
##################################################   _- `~~~', #########
################################################  _-~       )  #########
#############################################  _-~          |  #########
##########################################  _-~            ;  ##########
################################  __---___-~              |   ##########
#############################   _~   ,,                  ;  `,,  #######
###########################  _-~    ;'                  |  ,'  ; #######
#########################  _~      '                    `~'   ; ########
##################   __---;                                 ,' #########
##############   __~~  ___                                ,' ###########
###########  _-~~   -~~ _          N                    ,' #############
########### `-_         _           C                  ; ###############
#############  ~~----~~~   ;         S                ; ################
###############  /          ;         U               ; ################
#############  /             ;                      ; ##################
###########  /                `                    ; ###################
#########  /                                      ; ####################
#######                                            #####################
########################################################################

------------------------------------------------------------------------
- GENO-DIVER                                                           -
- Complex Genomic Simulator                                            -
```

- Complex Genomic Simulator
  - Historical Population generated based on coalescence methods (MaCS; Chen et al. 2009) and then generations created forward-in-time.
  - Quantitative and Fitness traits (i.e. lethal or sub-lethal mutations).
  - Ability to utilize prediction models based on genomic information.

Materials and Methods

Simulation Scenarios

- Replacement Rate:
  - Low: male (30 %) and female (24 %); maximum of 6 generations.
  - High: male (50 %) and female (40 %); maximum of 3 generations.
- Heritability:
  - Low: 0.20 (dominance variance: 0.025).
  - High: 0.40 (dominance variance 0.05).
- Prediction Method:
  - All markers have an effect (GBLUP)
  - Variable selection (BayesC).

Simulation Scenario Summary

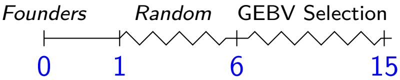
low_rep high_h2	high_rep high_h2
low_rep low_h2	high_rep low_h2

Materials and Methods

Genome and Population Design

- Five chromosomes with a length of 100 Mb.
- Each chromosome contains 105 QTL and 2100 markers.
- The breeding population consisted of 50 males and 400 females.
- Progeny were selected and parents culled based on high GEBV.
- Mating was at random and 1 progeny produced.

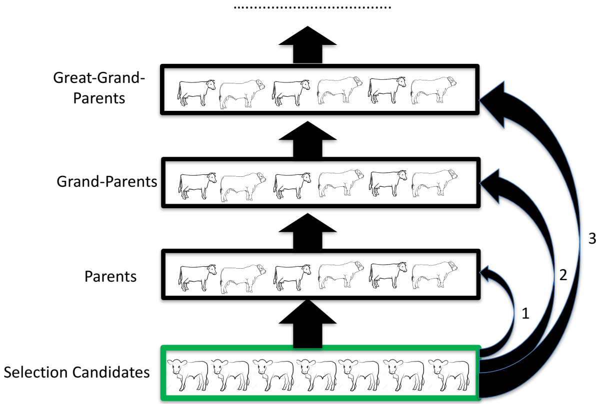
Simulation Overview



Materials and Methods

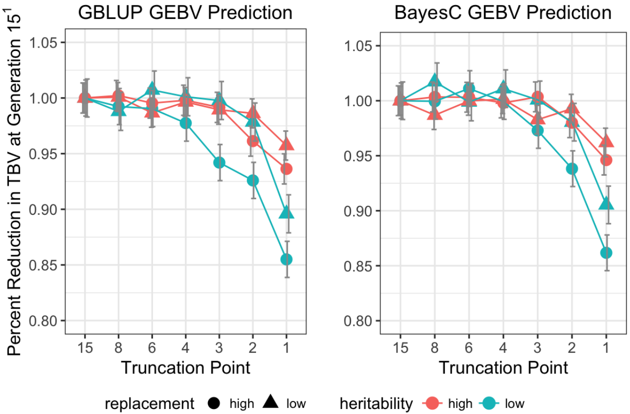
Truncating Data

- Truncate data 15, 8, 6, 4, 3, 2 and 1 generation back from the selection candidates.



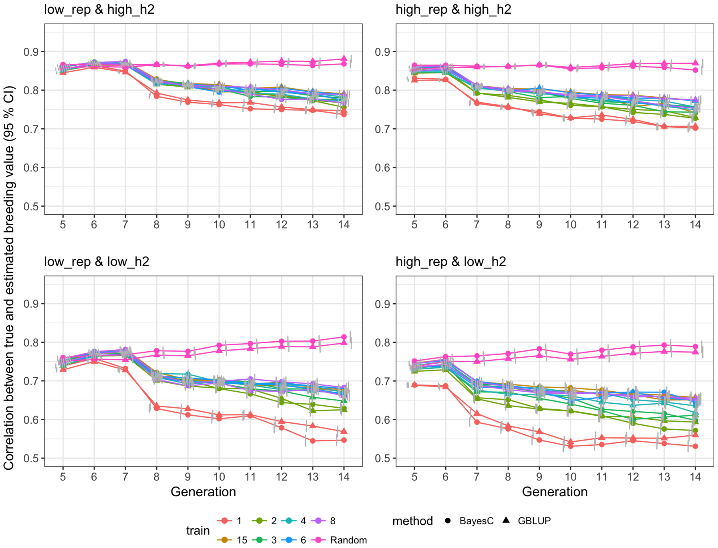
Results

- Truncating data from 4 or more generations back did not impact TBV at generation 15.

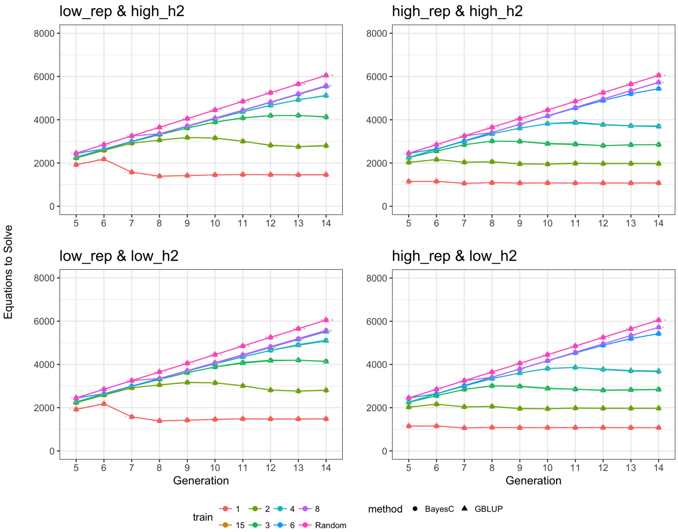


Reduction in True Breeding Values (TBV) at generation 15 was calculated as the ratio of mean TBV at generation 15 for a given cutoff point over the mean TBV at generation 15 when utilizing all animals.

Results

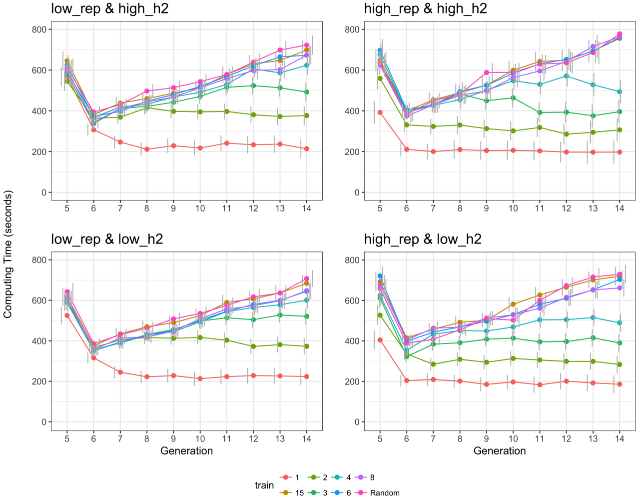

low_rep & high_h2
high_rep & high_h2
low_rep & low_h2
high_rep & low_h2
Correlation between true and estimated breeding value (95 % CI)
0.9
0.8
0.7
0.6
0.5
5
6
7
8
9
10
11
12
13
14
Generation
train
1
2
4
8
15
3
6
Random
method
BayesC
GBLUP

Results


low_rep & high_h2
high_rep & high_h2
low_rep & low_h2
high_rep & low_h2
Equations to Solve
0
2000
4000
6000
8000
5
6
7
8
9
10
11
12
13
14
Generation
train
1
2
4
8
15
3
6
Random
method
● BayesC
▲ GBLUP

Results

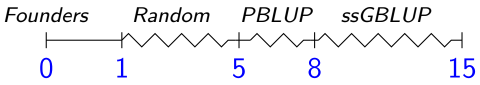
- When the GEBV are predicted using a marker effects model, running time is reduced.



What about ssGBLUP?

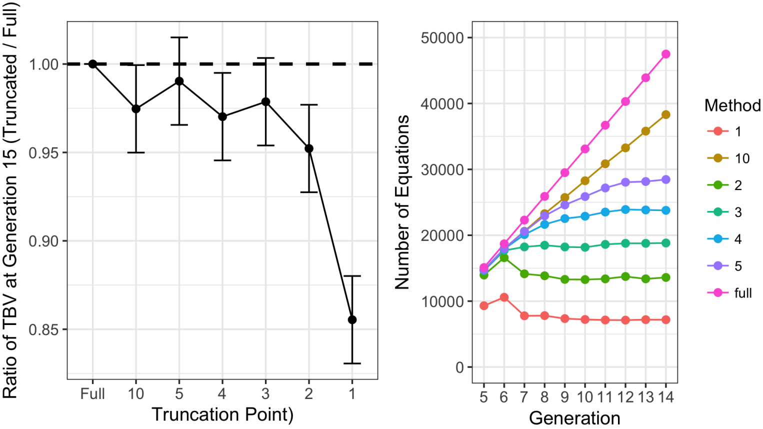
- Similar simulation parameters as outlined previously.
- The breeding population consisted of 100 males and 600 females.
- Mating was at random and 6 progeny produced.
- Started genotyping selected parents at generation 8.
- Truncation data 15, 10, 5, 4, 3, 2 and 1 generation back from the selection candidates

Simulation Overview



What about ssGBLUP?

- Similar to the previous results, truncating data 3 or more generations back did not impact TBV at generation 15.



Conclusions

- Genotype data continues to be collected and rate of genotyping within livestock population will continue to increase.
  - Removal of genotypes and phenotypes of distantly related animals does not negatively impact the prediction of young selection candidates.
  - Larger reference population does not necessarily mean higher GEBV accuracy!!
  - Looking at a genotyped animals relationship to the current group of selection candidates provides an idea of whether to include or exclude an animal.
  - Number of equations to solve as generations proceed stabilizes.

Acknowledgments

- Dr. Matt Spangler
- Dr. Steve Kachman
- DNA Genetics

Questions