



HAL
open science

Bayes Cpi versus GBLUP, OLS regression, sparse PLS and elastic net methods for genomic selection in french dairy cattle

Carine Colombani, Andres Legarra, Pascal Croiseau, S. Fritz, F. Guillaume,
Vincent Ducrocq, Christèle Robert-Granié

► To cite this version:

Carine Colombani, Andres Legarra, Pascal Croiseau, S. Fritz, F. Guillaume, et al.. Bayes Cpi versus GBLUP, OLS regression, sparse PLS and elastic net methods for genomic selection in french dairy cattle. 62. Annual Meeting of the European Federation of Animal Science (EAAP), Aug 2011, Stavanger, Norway. Wageningen Academic, Abstract, p.7, 2011, 978-90-8686-177-4. hal-01193657

HAL Id: hal-01193657

<https://hal.science/hal-01193657>

Submitted on 6 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bayes Cpi versus GBLUP, PLS regression, sparse PLS and elastic net methods for genomic selection in french dairy cattle

Colombani, C.¹, Legarra, A.¹, Croiseau, P.², Fritz, S.³, Guillaume, F.⁴, Ducrocq, V.² and Robert-Granié, C.¹, ¹INRA, Animal Genetics, UR631-SAGA, BP 52627, 31326 Castanet-Tolosan, France, ²INRA, Animal Genetics, UMR1313-GABI, 78352 Jouy-en-Josas, France, ³UNCEIA, 149 rue de Bercy, 75595 Paris, France, ⁴Institut de l'élevage, 149 rue de Bercy, 75595 Paris, France; christele.robert-granie@toulouse.inra.fr

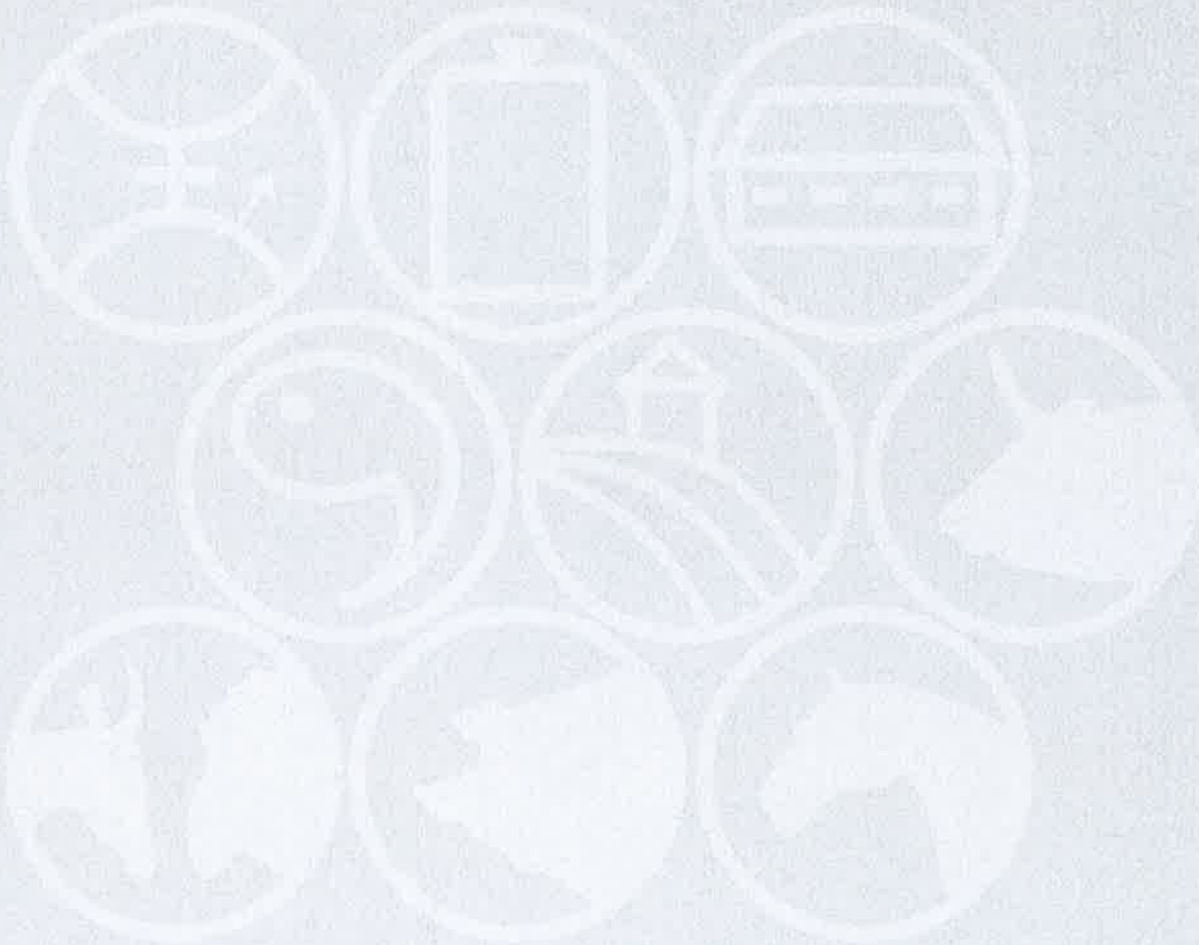
Genomic Selection (GS) relies on genomic enhanced breeding values (GEBV) derived from dense markers information over the entire genome for the selection of young animals. GEBV predictions have to be accurate for a successful application of GS. The objective of this study was to compare the predictive ability of a Bayesian approach -Bayes C π - which performs SNP selection by stochastic search (including or not pedigree information) with methods already tested on French dairy cattle: Genomic BLUP (GBLUP) and dimension reduction or variable selection methods such as Partial Least Squares (PLS) regression, Sparse PLS regression and Elastic Net (EN). The prediction equation was estimated with training data sets composed of 2,976 Holstein bulls or 950 Montbéliarde bulls, genotyped for 38,462 or 39,738 SNP respectively, with a minor allele frequency of 3%. Then, phenotypes were predicted on validation data sets composed of 964 (222) young Holstein (Montbéliarde) bulls. Five milk production traits and conception rate were studied using DYD (Daughter Yield Deviation) weighted by Effective Daughter Contribution (EDC) as phenotypes. The EDC weighted correlation between GEBV and observed DYD was computed. Bayes C π and EN approaches gave slightly better accuracies of GEBV prediction. The convergence of Bayes C π was reached quickly with few samples for the Holstein data set while the prior probability that a marker has zero effect (π) had to be fixed in the Montbéliarde data set. Adding pedigree information to the BayesC π approach did not improve the correlation nor the slope of the regression of observed DYD on estimated DYD.

Genome-wide breeding value estimation with the Bayesian elastic net

Yu, X. and Meuwissen, T.H.E., Norwegian University of Life Sciences, Animal and Aquacultural Sciences, Arboretveien 12, 1432 Ås, Norway; xijiang.yu@unb.no

Genome-wide breeding value (GWEBV) estimation is often facing a situation where the number of predictors (genetic markers) is much more larger than the number of records (phenotypes), or $P \gg n$. Several regularization methods were proposed to reduce the predictors. The Bayesian lasso regularize the predictors with the first order function of marker effects, i.e. the sum of their absolute values, as penalty. Recently, Zou *et al.* proposed a elastic net (enet) method. In addition to the first order penalization, enet also penalizes the sum squares of marker effects. The enet can treat closely associated marker effects, which might be in a same physiological pathway, as a whole, such that it can include or exclude them from a model altogether. Enet method is currently mostly used in association studies on human diseases, where it often outperforms Bayesian lasso. In this study, we compared a Bayesian version of enet (Benet) with three other methods, including fast Bayes-B, GBLUP, and MixP, with simulated data on a real Norwegian red cattle pedigree. In Benet, the penalties on the absolute value of the marker effects is parameterized as $\pi\Lambda$, and on the squared value as $(1-\pi)\Lambda$, where Λ is a scaling parameter determining the prior variances of the marker effects, and π varies between 0 (all penalty on the squared effects) and 1 (all penalty on the absolute value). The results show that π does has little effect on the accuracy of GWEBV estimation. The accuracy goes up slowly as π increases. When π approaches 1, the results can vary a lot and are unpredictable. A value of π of 0.97 is used for final comparison, in which case the accuracy of Benet is similar to that of GBLUP, and in some cases performs slightly better. Possibly Benet may perform better in real data, where gene effects can be associated by being in the same pathway.

Book of Abstracts of the 62nd Annual Meeting of the European Federation of Animal Science



**Book of abstracts No. 17 (2011)
Stavanger, Norway
29 August - 2 September 2011**