Comparing frequentist and bayesian methods when analyzing the result data of official variety trials in Finland

Timo Hurme, MTT – Agrifood Research Finland, Information Services
Kaarina Vuori, MTT – Agrifood Research Finland, Animal Breeding

Introduction

In Finland MTT (Agrifood Research Finland) carries out official variety trials to determine the agricultural value of field crop varieties. These variety trials have been running for over 30 years and today variety trials are carried out at some 20 experimental sites and there are annually over 100 trials. Here we have the yield data on 15 common barley varieties from the last eight years (1997-2004). The experimental unit in this data is the estimated trialwise mean of a variety. Some earlier research results are presented in Kangas et al. (2005).

We were interested in studying the differences of barley varieties in yield by taking into account the environmental conditions. Also we were planning to compare different estimation methods and software in question of interest.

Description of the data

Hierarchy levels:

Year	= Year (8 years)
Town	= Area where experiment was arranged (14 towns)
expid	= Identification number of an individual experiment (124 experiments)
obs	= Observation id (846 observations)

Dependent variable:

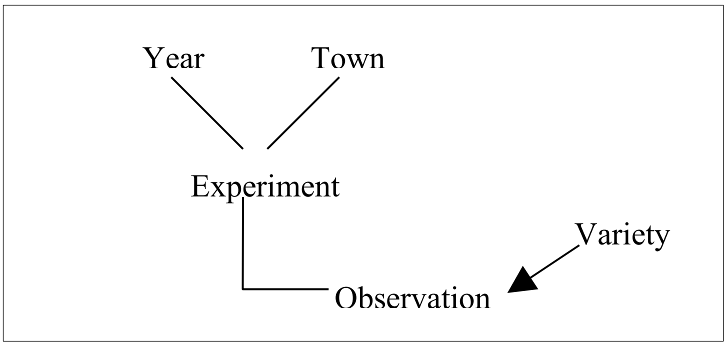
Yield_t	= Yield of the variety (t/ha) (range: 0.5-8.0)

Independent variables:

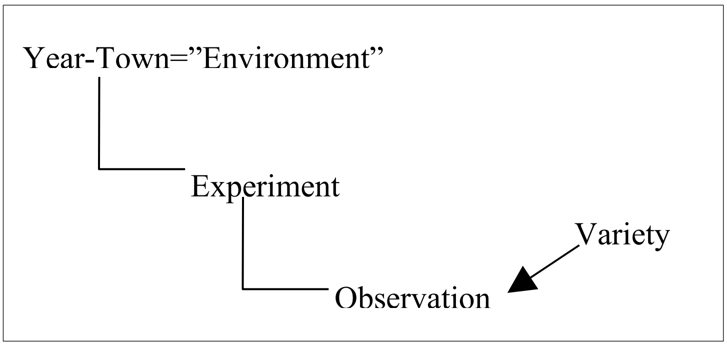
Variety_name	= Name of the variety (15 varieties)
dd_100_cent	= Days of degree divided by 100 and centered. Days of degree is calculated by summing daily the “average temperature of the day minus five” through the growing season. Negative values of daily differences are ignored. So, days of degree represents the environmental conditions of the growing season in terms of temperature.

Modelling process

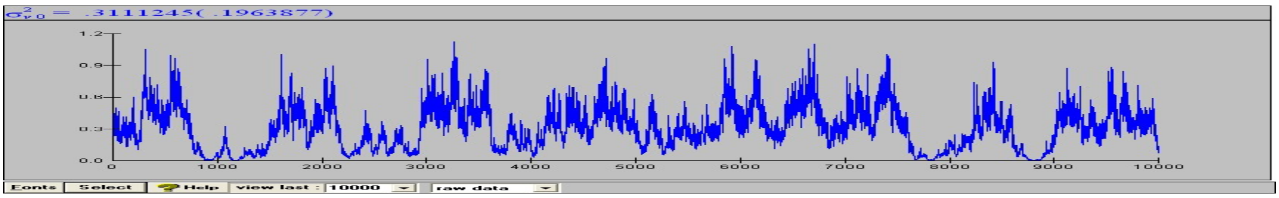
In our modelling process we used MLwiN and later we also fitted the final model using WinBUGS. The starting point for our modelling was a cross-classified hierarchical structure between the variables of interest. Because of the structure, we used MCMC methods (with 10000 iterations) in the modelling process. The structure of the initial model is presented in the following graph.



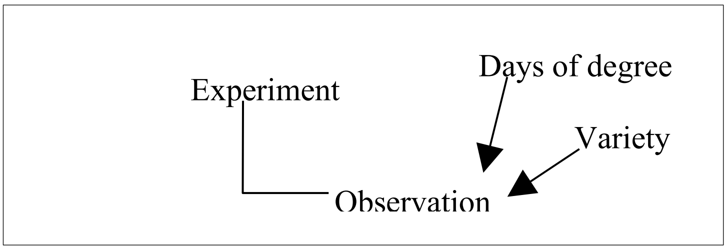
For this initial model we obtained the DIC value 1443.64 (on 133.65 pD). However, when examining the significance of factors Year and Town, we found out that they were not significant (estimated variances for year and town were 0.371 (SE=0.349) and 0.102 (SE=0.103), respectively). The weak effect of year and town may be explained by the low number of levels in them (8 and 14, respectively). So we decided to create a variable combining the information from year and town thereby representing the effect of environment in general.



The resulting combination variable consisted of 87 levels. This led to the fact that there were only 1-3 experiments per “Environment”. Moreover most “Environments” consisted only of one experiment. So, now “Environment” and Experiment reflected almost the same effect and this led into difficulties in estimation. The trace of MCMC iterations mixed very badly and tended to zero.



We decided to drop the highest level in hierarchy and add a new fixed effect Days of degree (centered and scaled) which would capture some of the environmental variation.



This simplified the model and it converged properly. Days of degree also proved to be a significant predictor: 0.323 (SE=0.109). The problem with days of degree was that there were many missing values (219 = 26%). Fortunately the values were missing at random and we decided to prefer this model even though it only used 627 observations. This model was used as our final model. However we wanted to evaluate the effect of missing values to the results by comparing this model to the model with imputed values for days of degree.

The imputation was carried out by using a Multivariate Normal Model, where there were four responses. The responses were days of degree and three other variables that were highly correlated with days of degree (1000-seed weight, hectolitre weight and length of the stand). We ran 5000 MCMC iterations on this model and saved five sets (every 1000 iterations) of imputed values for days of degree. These five sets were used in our model (multiple imputation) instead of the original days of degree to compare the results with the first model (with missing values).

Comparing methods and software in the final model

We ran our final model using IGLS estimation, RIGLS estimation, MCMC with 10000 and 25000 iterations in MLwinN and with 25000 iterations in WinBUGS. Also we ran our model with the five sets of imputed values using MCMC with 10000 iterations and calculated the parameter estimates and their standard errors ("Multiple imputation" model). The parameter estimates and their standard errors (in brackets) are presented in the following table.

Parameters	**IGLS MLwiN**	**RIGLS MLwiN**	**MCMC MLwiN (10000)**	**MCMC MLwiN (25000)**	**MCMC WinBUGS (25000)**	**MI for dd MLwiN (10000)**
Int	4.537 (0.165)	4.536 (0.166)	4.545 (0.169)	4.546 (0.166)	4.543 (0.169)	4.552 (0.161)
"Rolfi" (variety)*	0.136 (0.156)	0.136 (0.158)	0.133 (0.160)	0.135 (0.159)	0.141 (0.159)	0.034 (0.146)
"Saana"	-0.199 (0.130)	-0.198 (0.131)	-0.199 (0.133)	-0.199 (0.132)	-0.198 (0.132)	-0.121 (0.127)
"Kunnari"	0.432 (0.125)	0.432 (0.127)	0.431 (0.128)	0.432 (0.128)	0.433 (0.128)	0.353 (0.125)
"Jyvä"	0.107 (0.130)	0.107 (0.131)	0.105 (0.132)	0.106 (0.132)	0.110 (0.131)	0.061 (0.128)
"Pohto"	0.133 (0.127)	0.133 (0.129)	0.132 (0.130)	0.132 (0.130)	0.134 (0.129)	0.098 (0.127)
"Erkki"	0.179 (0.149)	0.179 (0.151)	0.179 (0.152)	0.179 (0.152)	0.182 (0.151)	0.105 (0.149)
"Gaute"	0.602 (0.160)	0.602 (0.162)	0.601 (0.164)	0.601 (0.163)	0.605 (0.162)	0.485 (0.148)
"Kustaa"	-0.500 (0.133)	-0.499 (0.135)	-0.499 (0.136)	-0.500 (0.135)	-0.500 (0.136)	-0.458 (0.131)
"Kymppi"	-0.499 (0.140)	-0.499 (0.142)	-0.498 (0.142)	-0.499 (0.142)	-0.500 (0.142)	-0.438 (0.142)
"Edel"	0.462 (0.175)	0.462 (0.177)	0.460 (0.177)	0.462 (0.177)	0.462 (0.176)	0.543 (0.153)
"Arve"	0.023 (0.134)	0.023 (0.136)	0.021 (0.138)	0.021 (0.137)	0.027 (0.137)	-0.093 (0.128)
"Tofta"	-0.015 (0.145)	-0.015 (0.147)	-0.015 (0.147)	-0.015 (0.147)	-0.016 (0.147)	0.018 (0.138)
"Scarlett"	0.078 (0.142)	0.078 (0.144)	0.076 (0.144)	0.076 (0.144)	0.077 (0.144)	0.101 (0.132)
"Optima"	0.079 (0.158)	0.079 (0.161)	0.080 (0.160)	0.079 (0.160)	0.076 (0.161)	0.140 (0.144)
dd_100_cent **	0.327 (0.106)	0.327 (0.108)	0.323 (0.109)	0.325 (0.108)	0.332 (0.110)	0.107 (0.080)
Var(experiment)	1.353 (0.209)	1.371 (0.212)	1.397 (0.223)	1.398 (0.222)	1.405 (0.224)	1.369 (0.184)
Var(residual)	0.248 (0.015)	0.255 (0.016)	0.256 (0.016)	0.256 (0.016)	0.256 (0.016)	0.275 (0.015)

* Reference category for varieties was "Botnia".
** dd_100_cent = Days of degree centered and divided by 100

Overall there seemed to be only minor differences between estimation methods. Especially in MLwiN the results from different estimation methods were very similar, only IGLS seemed to estimate slightly smaller standard errors for the parameters. We detected a little difference between the MCMC results in MLwinN and WinBUGS, but in practise the results were very similar. When comparing the model with the original days of degree and the imputed days of degree, there appeared to be differences between the parameter estimates. Biggest difference seemed to be in the estimate for days of degree, which was the variable that was largely imputed. Other parameter estimates have also changed, but not as much. The estimates for the variance of experiment seemed to be roughly in line with the earlier models. Opposite to the parameter estimates, the standard errors for all the parameter estimates seemed to be very much like the ones detected in other models, only slightly smaller (increased sample size).

We decided that we would base our inference on the RIGLS model, because it seemed to give results that were very much in line with the other models and it is also not computationally as intensive as the MCMC models. The full statistical model is presented below.

$$\begin{aligned}\text{yield\_t}_{ij} = {} & \beta_{0j} + \beta_1\text{ROLFI}_{ij} + \beta_2\text{SAANA}_{ij} + \beta_3\text{KUNNARI}_{ij} + \beta_4\text{JYVÄ}_{ij} + \beta_5\text{POHTO}_{ij} + \beta_6\text{ERKKI}_{ij} + \beta_7\text{GAUTE}_{ij} + \\ & \beta_8\text{KUSTAA}_{ij} + \beta_9\text{KYMPPI}_{ij} + \beta_{10}\text{EDEL}_{ij} + \beta_{11}\text{ARVE}_{ij} + \beta_{12}\text{TOFTA}_{ij} + \beta_{13}\text{SCARLETT}_{ij} + \beta_{14}\text{OPTIMA}_{ij} + \\ & \beta_{15}\text{dd\_100\_cent}_{ij} + e_{ij}\end{aligned}$$

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$u_{0j} \sim \text{N}(0, \sigma^2_{u0})$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{e})$$

$-2*loglikelihood$ = 1223,641(627 of 846 cases in use)

Results

We found out that the days of degree was a statistically significant predictor in the model. Therefore it explained some of the environmental effect. When we examined the differences between varieties, we found out that there were three varieties that outperformed the reference variety “Botnia”. These varieties were “Gaute” ($p<0.001$), “Edel” ($p<0.01$) and “Kunnari” ($p<0.001$). There were also two varieties that did not perform as well as “Botnia”. These varieties were “Kustaa” ($p<0.001$) and “Kymppi” ($p<0.001$).

Most of the variation in this data was between experiments (84%) compared to within experiments (16%). The large proportion of between experiments variation is most likely caused by the fact that for most cases there was only one experiment in a town each year. Therefore a lot of environmental variation was included in the experiment level. On the other hand the experiments were planned so that most of the variation within experiments was explained by the treatments i.e. the varieties.

Final Remarks

- There were no major differences between estimation methods.
- The use of imputed values in our predictor “days of degree” seemed to somewhat affect the results.
- There were significant differences between varieties in yield and also the days of degree was a significant predictor.

Acknowledgements

Thanks to our instructor William J. Browne and also to other teachers Henrik Stryhn and Rasmus Waagepetersen for all your help. We also thank Anders Ringgaard Kristensen for a successful course.

References

Kangas A., Laine A., Niskanen M., Salo Y., Vuorinen M., Jauhiainen L. and Nikander H. (2005). Results of official variety trials 1997-2004, *Agrifood Research Reports* **83**, MTT Agrifood Research Finland, Jokioinen, 193p.