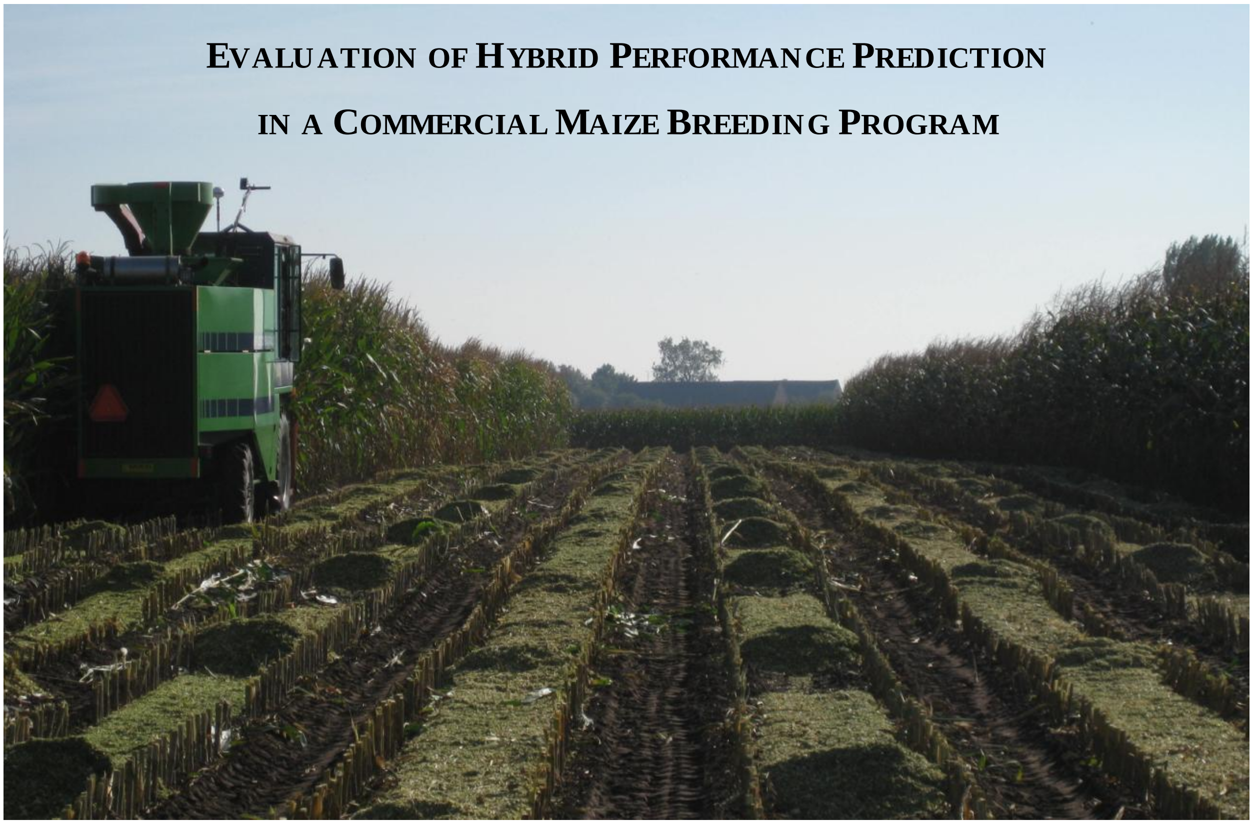
INTERNSHIP: PLANT BREEDING **Victoria Mungall**

PBR-70427, 24 ECTS **850825-591-070**

EVALUATION OF HYBRID PERFORMANCE PREDICTION IN A COMMERCIAL MAIZE BREEDING PROGRAM



ON-SITE SUPERVISORS:

Louis Vlaswinkel
Senior Maize Breeder, Limagrain Nederland

Cornélie Noordam-Boot
Assistant Maize Breeder, Limagrain Nederland

EDUCATIONAL SUPERVISORS:

dr.ir. Luisa Trindade
Group Leader Biobased Economy, Wageningen UR Plant Breeding

Prof.dr. Richard Visser
Department Chair, Wageningen UR Plant Breeding

PLANT BREEDING
WAGENINGEN UR

Limagrain
Nederland

ACKNOWLEDGEMENTS

Most importantly thanks to my supervisors at Limagrain Louis Vlaswinkel and Cornelie Noordam-Boot for sharing their expertise and passion for breeding with me. The entire maize breeding team in Rilland especially Roelf, René, Heleen, Jaimy, Jeroen and Theo. Cristina, Holger and Tingting for making my month or so spend at the Rhinefelderhof site a pleasure. In the Department of Plant Breeding, Richard Visser and Luisa Trindade for her feedback and dedication.

CONTENTS

1 Company Profile: Limagrain Nederland

Limagrain Nederland B.V. in Rilland, Zeeland is the headquarters for Limagrain Benelux. The research station in Rilland was founded as Van der Have in 1879. The company began with the breeding of sugarbeet in 1941 and maize in 1949 (Table 1). Van der Have then merged with Zeneca to become Advanta, the world's fifth largest seed company. The last change occurred in 2005 when Advanta became part of the French company Limagrain. The Van der Have name carries on in the sugar beet breeding company SES Van der Have, which is still located at the Rilland station, now a shared site with Limagrain. There are also breeding programs for barley, wheat, flax, rapeseed, peas and onions at this site.

The maize breeding program in Rilland focuses on early silage maize. This program has proved itself to be competitive with success in Western and Central Europe, with particularly large success in the very early markets such as the UK and Denmark. The germplasm of this program is of high quality (high digestibility), as a result the program is very focused on producing high quality as well as high yielding hybrids. This is made possible in part by the expertise and equipment for near infrared spectroscopy (NIRS) used for analysis of digestibility components also located at the station. Ten people in Rilland are directly involved with the maize breeding program (30 total including labs).

Table 1. History of maize breeding in Rilland.

1949	Breeding of maize begins as Van der Have
1972	Collaboration with Pioneer
1975	Development of own breeding material
End of 1980s	End of relationship with Pioneer
End of 1980s	Success in Netherlands, Belgium and UK
1996	Creation of Advanta with merge of VanderHave and Zeneca (SES)
Early 2000s	Increasing importance in Germany and France
2005	Became part of Limagrain

Currently, all maize varieties that reach the market from this program are developed through conventional breeding methods. However, this may change in the near future as MAS has begun in two smaller programs both of which were started three years ago. The first is a reverse breeding program where elite lines are crossed with hybrids of competitor companies and markers are used to try to determine the profile of a parent of the hybrid. The second is a recurrent selection program (MARS), but the preliminary results of this program are disappointing.

The maize breeding team is also responsible for the following activities for their own program as well as for other breeders in the company:

- Management of early silage maize breeding program
- Management of hybrid trials for North West Europe (silage)
    - Yield trials
    - Top-crosses
- Observation nurseries (Wouw, Netherlands)
- Doubled haploid production

2 Introduction: Hybrid Maize Breeding in Europe

One of the only major crops originating from the Americas, maize (*Zea mays* L.) was first introduced in Europe, by Columbus in Spain (1494) and a little later by Verrazano (1524) and Cartier (1534) in French Normandy (Barrière et al. 2006). Although the first use of maize in Europe was for human consumption, as it was traditionally in America, the value of maize as an excellent forage crop for livestock was quickly recognized. However for a long time breeding focused on the improvements of grain traits rather than silage. These varieties were either landraces or open pollinated varieties (mostly of the flint type) until the introduction of US hybrids after the Second World War.

Breeding maize as a forage crop really only began in Europe only 25 years ago, in the Netherlands (Barrière et al. 2005). While breeding in the US, France and Germany focused on grain, the main objective of Dutch breeders was to obtain high yielding silage maize. Today forage hybrids available are still largely based on grain maize germplasm due to this short breeding history (Barrière et al. 2006). The first generation of early European maize hybrids were based on crosses between two distinct heterotic groups: flint and dent. The European flint lines were introduced to Europe more than 500 years ago, while the dents were derived from US dent lines during the past 50 years (Schrag et al. 2010). Thus, these two heterotic pools have been genetically separated for a long time. The first flint x dent hybrids (INRA258, Brillant DK202, Capella, LG11, Blizzard G188) were earlier than most varieties previously available and resulted in significant extension of silage maize cropping (Barrière et al. 2005). These two heterotic pools continue to make up most of the genetic basis for maize hybrids in Europe with the larger dents generally being used as the mother in hybrid production (Figure 1).

Figure 1. Flint-dent hybrids are produced from crossing American dent and European flint inbred lines. These hybrids are grown throughout North and Central Europe; in contrast to Southern Europe and the US, where dent-dent hybrids are prevalent.

3 Evaluation of Accuracy of Hybrid Performance Prediction

3.1 General and specific combining ability

The methodology for breeding hybrid cultivars has primarily been developed in maize and as a result has reached a relatively high level of sophistication in this crop. The concepts of general and specific combining ability were introduced by Sprague and Tatum (1942) to assist in identifying the best parental combinations in a maize hybridization program. General combining ability (GCA) is the average performance of a line in hybrid combination and is associated with additive genetic effects. Specific combining ability (SCA) is the deviation of performance in individual crosses from these average performances of the lines involved and is associated with non-additive genetic effects.

In studies, diallel analysis is often used to estimate GCA and SCA effects (Beck et al. 1990). Within breeding programs, as the number of inbred lines increases the number possible crosses between lines from different heterotic pools grows rapidly. Since producing and evaluating all potential combinations in the field would be very expensive and cumbersome, in practice only a small proportion of all possible experimental hybrids are evaluated in field trials. Still the number of yield trials for the Rilland program, for example, is 30 000 silage and 10 000 grain plots. Thus, estimates of GCA and SCA can be used in prediction of promising hybrids to assist with the breeder's decision of which combinations of lines to test. In addition, the *per se* performance of an inbred line does not correlate well with performance of the resulting hybrids for many important traits including yield (L. Vlaswinkel, *personal communication*). Hence, the breeder will tend to test as many lines as possible and predictions of performance have the potential to greatly improve the efficiency of commercial breeding programs.

Typical testing of hybrid performance is as follows:

i. Many promising inbred lines are tested with lines that are well known by the breeder, such as lines currently in or nearing commercial application. Usually new lines are tested with two of these 'tester' lines in a top cross. From the resulting progenies, the general ability of each inbred to combine with other lines, or the general combining ability (GCA), can be assessed.

ii. Inbreds with a good GCA are selected to be tested again with several promising or well established inbreds (of the other heterotic group), which will be evaluated in multi-year, multi-location trials. In this way, specific combining ability (SCA) can be assessed as the performance of the inbred line when combined with a specific tester line. Thus, each inbred combination will have a unique SCA value for a given trait.

In order to speed up the breeding process, breeders often try to obtain an impression of the combining ability of inbred lines before they have reached complete homozygosity. Top cross trials may be performed on generation S3 or S4 inbreds with comparable results to trials using later generation lines (Bernardo 2003). This process, known as early testing, may reduce costs since the least promising lines can be removed from the program at an earlier stage. Now however, with the almost routine use of the double haploidization process, early testing has become less relevant since inbred lines will become homozygous in one step using double haploids. In flint-dent hybrids, the double haploid procedure is very efficient for dent, but has been more difficult to optimize for flint. As a result, breeders may rely more on conventional methods to obtain homozygosity in flint lines.

3.2 Predicting hybrid performance: the BLUP method

As mentioned previously, in theory combining abilities should be determined using arithmetic mean values from a half diallel crossing block, however this is not done in practice since the number of crosses may be prohibitive. In practice, only a small amount of all possible crosses are tested and are often tested on few locations, for few years within very unbalanced designs. It is also possible that an inbred line will only be observed in combination with one tester and in this case it is impossible to separate the GCA of the inbred from the SCA for this hybrid combination.

As a result of these problems, breeders use a more complex calculation method to obtain more accurate estimates of combining abilities of lines. The most common approach, derived from animal genetics, is the best linear unbiased prediction (BLUP) method. This method uses observations of the relatives of a genotype to estimate its breeding value. The breeding value of a genotype (line) is calculated as the weighted sum of the performance of that line and the performance of its relatives, the closer the relative, the larger the weight. Both of the two predictors analyzed in this report are based on the BLUP method.

Currently, Limagrain has the possibility of using two predictors in deciding which combinations of inbreds to test: the GCA prediction and the hybrid prediction (Equation 1 and 2).

Equation 1. GCA prediction = mean + GCA-female + GCA-male

Equation 2. Hybrid prediction = mean + GCA-female + GCA-male + SCA-hybrid

Example: Hybrid prediction.
DL26 x TC33 ➔ 2-05R00042

Hybrid prediction = GCA(DL26) + GCA(TC33) + SCA(2-05R00042)

The SCA is suspected not to contribute significantly to hybrid performance. Therefore, the GCA prediction alone is expected to be an adequate predictor.

3.3 Objectives

The objective of this chapter was to investigate the accuracy and efficiency of the hybrid predictions currently used at Limagrain Nederland. More specifically:

- To determine whether GCA or hybrid predictions are better predictors of hybrid performance in the coming year.
- To determine if predictions can be improved by basing predictions on trial results of several previous years.

3.4 Evaluating GCA and Hybrid Prediction Accuracy

Prediction of hybrid performance using the GCA and hybrid prediction was evaluated for the two most important traits: i) silage yield and ii) grain yield. Grain yield remains important for silage maize breeding due to the prevalence of grain maize over silage, the demand of farmers for mixed cultivars and the large financial investments of a breeding program. Section 3.4 investigated accuracy for hybrid trial results of 2009 only; the most recent results available.

I. Silage Yield

For the most recent hybrid trial results (2009), both GCA and hybrid predictions were relatively accurate (Table 2). Paired t-tests found that the predictors were not significantly different from the trial results (p-values=0.1566, 0.1910 for the GCA and hybrid predictions respectively, n=651). GCA and hybrid predictions were also not significantly different from each other, suggesting that the SCA did not have a significant contribution to hybrid performance (p-value=0.2048). This is in contrast to hybrid theory that the SCA is an important component of performance (Sprague and Tatum 1942).

However this is likely the result of the breeding scheme rather than a generalization for silage yield. Through the top cross procedure, selections are made based on hybrid performance with multiple tester lines (normally two). The result is an indirect selection for lines with a high GCA and an SCA which approaches zero. Since the breeding process is so expensive, the aim is for lines with a good performance in multiple combinations (high GCA), so that in future they may be used again. On the other hand, lines that have an exceptional performance in only one hybrid combination would not be discarded from the program since one superior hybrid also has a large financial value.

Table 2. Results of 2009 hybrid trial results for silage yield and the corresponding GCA and hybrid predictions for 2009 (n=651).

	Trial Results (tonnes/hectare)	**GCA prediction** (GCA1 + GCA2)	**Hybrid prediction** (GCA1 + GCA2 + SCA)
Mean	18.35	18.41	18.40
Max	21.24	21.71	21.82
Min	14.45	13.45	13.09
Range	6.78	8.26	8.72
Standard Deviation	1.24	1.33	1.37

Residuals of predictions versus trial results were relatively randomly distributed (Figure 2). 9.5% of GCA predictions deviated from trial results by greater than 10% and these large residuals were unrelated to the number of observations or earliness of a hybrid (Figure 3).

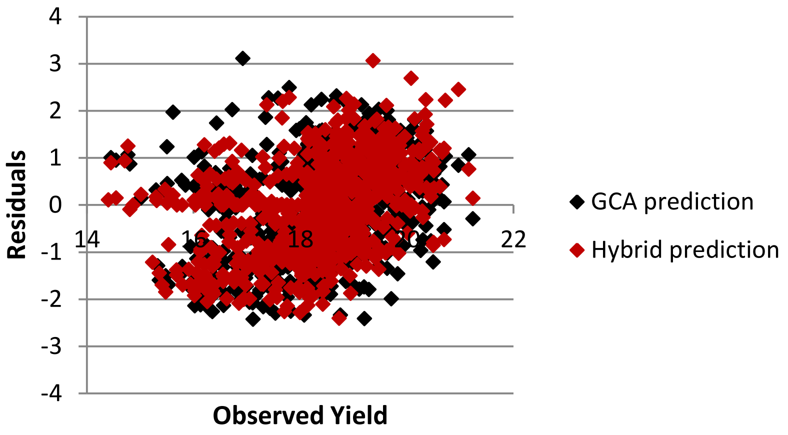


Figure 2. Residuals of GCA and hybrid predictions for silage yield results of 2009. Residuals were calculated as observed yield – predicted yield.

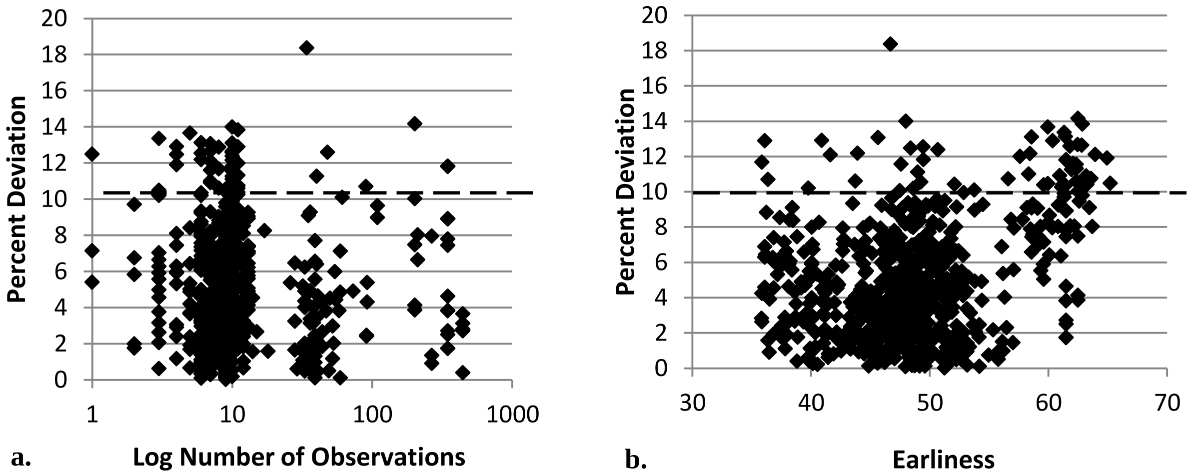


Figure 3. Percent deviation of GCA predictions from 2009 silage yield results shown against a. log number of observations and b. earliness for each hybrid. Dashed lines indicate 10% deviation from trial results.

II. Grain Yield

In contrast to silage yield, predictions for grain yield were significantly different from 2009 trial results (p-values<0.001). However, GCA and hybrid predictions were not significantly different from each other again suggesting that the SCA does not have an important contribution to yield (p-value=0.9872; Appendix 1). This result is in contrast to claims that up to 76% of hybrid yield can be accounted for by mid-parent heterosis (the difference in performance between a hybrid and the mean of its parents; Smith 1986). On the other hand, in a study of prediction efficiency Schrag et al. (2007) also concluded that GCA effects dominated over SCA effects for grain yield of hybrids. They concluded that GCA effects alone were efficient in identifying promising hybrids.

Nevertheless, although the two predictors are similar to each other, they are both relatively inaccurate (Table 3). For GCA predictions for example, 19% of predictions deviated more than 10% from observed yields. Residuals between predicted and observed yields were not randomly distributed and predictions were less accurate when yield is low (under 100, Figure 4). For those hybrids, yield was almost always over predicted, with the exception of a three-way hybrid. Accuracy of predictions was unrelated to number of observations or earliness for a hybrid (Figure 5). Many hybrids with 8-10 observations had smaller residuals than those with 100 observations.

Table 3. Results of 2009 hybrid trial results for grain yield and the corresponding GCA and hybrid predictions for 2009 (n=641).

	Trial Results (tonnes/hectare)	**GCA prediction** (GCA1 + GCA2)	**Hybrid prediction** (GCA1 + GCA2 + SCA)
Mean	110.69	115.42	115.43
Max	132.45	137.12	141.16
Min	82.16	89.04	86.05
Range	50.29	48.08	55.12
Standard Deviation	8.34	7.07	7.74

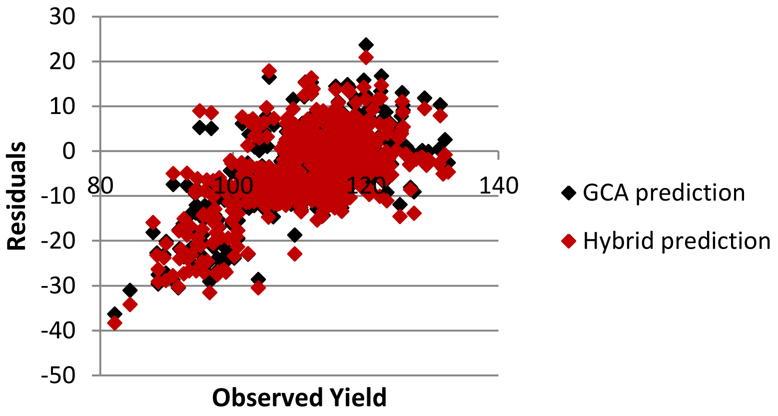


Figure 4. Residuals of GCA and hybrid predictions for grain yield results of 2009. Residuals were calculated as observed yield – predicted yield.

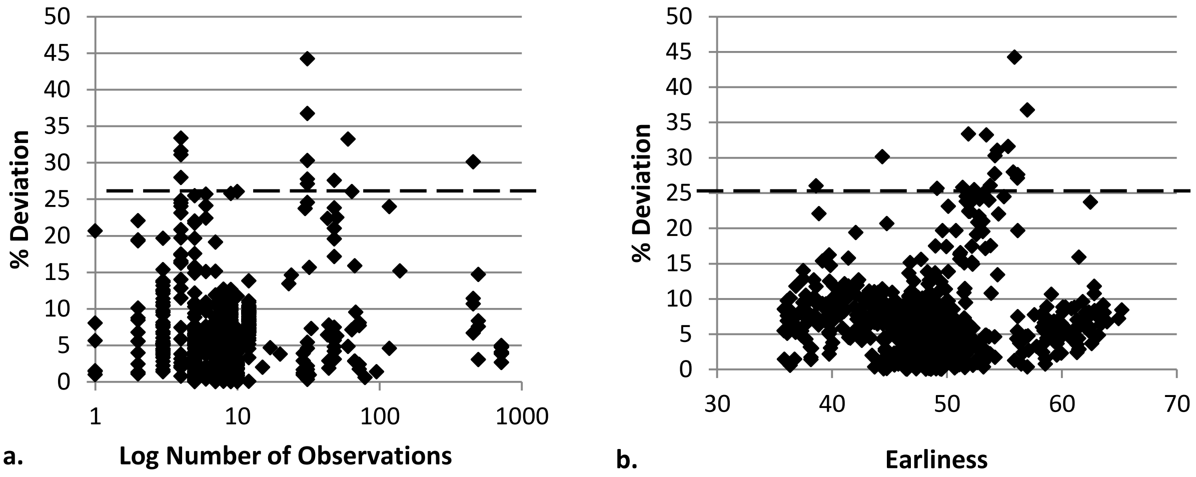


Figure 5. Percent deviation of GCA predictions from 2009 grain yields shown against a. log number of observations and b. earliness for each hybrid. Dashed lines indicate 10% deviation from trial results.

Table 4 lists the hybrids with the least accurate predictions. IVQ835 & QR15 (208EI11389) had the largest percent deviation (44%) and interestingly was also the lowest yielding hybrid tested. Since these results are for a single year, the apparently poor prediction of DA32 & QR15 (2-5021) can be explained by a seed quality problem in 2009. Lastly, although the inbred line QR15 appears to have many poor predictions associated with its hybrids, this line was tested frequently and its frequency of poor predictions was proportional to its overall occurrence.

Table 4. Grain yield predictions with the largest deviation from hybrid trial results (2009).

Hybrid	Pedigree	Trial Results	GCA Prediction	Residual	% Deviation	NbrObs
208EI11389	IVQ835 & QR15	82.16	118.51	-36.34	44.23	31
207EI11191	ISF12 & IVX77	84.51	115.56	-31.06	36.75	31
2-04R03605	DA36 & QR15	91.78	122.26	-30.48	33.21	60
207EP02573	DL26 & NP1941	89.88	117.10	-27.22	30.28	31
2-5021	DA32 & QR15	96.48	125.55	-29.07	30.13	456
207RI11904	IFW91 & RVL11	95.61	122.13	-26.52	27.74	31
208EI09508	IFF801 & YIL05420	103.82	132.45	-28.64	27.58	48
208CE11575	BFC07 & DL26	97.76	124.24	-26.48	27.09	31
303EN01318	IFS4 & IVX77	96.14	121.18	-25.04	26.05	64
2-03R03233	DA36 & TM01	93.40	117.72	-24.32	26.04	10
207RI00094	KT94 & RFF92	97.63	121.59	-23.96	24.54	31

Residuals and percent deviation based on GCA predictions only. Hybrids with less than 10 observations were excluded.

Grain yield predictions may be less accurate than silage yield because grain yield is more influenced by location and year effects. Moisture content of grain was also investigated briefly as moisture content is thought to have a higher heritability than grain yield and is the most important trait following yield for grain maize (L. Vlaswinkel, *personal communication*). Unexpectedly, variability of moisture content was large, with approximately 20% of GCA predictions deviating more than 10% from measured moisture content, similar to the level of inaccuracy in grain predictions (Figure 6). Interestingly, moisture content data fell into three distinct groups with measured contents of approximately 25, 28 or 32%. This may be the result of the near infrared spectroscopy (NIR) methods used to measure moisture contents, though this method is calibrated and checked using conventional ovens regularly.

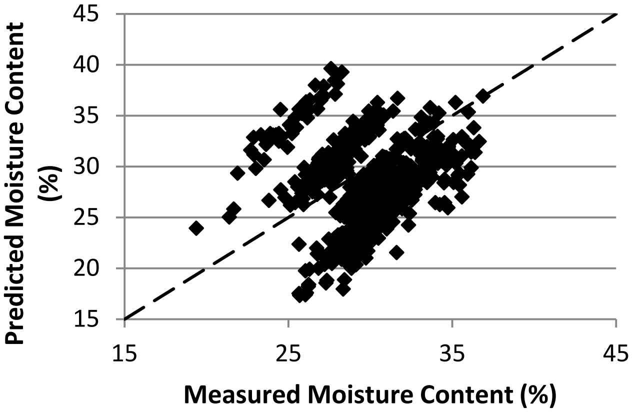


Figure 6. Accuracy of predictions of moisture content of grain maize. Dashed line shows a 1:1 relationship between observed and predicted.

4 Predictions based on Single and Multiple Years

In Section 3 hybrid performance predictions were based on results of a single year, however it is also possible to base predictions on trial results of several previous years. Especially for grain yield prediction it is interesting to see if predictions can be improved by enlarging the dataset. Since GCA predictions and hybrid predictions were not significantly different predictors of performance, only GCA predictions were considered in this section.

I. Silage Yield

In Figure 7, GCA predictions for 2009 were calculated based on an increasing number of previous years trial data. As expected, predictions for 2009 had the highest correlation with 2009 yields (Figure 7a), however contrary to logic, this correlation was not improved by increasing the number of observations. This may indicate that predictions based on data from multiple years inevitably include some year effects, which in turn increases variability. Despite this increased variability, a one-way ANOVA test showed that as expected predictions were improved by basing them on multiple years (Table 5). Predictions based on multiple years were more accurate than those based on a single year for 2009, with the exception of those based on two and three years prior. Thus, it was advantageous to incorporate data from previous years into predictions for the coming year.

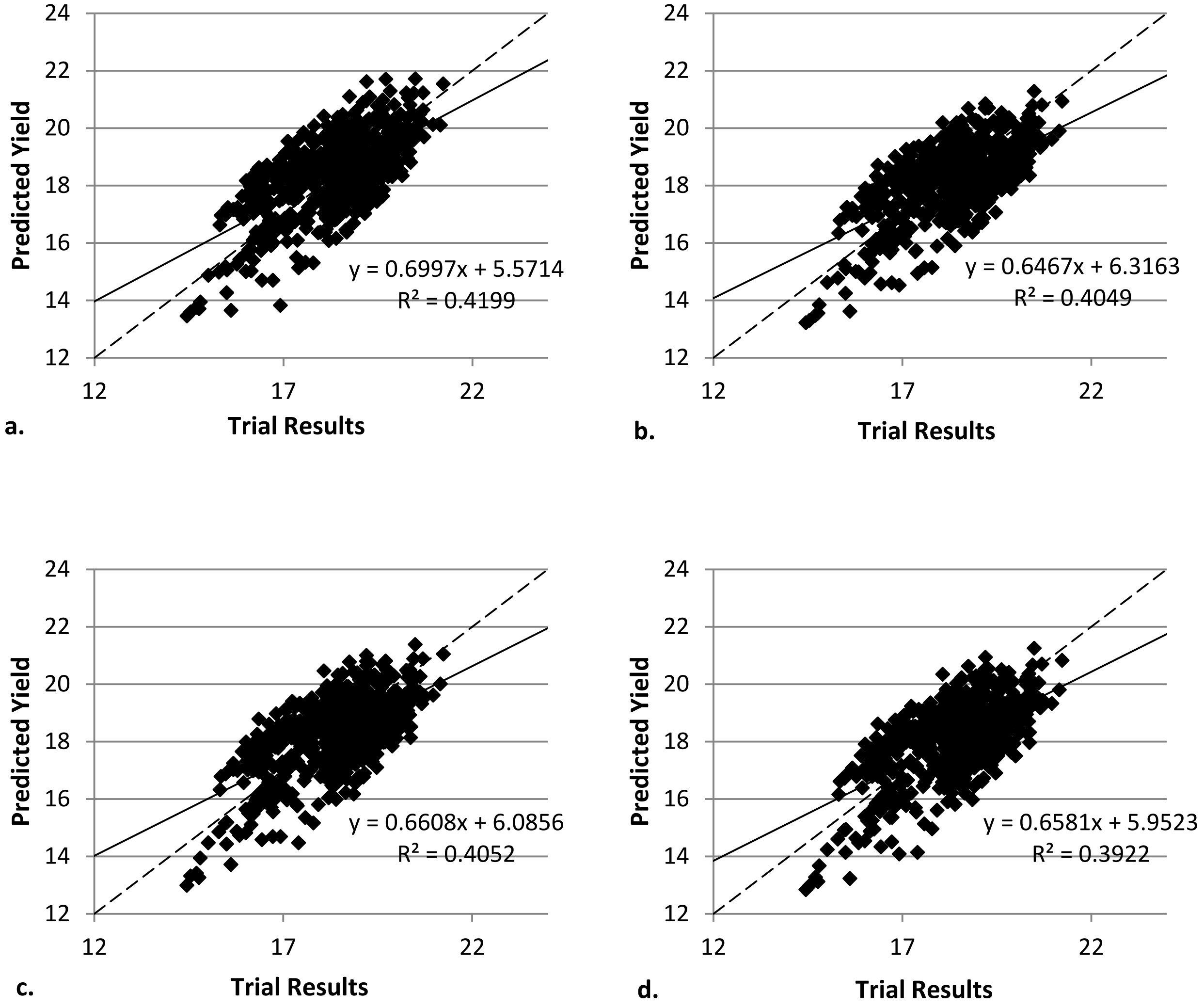


Figure 7. Silage GCA predictions based on a varying number of previous years compared to 2009 trial results. a. 1 year (2009) b. 2 years (2008, 2009) c. 3 years (2007, 2008, 2009) d. 4 years (2006, 2007, 2008, 2009). Dashed line represents a 1:1 relationship between predicted and observed yields.

Table 5. Mean predicted grain yields based on trial results from a varying number of years.

	Mean (tonnes/hectare)	**Standard Deviation**
1 Year	18.426	1.322
2 Year	18.196	1.243
3 Year	18.220	1.269
4 Year	18.036	1.284
Overall mean	18.219	
Least significant difference (LSD)	0.1371	

II. Grain Yield

Similar to silage yield, the correlation between predicted and observed grain yields did not improve as trial data from more years was added (Figure 8). At the same time incorporation of trial data from previous years significantly improved predictions, however there was no added advantage of basing predictions on data from more than two years prior (Table 6). For accuracy of both grain and silage yield predictions it is advisable to repeat this analysis in future since differences between means are so low.

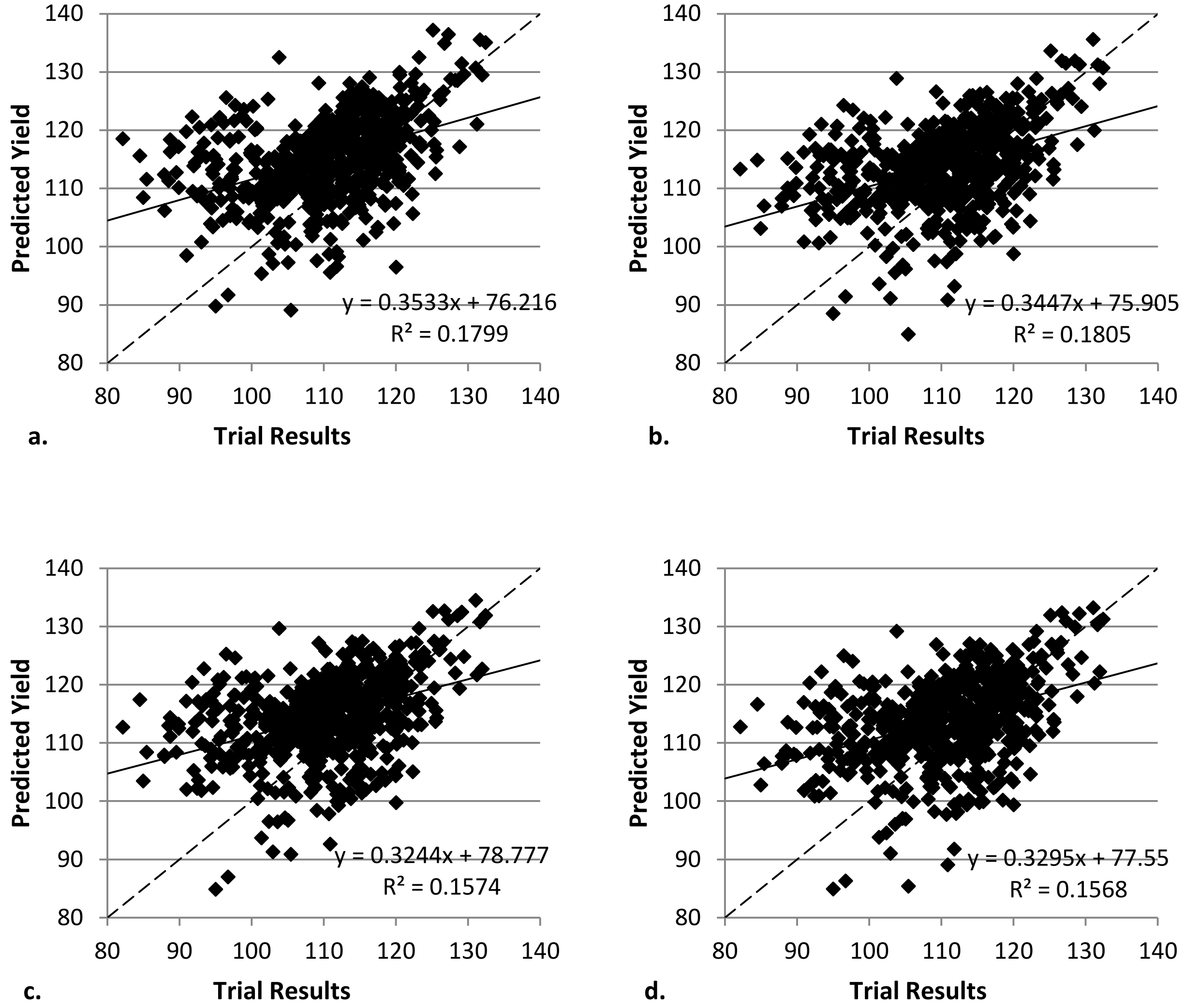


Figure 8. Silage GCA predictions based on a varying number of previous years. a. 1 year (2009) b. 2 years (2008, 2009) c. 3 years (2007, 2008, 2009) d. 4 years (2006, 2007, 2008, 2009). Dashed line represents a 1:1 relationship between predicted and observed yields.

Table 6. Mean predicted grain yields based on trial results from a varying number of years.

	Mean (tones/hectare)	**Standard Deviation**
1 Year	115.08	7.00
2 Year	113.70	6.92
3 Year	114.27	7.03
4 Year	113.61	7.15
Overall mean	114.18	
Least significant difference (LSD)	0.739	

5 Preliminary 2010 Silage Results

As the 2010 harvest came to an end, results quickly became available. A small subset of 29 hybrids tested in 2010 throughout North-western Europe was analyzed to confirm that the SCA can be omitted from future yield predictions. Only silage yield and dry matter content were considered.

Similar to findings of Section 3, silage yield of hybrids were generally overestimated by predictors however these differences were not significant for GCA or hybrid predictions and both were reliable predictors (Table 7; Figure 9). Again, GCA predictions were not significantly different from hybrid predictions (p-value<0.001)). Although, GCA predictions were not significantly different than hybrid predictions, inclusion of the SCA in the hybrid prediction usually increased deviation from actual yields slightly, especially in those with the largest SCA values (Figure 10), although overall percent deviation from results was no worse for those hybrids with larger contributions of SCA than all other predictions.

Table 7. Results of 2010 hybrid trial results for silage yield and the corresponding GCA and hybrid predictions for 2009 (n=29).

	Trial Results (tonnes/hectare)	**GCA prediction** (GCA1 + GCA2)	**Hybrid prediction** (GCA1 + GCA2 + SCA)
Mean	17.52	18.97	19.17
Max	16.12	16.32	16.77
Min	18.73	21.04	21.31
Range	2.62	4.71	4.55
Standard Deviation	0.52	1.26	1.21

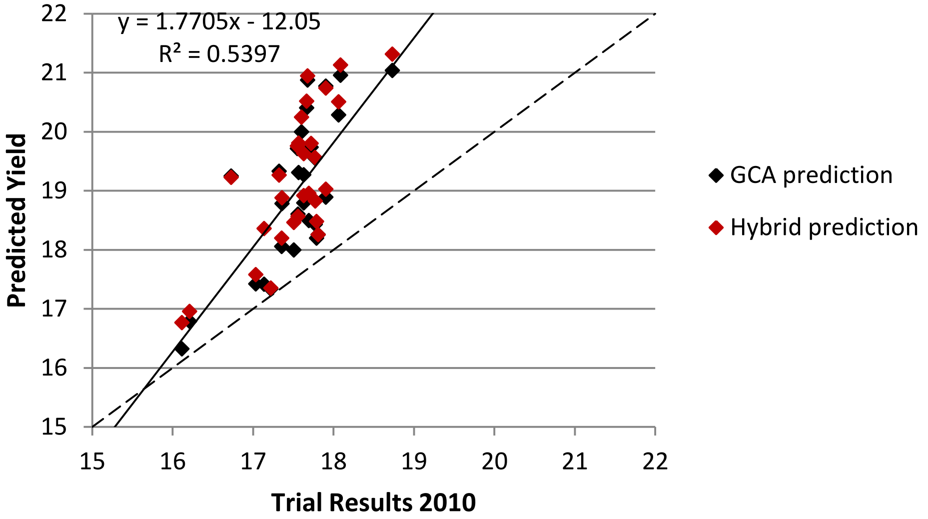


Figure 9. Accuracy of predictions of silage yield (tonnes/hectare) in 2010. Dashed line represents a 1:1 relationship between predicted and observed yield.

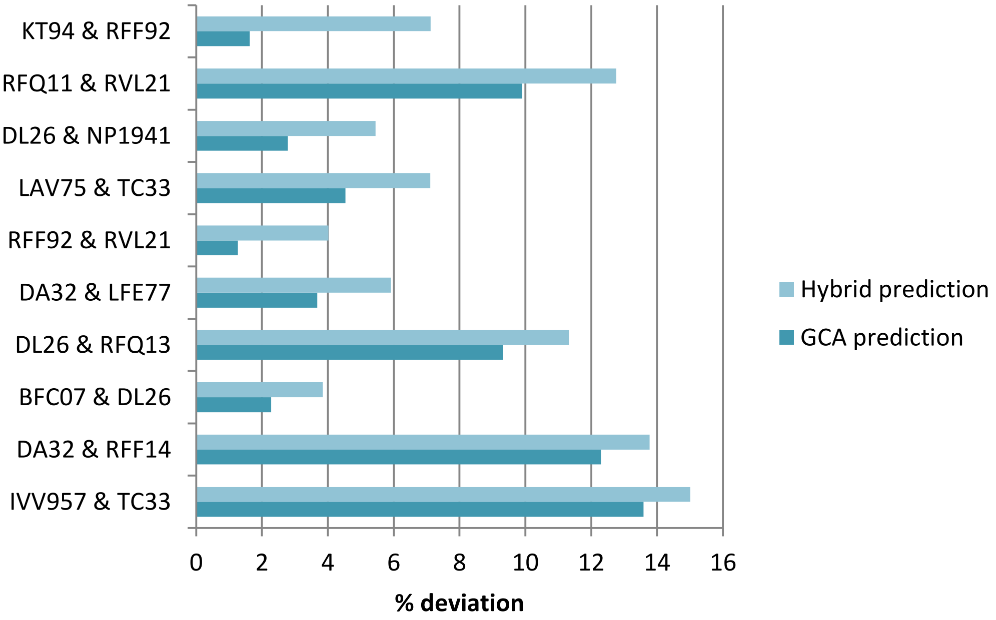


Figure 10. Parents of ten hybrids with the largest predicted SCA values and percent deviation of GCA prediction and hybrid prediction from 2010 silage yields.

Comparable to moisture content for grain maize, dry matter (DM) content of silage maize was expected to have a better predictability than silage yield due to a higher heritability. This hypothesis was confirmed by the 2010 data and suggested that DM content is influenced less by the environment than silage yield (Figure 11). DM content is of interest to the breeder because it can be used as an estimation of relative earliness of a hybrid.

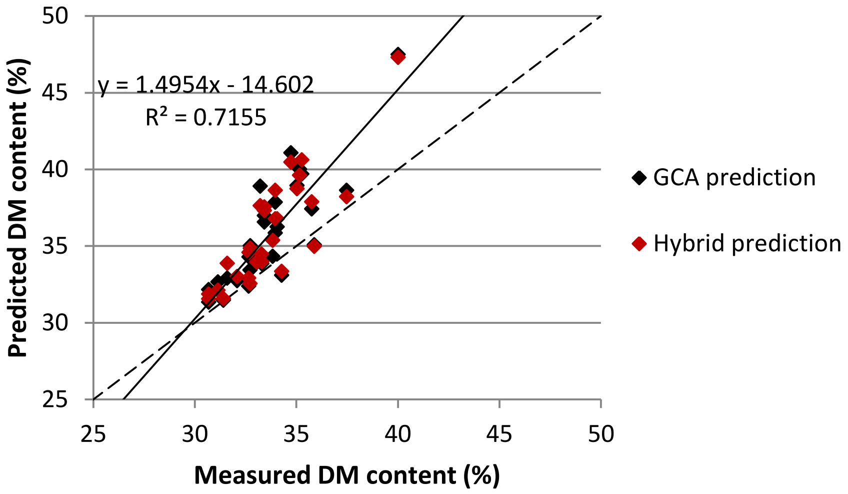


Figure 11. Accuracy of predictions of dry matter content (%) in 2010. Dashed line represents a 1:1 relationship between predicted and observed.

Overall for this group of 2010 silage yields, the GCA and hybrid prediction both reliably predicted trial yields. Not surprisingly, yield was associated with large variability and DM content predictions were more accurate than yields.

6 Conclusions

- For silage yield prediction, GCA predictions are relatively accurate.
- Grain yield prediction was less accurate than silage yield and hybrids with lower yields were the least accurate. Prediction methods for this trait need to be re-evaluated.
- For both yield traits, the SCA effect was not significant, therefore to simplify and increase efficiency the SCA may be excluded from future prediction calculations.
- Both silage and grain yield predictions were improved significantly by basing predictions on trial data from up to 4 prior years. However this was not consistent, particularly for grain where there was no added advantage of using data from more than two years prior. This should be re-evaluated in future.

For Limagrain, learning that the SCA does not have a significant contribution to hybrid yields in practice within a their commercial program has the potential to save resources. It is not necessary to test inbred lines with more than two tester lines, for instance, since the SCA from one specific combination is unlikely to result in a superior hybrid. In addition, although it is important to test inbreds in multiple locations and for several years, it is possible that year effects contribute to apparent contributions of SCA.

The evaluation of hybrid performance prediction performed here was a relatively small, easy project. To improve evaluation of accuracy of various prediction methods cross-validation could be used such as that performed by Schrag et al. (2010). This involves dividing data into a calibration set and a test set for validation of the prediction methods. Normally, multiple rounds of cross-validation are performed using different divisions in these two data sets to decrease variability. More importantly, in order to determine the sources of error in the prediction methods, for grain yields for example, the BLUP models used should be examined by a statistician. Currently, these complex calculations are performed by software called Abacus developed by Limagrain's Bioinformatics team.

7 Hybrid Performance Prediction in the Future

Marker-based approaches allow predictions of hybrid performance to be made for hybrids which have not been tested in the field to be made on the basis of field trial data available from related crosses (Schrag et al. 2010). Bernardo (1994) applied the best linear unbiased prediction (BLUP) approach to maize, which used a combination of phenotypic trait data and information about genetic relationships among parental inbreds, based on coancestry coefficients estimated from pedigree records or molecular marker data. The results of this approach were relatively good and it was further improved by estimation of genetic relationships using quantitative trait loci (QTLs).

Schrag et al. (2007) compared the GCA prediction method with several marker-based predictions of hybrid performance. They found that only for a large number of inbred combinations (98), marker-based prediction approaches were more efficient than a GCA-based approach. For smaller experiments the GCA-based approach had the best prediction efficiency. However, they remark that haplotype block length estimates were affected by missing marker observations and hybrid performance prediction using this method could be improved by using a procedure to estimate haplotypes which is unaffected by missing observations. Comparing the two marker-based methods, predictions based on haplotype blocks resulted in improved prediction over the use of single AFLP markers. In conclusion, hybrid performance prediction has become sophisticated and relatively accurate, however difficulties still lie in application of these methods to breeding programs due to their complexity and cost. At Limagrain, although there is some marker data and certainly pedigree data available across the company, currently only phenotypic data is used to calculate combining abilities due to data gaps and lack of confidence in the additional methods. This may be the best option for efficient hybrid prediction at this time, as Schrag et al. (2010) concluded that prediction efficiency was highest using pedigree and line per se data, and that molecular markers were more efficient only if neither of these data were available.

8 Internship at Limagrain: Summary

The most important thing my time at Limagrain provided was an idea of the day-to-day activities of a commercial breeding program. This proved to be both an excellent experience for me and sometimes a less than academic experience as well.

Internship activities:

- Selfing (hand pollination) and harvesting of double haploid plants in the greenhouse
- Selfing and crossing of inbreed lines in the field
- Estimation of silk and tassle flowering time in nurseries
- Scoring of various agronomic traits (using UPOV criteria)
- Inoculation of cobs with *Fusarium*
- Collecting and preparing leaf and stem samples
- Analysis of GCA data
- Harvesting of diallel crosses in the field
- Shelling and processing of seed
- Introduction to NIR techniques

Although these tasks often required little knowledge of plant breeding, I learned a great deal about how breeding programs work in practice and got some much needed field experience. My supervisors always informed me which program I was working in and how the work would contribute.

It was eye opening to see all programs (and all stages of all programs) running at once, as opposed to one single breeding scheme step-by-step. For example, a nursery for breeding material is much more complicated than I expected (Appendix 2). I also realized that it takes an incredible amount of meticulous work to run a successful breeding program and that good people are critical. You also need to be very organized not only to organize the logistics of shipping seed, labelling, etc. but also the people.

Breeders

Despite the fact that I was involved in the two busiest times of year for maize breeders: flowering in July and harvest in early October, I was overwhelmed by how much breeders travel. My supervisors were constantly on the road and many times I was as well. Mainly due to the travelling I come to the opinion that you cannot work as a breeder 40 hours per week, it is a much larger commitment, but this is generally the case in agriculture. My colleague described this as a lifestyle choice. Despite the work of the breeder, success of a program is more determined by the germplasm of the company and due to long timelines, the work of the breeders prior to yourself. Thus, I have realized that a choosing a company and a crop may be more important than I thought in the past.

Breeding Companies

Breeding companies tend to be large and their sizes are growing. Logically, working for large companies is associated with a different company atmosphere and stricter rules, but this is unlikely to change.

It was interesting for me to see how the people here have dealt with the sale of Advanta to Limagrain in 2005. Especially while I was in Colmar, where a LG team and our team (formerly Advanta) worked out of the same building, the different approaches were very evident. Part of this was no doubt due to cultural differences between France and the Netherlands or Germany, but nevertheless after 5 years of working for Limagrain, my team generally stuck to the same methods used when they were part of Advanta.

8.1 Recommendations

Overall, although I gained some valuable experience in my time at Limagrain, I would not recommend this particular internship to a peer. Certainly, housing an intern is a lot of work for the hosting team. My supervisors were in general too busy to provide the time and instruction necessary to optimally use my time here. At the same time the breeding team is very good; constantly updating their strategies and applying the newest techniques. It was good to see that breeding is such a dynamic career, not only throughout different times of the season, but over several years. The company also provided me with a place to live while working there, for which I am very grateful. My outlook on the research the company does as a whole is positive, and surely other internships within the company may be more positive.

9 References

Barrière, Y, D Alber, O Dolstra, C Lapierre, M Motto, A Ordas, J Van Waes, L Vlaswinkel, C Welcker and JP Monod. 2005. Past and prospects of forage maize breeding in Europe. I. The grass cell wall as a basis of genetic variation and future improvements in feeding value. Maydica. 50:259-274.

Barrière, Y, D Alber, O Dolstra, C Lapierre, M Motto, A Ordas, J Van Waes, L Vlaswinkel, C Welcker and JP Monod. 2006. Past and prospects of forage maize breeding in Europe. II. History, germplasm evolution and correlative agronomic changes. Maydica. 51: 435-449.

Beck, DL, SK Vasal and J Crossa. 1990. Heterosis and combining ability of CIMMYT's tropical early and intermediate maturity maize (*Zea mays* L.) germplasm. Maydica. 35:279-285.

Bernardo, R. 1994. Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Science. 34:20-25.

Bernardo, R. 2003. On the effectiveness of early generation selection in self-pollinated crops. Crop Science. 43(4):1558-1560.

Gardner, CO and SA Eberhart. 1966. Analysis and interpretation of the variety cross diallel and related populations. Biometrics. 22:439-452.

Schrag, TA, HP Maurer, AE Melchinger, HP Piepho, J Peleman and M Frisch. 2007. Prediction of single-cross hybrid performance in maize using haplotype blocks associated with QTL for grain yield Theoretical and Applied Genetics. 114:1345-1355.

Schrag TA, J Mohring, AE Melchinger, B Kusterer, BS Dhillon, HP Piepho and M Frisch. 2010. Prediction of hybrid performance in maize using molecular markers and joint analyses of hybrids and parental inbreds. Theoretical and Applied Genetics. 120(2):451-461.

Smith, OS. 1986. Covariance between line *per se* and testcross performance. Crop Science. 26:540-543.

Sprague, GF and LA Tatum. 1942. General versus specific combining ability in single crosses of corn. Journal of American Society of Agronomy. 34:923-932.

10 Appendix 1: Genstat Output

1) Example of two-sample paired t-test: comparison of GCA and hybrid predictions for grain yield

```
Summary

   Sample        Size           Mean        Min        Max
        1         642          115.4         89        137
        2         642          115.4         86        141

Difference between means: -0.00424
Approx s.e. of difference:    0.600

Test of null hypothesis that mean 1 is equal to mean 2

Probability                                 =        0.9872

95% confidence interval for difference: (-1.180, 1.171)
```

2) One-way ANOVA for silage yield prediction based on single and multiple years

```
Analysis of variance

Variate: YLDS

Source of variation              d.f.          s.s.         m.s.      v.r.    F pr.
C1                                  3        51.306       17.102     10.44    <.001
Residual                         2676      4383.391        1.638
Total                            2679      4434.697


Tables of means

Variate: YLDS

Grand mean  18.219

        C1       1 Year Prediction      2 Year Prediction      3 Year Prediction
                            18.426                 18.196                 18.220

        C1       4 Year Prediction
                            18.036

 Standard errors of means

Table                            C1
rep.                            670
d.f.                           2676
e.s.e.                       0.0494

Least significant differences of means (5% level)

Table                            C1
rep.                            670
d.f.                           2676
```

```
l.s.d.                        0.1371

Fisher's protected least significant difference test

                                 Mean
     4 Year Prediction           18.04 a
     2 Year Prediction           18.20 b
     3 Year Prediction           18.22 b
     1 Year Prediction           18.43
```

3) Linear regression for three year GCA predictions and 2010 silage yield results

```
Summary of analysis

Source            d.f.          s.s.          m.s.        v.r.    F pr.
Regression           1         24.10       24.1019       31.66    <.001
Residual            27         20.55        0.7613
Total               28         44.66        1.5949

Percentage variance accounted for 52.3
Standard error of observations is estimated to be 0.873.

Estimates of parameters

Parameter               estimate           s.e.       t(27)     t pr.
Constant                  -12.05           5.52       -2.18     0.038
Trial_Results              1.770          0.315        5.63     <.001
```

11 Appendix 2: Field map of Nursery in Rheinfelderhof, France