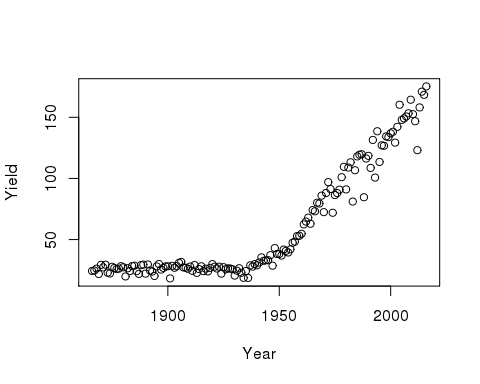
**Methods for measuring genetic gain**

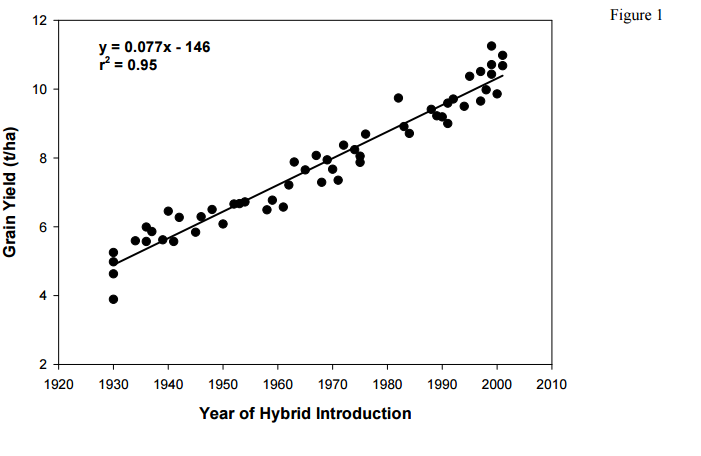
## Literature review

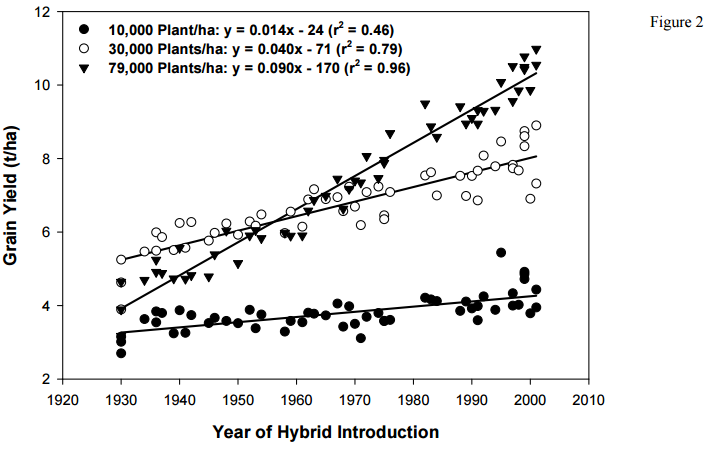
# 1. USDA or country wide yield averages

If there was a single graphic that is almost universal in corn breeding it is the historical corn yields by year for the united states. The USDA has been keeping track of yield in bu/ac since 1866. The slope of the line is a testament to genetic and agronomic improvements made since plant breeding of maize began in the 1920s. [[1]](#footnote-1) This trend is seen in maize, rice and wheat across many countries. [[2]](#footnote-2) 

The dramatic increases of yield from stagnant OPVs that averaged 20-30 Bu/Ac from 1866-1920 to modern hybrid maize that in 2016 yielded 174 Bu/ac on average is impressive. This measurement of genetic gain is remarkable also because of its simplicity of measurement and calculation. The USDA calculates harvested acres and collects grain sales data to estimate the average yield per acre. There are years where the yield drops - notice particularly 2012, when we had a severe drought across most of the midwest. No adjustment is made for year effects, location effects, management differences, or biotech trait differences. It is a simple mean regressed across years. Yet it is the gold standard for demonstrating genetic gain. This is the gain that our customers see.

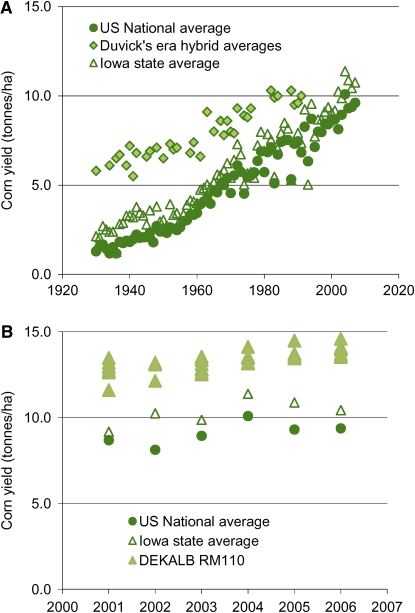
# 2. ERA studies

Era studies measure in the same year in a balanced replicated field experiment hybrids from the past to the present. They have been used to extensively by Duvick [[3]](#footnote-3) and Pioneer[[4]](#footnote-4) to show the realized genetic gain from selection in commercial maize breeding projgrams. Typically yield is plotted by year of hybrid release and the regression of yield x time is used to estimate realized genetic gain as BLUPs or LSmeans. 

These studies monitor genetic progress and are often replicated in different agronomic conditions - population density is the most common, but also nitrogren, and drought. [[5]](#footnote-5) 

Other references [[6]](#footnote-6)

Edgerton (2009) compared Era studies, average yield of 110 RM precommercial hybrids, and P3-P4 advanced hybrids. [[7]](#footnote-7). This is one of the few published reports that directly compares the National/regional averages with the results of era studies (A) and Monsanto P4 yield trials (B).



Comparison of Era, average yields from USA, Iowa, and 110 Monsanto P3-P4 trials.

The analysis of the hybrid data across years is not complicated, because ERA studies are selected hybrids grown in balanced, replicated experiments across years and locations. Notice for most of the years there are only a few hybrids chosen in these published studies. This is to keep the experiment size down to a reasonable number to test across as many years as possible. Selection of the hybrids to represent each year should be consistent across the eras under scrutiny - either chosen based on sales data, impact, yield, or random choice. The model within the analysis is pretty much simple regression. Not a lot of worry about year or location effects.

Surprisingly, there is remarkable agreement between published Era studies and average corn yield improvements.

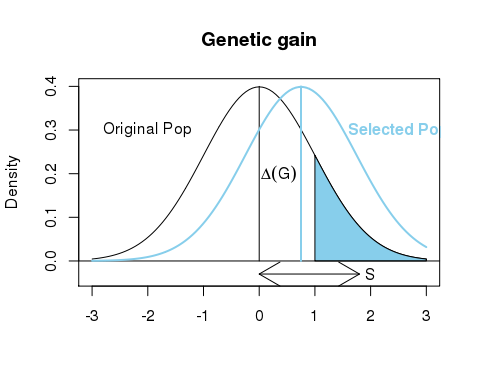
# 3. Yield contest and max yield studies

As seen by Edgerton (2009) there is a drop in yields from Monsanto precommercial trials to average yields seen by farmers. There is also a gap between the maximum pos yields and those seen in our trials and average yields. For example, in 2016 national yield winners were as high as 532 bu/ac and no winner was lower than 302 bushels. [[8]](#footnote-8) National average yields remain below 180 bushels. This yield gap between maximum potential and observed yield is a function of management practices, fertilizer amounts, ideal moisture and environmental conditions.[[9]](#footnote-9) Improved genetics has not increased the maximum yield much - argued by Tollenaar and Lee (2002) because of only minor differences in yield at 10,000 plants/ha, but as density increased, so did the response in yield. Genetic improvement made yield gains possible with more intensive management. Older, lower yielding varieties in more intensiver production environments do not show improved yields.

Some physiologists have argued that we are approaching the maximum potential yields.[[10]](#footnote-10) This is in contrast to the relative linear genetic gain observed across many countries and years reported by Sasha Hafner. [[11]](#footnote-11) What is clear is that yield improvement is not independent from management practices. Genetic gain and improved agronomic practices are dependent on each other. We will struggle to increase yields in marginal environments or where agricultural practices are poor with just improved genetics because the limiting factor is not genetics, but environmental.

# 4. Predicted genetic gain within a year

Theoretical genetic gain for truncating selection is a function of heritability, selection intensity, and mating control.



If heritability is estimated correctly, then theoretically we could estimate genetic gain from the selection differential. This was tested by Jesus Saldivar (1998)[[12]](#footnote-12) Saldivar was working with Tom Hoegemeyer with SC1, SC2, PCM1 and PCM2 data from two years of a commercial seed company based in NE, presumable Hoegemeyer seeds. Using the model

with heritability estimated as

He calculated realized genetic gain simply as

where

# 5. Realized genetic gain - selections performance in a second year

Within Monsanto we have generally followed the procedure from St. Martin and McBlain[[13]](#footnote-13), which is also used by Saldivar.

This method uses the mean of common checks from the current year and checks from the previous year to normalize year and location effects. The mean difference between the two years after adjusting for common check effects is the estimate of realized genetic gain. Another source for this approach is Khalil et al. (1995)[[14]](#footnote-14). Saldivar found very little correlation between the estimate within a year and the estimate including performance in a second year. Presumably the estimate of heritability within a year may not be that good, which for his dataset makes sense the early testing was highly variable with only 2% of the variance from genetic components. Heritability estimates were from 0.00 to 0.8. Heritability improves and stabilied in later stage trials 0.5-0.8. I suspect the low correlation may be due to poor estimation of heritability.

Two step analysis method from Minghui

proc hpmixed data=dat\_chk; by sample;  
class YEAR loc\_id PEDIGREE;  
model yld\_fld=PEDIGREE;  
random YEAR loc\_id(YEAR) / solution;   
ods output solutionR=YEAR\_eff(keep=Effect sample YEAR loc\_id estimate);   
run;   
  
proc sql;  
create table dat\_cn as select distinct a.\*, b.estimate as eff\_yr 'eff\_yr' from k.dat\_all a, year\_eff b where a.year=b.year and a.sample=b.sample and b.effect='year';  
create table dat\_cn\_1 as select distinct a.\*, b.estimate as eff\_loc 'eff\_loc' from dat\_cn a, year\_eff b   
where a.year=b.year and a.sample=b.sample and a.loc\_id=b.loc\_id and b.effect='LOC\_ID(year)' order by sample, year, loc\_id, pedigree;  
Quit;  
Data FORGG; set dat\_cn\_1; yld\_adj = yld\_fld - (eff\_yr+eff\_loc); run;  
  
proc hpmixed data=FORGG; where is\_chk='F’; By sample;  
class PEDIGREE;  
model yld\_adj = / S ;  
random PEDIGREE/ s;   
ods output solutionR =BlupEff;   
run; quit;

SAS code for one step approach from Minghui

Proc hpmixed data=dat\_all; by sample;  
Class year pedigree loc\_id;  
Model yld\_fld =;  
Random year pedigree loc\_id/solution;  
ods select SolutionR CovParms;  
ods output SolutionR=blups CovParms=vars;   
run; quit;

For PCM3, the hybrids tested are on the cusp of being the next commercial class for Dekalb and licensees. The source of PCM3 entries could be coded x coded hybrids from previous years inbreds, or recently advanced hybrids from PCM2.5, or diamond mined - older coded x coded. There is not an attempt to balance contribution of individual inbreds are hybrids are not selected based on GCA across balanced testers, and we are trying to leverage the very best genetics from the past and the present. This helps make sure that there is continual improvement in the potential commercial releases, even if there is a weakness in the current class. St. Martin and McBlain's proposed method was based on soybean regional variety trials where the progression from one stage to another was a much more orderly process where each entry is represented a single time and not used as a tester or used across multiple years. This scenario violates the assumptions of the theory underpinning the genetic gain equation. It may overestimate genetic gain due to step change lines and it confounds using the difference between years to measure changes in the breeding pipeline and its effect on genetic gain. At PCM3, the difference between cohorts gives a good estimate of potential commercial classes, but may not reflect completely the genetic gain of each new generation because of the complexity of the coded x coded and diamond mining sets being included together with new P2.5 advancements. [[15]](#footnote-15)

"It is clear that the effect of assumed heritability on estimated genetic gain depends on several factors even in this very simple case. In animal breeding practice EBVs are commonly calculated using multi-trait BLUP, so that genetic correlations are also involved, and data structures can be very complex. It would be of interest to know to what extent estimates of genetic gain are sensitive to assumed parameters in a real breeding programs."

It may be that analyzing the P1 or P2 class that has more balanced advancement based on GCA for the last 5 or 6 years violates fewer parameters.

Other references.

<http://www.sciencedirect.com/science/article/pii/S0378429006001353> - paper from sunflower. They had 122 trials across 10 years and estimated BLUP performance of each year hybrids with the model:

<http://www4.ncsu.edu/~fisik/course/lecture9_%20blp.pdf> fisik course BLUP and genetic gain equations

<http://www.nature.com/articles/hortres201615> - Genetic gain simulations for clonally propogated crops.

<https://dl.sciencesocieties.org/publications/meetings/download/pdf/2013am/81009>

1. <https://www.nass.usda.gov/Charts_and_Maps/Field_Crops/cornyld.php> [↑](#footnote-ref-1)
2. Hafner, S. 2003. Trends in maize, rice, and wheat yields for 188 nations over the past 40 years: a prevalence of linear growth. Agric, Ecosystems, and Env 97:275-283. <http://www.sciencedirect.com/science/article/pii/S0167880903000197> - Hafner [↑](#footnote-ref-2)
3. Duvick, D.N., J.S.C. Smith and M. Cooper. 2004. Plant Breeding Reviews, 24(2) Ed. J. Janick, John Wiley and Sons, Inc. [↑](#footnote-ref-3)
4. [https://www.pioneer.com/home/site/us/agronomy/library/template.CONTENT/guid.F05C40A8-61A9-6998-03D8-531522BA2117#](https://www.pioneer.com/home/site/us/agronomy/library/template.CONTENT/guid.F05C40A8-61A9-6998-03D8-531522BA2117) [↑](#footnote-ref-4)
5. <http://www.cropscience.org.au/icsc2004/symposia/3/4/1968_nieburw.htm> [↑](#footnote-ref-5)
6. <https://www.researchgate.net/profile/J_Smith6/publication/228032721_Long-Term_Selection_in_a_Commercial_Hybrid_Maize_Breeding_Program/links/55cb382808aeca747d6a822f.pdf> [↑](#footnote-ref-6)
7. Edgerton, M. D. 2009. Increasing crop productivity to meet global needs for feed, food, and fuel. Plant PHysiol. 149: 7-13. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2613695/> [↑](#footnote-ref-7)
8. <http://www.ncga.com/upload/files/documents/pdf/ncyc/2015-National-Winners.pdf> [↑](#footnote-ref-8)
9. Tollenaar, M. and E.A. Lee. 2002. Yield potential, yield stability, and stress tolerance in maize. Field Crops Research 75:161-169 [↑](#footnote-ref-9)
10. Lobell, D.B., K.G. Cassman, and C.B. Field. 2009. Crop yield gaps: their importance, magnitueds, and causes. Annu. Rev. Encion. Resour. 34:179-204. [↑](#footnote-ref-10)
11. Hafner, S. 2003. Trends in maize, rice, and wheat yields for 188 nations over the past 40 years: a prevalence of linear growth. Agric, Ecosystems, and Env 97:275-283. [↑](#footnote-ref-11)
12. Saldivar, J.D. 1998. Evaluation of performance of corn yield trials in a commercial corn breeding program in the US corn belt: Measurement of genetic gain. Dissertation. Univ of Nebraska. [↑](#footnote-ref-12)
13. St. Martin, S.K. and B.A. McBlain. 1991. Procedure to estimate genetic gain by stages in multistage testing programs. Crop Sci 31:1367-1369 [↑](#footnote-ref-13)
14. Khalil, I.H., B.F. Carver, E.L. Smith. 1995. Genetic gains in two selection phases of a wheat-breeding programme. Plant Breeding 114:117-120 [↑](#footnote-ref-14)
15. James, J.W. 1995. Predicted and observed responses in BLUP estimates of genetic gain. Computing techniques: Developments and validations. Ed. by the Association for the Advancement of Animal Breeding and Genetics. [↑](#footnote-ref-15)