



# Empirical results and interpretation

Jinliang Yang  
Nov. 7th, 2018

# Breeder's equation

The complex version

$$R = \frac{ih^2\sigma_P}{L}$$

# Breeder's equation

The complex version

$$R = \frac{ih^2\sigma_P}{L}$$

The simple version

$$R = h^2S$$

# Breeder's equation

The complex version

$$R = \frac{ih^2\sigma_P}{L}$$

The simple version

$$R = h^2 S$$

The heritability, therefore

$$h^2 = R/S$$

# Realized heritability

$$h_R^2 = R/S$$

Shows how the response is related to the selection differential

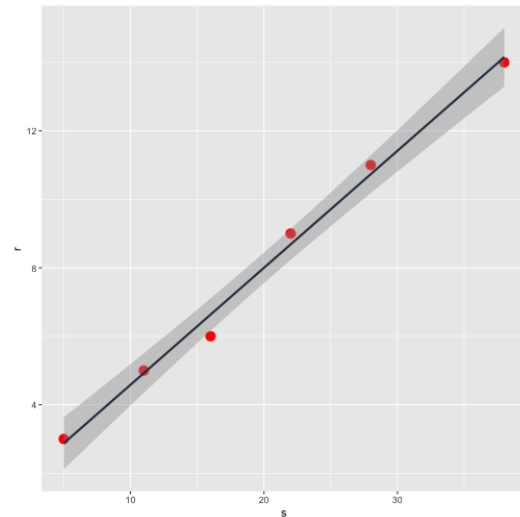
# Realized heritability

$$h_R^2 = R/S$$

Shows how the response is related to the selection differential

```
S <- c(5, 6, 5, 6, 6, 10)
R <- c(3, 2, 1, 3, 2, 3)
df <- data.frame(s=cumsum(S),
                 r=cumsum(R))
```

```
library(ggplot2)
ggplot(df, aes(x=s, y=r)) +
  geom_point(color='red', size = 4)
  geom_smooth(method=lm, color='#2C3E50')
```



# Realized heritability

$$h_R^2 = R/S$$

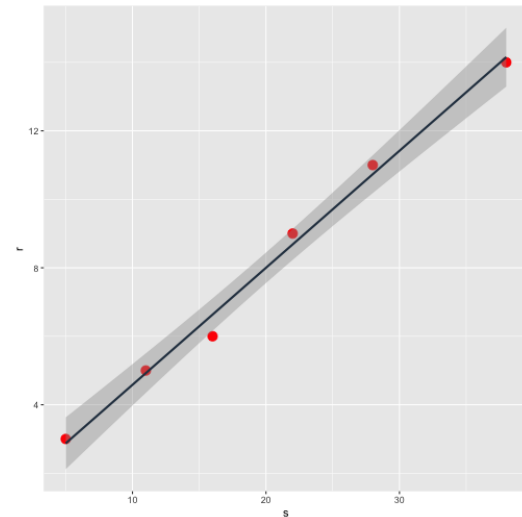
Shows how the response is related to the selection differential

Be cautious:

1. Reduces response to selection after first generation for high heritability traits (**Bulmer effect**)

```
S <- c(5, 6, 5, 6, 6, 10)
R <- c(3, 2, 1, 3, 2, 3)
df <- data.frame(s=cumsum(S),
                 r=cumsum(R))
```

```
library(ggplot2)
ggplot(df, aes(x=s, y=r)) +
  geom_point(color='red', size = 4)
  geom_smooth(method=lm, color='#2C3E50')
```



# Realized heritability

$$h_R^2 = R/S$$

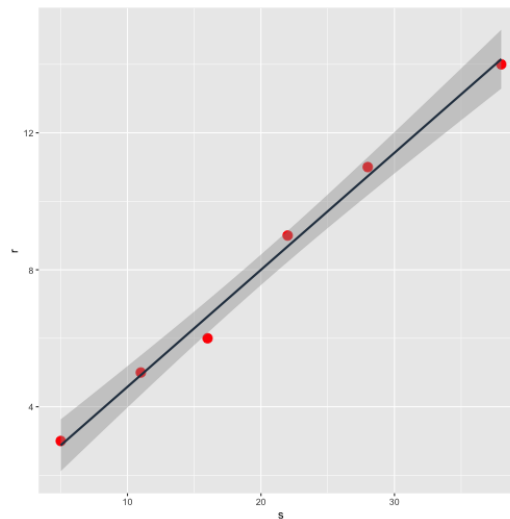
Shows how the response is related to the selection differential

Be cautious:

1. Reduces response to selection after first generation for high heritability traits (**Bulmer effect**)
2. Systematic changes in environment or inbreeding depression will affect response.
3. Random drift affect response.

```
S <- c(5, 6, 5, 6, 6, 10)
R <- c(3, 2, 1, 3, 2, 3)
df <- data.frame(s=cumsum(S),
                 r=cumsum(R))
```

```
library(ggplot2)
ggplot(df, aes(x=s, y=r)) +
  geom_point(color='red', size = 4)
  geom_smooth(method=lm, color='#2C3E50')
```





# Questions about the consequences of selection

1. For **how long** does the response continue?
2. By **how much** the population mean can be changed?
3. What is the **genetic nature** of the limit to further progress?

# Illinois long-term selection experiment

- The experiment started in 1896 by C.G. Hopkins and is [still active!](#)
- They are selecting lines for higher or lower concentration of protein or oil in the maize kernel.
- **Mass selection** was used with a selection intensity of approximately 1 out of 5 for most of the experiment

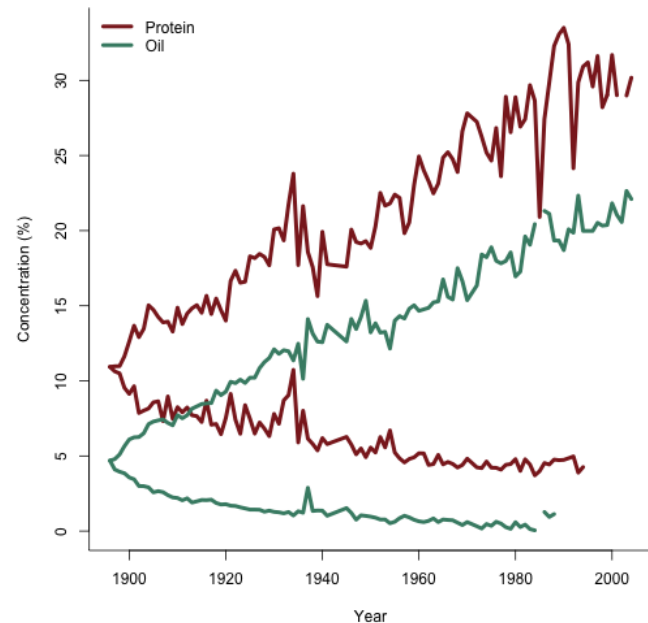
Dudley and Lambert,  
2004, "100 Generations  
of Selection for Oil and  
Protein in Corn"

# Illinois long-term selection experiment

- The experiment started in 1896 by C.G. Hopkins and is [still active!](#)
- They are selecting lines for higher or lower concentration of protein or oil in the maize kernel.
- **Mass selection** was used with a selection intensity of approximately 1 out of 5 for most of the experiment

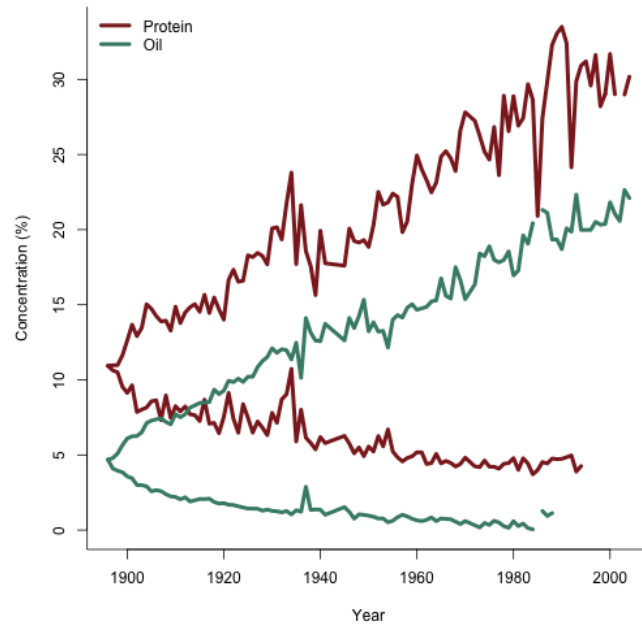
Dudley and Lambert, 2004, "100 Generations of Selection for Oil and Protein in Corn"

```
d <- read.csv(file="https://jyanglab.org/data/ILP_IHP.csv")
par(bty="l", pty="m", mar=c(5, 4, 1, 1))
matplot(x=d$YR, y=d[, c("IHP", "ILP")],
        xlab="Year", ylab="Concentration (%)",
        legend("topleft", c("Protein", "Oil")))
```



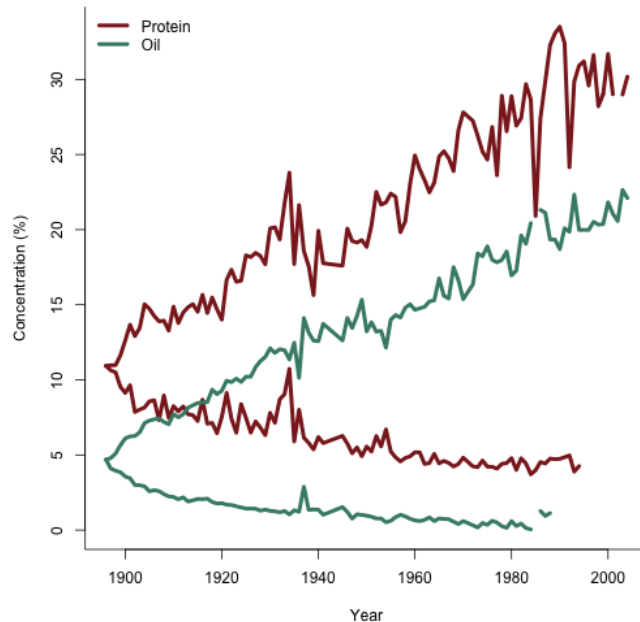
# Illinois long-term selection experiment

## Variability of the responses



# Illinois long-term selection experiment

## Variability of the responses

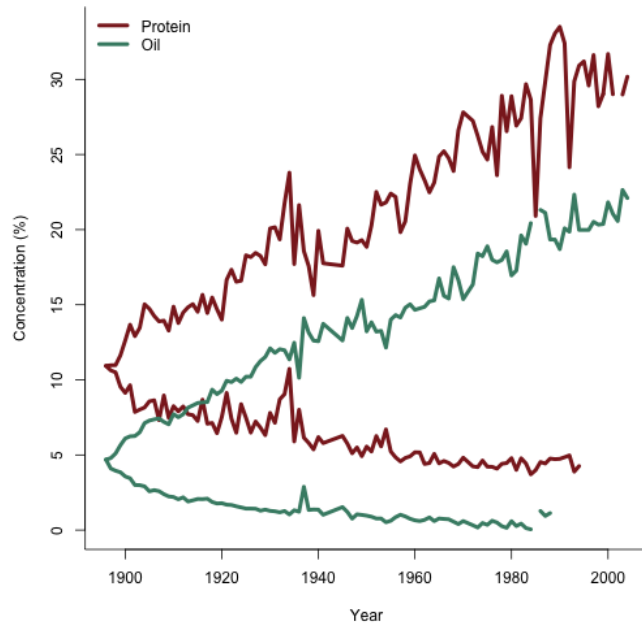


## Possible reasons?

1. **Random drift** due to the restricted number of parents
2. **Sampling error** in estimating the generation mean
3. Variation of **selection differentials**
4. **Environmental factors**

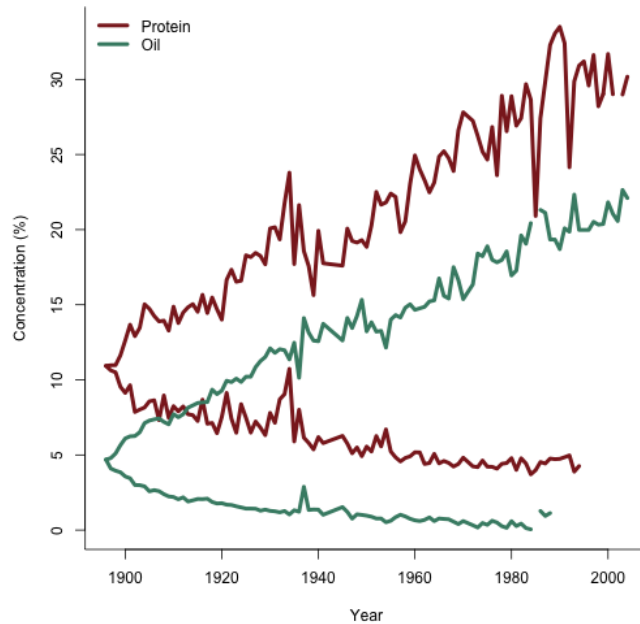
# Illinois long-term selection experiment

## Asymmetry of responses



# Illinois long-term selection experiment

## Asymmetry of responses

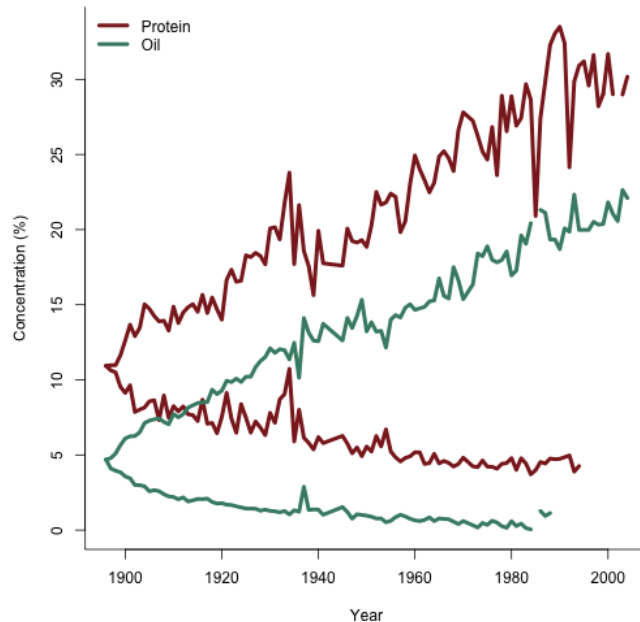


Discussion for possible reasons?

whether it is real (need replications)

# Illinois long-term selection experiment

## Asymmetry of responses



## Discussion for possible reasons?

whether it is real (need replications)

1. Random drift?
2. Selection differentials?
3. Inbreeding depression?
  - reduce the rate in the upward and increase it in the downward direction
4. Others?



# Asymmetry of responses

Discussion for possible reasons?

- Genetic asymmetry
- Genes with large effects

# Asymmetry of responses

Discussion for possible reasons?

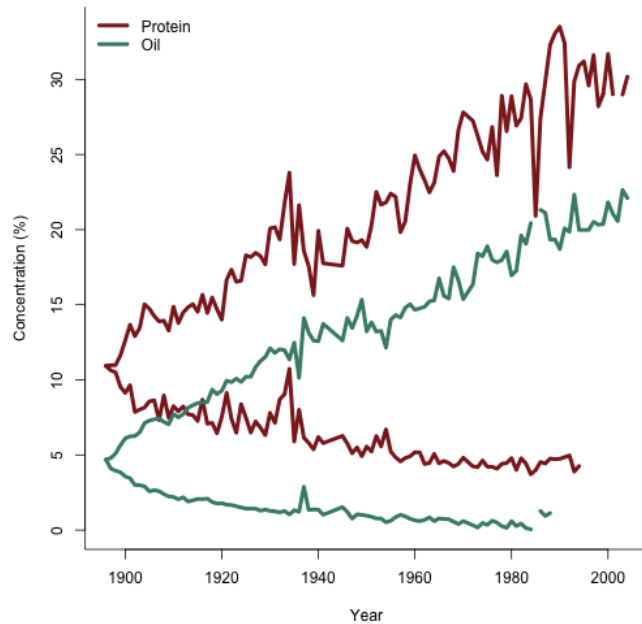
- Genetic asymmetry
- Genes with large effects

**Other reasons (F&M p211-215):**

- Scalar asymmetry
- Indirect selection
- Maternal effects

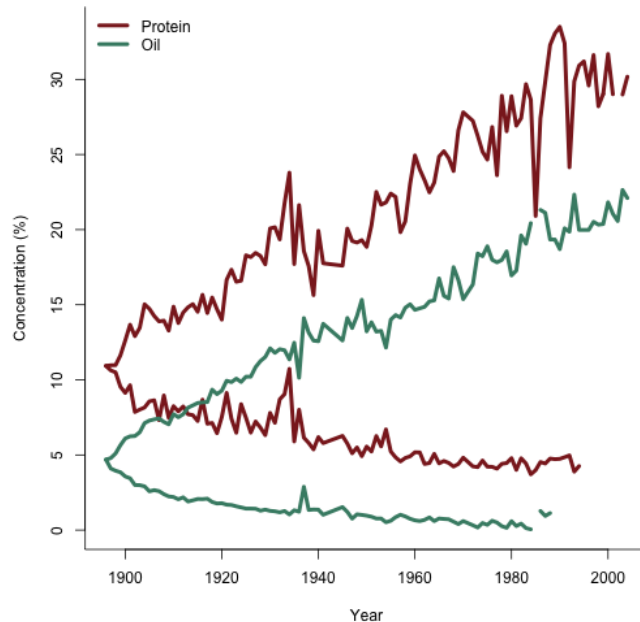
# Illinois long-term selection experiment

## Selection limits



# Illinois long-term selection experiment

## Selection limits



## Questions about selection limits?

1. For **how long** does the response continue?
2. By **how much** the population mean can be changed?

# Selection limits

Total response (  $R_T$  ) = selection limit in high direction - selection limit in low direction

# Selection limits

Total response (  $R_T$  ) = selection limit in high direction - selection limit in low direction

## Genetic architecture

- Total response depends on **number of loci** contributing to a given amount of variation.
- For a given amount of variation, the number of loci is inversely related to **average size of effects**.

# Selection limits

Total response (  $R_T$  ) = selection limit in high direction - selection limit in low direction

## Genetic architecture

- Total response depends on **number of loci** contributing to a given amount of variation.
- For a given amount of variation, the number of loci is inversely related to **average size of effects**.

## Mutation

Long term results become unpredictable, because **mutation** produces new variation whose nature we cannot predict.

# Theoretical maximum limit

Two assumptions:

1. All the loci have the same magnitude of additive effect:  $a$
2. All alleles start at the frequencies of 0.5.



# Theoretical maximum limit

Two assumptions:

1. All the loci have the same magnitude of additive effect:  $a$
2. All alleles start at the frequencies of 0.5.

Then the range between two homozygotes is  $2a$  for one locus. For more than one loci:

$$R_T = \sum_{i=1}^n 2a = 2na$$

# Theoretical maximum limit

Two assumptions:

1. All the loci have the same magnitude of additive effect:  $a$
2. All alleles start at the frequencies of 0.5.

Then the range between two homozygotes is  $2a$  for one locus. For more than one loci:

$$R_T = \sum_{i=1}^n 2a = 2na$$

The additive variance, assuming all alleles start at a frequency of 0.50, then

$$\sigma_A^2 = 1/2 \sum_{i=1}^n a^2 = 1/2na^2$$

# Theoretical maximum limit

Two assumptions:

1. All the loci have the same magnitude of additive effect:  $a$
2. All alleles start at the frequencies of 0.5.

Then the range between two homozygotes is  $2a$  for one locus. For more than one loci:

$$R_T = \sum_{i=1}^n 2a = 2na$$

The additive variance, assuming all alleles start at a frequency of 0.50, then

$$\sigma_A^2 = 1/2 \sum_{i=1}^n a^2 = 1/2na^2$$

The relationship between the range and the additive variance is obtained as

$$\frac{R_T^2}{\sigma_A^2} = \frac{4n^2a^2}{1/2na^2} = 8n$$

# Example:

<b>Observation</b>	<b>Exp1</b>	<b>Exp2</b>	<b>Exp3</b>
$R_T/\sigma_A$	100	50	10
$n$	?	?	?

# Example:

Observation	Exp1	Exp2	Exp3
$R_T/\sigma_A$	100	50	10
$n = R_T^2/8\sigma_A^2$	1250	50	12.5

# Example:

Observation	Exp1	Exp2	Exp3
$R_T/\sigma_A$	100	50	10
$n = R_T^2/8\sigma_A^2$	1250	50	12.5

## Theory of limits:

The total response relative to the initial genetic variation, depends primarily on **the number of loci**.

Addressed the third question:

- What is the **genetic nature** of the limit to further progress?

# Theoretical maximum limit

$$R_T = \sum_{i=1}^n 2a = 2na$$

Only could be achieved if the effective population size being very large.

# Theoretical maximum limit

$$R_T = \sum_{i=1}^n 2a = 2na$$

Only could be achieved if the effective population size being very large.

## In practice

Limited parents used => unfavourable alleles are fixed by genetic drift.



# Theoretical maximum limit

The chance of fixation of a favourable allele:

$$N_e s$$

- $s$ : selection coefficient.  $s = i(2a/\sigma_P)$  given by F&M eq. 11.8.
- $N_e$ : effective population size

# Theoretical maximum limit

The chance of fixation of a favourable allele:

$$N_e s$$

- $s$ : selection coefficient.  $s = i(2a/\sigma_P)$  given by F&M eq. 11.8.
- $N_e$ : effective population size

Therefore, the chance of fixation is a function of  $N_e i$ .

# Theoretical maximum limit

The chance of fixation of a favourable allele:

$$N_e s$$

- $s$ : selection coefficient.  $s = i(2a/\sigma_P)$  given by F&M eq. 11.8.
- $N_e$ : effective population size

Therefore, the chance of fixation is a function of  $N_e i$ .

The total response should be greater:

1. with larger population size ( $N_e$ )
2. with more intense selection ( $i$ )

# Theoretical maximum limit

Breeder's equation:

$$R = ih^2\sigma_P$$

This is the predicted response in one generation.

# Theoretical maximum limit

Breeder's equation:

$$R = ih^2\sigma_P$$

This is the predicted response in one generation.

The theoretical half-life of response:

Up to about  $2N_e$  generations.

# Theoretical maximum limit

Breeder's equation:

$$R = ih^2\sigma_P$$

This is the predicted response in one generation.

The theoretical half-life of response:

Up to about  $2N_e$  generations.

$$R_{max} = 2N_e ih^2\sigma_P$$

Robertson 1960

# Theoretical maximum limit

Breeder's equation:

$$R = ih^2\sigma_P$$

This is the predicted response in one generation.

The theoretical half-life of response:

Up to about  $2N_e$  generations.

$$R_{max} = 2N_e ih^2\sigma_P$$

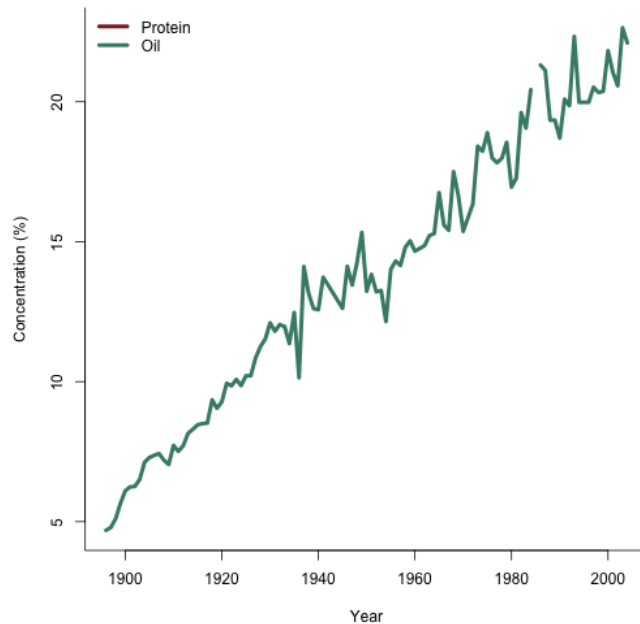
Robertson 1960

Addressed the first and second question:

For **how long** does the response continue? By **how much** the population mean can be changed?

# Illinois long-term selection experiment

Predict the theoretical limit?



From Moose et al., 2004 and  
Dudley and Lambert, 2004:

- In 1896, 163 ears from the open-pollinated variety Burr's White, 24 highest ears formed the Illinois high oil strain (IHO) (**Ne ~ 4-12**)
- Later, 300-500 kernels => **Ne raised to a maximum of 96**
- About 20% selected in each generation
- Realized  $h^2 = 0.25$
- $\sigma_P$  about 0.9

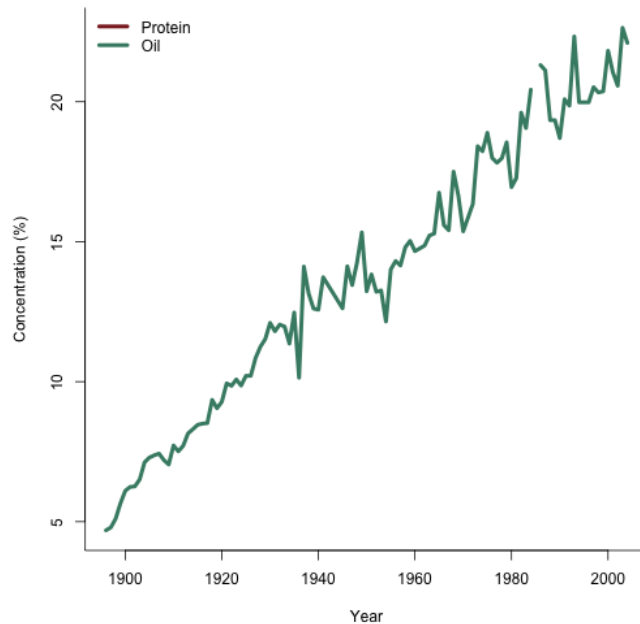
```
ifun(0.2)
```

```
## [1] 1.39981
```



# Illinois long-term selection experiment

Predict the theoretical limit?



$$R_{max} = 2N_e i h^2 \sigma_P$$

```
2*12*ifun(0.2)*0.25*0.9
```

```
## [1] 7.558972
```

```
2*96*ifun(0.2)*0.25*0.9
```

```
## [1] 60.47177
```