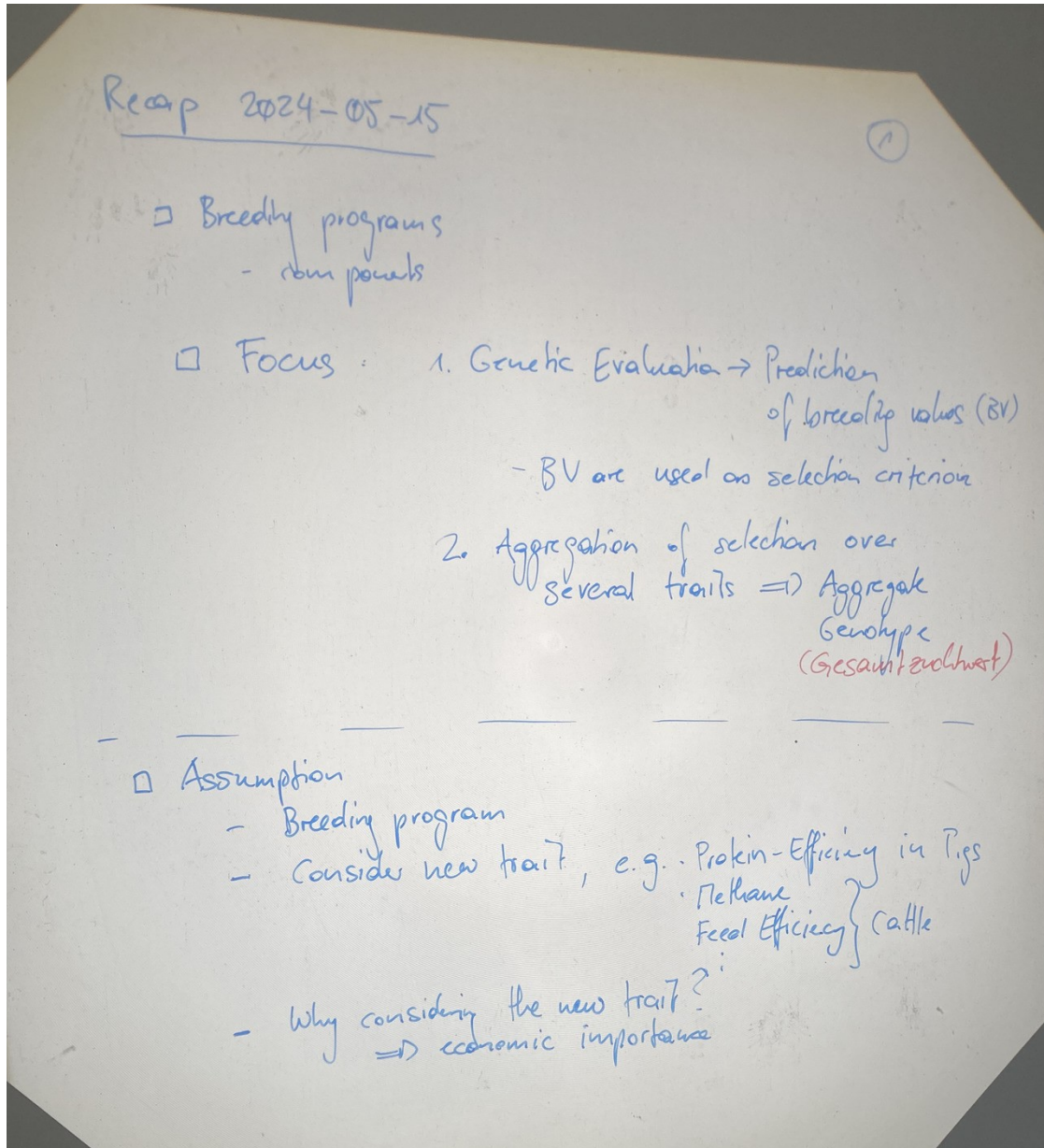


OHP Picture 1



OHP Picture 2

- Reason why consider a new trait in breeding program and thereby improving the population for this trait: Long-term goals. ③

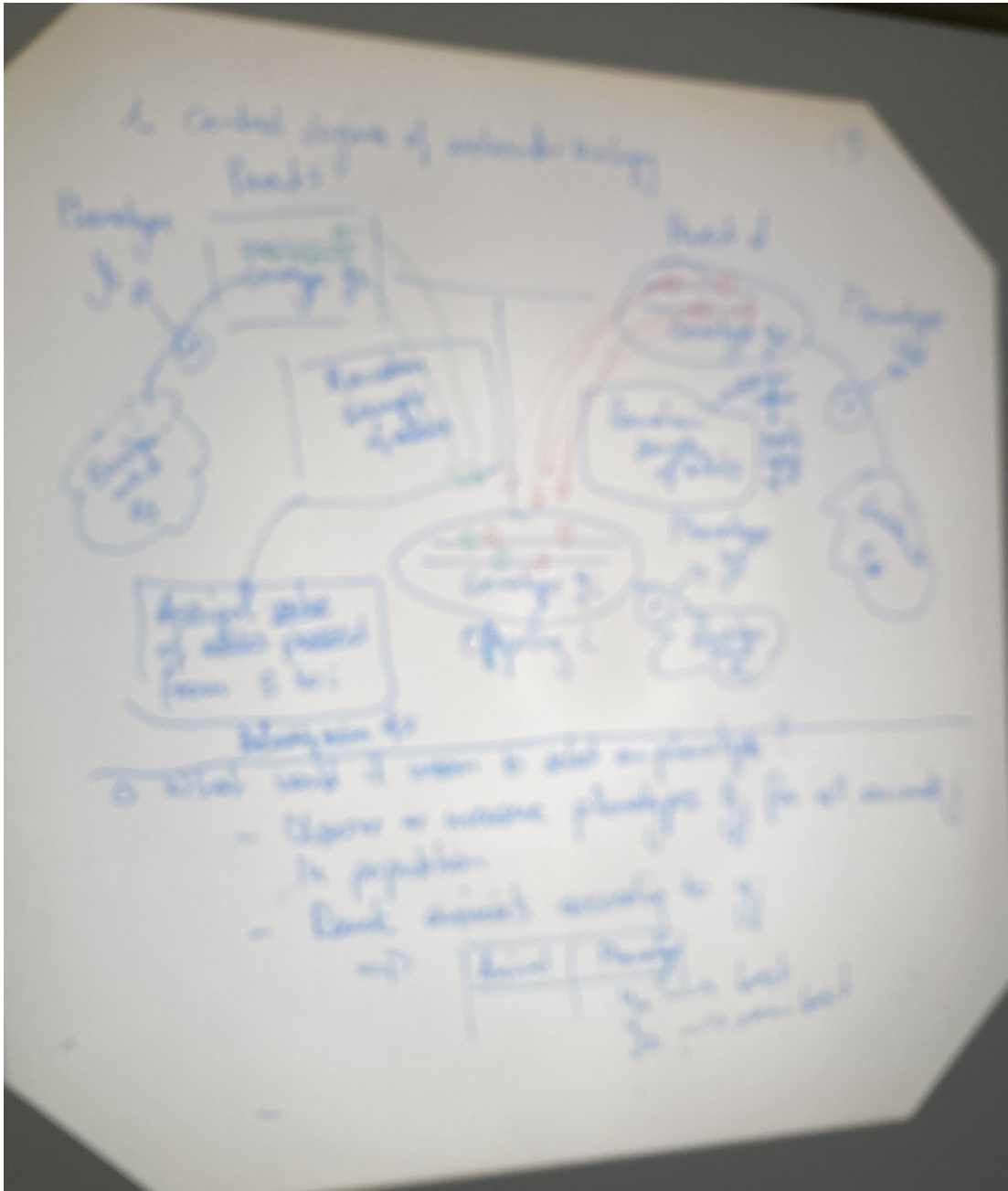
Because improvement at the genetic level via breeding (selection mating) happens from generation to the next generation \Rightarrow Planning horizon generation intervals (Pigs 1.5-2yrs) (Cattle: 5-7 years)

- Short-term problems usually require ^{months to 1 year} management solutions outside of the scope of breeding activities.
 Work at the farm-level
 2-3 years with genomic selection)

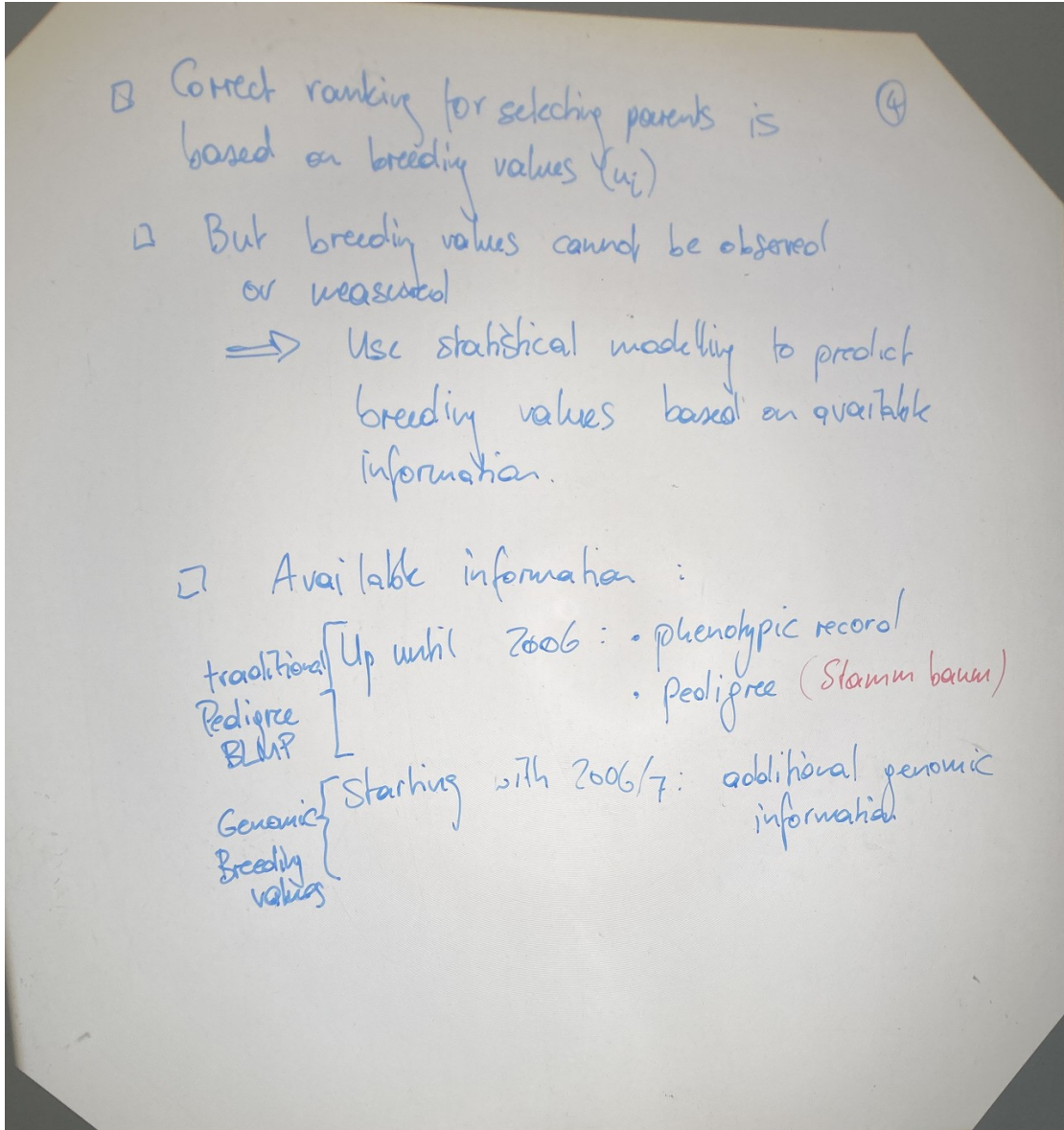
Selection:

- Selection on phenotypes is inefficient
- Take into account
 - phenotype $y_i = g_i + e_i$; because of Central Dogma of molecular Biology
 - inheritance between parents and offspring

OHP Picture 3



OHP Picture 4



OHP Picture 5

□ Statistical Model is based on the decomposition of the phenotype y_i into different components (5)

□ First step: genetic/environment

$$y_i = g_i + e_i$$

□ Second step: decomposition of genetic part

$$\begin{aligned} g_i &= \mu + \text{breeding value} + \text{dominance} + \text{interaction} \\ &= \mu + \text{breeding value} + \text{rest} \end{aligned}$$

Insert into y_i :

$$y_i = \mu + \text{breeding value}_i + e_i^*$$

Herd, age, height, ...

□ Third step: Environment into known fixed part and unknown random part

$$y_i = \mu + X_i^T \cdot b + u_i + e_i^{**}$$

□ Matrix-vector: $y = X \cdot \underline{b} + \underline{Z} \cdot \underline{u} + \underline{e}$

□ Statistical Model

$$y = Xb + \sum u + e$$
 → random predictor (breeding value) ^③
 → random residuals
 → predictors (x_1, x_2, \dots); vector b : unknown fixed effects
 → vector of response variables (trait measurements)

□ function $m(x, u) \Rightarrow$ additive function
 linear in parameters b, u
 not terms like $b^2, u^3, \exp(b)$
 non-linear

One trait y

□ Application to genetic evaluation :
 - very many predictors, where only a restricted number is meaningful for responses

□ Example :

- Body Weight as response $\Rightarrow y$
- 2 predictors :
 - Breast Circumference
 - Random Predictor

□ Full Model :

(Intercept)	: -1218
BC	: 8.53

Red Model :

(Intercept)	: -1065
BC	: 8.67

OHP Picture 7

□ Systematic Approach to select predictors: Model Selection (?)

□ Including predictors which are not associated with response, they lead to increased variation in estimates of other predictors (example: include Rand Pred into model changed estimate of intercept and BC)

⇒ Bias-Variance-Tradeoff

□ Too few predictors lead to bias in estimates

Too many predictors ⇒ increased variance of estimates

$$C_p = \frac{SSE(M)}{\hat{\sigma}^2} - \frac{n + 2|M|}{n} \rightarrow \# \text{ pred.} \quad (8)$$

$\hat{\sigma}^2$ → residual error variance for the full model
 n → number of observations
 $|M|$ → # pred.

□ Model selection is used to determine meaningful set of fixed predictors

□ $y = Xb + Zu + e$

Xb → determined by model selection

Why is u random?
 → Because the structure of variation of breeding values has to be taken into account.

□ For random parts: important to have detectable variation in the data

□ In livestock breeding estimates of variance components are done only rarely and in separate evaluation.

OHP Picture 9

□ Prediction of breeding values

$$y = Xb + Zu + e$$

Annotations:
- y : known observations
- X : known from dataset
- Z : known from pedigree
- e : error term

- vectors b , u and e are unknown
where e is not of interest

Goal : Use known parts y, X, Z to
compute estimates \hat{b} for unknown b
and compute prediction \hat{u} for u

- Fixed Least Squares
Rand BLUP

Livestock Breeding Statistics
Fr 9-12