59th Annual EAAP Meeting
August 24-27, 2008, Vilnius, Lithuania
Session 23, Abstract 2580
Email: birgit.zumbach@norsvin.no

Use of serial pig body weights for genetic improvement

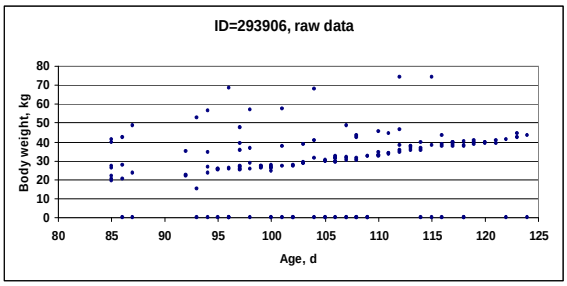
B. Zumbach, I. Misztal, C. Y. Chen, S. Tsuruta,
W. O. Herring, T. Long, M. Culbertson

norsvin

Smithfield Premium Genetics Group

Introduction

Sample weight data from electronic feeders for swine



Objective

- Find algorithms to „clean“ the data

- Find whether the data is useful for genetic analyses of short-term gains

Robust Regression

- Weighted least squares
- Weights=f(residual, SD)
- Determines outliers
- ROBUSTREG in SAS

Methods of Robust Regression

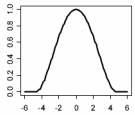
M- estimation (Huber, 1973):

$$\min \sum_{i=1}^{n} w_i^2 e_i^2$$

$$Bisquare\ weight\ \ function: w(e_i, \sigma, wf) = \left[1 - \left(\frac{e_i}{\sigma\ wf}\right)^2\right]^2$$

wf – weight factor

$$OUTLIER = \begin{cases} 1\ if\ |e_i| > c\ wf \\ 0\ \ otherwise \end{cases}$$



c – cutoff value

Data

3,888 Duroc pigs,

Osborne FIRE ®
feeder station

↓

884,132 weight data
(June 2004-Feb. 2007)

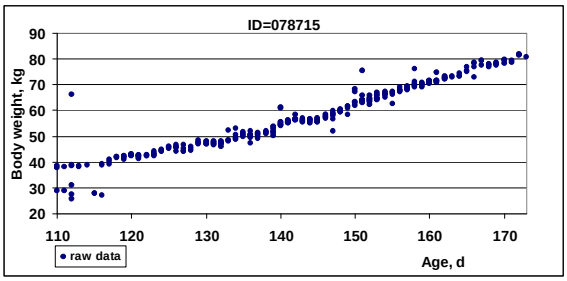
Age at beginning of test: 77-149 days
Age at end of test: 95-184 days

Statistical model for robust regression:

Body weight = $a_0 + b_1$*age + b_2*age^2 + e

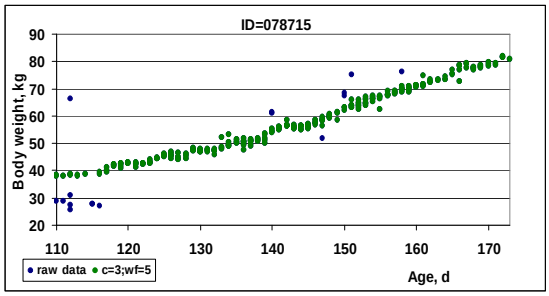
Results RR

Effect of cutoff (c) values on outlier detection



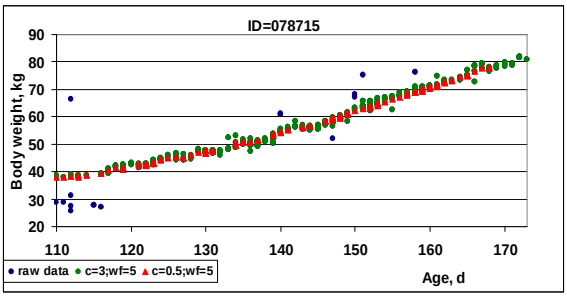
Results RR

Effect of cutoff (c) values on outlier detection



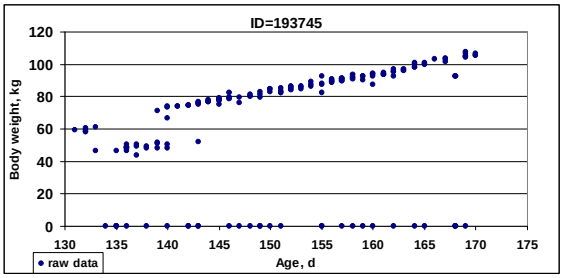
Results RR

Effect of cutoff (c) values on outlier detection



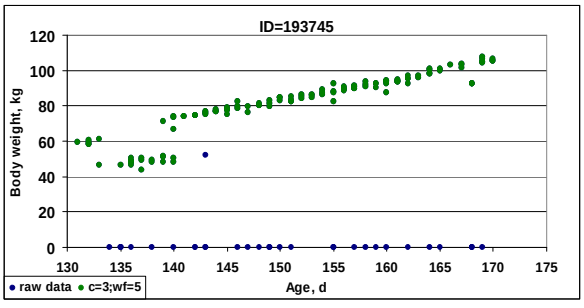
Results RR

Effect of cutoff (c) values on outlier detection



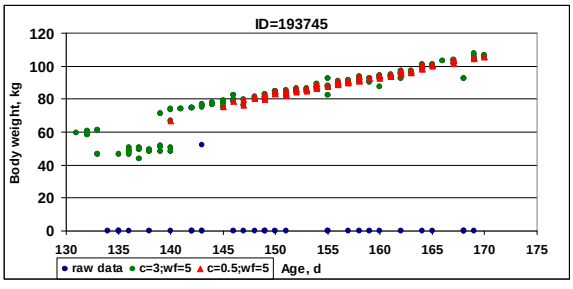
Results RR

Effect of cutoff (c) values on outlier detection



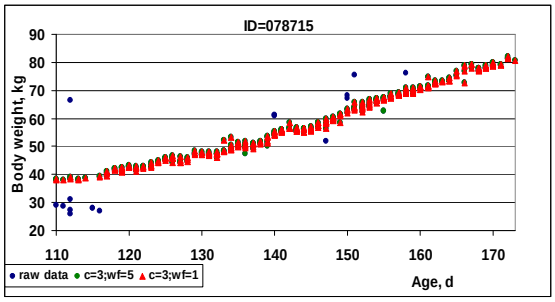
Results RR

Effect of cutoff (c) values on outlier detection



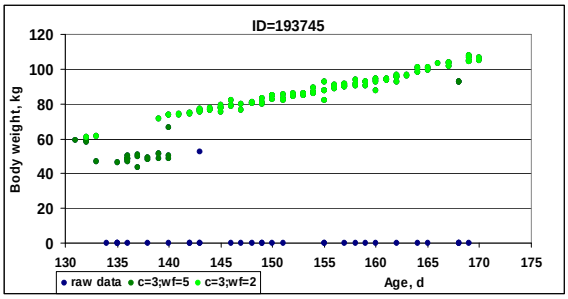
Results RR

Effect of wf-values (weight function) on outlier detection



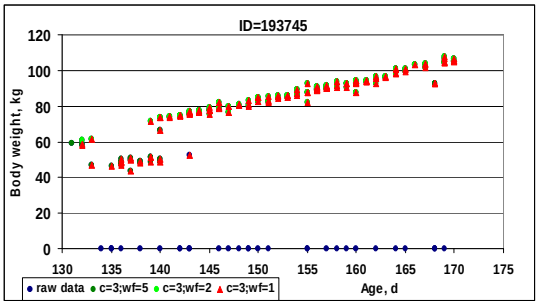
Results RR

Effect of wf-values (weight function) on outlier detection

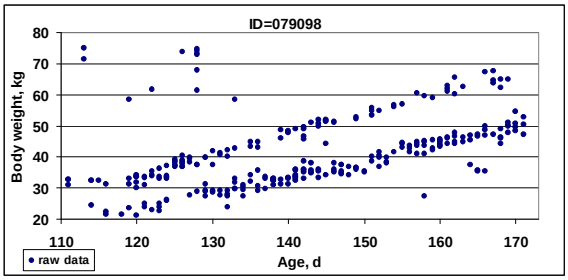


Results RR

Effect of wf-values (weight function) on outlier detection

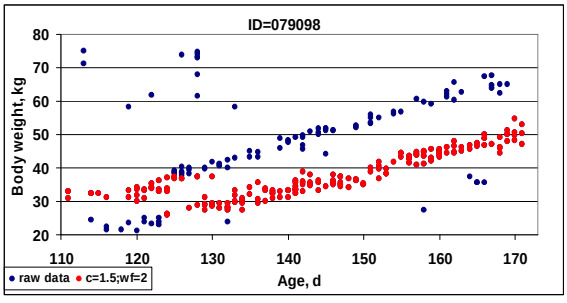


Abnormal growth patterns



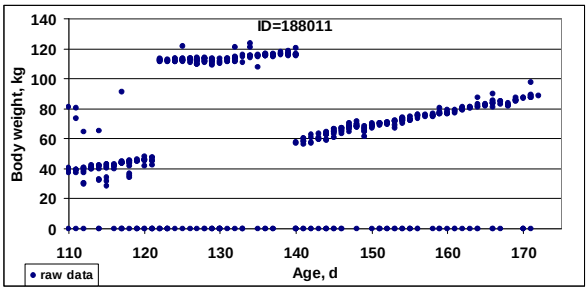
Results RR

Abnormal growth patterns



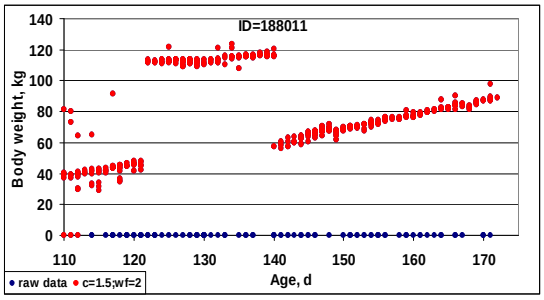
Results RR abnormal

Abnormal growth patterns



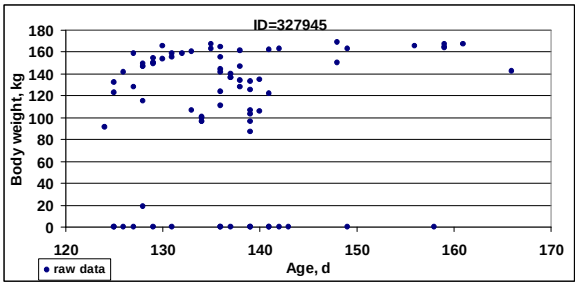
Results RR abnormal

Abnormal growth patterns



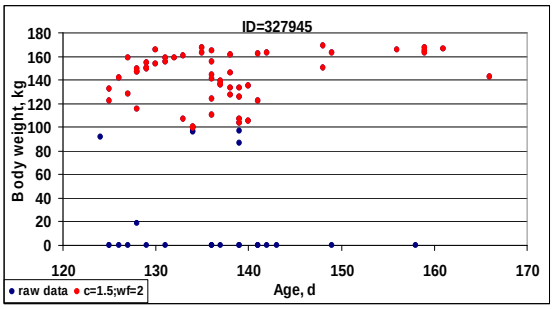
Results RR abnormal

Abnormal growth patterns



Results RR abnormal

Abnormal growth patterns



Data Cleaning

884,132 BW rec. (3,888 animals)

1

Robust regression

M estimation
k=1.5; c=2

607,597 BW rec.

2

Animals ≥ 50 BW rec.
SE residual DG ≤ 2 kg

Calculation of
daily gain (DG)

89,443 DG rec.

Discard records >,< $\bar{x}$ ± 3 SD

69,068 DG rec. (1,921 animals)

Pedigree

# Animals	Daily	Weekly	Bi-weekly
In pedigree	3,003	2,961	2,785
With records	1,921	1,884	1,735

Simple Statistics

Mean (SD) of daily gain and age

	Daily	Weekly	Bi-weekly
N	69,068	9,057	4,232
Age, d	129 (21)	127 (21)	126 (20)
Daily gain, g	861 (998)	869 (357)	872 (318)

Genetic analyses – Repeatability model

$$y = X\beta + Z_1a + Z_2pe + Z_3l + e$$

ß: - CG: pen - test date (Daily)
pen - year - test week (Weekly)
pen - year - test bi-week (Bi-weekly)
- Age (linear covariate)

a: additive effect
pe: permanent environmental effect
l: litter effect
e: residual effect

Parameters of short-term gain

DG	V_{animal}	V_{litter}	V_{pe}	V_e	h^2
Weekly	3,241 (1,269)	1,251 (733)	4,171 (1,023)	45,804 (852)	0.06 (0.02)
Bi-weekly	3,102 (1,371)	1,369 (924)	5,982 (1,256)	23,992 (777)	0.09 (0.04)

Gains over n days and parameters

- Using repeatability models

Daily gains:

$$y_{ijk} = \mu + a_i + pe_i + l_j + e_{ijk}$$

$$\sigma_a^2 \quad \sigma_{pe}^2 \quad \sigma_l^2 \quad \sigma_e^2$$

Daily gains averaged over n days:

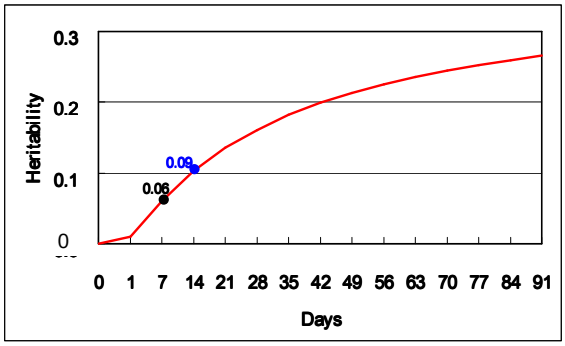
$$\frac{\sum y_{ijk}}{n} = \mu + a_i + pe_i + l_j + \frac{\sum e_{ijk}}{n}$$

$$\sigma_a^2 \quad \sigma_{pe}^2 \quad \sigma_l^2 \quad \frac{\sigma_e^2}{n}$$

Parameters of short-term gain

DG	V_{animal}	V_{litter}	V_{pe}	V_e	h^2
Daily-Expected	3,100	1,300	5,000	330,000	0.009
Daily-Estimated	328 (365)	0.1 (0.4)	0.1 (0.2)	773,000 (4,639)	0.0004 (0.0005)

Expected heritabilities of Daily gain averaged over n days



Conclusions

- Robust regression is an efficient method for 'cleaning' serial BW data
- BW data from automatic feeder stations (Weekly, Bi-weekly) are useful for genetic analyses
- Random regression analyses of ADG may be successful for longer intervals