Boosted regression trees to predict pneumonia, growth and meat percentage of slaughter pigs

August 29th, 2018 – EAAP 2018, Dubrovnik

Erwin Mollenhorst, Karel de Greef, Bart Ducro, Ina Hulsegge, Rita Hoving, Roel Veerkamp, Claudia Kamphuis

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

Pig production system

- Sow producing a litter of 10 - 20 piglets (1 - 2 kg)
- Weaning at 3.5 - 4 weeks (7 - 9 kg)
- Start finishing phase at ± 2 months (20 - 25 kg)
- Slaughtered at ± 5.5 months (110 - 130 kg)
    - Carcass weight 80 - 100 kg

Introduction

Variability in slaughter pigs

Session 29: Variability in the pig production chain

Difference in growth -> multiple deliveries

Abnormalities, e.g. pneumonia

Deviating meat percentage

Data available

Historic slaughter data

Production cycle data

Idea: adapt management based on early prediction

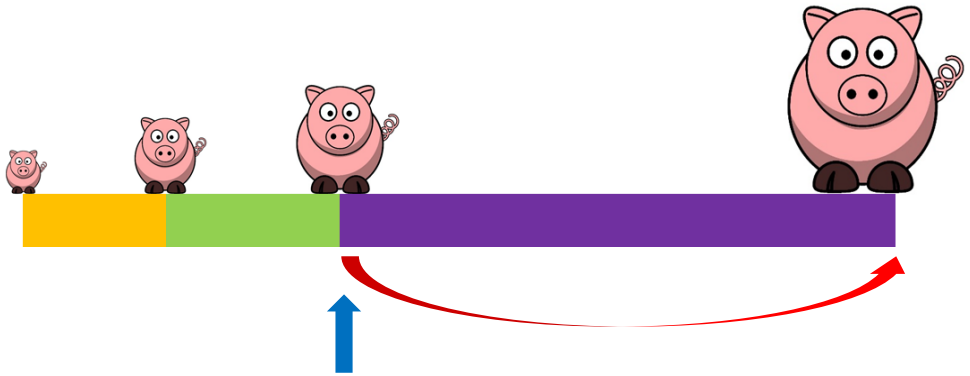
-> decreased costs or higher revenues

Goal

To predict deviant slaughter pigs

based on routine data available at

the onset of the growing-finishing phase



Dataset from VIC Sterksel

65,208 records of individual pigs

Born between 2004 – 2016

Information on:

- Offspring, litter
- Locations, transfer dates, weights
- Slaughterhouse data

Predicted variables

Binary traits (0/1) on individual pigs

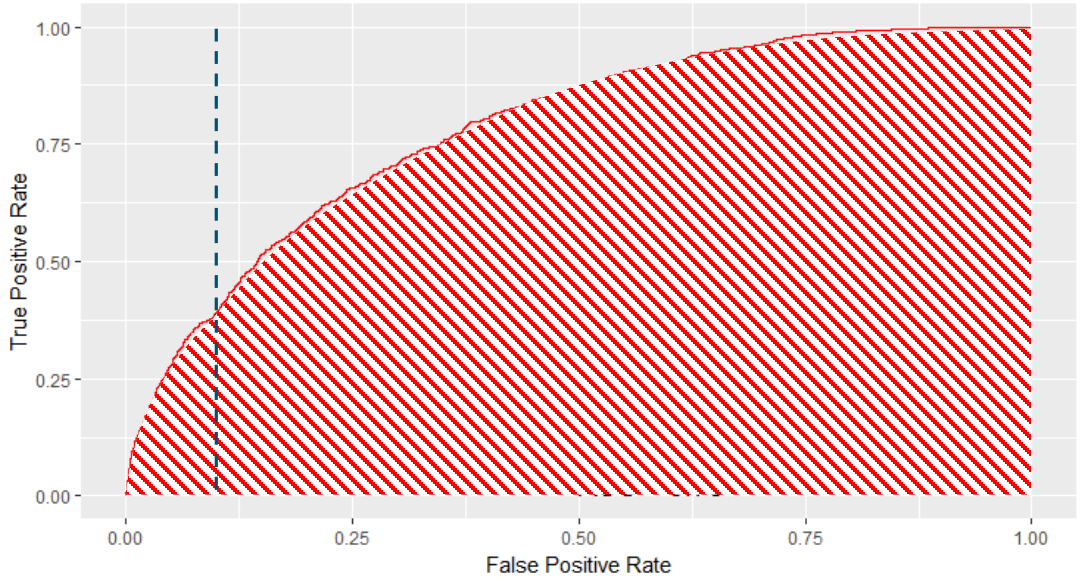
- Pneumonia (no/yes)
- 10% lowest growth rate (normal/low)
- 10% lowest meat percentage (normal/low)

Generalized boosted regression models

- gbm package in R

Performance criteria

ROC curve



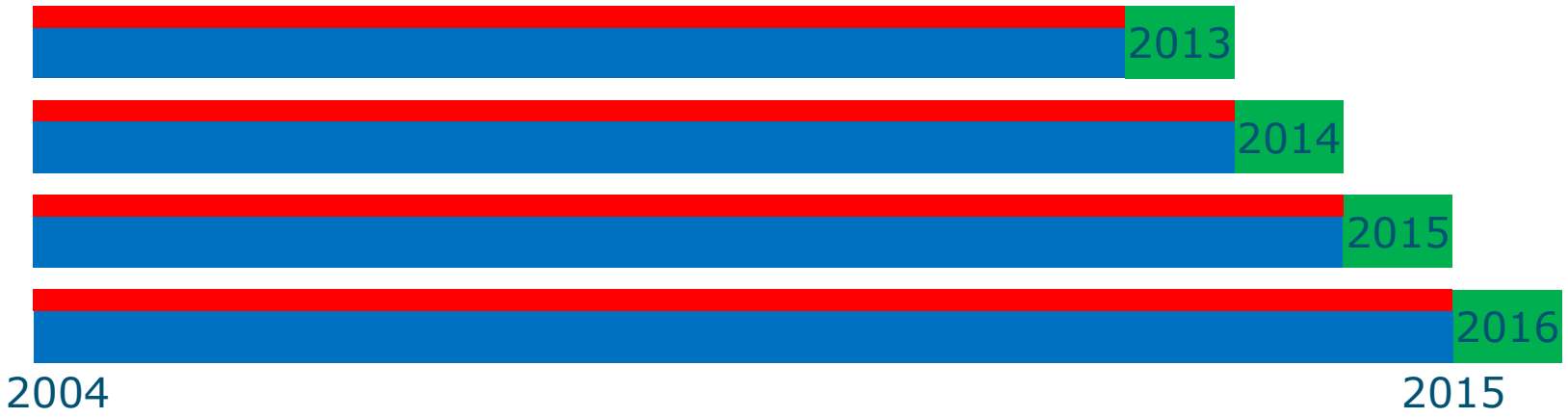
		True class		
		Pos	Neg	
Pred. class	Pos	A	C	A+C
	Neg	B	D	B+D
		A+B	C+D	Total
		SN	SP	
		A/(A+B)	D/(C+D)	

Area under ROC curve (AUC)

Sensitivity at 90% specificity (SN_{SP90})

Validation

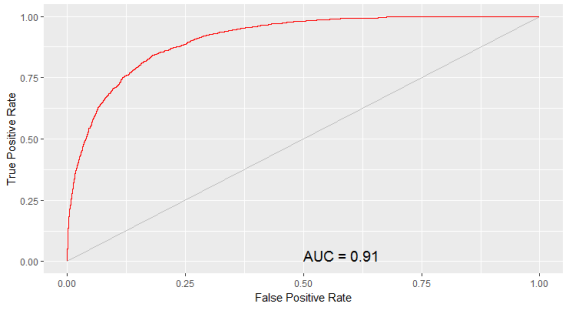
70% train, 30% test, 1 year validation



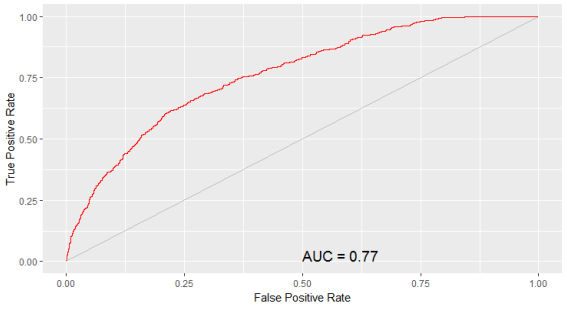
Final performance: weighted average of 4 years

Pneumonia 2013 - AUC

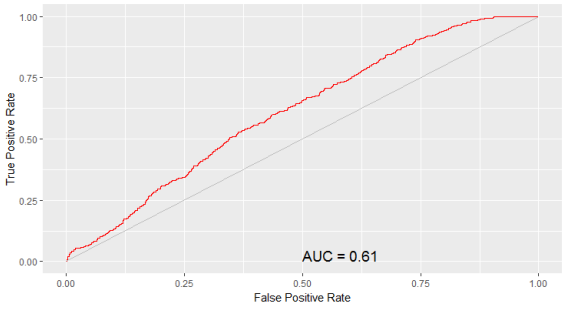
Train



Test



Validation

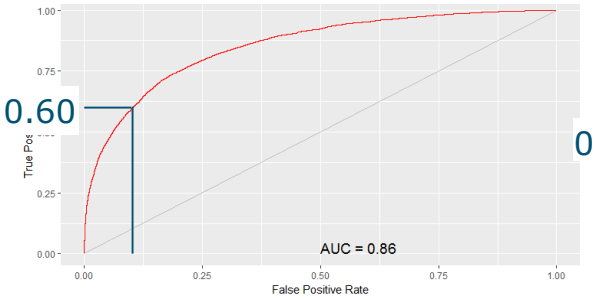


5 most important variables (relative influence)

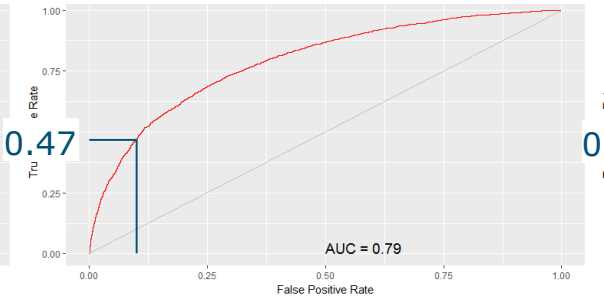
Piglet section (19%)	Location
Growing-finishing section (17%)	
Moving average pneumonia incidence of pen (13%)	Pen/trait history
Birth month (9%)	Season
Birth section (7%)	Location

Growth 2013 – AUC - sensitivity

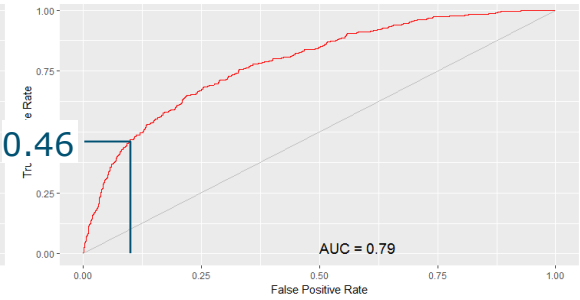
Train



Test



Validation



5 most important variables (relative influence)

- Growing-finishing section (17%) — Location
- Piglet section (12%) — Location
- Weight at start growing-finishing (10%) — Growth
- Age at start growing-finishing (10%) — Growth
- Moving average slaughter weight of pen (7%) — Pen/trait history

Average results 4 years

	Pneumonia			Growth			Meat perc.		
	Train	Test	Valid.	Train	Test	Valid.	Train	Test	Valid.
AUC	0.90	0.77	0.61	0.86	0.80	0.77	0.82	0.70	0.53
SN_{SP90}	0.67	0.39	0.15	0.61	0.47	0.44	0.51	0.32	0.11

Performance train > test > validation

Differences smallest for growth (least overfit)

AUC_{valid} for growth 0.77, fairly good

Sensitivity at 90% specificity, however, below 50%

Low prevalence

Practical implications for growth

		True class			
		Pos	Neg		
Pred.	Pos				
class	Neg	56	810	866	
		100	900	1000	PPV = 100 / 1000 = 10%

Three times increase in positive predicted value

Useful when no costs (time/money) are involved,

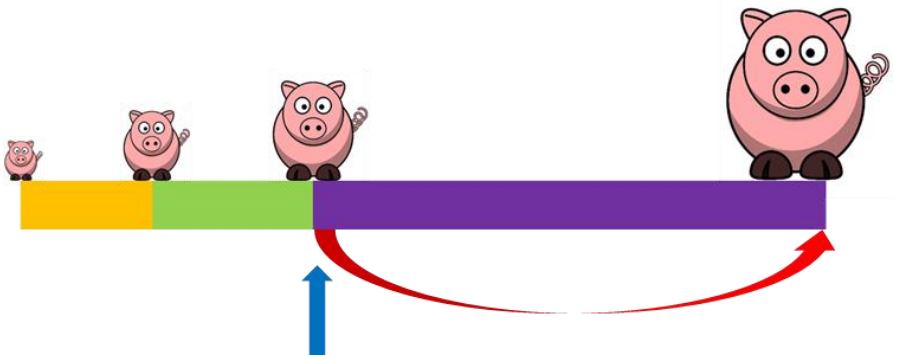
e.g. at start fattening phase

Conclusions

No reasonable prediction pneumonia / low meat percentage

Better identifying pigs with low growth rate

First step towards early warning system



Acknowledgements

Swine Innovation Centre (VIC) Sterksel

Gisabeth Binnendijk

This project was part of

KB-27-01-0013