## HOUSING AND MANAGEMENT

# Boosted trees to predict pneumonia, growth, and meat percentage of growing-finishing pigs[1]

Herman Mollenhorst[2], Bart J. Ducro, Karel H. de Greef, Ina Hulsegge, and Claudia Kamphuis

Wageningen University and Research, Animal Breeding and Genomics, PO Box 338, 6700 AH Wageningen, the Netherlands

[2]Corresponding author: erwin.mollenhorst@wur.nl

ORCiD number: 0000-0002-3633-2176 (Herman Mollenhorst).

## Abstract

In pig production, efficiency is benefiting from uniform growth in pens resulting in single deliveries from a pen of possibly all animals in the targeted weight range. Abnormalities, like pneumonia or aberrant growth, reduce production efficiency as it reduces the uniformity and might cause multiple deliveries per batch and pigs delivered with a low meat yield or outside the targeted weight range. Early identification of pigs prone to develop these abnormalities, for example, at the onset of the growing-finishing phase, would help to prevent heterogeneous pens through management interventions. Data about previous production cycles at the farm combined with data from the piglet's own history may help in identifying these abnormalities. The aim of this study, therefore, was to predict at the onset of the growing-finishing phase, that is, at 3 mo in advance, deviant pigs at slaughter with a machine-learning technique called boosted trees. The dataset used was extracted from the farm management system of a research center. It contained over 70,000 records of individual pigs born between 2004 and 2016, including information on, for example, offspring, litter size, transfer dates between production stages, their respective locations within the barns, and individual live-weights at several production stages. Results obtained on an independent test set showed that at a 90% specificity rate, the sensitivity was 16% for low meat percentage, 20% for pneumonia and 36% for low lifetime growth rate. For low lifetime growth rate, this meant an almost three times increase in positive predictive value compared to the current situation. From these results, it was concluded that routine performance information available at the onset of the growing-finishing phase combined with data about previous production cycles formed a moderate base to identify pigs prone to develop pneumonia (AUC > 0.60) and a good base to identify pigs prone to develop growth aberrations (AUC > 0.70) during the growing-finishing phase. The mentioned information, however, was not a sufficient base to identify pigs prone to develop low meat percentage (AUC < 0.60). The shown ability to identify growth aberrations and pneumonia can be considered a good first step towards the development of an early warning system for pigs in the growing-finishing phase.

**Key words:** boosted trees, growth, machine learning, pig production, pneumonia

## Introduction

Animal production is more and more confronted with new challenges. On the one hand, there is the increasing demand for animal protein in the future due to population growth and increase of income. On the other hand, to meet the increasing demand for animal protein, production systems should not compromise health and welfare of livestock and should have a minimum impact on environment and land use. According to FAO (2011), these challenges will only be achieved by increasing efficiency of production. Increase of production efficiency is mainly focused on improvement of animal growth preferably on less feed through breeding and optimizing management (Brameld and Parr, 2016). Delivery of finisher pigs of about 120 kg live weight to the slaughterhouse takes place at about 5.5 mo of age. In pig production, efficiency is benefiting from uniform growth in pens resulting in single deliveries from a pen of possibly all animals in the targeted weight range (Patience et al., 2004; Alfonso et al., 2010). Abnormalities, like pneumonia, represent a considerable problem for the swine industry primarily due to the reduction of daily weight gain (Merialdi et al., 2012). This aberrant growth reduces uniformity and might cause multiple deliveries per batch and pigs delivered with a low meat yield or outside the targeted weight range.

Identifying signs of emerging production deviations at an early stage, for example, at the onset of the growing-finishing phase, when piglets are about 2 mo of age, would help to prevent heterogeneous pens through management intervention. Prediction of future performance, required to identify early signs of deviations, are traditionally based on early body weight recordings. The prediction models are mainly based on nonlinear regression models (i.e., growth models), which appeared to be good descriptors of the growth, but their predictive power is often limited (Leen et al., 2017). Moreover, extending prediction models with additional factors, like breed and sex as well as environmental factors that affect future performance (e.g., Green and Whittemore, 2005) is rather complex and requires advanced mathematical modeling.

Early signs of production deviations might be retrieved from historic data about previous production cycles at the farm or from previous deliveries. These data, however, are often incomplete, especially at individual level. Machine-learning techniques are able to deal with incomplete data, irrelevant input variables and are less vulnerable for assumptions concerning, for example, (co)linearity and distributions than classical regression techniques (Breiman, 2001; Friedman, 2001). Furthermore, machine-learning techniques proofed to be competitive in various studies in the animals sciences domain in which future performance was predicted using regression or machine-learning techniques (e.g., Roush et al., 2006; Felipe et al., 2015; Alsahaf et al., 2018; Alves et al., 2019). To predict future performance based on the integration of animal and environmental information, sometimes being incomplete and noisy, machine-learning techniques appear to be a valuable and suitable technique. One of the basic and probably most studied machine-learning techniques is decision tree induction (Witten and Frank, 2005). The main disadvantage of decisions trees, however, is their inaccuracy. To alleviate this problem, ensemble methods have been developed that can combine multiple models. Examples of ensemble methods include bagging, boosting, and stacking, of which boosting is considered the most powerful (Witten and Frank, 2005). Boosting is an iterative method. At each iteration, it puts more emphasis on the instances predicted wrongly in previous iterations, when classification is the aim (Witten and Frank, 2005). When evaluated together with other machine-learning techniques, boosted trees are often among the best performing machine-learning methods in different fields (e.g., Ahmad et al., 2019; Knoll et al., 2019; Song et al., 2019).

To demonstrate whether it is possible to predict growth of pigs based on animal and environmental information, the aim of this study was to predict deviant slaughter pigs based on routine data available at the onset of the growing-finishing phase with a machine-learning technique called boosted trees.

## Materials and Methods

Routine data used in this study were acquired from the farm management program of a research farm. Since no animal experiments were performed for this study, approval by an Animal Care and Use Committee was not necessary.
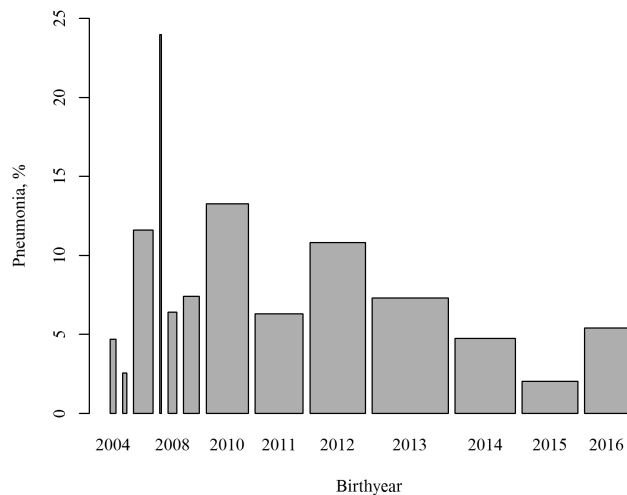
### Data Sets

The Swine Innovation Center "VIC Sterksel" is a research center of Wageningen University and Research, located in Sterksel, the Netherlands. At VIC Sterksel, detailed information of all pigs is recorded in the farm management system, for example, on offspring, litter size, transfer dates between production stages, their respective locations within the barns, and individual live-weights at several production stages. The dataset used for this study was extracted from this farm management system and contained over 70,000 records of individual pigs born between 2004 and 2016. Since the aim was to predict which pigs in a batch, defined as group of piglets starting on the same date in the growing-finishing phase, would grow least or would have the lowest meat percentage, we decided to focus on batches at the start of the growing-finishing phase of at least 100 piglets. Furthermore, only pigs from 2 regular growing-finishing stables that resemble commercial husbandry, were included in the study. This resulted in 325 batches containing 61,041 pigs in total. Based on slaughter data, 3 binary traits at the pig level were defined: 1) pneumonia status; 2) belonging to the 10% animals of a batch with the lowest lifetime growth rate; and 3) belonging to the 10% animals of a batch with the lowest meat percentage. Pneumonia status and meat percentage were based on regular slaughterhouse recordings. Lifetime growth rate was calculated as carcass weight (in kg) divided by age at slaughter (in days).
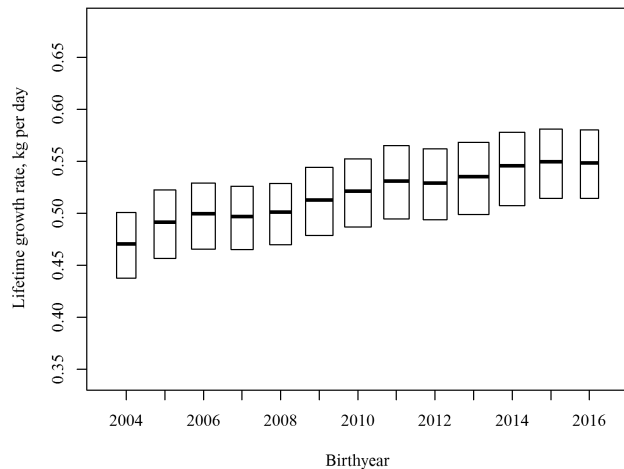
The pneumonia prevalence was quite variable over the whole period (Fig. 1). As pneumonia status was a binary variable on animal level, by definition no variation within a year is present. For some years before 2010, the number of nonmissing values was low, which could be caused by incomplete recordings of pneumonia status or only recording positive cases. This could also explain the outlier in 2007. Lifetime growth rate increased over time, while variation was rather constant (Fig. 2). Meat percentage was rather constant over time (Fig. 3).

### Data Preprocessing

Each individual pig's record included litter information, like the number of live, dead, and mummified piglets born and number of male and female piglets born alive. The latter 2 variables were combined to one stating the percentage of males in the litter. The quality of the litter in which a pig was born was described by the mean, standard deviation, and median of the birthweights per litter. The piglets "position" in the litter was described by the quartile of birthweight it belongs to within its litter and by the deviation of the piglet's individual birthweight from the median of the litter.
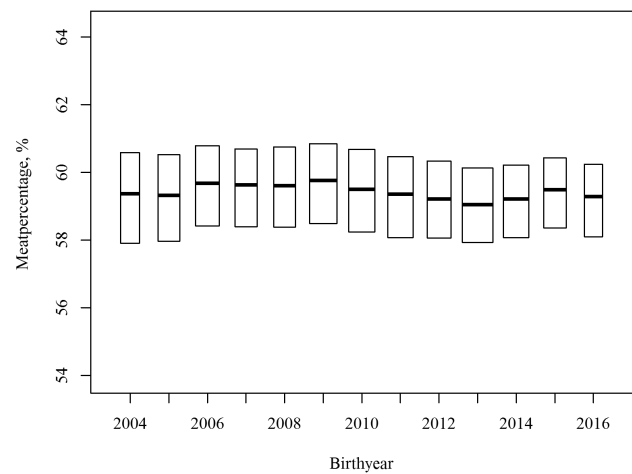
**Figure 1.** Mean pneumonia prevalence per year. Bar width represents the number of nonmissing values for pneumonia status. Total number of records is 35,125.



**Figure 2.** Median and interquartile range of lifetime growth rate (kg/d) per year. Box width represents the number of nonmissing values for lifetime growth rate. Total number of records is 61,041.

Organ and carcass deviations were routinely scored at the slaughter line (Elbers et al., 1992; European Community, 2004). Organ deviations considered pneumonia and affected livers (light and medium were merged together for the analysis) and carcass deviations considered pleuritis and skin and leg inflammations. Next to these binary variables per slaughtered pig, moving averages were calculated at pen level over the last 5 batches that were raised in that pen.

Environmental factors such as pen conditions might affect growth. This is a high dimensional factor of which the effect is believed to be concentrated in an underlying (variance-covariance) structure of lower dimension. High dimensionality can result in less stable predictions and to avoid overfitting dimension reduction is often required (Darnell et al., 2017). For dimension reduction, we applied a linear mixed model (Proc Mixed, SAS V9.3) to estimate the underlying variance-covariance structure of pen-batch effects on lifetime growth performance. The obtained Best Linear Unbiased Predictor (**BLUP**) of the previous batch in a pen was included as prediction variables for the current batch in that pen. Additionally, the moving average



**Figure 3.** Median and interquartile range of meat percentage per year. Box width represents the number of nonmissing values for meat percentage. Total number of records is 60,889.

of BLUP-estimates of the last 2 batches in a pen were included. An overview of all prediction variables is given in Tables 1 and 2.

## Model Development

In this study, we wanted to demonstrate the opportunities of boosted trees, as an example of a machine-learning technique. There are different boosted trees algorithms available, and the Gradient Boosting Machine (**GBM**) offered by the h2o.gbm R package (h2o version 3.22.1.1) is one of them. The GBM was used to predict traits at the onset of the growing-finishing phase. The Gradient Boosting Machine is extensively described by Hastie et al. (2009). Boosting is a forward learning iterative method. At each iteration, it puts more emphasis on the instances predicted wrongly in previous iterations, when classification is the aim (Witten and Frank, 2005). Some default model parameters were adapted, the number of trees (ntrees) was set at 1,000, the maximum number of splits per tree (max_depth) was set at 3, and the learning rate (learn_rate) was set at 0.01, in order to speed up the analysis. The large number of models used to calculate the average performance could compensate for the reduction in number of trees and interaction depth per model. All analyses were performed in RStudio (version 1.1.423 running R version 3.5.0).

## Model Testing

In order to test the model on independent data, for each batch ($n$ = 98) in the years 2013 to 2016, a new model was trained on a training dataset and tested on the batch under consideration (Test$_{New}$). The training dataset was each time a 70% random sample at batch level from all batches from the years 2004 to 2012 ($n$ = 227 batches for growth and meat percentage and $n$ = 128 batches for pneumonia) enlarged with one batch at every repetition ($n$ = 98) of model development and testing. The remaining 30% of the training dataset was used as test set to obtain the performance of the model on data from the same time period as the training set (Test$_{Train}$). The weighted average performance over 98 repetitions was considered as final performance.

## Model Performance Criteria

Sensitivity or true positive rate is the fraction of real positive cases that is predicted to be positive. In our study, the true

**Table 1.** Mean, first and third quartile and percentage of missing values (of 61,041 records) for numerical prediction variables used as input for training boosted trees to predict pneumonia, growth, and meat percentage

| Variable name | Mean | Q1 | Q3 | Missing, % |
|---|---|---|---|---|
| Birth weight, kg | 1.45 | 1.22 | 1.65 | 0.4 |
| Number of live born piglets per litter | 14.7 | 13 | 17 | 15.4 |
| Percentage males in litter | 0.52 | 0.43 | 0.60 | 15.4 |
| Number of dead born piglets in litter | 0.91 | 0 | 1 | 15.4 |
| Number of mummified piglets in litter | 0.29 | 0 | 0 | 15.4 |
| Median birth weight litter | 1.43 | 1.26 | 1.57 | 0.1 |
| Standard deviation birth weight litter | 0.27 | 0.21 | 0.32 | 0.1 |
| Deviation from median birth weight of litter | 0.02 | −0.12 | 0.17 | 0.4 |
| Weight at weaning, kg | 7.87 | 6.80 | 9.00 | 1.9 |
| Age at weaning, days | 26.6 | 25 | 28 | 21.8 |
| Number of weaned piglets in litter | 12.1 | 11 | 13 | 15.4 |
| Number of deaths in litter till weaning | 1.84 | 0 | 3 | 15.4 |
| Growth rate till start growing-finishing phase, kg/d | 0.37 | 0.33 | 0.41 | 11.8 |
| Number of growing-finishing pigs in pen[1] | 17.0 | 11 | 12 | 0.0 |
| Moving average slaughter weight, kg | 91.0 | 89.4 | 92.6 | 12.5 |
| Moving average pneumonia | 0.05 | 0.00 | 0.07 | 12.5 |
| Moving average affected liver | 0.01 | 0.00 | 0.02 | 12.5 |
| Moving average pleuritis | 0.18 | 0.11 | 0.24 | 12.5 |
| Moving average skin inflammations | 0.01 | 0.00 | 0.02 | 12.5 |
| Moving average leg inflammations | 0.02 | 0.00 | 0.03 | 12.5 |
| BLUP estimator growth per day previous batch | 0.03 | −1.92 | 1.98 | 2.9 |
| Moving average of BLUP estimator growth per day 2 previous batches | 0.04 | −1.50 | 1.56 | 6.4 |

[1]High mean due to some very large groups, probably due to classifying whole section as one pen (All stables were subdivided in sections and within sections in pens).

**Table 2.** Frequencies in five most frequent categories per variable and percentage of missing values (of 61,041 records) for categorical and ordered prediction variables used as input for training boosted trees to predict pneumonia, growth, and meat percentage

| Variable name | 1 | 2 | 3 | 4 | 5 | >5/other | Missing, % |
|---|---|---|---|---|---|---|---|
| Litter number of mother | 10,627 | 10,967 | 9,982 | 8,651 | 7,182 | 13,632 | 0.0 |
| Piglet belongs to which quartile of birthweight within litter | 13,556 | 15,039 | 17,294 | 14,759 | | | 0.4 |
| Number of (foster) sows till weaning | 48,898 | 10,796 | 1,156 | 118 | 7 | | 0.1 |
| Sex[1] | 29,612 | 29,638 | | | | | 2.9 |
| Boar line[2] | 50,226 | 881 | 693 | 674 | 16 | | 14.0 |
| Sow line[3] | 32,486 | 16,461 | 4,080 | 1,328 | | | 11.0 |
| Nursing stable[4] | 31,216 | 28,372 | 143 | 26 | | | 2.1 |
| Weaners stable[4] | 25,495 | 11,124 | 6,192 | 1,482 | 985 | 2,430 | 21.8 |
| Growing-finishing stable[4] | 59,607 | 1,434 | | | | | 0.0 |

[1]Sex: 1 = female, 2 = male.
[2]Boar line: 1 = synthetic, 2 = large white, 3 = Duroc, 4 = landrace, 5 = Pietrain.
[3]Sow line: 1 = landrace × large white; 2 = large white × landrace, 3 = large white, 4 = landrace.
[4]All stables were subdivided in sections and within sections in pens. Section information was also used as input for the models.

positives were pigs with pneumonia or pigs belonging to the group with 10% lowest lifetime growth rate or meat percentage. Specificity or true negative rate is the fraction of real negative cases that is predicted to be negative. Sensitivity and specificity could be calculated at all levels of probability thresholds of being positive as produced by the GBM model. The trade-off between sensitivity and specificity could be shown in a receiver operating characteristic (**ROC**) curve, in which sensitivity is plotted against the false-positive rate (1—specificity) (Metz, 1978). The overall performance of a classifier can be characterized by the area under the ROC curve (AUC) (e.g., Hanley and McNeil, 1982; Detilleux et al., 1999), and these AUC values were reported. Moreover, as we were not interested in the entire ROC curve, but in selecting a rather small part of approximately 10% of the animals per batch, we also evaluated models on their sensitivity at a fixed specificity of 90%. Next to the average performance over 98 repetitions, we also stored all predicted probabilities of

the observations in the 98 Test$_{New}$ datasets and used these to plot an aggregated ROC curve for each output variable.

### Variable Importance

Next to the model performance, the GBM model also reproduces information on the relative influence of each variable in the prediction model, based on the reduction of the squared error in each node. This variable importance was expressed as the percentage contribution of each variable in the prediction of the outcome variable.
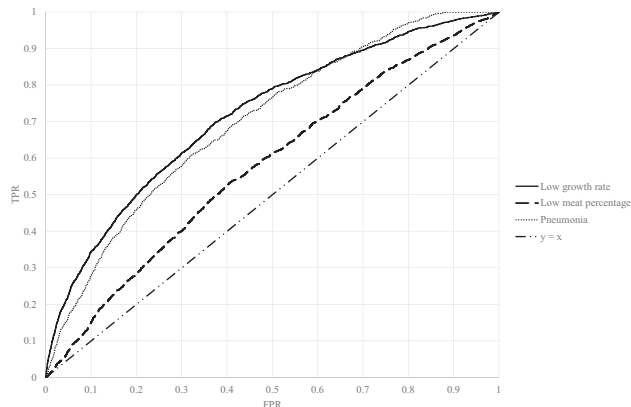
## Results

### Model Performance

Table 3 shows the average performance from 98 model repetitions of the prediction models for pneumonia, low meat

**Table 3.** Performance characteristics, area under the receiver operating characteristic curve (AUC), and sensitivity at 90% specificity, for pneumonia, low meat percentage, and low lifetime growth rate prediction as averaged over 98 model repetitions (including SD)[1]

| | AUC (SD) | | | Sensitivity at 90% specificity (SD) | | |
|---|---|---|---|---|---|---|
| Predicted variable | Train | Test$_{Train}$ | Test$_{New}$ | Train | Test$_{Train}$ | Test$_{New}$ |
| Pneumonia | 0.83 (0.01) | 0.73 (0.03) | 0.64 (0.17) | 50 (2) | 30 (5) | 20 (27) |
| Low meat % | 0.71 (0.01) | 0.63 (0.01) | 0.58 (0.09) | 31 (1) | 19 (1) | 16 (12) |
| Low lifetime growth rate | 0.74 (0.00) | 0.69 (0.01) | 0.73 (0.09) | 38 (1) | 31 (1) | 36 (17) |

[1]In each model repetition, the Test$_{New}$ set was an entirely independent dataset containing the next batch, while the Train and Test$_{Train}$ set were a 70/30% random sample at batch level of all previous batches.



**Figure 4.** Aggregated receiver operating characteristic curves, based on all predicted probabilities of the observations in the 98 Test$_{New}$ datasets, for pneumonia [dotted line, area under the curve (AUC) = 0.70, sensitivity at 90% specificity = 28%], low lifetime growth rate (solid line, AUC = 0.72, sensitivity at 90% specificity = 34%), and low meat percentage (dashed line, AUC = 0.58, sensitivity at 90% specificity = 15%).

percentage and low lifetime growth rate. Testing on independent data showed that AUC were poor (<0.60) to fairly good (>0.70) for prediction of low meat percentage (0.58), pneumonia (0.64), and low lifetime growth rate (0.73). At a 90% specificity rate, sensitivity was 16, 20, and 36% for low meat percentage, pneumonia, and low lifetime growth rate, respectively. The models for pneumonia and low meat percentage somewhat overfitted on the train set, whereas differences in performance on Test$_{Train}$ and Test$_{New}$ were rather low. For the best performing model, the one for predicting low lifetime growth rate was least overfit and performed equally well on Test$_{Train}$ and Test$_{New}$. Variation in performance, however, was much higher between Test$_{New}$ batches (Table 3). This means that, at least on average, the model was able to predict lifetime growth rate of pigs in future batches (Test$_{New}$) as well as it was able to predict lifetime growth rate of pigs in random batches in time (Test$_{Train}$). The performance of the pneumonia model was underestimated, because no performance could be calculated when the real incidence of pneumonia in a batch was zero. These batches, however, could be expected to be more accurately predicted than batches with high incidences of pneumonia. Leaving out performance of batches with high performance lowers the reported average performance in Table 3. This could also explain the larger drop in performance between Test$_{Train}$ and Test$_{New}$ for pneumonia, as compared to low meat percentage and low lifetime growth rate.

The aggregated ROC curves based on all predicted probabilities of the observations in the 98 Test$_{New}$ datasets are shown in Fig. 4. The performance metrics for low lifetime growth rate and low meat percentage agree with those in Table 3, whereas performance

for pneumonia is better in Fig. 4, which supports that Test$_{New}$ performance for pneumonia in Table 3 is underestimated. The AUC of low lifetime growth rate and pneumonia are almost equal in the aggregated ROC curve, while the curves cross each other. At a 90% specificity rate, however, sensitivity is still considerably higher for low lifetime growth rate than for pneumonia.

### Variable Importance

The 10 most important variables per prediction model are shown in Table 4. For predicting low growth rate over the whole lifetime, growth rate till the start of the growing-finishing phase, the moment of prediction, appeared to be the most important variable and accounted for 50% of the reduction of the squared error. Two other weight-related variables, birth weight and weight history of the pen, were ranked second and third. The latter one, however, is more related to the pen than to the individual pig. Furthermore, other variables in the top 10 represented information about locations (like nursing section and the ones related to the BLUP-estimators of pen effects), genetics (boar line), and birth year, which showed that combining a large variety of data was useful.

For predicting pneumonia, the history of the pen was rather important, as shown by the most important variable expressing the moving average of previous batches, but also variables like weaners stable, nursing section, and variables related to lifetime growth rate and pleuritis of previous batches. Furthermore, birth year and month were rather important.

For predicting low meat percentage, which appeared to be the most difficult one, many variables contributed interchangeably, as shown by the rather low mean percentages (highest one 18%). The most important variables related to sex, genetics (boar line), pen history, and many different weight-related variables.

### Discussion

The aim of the study was to identify pigs that were prone to develop aberrant growth rate, meat percentage or to develop pneumonia during the growing-finishing phase. For that purpose, predictions were made based on performance data of the pigs until start of the growing-finishing phase. The prediction was based only on existing routine data from an experimental farm and no additional data have been collected for this study. The reason for that was 2-fold. A first reason was that we wanted to demonstrate the value of existing information in prediction. Nowadays, the focus in prediction is merely on collecting data using new (sensor) techniques (e.g., Ferrari et al., 2008; Maselyne et al., 2018; Pezzuolo et al., 2018) and the step to integrate with existing information is often neglected (Rutten et al., 2013). According to Cornou and Kristensen (2013), decision making is based on a combination of observations of the animals and their environment, as well as production results. The added value

**Table 4.** Top 10 important variables for predicting pigs with low lifetime growth rate, pneumonia, and low meat percentage expressed in contribution to the reduction of the loss function as averaged (including standard deviations) over batches from 98 model runs

| Variable name | Low lifetime growth rate | Pneumonia | Low meat percentage |
|---|---|---|---|
| Growth rate till start growing-finishing phase | 0.50 (0.03) | | 0.07 (0.01) |
| Birth weight | 0.15 (0.02) | | 0.02 (0.01) |
| Moving average lifetime growth rate of previous batches in pen | 0.04 (0.01) | 0.04 (0.01) | |
| Nursing section | 0.03 (0.01) | 0.04 (0.01) | 0.06 (0.01) |
| Deviation from median birth weight of litter | 0.03 (0.01) | 0.02 (0.01) | 0.02 (0.01) |
| Weight at weaning | 0.03 (0.01) | 0.02 (0.01) | 0.03 (0.01) |
| Boar line | 0.03 (0.01) | | 0.13 (0.01) |
| Birth year | 0.02 (0.01) | 0.19 (0.06) | 0.07 (0.02) |
| Moving average of BLUP estimator of lifetime growth rate | 0.02 (0.01) | | |
| BLUP estimator of lifetime growth rate previous batch in pen | 0.02 (0.01) | 0.02 (0.01) | |
| Moving average pneumonia of previous batches in pen | | 0.36 (0.03) | |
| Birth month | | 0.07 (0.03) | |
| Weaners stable | | 0.05 (0.03) | |
| Moving average pleuritis of previous batches in pen | | 0.03 (0.01) | |
| Sex | | | 0.18 (0.04) |
| Moving average meat percentage of previous batches in pen | | | 0.13 (0.02) |
| Median birth weight litter | | | 0.03 (0.01) |

of new sensor technology or monitoring strategies should be considered in combination with information already available in the daily registrations in the farm management system.

A second reason concerned the time window of prediction. Most often the time window considered in predicting growth is as short as a week (Yu et al., 2006) or 1 d (Roush et al., 2006). From the information available at the onset of the growing-finishing phase, we wanted to predict the outcome of target traits at moment of slaughter, which is about 3 mo later. This was more challenging and required larger training sets (Alsahaf et al., 2018). Permanent environmental effects, that is, farm-specific effects, were considered to be important in the prediction and could be retrieved from historical data of the farm. Therefore, all data stored in the farm management system of VIC Sterksel about previous batches were offered to the machine-learning procedure.

Machine learning is applied often nowadays, and has shown to be competitive with logistic regression in previously conducted studies within the animal domain (e.g., Roush et al., 2006; Felipe et al., 2015; Alsahaf et al., 2018; Alves et al., 2019). To confirm these results with our own data, we also applied logistic regression (using h20.glm) to our data. To enable logistic regression, we first imputed missing values (using mice R-package, version 3.5.0) to create datasets with the same number of records as in the main analyses. Results showed that GBM equaled or slightly outperformed logistic regression for all 3 predictions. The largest difference on Test$_{New}$ was for the prediction of pneumonia with an average sensitivity of 21% for GBM and 17% for logistic regression at 90% specificity rate and, respectively, an AUC of 0.65 and 0.60. Second, we excluded records with missing values in any of the prediction variables. This resulted in smaller datasets with up to 44% less records. Again, GBM performed equally or slightly better than logistic regression, although differences were smaller than on the imputed datasets. These results confirmed that, also in our study, GBM is at least competitive with logistic regression. Lastly, we compared GBM results on imputed datasets or datasets without missing values with GBM results on the full datasets (from the main analyses in this study). It appeared that the additional effort of imputing or reducing the dataset did result in consistently better prediction performance, while up to 44% less records received a predicted value when records

with missing values were excluded. So, in our study, GBM was competitive with logistic regression, without the necessity of data handling that is required for logistic regression.

One of the major complicating factors in predicting production performance in pigs is the uncertainty of age at slaughter. Production performance, like slaughter weight and meat percentage, are largely affected by age at slaughter. Age at slaughter is primarily a decision made by the farm manager, who is considering, among others, market prices, contract obligations, and additional management aspects. Management decisions are taken somewhere during the growing-finishing phase and are unknown at moment of prediction. To overcome the influence of decision making on our outcome variable, we decided to predict growth during the growing-finishing phase. Growth rate is less affected by management decisions and was, therefore, expected to be more predictable.

Environmental conditions at pen, section, or stable level might affect performance and is, therefore, usually included in analysis of field experiments (e.g., Rehfeldt et al., 2008). To account for systematic environmental effects of pen in prediction of the current batch, we have chosen to calculate the average performance of the 5 batches preceding the current batch in the pen and offered that to the machine-learning procedure. In this way, the systematic effect of location could be accounted for in the prediction. From the variable importance metric, it was seen that pneumonia history of a pen was the highest contributor to pneumonia prediction, and meat percentage history of a pen was the second highest contributing variable (together with boar line) to prediction of low meat percentage. Contribution of pen history to prediction of pneumonia might point to systematic less optimal conditions in certain parts of the stable, but the contribution of pen history to meat percentage is less clear, although temperature is known to have some effect on meat percentage (Arkfeld et al., 2017).

Growth up to the onset of the growing-finishing phase was largely dominating the prediction of low growth as indicated by a variable importance of 0.50. This high importance was in accordance with the strong relationship of early growth to later growth as has been established in various studies (e.g., Quiniou et al., 2002). Additionally, birth weight was a good predictor as well and this corresponded with a reasonable relation of birth weight with growth later in life (Rehfeldt et al., 2008).

Meat percentage is a trait that shows in general relatively little variation and is, therefore, difficult to predict. In this study, the coefficient of variation was 3.1%, whereas a value of 2% was reported in literature (Shirali et al., 2017). According to Shirali et al. (2017), meat percentage is mainly influenced by genetics, sex, and age at slaughter. Sex indeed had largest variable importance followed by boar line, that is, genetics, which corresponded to the results of Calderon Diaz et al. (2017). Age at slaughter is highly subject to management decisions, and the dataset did not contain variables that hold information from which moment of delivery could be learnt. Variation in age at slaughter can cause a 1 to 4% difference in meat percentage (Weatherup et al., 2010). In planning batches for delivery, the farmer selects pigs based on live weight by visual inspection. Because of the weak relation between live weight and meat percentage, this way of selection does not guarantee that the selected pigs also have optimal meat percentage.

Quality of predictions is often assessed using ROC curves, being a method that is helpful to visualize the performance of classifiers and describe the trade-off between sensitivity and specificity. These curves are particularly useful in areas of cost-sensitive learning and learning in the presence of unbalanced cases (Fawcett, 2006). The AUC is often considered as it reflects the expected performance of a classifier irrespective of the chosen threshold, and it indicates the probability that an aberrant animal can be distinguished from a well performing animal. A larger AUC indicates a better average performance, although it might be that another classifier might perform better at specific combinations of sensitivity and specificity.

Both pneumonia and low growth had an AUC of about 0.70, but sensitivity of pneumonia was higher only at specificity levels of 0.30 and lower. At specificity levels above 0.40, the sensitivity of lifetime growth rate was higher.

From the results of AUC for the 3 target traits, it became clear that the AUC of low meat percentage was too low to be considered a sufficient prediction and the outcome was hardly any better than random prediction. The recording of performance till the onset of the growing-fattening phase had apparently no predictive power with respect to meat percentage, neither had the pen performance history. For prediction of meat percentage, additional variables should be recorded, for example, related to body composition of the live animal preferably available at moment of prediction or early in the growing-finishing phase.

The AUC results of pneumonia and low lifetime growth rate were better and were comparable to detection of, for example, lame cows (Kamphuis et al., 2013) and predicting of insemination outcomes in dairy cattle using random forest methodology (Shahinfar et al., 2014).

Intervention requires close monitoring and more intense management which cannot be given to the whole production unit. Therefore, an early indication of pigs at risk to develop an aberration would help the farmer to concentrate on a smaller unit only. For practical applications, the sensitivity is therefore often considered at a fixed specificity. For example, at a 90% specificity rate (i.e., only 10% of false positives are allowed), the sensitivity rates were 20% for pneumonia, 16% for low meat percentage, and 36% for low lifetime growth rate (Table 3). According to Kamphuis et al. (2013), this would mean that in a stable with 1,000 finishing pigs, 14 (20%) out of the 72 pneumonia cases (i.e., prevalence is 7.2%) will be detected, whereas 93 false alerts out of 928 healthy pigs can be expected. In other words, 107 piglets would receive an indication prior to the growing-finishing phase, of which 14 indeed will develop pneumonia (when no action is taken). For slow growth with a sensitivity of 36% at 90% specificity, 126 piglets would receive

an indication at start of the growing-finishing phase of which 36 indeed would grow too slow; 2 out of 7 alerts will be correct. This is almost 3 times increase in success compared to the current situation which has a positive predicted value, equivalent to prevalence, of 10%. This result was achieved using only routinely collected data from an experimental farm. Results, thus, can be considered as a first step towards an early warning system for slow growing pigs and development of pneumonia.

## Conclusions

Routine performance information available at the onset of the growing-finishing phase combined with data about previous production cycles formed a moderate base to identify pigs prone to develop pneumonia (AUC > 0.60) and a good base to identify pigs prone to develop growth aberrations (AUC > 0.70) during the growing-finishing phase. The mentioned information, however, was not a sufficient base to identify pigs prone to develop low meat percentage (AUC < 0.60). The shown ability to identify growth aberrations and pneumonia can be considered a good first step towards the development of an early warning system for pigs in the growing-finishing phase.

## Acknowledgments

## Literature Cited

Ahmad, A. K., A. Jafar, and K. Aljoumaa. 2019. Customer churn prediction in telecom using machine learning in big data platform. *J Big Data*. **6**. Art. no: 28. doi:10.1186/s40537-019-0191-6

Alfonso, L., G. Zudaire, M. V. Sarries, J. Viguera, and F. Flamarique. 2010. Investigation of uniformity in pig carcass and meat quality traits. *Animal* **4**:1739–1745. doi:10.1017/S1751731110000819

Alsahaf, A., G. Azzopardi, B. Ducro, E. Hanenberg, R. F. Veerkamp, and N. Petkov. 2018. Prediction of slaughter age in pigs and assessment of the predictive value of phenotypic and genetic information using random forest. *J. Anim. Sci.* **96**:4935–4943. doi:10.1093/jas/sky359

Alves, A. A. C., A. Chaparro Pinzon, R. M. d. Costa, M. S. d. Silva, E. H. M. Vieira, I. B. d. Mendonça, V. d. S. S. Viana, and R. N. B. Lôbo. 2019. Multiple regression and machine learning based methods for carcass traits and saleable meat cuts prediction using non-invasive in vivo measurements in commercial lambs. *Small Rumin. Res.* **171**:49–56. doi:10.1016/j.smallrumres.2018.12.008

Arkfeld, E. K., D. A. Mohrhauser, D. A. King, T. L. Wheeler, A. C. Dilger, S. D. Shackelford, and D. D. Boler. 2017. Characterization of variability in pork carcass composition and primal quality. *J. Anim. Sci.* **95**:697–708. doi:10.2527/jas.2016.1097

Brameld, J. M., and T. Parr. 2016. Improving efficiency in meat production. *Proc. Nutr. Soc.* **75**:242–246. doi:10.1017/S0029665116000161

Breiman, L. 2001. Random Forests. *Mach Learn*. **45**:5–32. doi:10.1023/a:1010933404324

Calderón Díaz, J. A., L. A. Boyle, A. Diana, F. C. Leonard, J. P. Moriarty, M. C. McElroy, S. McGettrick, D. Kelliher, and

E. García Manzanilla. 2017. Early life indicators predict mortality, illness, reduced welfare and carcass characteristics in finisher pigs. *Prev. Vet. Med.* **146**:94–102. doi:10.1016/j.prevetmed.2017.07.018

Cornou, C., and A. R. Kristensen. 2013. Use of information from monitoring and decision support systems in pig production: Collection, applications and expected benefits. *Livest. Sci.* **157**:552–567. doi:10.1016/j.livsci.2013.07.016

Darnell, G., S. Georgiev, S. Mukherjee, and B. E. Engelhardt. 2017. Adaptive randomized dimension reduction on massive data. *J Mach Learn Res.* **18**:1–30.

Detilleux, J., J. Arendt, F. Lomba, and P. Leroy. 1999. Methods for estimating areas under receiver-operating characteristic curves: Illustration with somatic-cell scores in subclinical intramammary infections. *Prev. Vet. Med.* **41**:75–88. doi:10.1016/S0167-5877(99)00054-9

Elbers, A. R. W., M. J. M. Tielen, J. M. A. Snijders, W. A. J. Cromwijk, and W. A. Hunneman. 1992. Epidemiological studies on lesions in finishing pigs in the Netherlands. I. Prevalence, seasonality and interrelationships. *Prev. Vet. Med.* **14**:217–231. doi:10.1016/0167-5877(92)90018-B

European Community. 2004. Regulation (EC) No 854/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific rules for the organization of official controls on products of animal origin intended for human consumption. *Official Journal of the European Union.* **L139**:206–320.

FAO. 2011. *World livestock 2011 - livestock in food security.* FAO, Rome, Italy.

Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recognit Lett.* **27**:861–874. doi:10.1016/j.patrec.2005.10.010

Felipe, V. P., M. A. Silva, B. D. Valente, and G. J. Rosa. 2015. Using multiple regression, Bayesian networks and artificial neural networks for prediction of total egg production in European quails based on earlier expressed phenotypes. *Poult. Sci.* **94**:772–780. doi:10.3382/ps/pev031

Ferrari, S., M. Silva, M. Guarino, J. M. Aerts, and D. Berckmans. 2008. Cough sound analysis to identify respiratory infection in pigs. *Comput Electron Agric.* **64**:318–325. doi:10.1016/j.compag.2008.07.003

Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. *Ann Stat.* **29**:1189–1232.

Green, D. M., and C. T. Whittemore. 2005. Calibration and sensitivity analysis of a model of the growing pig for weight gain and composition. *Agric Syst.* **84**:279–295. doi:10.1016/j.agsy.2004.06.017

Hanley, J. A., and B. J. McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* **143**:29–36. doi:10.1148/radiology.143.1.7063747

Hastie, T., R. Tibshirani, and J. H. Friedman. 2009. *The elements of statistical learning - Data mining, inference, and prediction.* 2nd edn. Springer-Verlag, New York, NY.

Kamphuis, C., E. Frank, J. K. Burke, G. A. Verkerk, and J. G. Jago. 2013. Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological characteristics of dairy cows to detect lameness. *J. Dairy Sci.* **96**:7043–7053. doi:10.3168/jds.2013-6993

Knoll, L., L. Breuer, and M. Bach. 2019. Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning. *Sci. Total Environ.* **668**:1317–1327. doi:10.1016/j.scitotenv.2019.03.045

Leen, F., A. Van den Broeke, B. Ampe, L. Lauwers, J. Van Meensel, and S. Millet. 2017. Evaluation of performance models for farm-specific optimization of pig production. *Livest Sci* **201**:99–108. doi:10.1016/j.livsci.2017.05.006

Maselyne, J., A. Van Nuffel, P. Briene, J. Vangeyte, B. De Ketelaere, S. Millet, J. Van den Hof, D. Maes, and W. Saeys. 2018. Online warning systems for individual fattening pigs based on their feeding pattern. *Biosyst Eng* **173**:143–156. doi:10.1016/j.biosystemseng.2017.08.006

Merialdi, G., M. Dottori, P. Bonilauri, A. Luppi, S. Gozio, P. Pozzi, B. Spaggiari, and P. Martelli. 2012. Survey of pleuritis and pulmonary lesions in pigs at abattoir with a focus on the extent of the condition and herd risk factors. *Vet. J.* **193**:234–239. doi:10.1016/j.tvjl.2011.11.009

Metz, C. E. 1978. Basic principles of ROC analysis. *Semin. Nucl. Med.* **8**:283–298. doi:10.1016/S0001-2998(78)80014-2

Patience, J. F., A. D. Beaulieu, and D. A. Gillis. 2004. The impact of ground water high in sulfates on the growth performance, nutrient utilization, and tissue mineral levels of pigs housed under commercial conditions. *J. Swine Health Prod.* **12**:228–236.

Pezzuolo, A., M. Guarino, L. Sartori, L. A. González, and F. Marinello. 2018. On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera. *Comput Electron Agric* **148**:29–36. doi:10.1016/j.compag.2018.03.003

Quiniou, N., J. Dagorn, and D. Gaudré. 2002. Variation of piglets' birth weight and consequences on subsequent performance. *Livest. Prod. Sci.* **78**:63–70. doi:10.1016/S0301-6226(02)00181-1

Rehfeldt, C., A. Tuchscherer, M. Hartung, and G. Kuhn. 2008. A second look at the influence of birth weight on carcass and meat quality in pigs. *Meat Sci.* **78**:170–175. doi:10.1016/j.meatsci.2007.05.029

Roush, W. B., W. A. Dozier, 3rd, and S. L. Branton. 2006. Comparison of Gompertz and neural network models of broiler growth. *Poult. Sci.* **85**:794–797. doi:10.1093/ps/85.4.794

Rutten, C. J., A. G. J. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. Invited review: Sensors to support health management on dairy farms. *J. Dairy Sci.* **96**:1928–1952. doi:10.3168/jds.2012-6107

Shahinfar, S., D. Page, J. Guenther, V. Cabrera, P. Fricke, and K. Weigel. 2014. Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms. *J. Dairy Sci.* **97**:731–742. doi:10.3168/jds.2013-6693

Shirali, M., A. B. Strathe, T. Mark, B. Nielsen, and J. Jensen. 2017. Joint analysis of longitudinal feed intake and single recorded production traits in pigs using a novel Horizontal model. *J. Anim. Sci.* **95**:1050–1062. doi:10.2527/jas.2016.0606

Song, Y., H. Zhou, P. Wang, and M. Yang. 2019. Prediction of clathrate hydrate phase equilibria using gradient boosted regression trees and deep neural networks. *J Chem Thermodyn* **135**:86–96. doi:10.1016/j.jct.2019.03.030

Weatherup, R. N., V. E. Beattie, B. W. Moss, D. J. Kilpatrick, and N. Walker. 2010. The effect of increasing slaughter weight on the production performance and meat quality of finishing pigs. *Anim. Sci.* **67**:591–600. doi:10.1017/s1357729800033038

Witten, I. H., and E. Frank. 2005. *Data mining: Practical machine learning tools and techniques.* 2nd edn. Elsevier/Morgan Kaufmann, San Fransisco, CA.

Yu, R., P. Leung, and P. Bienfang. 2006. Predicting shrimp growth: Artificial neural network versus nonlinear regression models. *Aquac Eng* **34**:26–32. doi:10.1016/j.aquaeng.2005.03.003