Norwegian University of Life Sciences
Faculty of Chemistry, Biotechnology
and Food Science

Philosophiae Doctor (PhD)
Thesis 2024:13

Objective methods of classification for beef- and pork carcases as a basis for evaluation of carcase yield

Objektive metoder for klassifisering av storfe- og griseslakt som grunnlag for vurdering av slakteutbytte

Andrew Russell Thee Heggli

Objective methods of classification for beef- and pork carcases as a basis for evaluation of carcase yield

Objektive metoder for klassifisering av storfe- og griseslakt som grunnlag for vurdering av slakteutbytte

Philosophiae Doctor (PhD) Thesis
Andrew Heggli

Norwegian University of Life Sciences
The PhD programme in Life and Food Science
at the
Faculty of Chemistry, Biotechnology and Food Science

Ås 2023

Thesis number 2024:13
ISSN 1894-6402
ISBN 978-82-575-2131-8

Supervisors and Evaluation Committee

Supervisors

Dr. Hilde Vinje
Associate professor
Norwegian University of Life Sciences (NMBU), Ås, Norway

Dr. Lars Erik Gangsei
Researcher
Animalia AS, Oslo, Norway

Dr. Ole Alvseike
Head of development, Research and Development
Animalia AS, Oslo, Norway

Evaluation committee

Dr. Mette Christensen
Senior Specialist in Meat Science
Front Matec, Frederikssund, Capital Region, Denmark

Dr. Thea Bjørnland
Associate professor
Norwegian University of Science and Technology, Trondheim, Norway

Dr. Thore Egeland
Professor
Norwegian University of Life Sciences (NMBU), Ås, Norway

Acknowledgements

The work presented in this thesis was carried out at the Faculty of Chemistry, Biotechnology and Food Science (KBM) at the Norwegian University of Life Sciences (NMBU). The project was funded by the Research Council of Norway through grant number 311394.

I would first like to extend a heartfelt thank you to my two main supervisors Lars Erik Gangsei and Hilde Vinje. A better dream team of supervisors would not have been possible. They have mentored me through a doctorate that started only two months before a global pandemic hit, and made adjustments immediately to all the challenges that followed.

Before I ever met Lars Erik, people had spoken very highly of him, and that reputation is well earned. I am especially appreciative of the incredible patience he has had trying to teach coding, mathematics and statistics to someone who does not have a background in coding, statistics or mathematics. Lars Erik embodies not only what I consider to be the markings of a great scientist, but a good person. He has not only shared his knowledge liberally, but also made sure that aspects during his own doctorate that could have been made easier, were made easier for me.

Hilde Vinje, like Lars Erik, has a stellar, well earned, reputation. Hilde is especially pedagogical, always beaming, and a well of information. Hilde has made sure the doctorate stayed on track while handling increased responsibility at NMBU, having twins, and a broken foot. The latter two at the same time! Hilde was also responsible for the first statistical subject I took during the doctorate, and was always willing to share her expertise and time with me.

A special thanks to my esteemed colleagues at Animalia AS. Especially Ole Alvseike for moral support, ideas, and help with writing when writers block reared its head, and Morten Røe for sharing his vast knowledge of classification, and his love of song and poetry. A special thanks to the incredible professors and students at KBM for stimulating conversations, and genuine enthusiasm for expanding our understanding of the world, through science, and specifically biostatistics.

A warm, loving thank you to my closest friends and family for their support through a challenging time. Thank you to my son, for the laughter and joy you've given since we had you. Last, but definitely not least, boundless gratitude to my wife, Susanna. Your support, love and sacrifice has been invaluable.

Table of Contents

1 Abbreviations and definitions

CB: Crossbreed
CCP: Commercial Cutting Pattern
CT: Computed Tomography
DNA: Deoxyribonucleic Acid
EU: European Union
EUROP: The letters used in classification of bovine and porcine carcasses in Europe. For bovine carcasses classification from P- to E+ signify respectively concave to convex carcasses. For porcine carcasses P- to E+ signify respectively lower to higher lean meat percentage.
HGP7: Hennessy Grading Probe 7
HVC: High Value Cut
LMP: Lean Meat Percentage
LMY: Lean Meat Yield
NIR: Near-Infrared
PB: Purebred
PCA: Principal Component Analysis
PLS: Partial Least Squares
PLS-QDA: Partial Least Squares - Quadratic Discriminant Analysis
SNP: Single Nucleotide Polymorphism
UNECE: United Nations Economic Commission for Europe
VIA: Video Image Analysis

2 List of papers

Paper I

Objective carcass grading for bovine animals based on carcass length

Andrew Heggli, Lars Erik Gangsei, Morten Røe, Ole Alvseike, Hilde Vinje

Acta Agriculturae Scandinavica, Section A – Animal Science, 70 (2021), pp. 113-121.

DOI: 10.1080/09064702.2021.1906940

Paper II

Carcase grading reflects the variation in beef yield – a multivariate method for exploring the relationship between beef yield and carcase traits

Andrew Heggli, Ole Alvseike, Frøydis Bjerke, Lars Erik Gangsei, Jørgen Kongsro, Morten Røe, Hilde Vinje

animal, Volume 17, Issue 6 (2023), 100854.

DOI: 10.1016/j.animal.2023.100854

Paper III

Classification of breed combinations for slaughter pigs based on genotypes – Modelling DNA samples of crossbreeds as fuzzy sets from purebred founders

Hilde Vinje, Hilde Kjelgaard Brustad, Andrew Heggli, Claudia Maria Sevillano, Maren von Son, Lars Erik Gangsei

Frontiers in Genetics, Volume 14 (2023)

DOI: 10.3389/fgene.2023.1289130

Paper IV

Evaluation of the relationship between porcine carcass grading, breed and yield of primal cuts

Andrew Heggli, Ole Alvseike, Frøydis Bjerke, Hilde Vinje, Jørgen Kongsro, Morten Røe, Lars Erik Gangsei

Manuscript

3 Abstract

Yield within the meat industry refers to the ratio and weight of tissues from the carcase. Carcase classification gives each carcase grades that determine the price paid for the carcase. Two aspects are especially important in regard to classification; that the grades given are objective, in the sense of being affected as little as possible by human influence, and that the grades accurately reflect the yield. If these two aspects are not well executed producers or abattoirs could potentially suffer a biased or unprecise carcase grading, resulting in an unfair system. Yield has previously been measured by the amount of lean meat in the carcase, both for bovine and porcine carcases. As other aspects of the carcase affect the value of the carcase, new yield variables are needed to account for the variation that lean meat cannot. A new, objective method of classification has partially replaced the previous method of classification by expert human classifiers. The intent of this thesis is to compare the objective classification method for bovine carcases to expert human classifiers, as well as examine the relationship between classification and yield variables more complex than previous literature for both bovine and porcine carcases.

The first study examined the relationship between human classification for conformation and fat cover of bovine carcases and classification from the objective method implemented in Norway since 2019. The new objective method used length measurement of the carcase as well as carcase traits to predict the classification. This method would be cheaper than other popular objective methods within the industry, and thus be of particular interest to smaller abattoirs. The results showed high correlations and low biases for conformation classification, which indicates that the objective classification method functioned well compared to human classifiers. The model functioned poorly when predicting fat cover. The increase in precision when breed information was added was marginal for predicting conformation, but larger for fat.

The second study examined the relationship between classification and yield from bovine carcases. Compared to previous studies, this study used a more complex measurement of yield that, in addition to ratio of fat and bone to carcase weight, included a separation of meat into trimmings of different fat percentages and cuts

that can be used without further processing. This analysis also included breed groups consisting of breeds with similar traits. The average percentage residual per breed group was calculated to examine breed as a confounding factor. The analyses in this study would be of particular interest to the producers of the various breeds of cattle, abattoirs and breeding companies. Classification was shown to explain the majority yield variation, and breed had an effect on yield outside of the variation explained by classification.

The third study developed a method of statistically classifying crossbred pigs based on single nucleotide polymorphism (SNP) data. The method used partial least squares (PLS) with SNP data from purebred (PB) pigs as training data. The PLS scores and known aspects of the multivariate normal distribution were used in Partial Least Squares - Quadratic Discriminant Analysis (PLS-QDA) to statistically classify crossbreed (CB) pigs based on their SNP data. The predicted crossbreed combinations were used in Paper IV as independent variables.

The fourth study examined the relationship between classification and yield from finishing pigs. Classification was done using the optical probe HGP7, Hennessy Grading Probe 7 (Hennessy Grading Systems, Auckland, New Zealand. Web page: hennessy-technology.com), with a subset of the sample also classified using the Autofom III (FrontMatec, Kolding, Denmark. Web page: frontmatec.com). The yield variables used as a response were lean meat, fat, and bone percentage, as well as percentage primal cuts. As primal cuts are often sold without further processing, this would be of particular interest to breeding companies and the meat industry as a whole. This study also included cross breed information based on genetic data in the model. The combination of breed information based on genetic data and percentage primal cuts as the response is a novelty within the literature. The results showed that classification explains variation in lean meat and fat to a degree in line with previous literature, but poorly explains variation in primal cut percentages. Breed varied in its significance in regard to the various yield measurements used in the study.

4 Norsk sammendrag

Utbytte innenfor kjøttindustrien refererer til ratioen og vekt mellom vevstyper fra slaktet. Klassifisering av slakt gir hvert slakt karakterer som bestemmer prisen på slaktet. To aspekt er spesielt viktige ved klassifisering; at karakterene som gis er objektive, i den forstand at de er minst mulig påvirket av menneskelig innflytelse, og at karakterene så nøyaktig som mulig representerer utbytte som kommer fra slaktet. Hvis disse to aspektene ikke er oppnådd så kan klassifisering muligens gi en systematisk fordel til enten produsentene eller slakteriene, som gjør at systemet ikke er rettferdig. Utbytte, både hos storfe og gris, har tidligere blitt målt basert på kjøttprosent. Ettersom det er andre aspekt ved slaktet som påvirker slaktets verdi, trengs nye mål på utbytte som klarer å beskrive variasjonen som ikke forklares av kjøttprosent. En ny, objektiv metode for klassifisering har delvis erstattet den tidligere metoden, altså klassifisering utført av godkjente menneskelig klassifisører. Målet med denne avhandlingen er å sammenligne den objektive metoden for storfe klassifisering med klassifisering gjort av ekspert klassifisører, og analysere forholdet mellom klassifisering og utbytte variabler som er mer komplekse enn tidligere studier, både for storfe- og griseslakt.

Den første studien utforsket forholdet mellom klassifisering av konformitet og fettgruppe for storfeslakt utført av menneskelig klassifisører og klassifisering utført av den objektive metoden implementert i Norge i 2019. Den nye objektive metoden brukte lengdemåling og andre egenskaper ved slaktet til å predikere konformitet og fettgruppe. Den objektive metoden er billigere enn andre populære objektive metoder og burde være av interesse for mindre slakterier. Resultatene viste høye korrelasjoner og liten forventningsskjevhet for konformitet, som indikerer at den objektive metoden fungerte bra sammenlignet med menneskelige klassifisører. Prediksjonsmodellen fungerte dårlig til å predikere fettgruppe. Økningen i presisjon når rase informasjon var inkludert i modellen var marginal for konformitet, men større for fettgruppe prediksjon.

Den andre studien utforsket forholdet mellom klassifisering og utbytte fra storfeslakt. Sammenlignet med tidligere studier brukte denne studien et mer komplekst mål på utbytte. Utbytte ble målt ved fettprosent, beinprosent, og kjøttprosent ble delt i fire grupper. Tre sorteringsgrupper hvor mengde fettprosent skilte gruppene fra hverandre, og en gruppe med kjøtt som ikke trengte videre foredling. Denne analysen inkluderte også rasegrupper, hvor hver gruppe besto av raser med lignende egenskaper. Den gjennomsnittlige statistiske feilen per rasegruppe ble regnet ut for å utforske rase som en konfunderende variabel. Analysene i denne studien er muligens av interesse for storfe produserenter, slakterier og avlsselskaper. Klassifisering ble vist å forklar majoriteten av utbyttevariasjon, og rase hadde en effekt på utbytte utover det som kunne forklares av klassifisering.

Den tredje studien utviklet en metode for å statistisk klassifisere slaktegris, som er krysningsraser, basert på SNP data. Metoden brukte partial least squares (PLS) sammen med SNP data fra renraset griser som treningsdata. PLS scorene og kjente aspekt ved den multivariate normal fordeling ble brukt i Partial Least Squares - Quadratic Discriminant Analysis (PLS-QDA) til statistisk klassifisering av krysningsraset slaktegris basert på deres SNP data. De predikerte kryssningsrase kombinasjonene ble brukt i artikkel IV som uavhengige variabler.

Den fjerde studien utforsket forholdet mellom klassifisering og utbytte fra slaktegris. Klassifiseringen ble utført ved bruk av HGP7-instrumentet (Hennessy Grading Systems, Auckland, New Zealand. Web page: hennessy-technology.com), mens et subsett av utvalget også ble klassifisert av Autofom III (FrontMatec, Kolding, Denmark. Web page: frontmatec.com). Utbyttevariablene som ble brukt som respons i modellen var kjøtt-, fett-, og beinprosent, og andel stykningsdeler. Ettersom stykningsdeler er ofte solgt uten videre foredling, kan dette være av interesse for avlsselskaper og kjøttindustrien generelt. Denne studien brukte også rasekombinasjon informasjon basert på SNP data i modellen. Kombinasjonen av rasekombinasjon informasjon basert på SNP data som forklaringsvariabel og andel stykningsdeler som respons i modellen er ikke tidligere gjort i litteraturen. Resultatene i studien viste at klassifisering forklarer variasjon i kjøtt- og fettprosent som var på linje med det som tidligere har blitt forsket på, men at variasjon i stykningsdeler var dårlig forklart. Om effekten av rase var signifikant var avhengig av hvilken utbytte variabel det gjaldt.

5 Synopsis

5.1 [Introduction]

Classification of animal carcases for the purpose of facilitating trade between producers and abattoirs has a long history and is ubiquitous (Polkinghorne et al., 2010). Though classification systems used today vary depending on location and carcase genus (Polkinghorne et al., 2010), classification in general is intended to fulfil a few specific functions. The primary function of classification is to ensure a fair price per carcase between producers and abattoirs. The first secondary function is to give information to the abattoirs about how each carcase can best be further processed. An additional secondary function is as a feed-back system to producers (Dorleku et al., 2023) of how well their carcases are conforming to the desired ratio of tissues, commonly referred to as yield, in the given market they are in.

As classification was not developed specifically as a detailed measure of yield, but more so a method of being able to describe carcases of varying muscularity and fatness, these two variables alone might not necessarily sufficiently describe the variation in yield. There have therefore been decades of studies examining the relationship between classification and yield (Kempster et al., 1982, Conroy et al., 2010, Oliver et al., 2010, Craigie et al., 2012). This is important as a strong relationship between classification and yield fulfils the functions previously stated. If carcases with more desirable yield are classified higher, and therefore paid better, the system produces a fair price. In addition, this provides feed-back to producers in the sense that they can compare the effect different production strategies have had on classification, such as difference in feeding or choice of sires. Conversely, if the relationship between yield and classification is weak, the feedback to producers is less clear.

Historically, classification has been executed by humans looking at carcases and using visual aspects of the carcase that are important for trading. Over time there has been a need to move away from human involvement in the interest of classification being more objective. Objectivity is important because if classification has as little involvement from humans as possible, the system is less susceptible to human influence such as favouritism or changes in mood. This ensures a fairer price, as well as equal feed-back to producers and a stronger relationship between yield and classification, by reducing variation in yield across different classification grades. The ratio between tissues in the carcase is important as each carcase is eventually broken down into products for consumers. A consequence of a mismatch between the ratio of tissues in the carcases and the ratio of tissues in the products is more waste. This could lead to further negative views of the meat industry, which is already under intense scrutiny.

A theoretical, perfect classification system could, at an efficient cost, account for all yield variation relevant to that market, regardless of age, gender, castration, breed or any other variables that are not specifically a descriptor of the tissue ratio and amount in the carcase. In practice, this is not possible, so the goal of classification is to explain the majority of yield. Fortunately, variance outside of classification can be accounted for by other methods. For example, a carcase of a particularly sought after breed could receive higher pay than a carcase with the same classification grade, such as is the case with Angus carcases in Norway (Nortura, 2023). This is why it is still useful to examine how much of the variation in yield is explained by variables other than classification, as well as confounding variables. In addition, understanding the relationship between commonly registered traits and yield facilitates easier comparison between studies from different countries and is beneficial for abattoirs when further processing the carcases. This thesis will examine how well objective classification functions compared to human classification, as well as examine how well classification fulfils its intended functions, through its relationship to yield variables that are relevant to the potential value of the carcase. This is done through the use of sufficiently complex statistical models to analyse each problem in the respective papers.

5.1.1 Objective classification

Human classification based on visual assessment of animal carcases is generally considered to be susceptible to inconsistencies, although there is no definitive evidence of this (Craigie et al., 2012). An example of a particular inconsistency could be between classifier variance, meaning different classifiers would give the same carcase different grades. Different methods have been developed to move away from human classification, including, but not limited to, video image analysis (VIA), optical probes, and ultrasound instruments. European legislation specifies a methodology for approvement of such methods (European Commission, 2017). VIA is commonly used for bovine and ovine carcases, while optical probes and ultrasound instruments are commonly used for porcine carcases. This thesis concerns itself only with bovine and porcine carcasses.

In Europe, including Norway, the EUROP grid method specified by European Union (EU) legislation (European Commission, 2013) is used to classify bovine carcasses, with one grade for conformation and one grade for fat cover. Both conformation and fat cover are on a 15-point scale, with conformation from P (more concave) to E (more convex), and fat cover from 1 (thinnest fat layer) to 5 (thickest fat layer), with both conformation and fat cover having an additional -/+ variation for each grade. There is an additional S grade in EUROP that is above E, but as bovine carcases in Norway do not reach this grade, it is not included in this thesis. Porcine carcasses are also classified according to EU regulations, but with only one criterion, the lean meat percentage (LMP), which determines the class the carcase is given. In Norway there are five potential classes, P- (LMP <48%), R (48 - 49%), U (50 – 54%), E (55 – 59%), and S (60 - 68%) (Røe, 2021).

5.1.1.1 Porcine classification

Value determination of porcine carcasses during the 1930s in Norway was originally done by separating pigs into weight classes. Throughout time there have been several developments, including, but not limited to, separating gilts, sows and boars during the 1940s, and measuring the subcutaneous fat layer with a specialised knife during the 1960s. In 1989 LMP was introduced as a method of classification of porcine carcases, measured by the use of the Hennessy GP2Q optical probe. The common intention for all developments was to facilitate production of the preferred type of carcase at the time, and differentiate carcases of different value. In 1996 the EUROP system of classification was implemented in Norway and the LMPs were utilised to separate porcine carcases into EUROP classes.

Classification by optical probes is a semi-automated method that functions by puncturing the carcase with a probe that has a light emitter and a receiver (Engel et al., 2012, Gangsei et al., 2018). The receiver registers the refracted light and based on the amount, discerns fat from meat. The data from the instrument is then used in a regression model to determine the LMP. This method is more objective than human visual appraisal of the carcase, and more detailed than purely weight classes or gender categories. There is still room for improvement as it requires qualified human individuals to use the instrument. For this reason, other methods of classification have been developed, including ultrasound instruments, such as the Autofom III (Brøndum, Egebo, Agerskov, & Busk, 1998, Busk, Olsen & Brøndum, 1999, Font i Furnols & Gilbert, 2009) by Front Matec. It functions without active human involvement and is therefore arguably more objective than optical probes, or at least less prone to human error. As with optical process such as the HGP7, Autofom is approved for classification by European Legislation (European Commission, 2017), through a method of validating the LMP from the instrument compared to LMP calculated after total or partial dissection (Walstra & Merkus, 1996) or computed tomography (CT) (Olsen et al., 2017).

The Autofom has been approved for classification in Norway within recent years, but the HGP7 is still more widely used. This could potentially change in the future, but with the higher cost and need for a larger space, it is potentially less suitable for smaller abattoirs. The initial intent was to analyze how well a more objective method, the Autofom, could explain variation in yield compared to the HGP7, though this was decided against as the correlation between the classification from the instruments was high, and a much larger data set existed for carcases only classified with the HGP7 instrument. The focus was then shifted to examining the relationship between HGP7 classification, yield and confounding factors.

5.1.1.2 Bovine classification

For bovine carcases, VIA is a popular method with several alternative manufacturers (Borggaard et al., 1996, Allen & Finnerty, 2000, Craigie et al., 2012, Wnęk et al., 2017) within Europe. This includes the VBS 2000 system from the German manufacturer E + V GmbH, the BCC-2/BCC-3 systems from the Danish manufacturer FrontMatec and the MAC-2/MAC-S systems from the French manufacturer Normaclass. All VIA systems function similarly using cameras to take pictures of the carcase, and use data from the pictures to determine classification of the carcases, or qualitative aspects of the carcase or cuts (Allen & Finnerty, 2000, Craige et al., 2012, Delgado-Pando et al., 2021). For example, the BCC-3 uses time-of-flight near-infrared cameras to produce 3D images of the carcase (Nisbet et al., 2023). Whole-carcase VIA systems are popular as they are automated, can grade hundreds of carcases per hour (Craige et al., 2012) at or close to line speed, with results comparable to expert human classifiers (Allen & Finnerty, 2000), and can be approved by the EU legislation (European Commission, 2017). The potential downside to VIA systems is high cost, which might not be suitable for smaller abattoirs, and though VIA can grade fat cover, it functions less well for fat cover classification than for conformation.

To offer a more economical alternative, The Norwegian Meat and Poultry Research Centre developed a semi-automated method that used an instrument for length measurement (Heggli et al., 2021). The length measurement is used in a model combined with age, weight, category and breed to predict conformation. This method was utilized from 2019, while fat cover is still graded by authorized human classifiers. This thesis used a similar, but modified, version of the model that was used in the industry. The purpose was to examine if the method could function at least as well as expert human classifiers. Comparing new methods to the previous method based on human classifiers enables verification that the objective methods are at least as good as human classification, but as humans are not objective, it would be good to have a more objective measure to compare against.

5.1.2 Measuring potential carcase value

A natural measure to compare different classification methods would be the monetary value of the carcase. Although possible (Janiszewski et al., 2018), this offers some difficulty as the monetary value of a carcase can be different across markets, and across time (Marcoux et al., 2007). Yield from carcases is the basis for the products that are made from the carcase, and thus determines the potential value of the carcase. As yield is comparable between markets and across time, it is preferable that any study examining classification and potential carcase value should use yield as a proxy for value of the carcase, and monetary value in a specific market can potentially be an additional measure.

The relationship between yield and classification has been examined previously (Kempster et al., 1982, Conroy et al., 2010, Oliver et al., 2010, Craigie et al., 2012), often with the measure of yield being lean meat yield (LMY) or saleable meat yield (SMY) (Craigie et al., 2012). LMY is usually defined as grams of lean meat (i.e., excluding bone and fat) per kilo of carcase weight (Bergen et al., 2003, Craigie et al., 2012, Delgado-Pando et al., 2021), while the definition of SMY varies depending on the market, as it is lean meat with some fat, and potentially bone included (Craige et al., 2012, Delgado-Pando et al., 2021). A measure that in addition to LMY includes percentage fat and percentage bone, is carcase composition (Drennan et al., 2008, Conroy et al., 2010). These are welcome additions as the ability of classification to predict SMY and LMY is often examined on the basis of SMY and LMY being proxies for carcase value, and fat percentage and bone percentage affect the potential value of the carcase.

Even with fat and bone percentage included, there is still room for improvement as livestock carcases are not valued purely based on carcase composition. For bovine carcases, depending on how the meat is distributed, some carcases can hold more weight in cuts that are of higher value, which is variation that will not be accounted for when purely using carcase composition. For porcine carcases, the ratio between primal cuts (Choi et al., 2018, Marcoux et al., 2003) particularly affects the value of the carcase (Marcoux et al., 2007), another attribute whose variation is not captured by SMY, LMY or carcase composition. This leads to the conclusion that when examining yield and classification, the measure of yield used should be more complex than just SMY, LMY or carcase composition, and should be specific to the genus of the carcase.

5.1.3 Measurements of yield

For this thesis, a multivariate yield measure was chosen for bovine carcases. The measurement includes fat percentage and bone percentage, while LMY was segregated into different types of meat based on potential value. This included one group labelled "high value cuts" (HVC) that included cuts that didn't require further processing, such as the tenderloin, and three groups of trimmings based on the amount of fat (5, 14 and 21%). Waste was included in the bone percentage as it is a very small percentage of total carcase weight and, as with bones, generally have negative value. For porcine carcases, the yield measure chosen for this thesis includes lean meat-, fat- and bone and waste percentage, as well as the ratio between the primal cuts (forepart, loin, belly, and hindpart).

These particular measures of yield offer several benefits. Though the value of each group within the yield definition could vary between markets, if all studies used these definitions, it would facilitate easier comparison between the studies while allowing for later calculation of the value within each market. Each yield measure is also specific to their respective genus of the carcase, and includes variables that are relevant to the potential value of the carcase.

5.1.4 Other variables and confounding factors

As classification cannot account for all variation in yield, it is important that other variables that could affect yield are taken into consideration in a statistical model/analysis. This is especially important when using a more complex definition of yield that is specific to the genus of the carcase. This gives producers, abattoirs, breeding companies and researchers' information about how yield is affected by variables outside of classification, and a basis for potential changes in compensation outside of classification differences. Commonly measured traits are preferable as they allow comparison between markets and are relevant to the potential value of the carcase. These variables could include the weight and category of the carcase, which are both included in this thesis for porcine and bovine carcases.

In addition to relevant variables outside of classification, confounding variables are important to take into account. Confounding variables in this thesis refer to variables that might explain variation in yield outside of the effect of classification, but would be preferable from an industry perspective to not have an effect. Thus, for this thesis, breed is considered a confounding variable as it would be preferable that classification captures enough variance that taking breed into account would be superfluous. It is therefore included for both porcine and bovine carcases. For porcine carcases, operator data was available and was also included in the model as a confounding variable for the same reason as breed.

5.1.5 Aim of thesis

The aim of the thesis was to examine the validity of objective methods of classification, with the specific aims of study being:

1. Examine the precision and bias between objective classification and traditional classification (Paper I)
2. Examine the relationship between classification and yield (Paper II & Paper IV)
3. Examine the relationship between classification and confounding factors (Paper I, Paper II, Paper III & Paper IV)

5.2 [Materials and Methods]

The following chapter provides an overview of the methods of the papers included in the thesis. This includes data collection and statistical analyses to provide a basis for understanding of the subsequent results and discussion. All sample sizes are after data cleaning. For sample sizes before data cleaning and a more detailed description of methods and statistical analyses, please see the individual papers supplied at the end of the thesis.

5.2.1 Data

5.2.1.1 Classification and carcase data (Paper I, II & IV)

Paper I

EUROP classification grades for bovine carcases (n = 3756) including conformation and fat cover grades were given by five of Animalia's expert classifiers at abattoirs between 2018 and 2020. The majority of carcases were classified by one classifier (n = 3622), but some were classified by two or up to all five of the expert classifiers (n = 124). In addition to classification grades, length and warm carcase weight were collected at the abattoirs. Length of the carcase was measured by a laser (item number MTS04700412 Meats AS, Bergen, Norway). Warm carcase weight was measured by each carcase hanging from a hook by the Achilles tendon and adjusted by a factor of 0.98 to approximate cold carcase weight. Age (in days) at time of slaughter and gender were collected from the animal livestock registry (Mattilsynet, 2022), a Norwegian registry with data registered by producers.

Paper II

As with Paper I, conformation and fat cover grades for bovine carcases (n = 1687) were registered at the abattoirs, but for Paper II conformation and fat cover grades were given by certified classifiers, not Animalia's expert classifiers. The classification grades were collected from 2011 to 2021. Conformation grades pre 2019 were given by human classifiers, while conformation grades post 2019 were given by a model developed by Animalia and Nortura that used length, weight, breed, age and gender as predictors. Fat cover grades were given by human classifiers for all years of data collection. In addition to classification grades, warm carcase weight adjusted by a factor of 0.98 to approximate cold carcase weight and category were registered at the abattoirs. Category refers to a combination of age and gender, and is almost identical to categories specified by the European commission (European Commission, 2013). The categories are calf, young bull, bull, steer, heifer, young cow and cow.

Paper IV

LMP values for porcine carcases (n = 3018) were collected from 2012 to 2023 at abattoirs by certified classifiers using the HGP7. A subset of 216 individuals were classified with the Autofom III in addition to the HGP7. In addition to LMP, warm carcase weight adjusted by a factor of 0.98 to approximate cold carcase weight, category and castration method were registered at the abattoirs. Category and castration method were used to create a new variable named gender/castration category which included immunologically castrated pigs, gilts and physically castrated pigs.

5.2.1.2 Yield data (Paper II & IV)

Animalia's (The Norwegian Meat and Poultry Research Centre) pilot plant processes carcases daily at the behest of the Norwegian market regulation authorities for the purpose of providing information to the regulation authorities, the classification system and other stakeholders. Carcases are deboned and all manner of tissues are weighed.

Paper II

Bovine carcases were processed at the pilot plant using two commercial cutting patterns (CCP). Both patterns were similar enough to be combined into one cohesive data set similar to the UNECE (United Nations Economic Commission for Europe) standards (UNECE, 2016). Each carcase was quartered, then processed into smaller cuts, trimmings, fat, bones and waste which were weighed and registered. These formed the basis for the creation of six product categories.

Cuts that required no further processing than being cut from the carcase and trimmed, and were considered of potentially high value (e.g., the tenderloin), were aggregated into a product category named high value cuts (HVC). Trimmings were separated into three disparate categories based on their fat content. The goal was to have 5, 14 and 21% fat in each category, respectively. The trimming categories was named meat5, meat14 and meat21 based on their respective fat percentages. Fat, and bone and waste were their own, self-named categories. This resulted in the six product categories HVC, meat5, meat14, meat21, fat, and bone and waste. In addition to the six product categories, the percentage the forepart made up of the whole carcase was also calculated, referred to as the percentage forepart.

Paper IV

The pilot plant processed porcine carcases using a CCP described in detail in Gangsei et al. (2018). This method is highly correlated (Gangsei et al., 2018) to the manual dissection ($\rho = 0.88$) and CT ($\rho = 0.91$) methods described in the EU legislation by the European Commission (2017). Each carcase was split into left and rights halves at the abattoirs before being sent to the pilot plant. Based on the end products of the CCP, and digital processing of these, the half carcases were segmented into three tissue types: lean meat, fat, and bone and waste. The sum of lean meat, fat, and bone and waste in each primal cut (forepart, loin, belly and hindpart) was calculated, and divided by the weight of the side carcase minus the feet and jowl to create four variables, referred to as the primal cut percentages (PCP). This resulted in seven yield variables: lean meat-, fat-, bone and waste percentage, forepart-, loin-, belly- and hindpart primal cut percentage.

5.2.1.3 Breed data (Paper I, II, III & IV)

Paper I

Breed data in Paper I was sourced from the animal livestock registry and was based on pedigree back to great-great-grandparents, meaning the lowest percentage breed an individual could belong to was 1/16th. The carcases in the data set had pedigrees from a total of 35 different breeds. These 35 breeds were sorted in 9 breed groups (Røe, 2019) based on shared qualitative traits. As breed populations vary in size, some breeds had to be combined to create a large enough sample size to be included in the analysis.

Paper II

Breed in Paper II came from two sources. The data from the animal livestock registry was as in Paper I, with breed distribution registered in 16th's. The other source of breed data was the abattoirs. The abattoirs only register the dominant breed, and as such a carcase that is e.g., 75% of a particular breed and 25% of another would be registered as only the first breed. Thus, the carcases with breed data from the abattoirs were treated as pure breeds in the analysis, though they in reality could have been crossbreeds. The breeds were aggregated into breed groups based on shared qualitative traits, similarly to Paper I, though Paper II had only seven breed groups.

Paper III & IV

Data in Paper III were supplied by Norsvin (Norway) and Topigs Norsvin (Netherlands). Animals were genotyped using a custom GeneSeek 50K (Illumina) SNP chip (Lincoln, NE, United States). Of the 50K SNPs, 23070 SNPs were used in the study, as is the standard routine by Topigs Norsvin. The PBs in the data set were either Landrace ("L", n = 1000), Large White (”W”, n = 1000), Duroc (”D”, n = 1000) or Hampshire (”H”, n = 14), which are the most common PBs in Norway. There was also an additional breed included that is not a part of the Norwegian pig population, the Pietrain ("P", n = 1000). In addition to the SNP data from PB pigs, SNP data from 1013 crossbred finishing pigs were collected. The predicted breed combinations for the CB finisher pigs were utilised in Paper IV.

5.2.2 Statistical analyses

All statistical analyses were coded and executed in the software R (RStudio Team, 2020). Coding scripts are openly available for the published papers (I, II and III) at https://doi.org/10.18710/T9SXVF, https://doi.org/10.18710/TMSSJP and https://github.com/LarsErikGangsei/CrossBreedPredictionPublic, respectively.

5.2.2.1 General theory

Statistical regression models are models that have a dependent variable, often referred to as the response, and independent variables, often referred to as predictor or explanatory variables. The information in the dependent and independent variables come from a data set with information sampled from a population.

Linear regression models

The current thesis uses linear regression models as the basic statistical tool in three of the articles (Paper I, II and IV). In vector form the most basic model can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e},$$

where the response $\mathbf{y}$ is a $n \times 1$ vector with one observation per individual and n individuals. $\mathbf{X}$ is a $n \times p$ design matrix of p independent variables, and $\mathbf{e}$ is $n \times 1$ vector with random errors. The $p \times 1$ vector $\boldsymbol{\beta}$ is the fixed effect regression parameters. The models used in Paper I are of this vector form. The predictor variables ($\mathbf{X}$) were strategically chosen to give both a high prediction precision and ease of interpretation. B-splines were used to model the non-linear relationships between the classification (response) and the predictor variables weight, length and age.

A slightly more advanced model, with a multivariate response, was applied for analysing the bovine yield data in Paper II. The general formula is the same as for the univariate response given above, but $\mathbf{y}$ and $\mathbf{e}$ in the multivariate model are of a different dimension and are both $n \times q$ matrices, where q is the number of variables in the response. $\mathbf{X}$ is a $n \times p$ design matrix with explanatory variables like in the univariate case. $\boldsymbol{\beta}$ is a $p \times q$ matrix whose elements are regression parameters. When standard ordinary least squares regression is used, the regression parameter estimates $\widehat{\boldsymbol{\beta}}$ will be the same whether a multivariate response model or q univariate models are applied for the elements in the response. However, a nice feature of the multi-response model is that a covariance structure can be added to the error term. In the univariate case it is common to assume that errors are independent, identical and normally distributed with zero mean and common variance σ^2, whereas in the multivariate case errors are assumed identical and normally distributed with zero mean and common covariance matrix $\boldsymbol{\Sigma}$. Analysis of the covariance matrix added further insight to the model and the relationship between the different variables in the multivariate response in Paper II.

In Paper I (the random effect model) and Paper IV mixed models were used. Mixed models have the general form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

where the only “new” aspect compared to ordinary linear regression are the random effects represented by the term **Zu**. This is the general form that was used in the R-package utilized in Paper IV and described in Bates et al. (2015). The $n \times m$ matrix $\mathbf{Z}$ is a design matrix, often but not necessarily with elements 0 and 1, and the $m \times 1$ vector $\mathbf{u}$ are random effect regression elements, assumed to be multivariate normal with zero mean. The random errors are assumed identical, independent, and normally distributed with common variance. An important goal for Paper IV was to construct a model including several random factors, operator, unknown breed and side of the carcase to get a better understanding of how much of the variance in each of the yield variables could be attributed to each of the random factors. This is done to understand how the random factor, rather than the specific levels of the random factor, affect the response.

Prediction vs understanding

Statistical models are often used for one of two purposes; to predict a response or to understand the relationship between the response and the independent variables. The difference between prediction and understanding is which aspect of the regression is focused on. In prediction, the focus is on reducing the distance between the observed and the predicted response. This distance is referred to as the residual, and the predicted variance associated with the error is the residual variance. When the purpose of the regression is prediction, it is important that the residual variance is minimised particularly for new individuals. This is often done by validating the methods on new data, e.g., by splitting the acquired data set into a training and a test set, such as with Paper I or training on a training set that is completely different from a set of new data in the test set as in Paper III. Other validation methods such as cross-validation exist, where one observation is taken out of the data set and the model is trained on the rest of the data before predicting the observation that was left out. This is then repeated for each observation in the data set, and is commonly referred to as leave-on-out cross validation. Another variant is k-fold cross validation (M. Stone, 1974), where the data set is split into k sub-samples, and the model is trained leaving out one of the k sub-samples, and the observations in the excluded sub-sample are predicted. This process is repeated for each sub-sample. Regardless of the validation method, the general idea is to fit the model and test its ability to predict new values on a new test set to obtain scores for prediction.

Another difference between prediction and understanding is in the choice of independent variables to include in the model. When prediction is the focus variable selection is often done by adding one variable after the other, or including all considered variables and excluding one by one. When understanding is the focus, independent variables should be decided on before fitting the model based on knowledge of which variables are relevant to the field of study in question. The complexity of the model also needs to be considered, as a model that is too complex might impede the understanding of the relationship between the dependent variable and independent variables. For Papers II and IV, the focus is on understanding how changes in independent variables relevant to the dependent variable, yield, affect change in yield.

Statistical classification

Statistical classification is a method of supervised learning that assigns a new individual to a single class out of a set of classes, while minimizing the probability of misclassifying, i.e., assigning an individual to a class to which they do not actually belong. Classification typically assigns the individual to belong to only one class in the set. If the individual belongs to a combination of classes, then the set is referred to as a fuzzy set. The individual will then be assigned a value between 0 and 1 for each class within the set. This was the focus of Paper III.

5.2.2.2 Statistical Models

Paper I

The focus of Paper I was on prediction of conformation and fat cover grade for bovine carcases, and as such, the data set was split into a training (n = 3622) and a test set (n = 124). The training set consisted of carcases that were only classified by one expert classifier, while the test set consisted of carcases that were classified by either two, three, four or all five classifiers. The training data was used to fit a linear regression model:

$$\mathrm{y} = \beta_0 + \beta_k x_k + \boldsymbol{\beta}_b^t \boldsymbol{x}_b + \boldsymbol{\beta}_w^t \boldsymbol{z}_w + \boldsymbol{\beta}_l^t \boldsymbol{z}_l + \boldsymbol{\beta}_a^t \boldsymbol{z}_a + e, \;\; e \sim N(0, \sigma_e^2) \qquad (1)$$

The response, y, is either the conformation or the fat cover grade given by the expert classifier for a single carcase. β_0 is the intercept. x_k is the relationship between weight (kg) and cubic length (m^3), known as Fulton's condition factor (Kurkilahti et al., 2002). $\boldsymbol{x}_b$ is a vector for breed groups, whose elements are in positive 16$^{\text{ths}}$, that sum to 1. B-splines were used to model potential non-linear relationships with weight, length, and age, represented by vectors $\boldsymbol{z}_w$, $\boldsymbol{z}_l$, and $\boldsymbol{z}_a$, respectively. Model (1) was fitted to males and females separately. Each of the models had two versions, one with conformation as the response, one with fat cover as the response. This resulted in four models. For each of the four models there were two versions, one that included breed as a predictor, and one that excluded breed as a predictor. These are referred to as the full- and the reduced model, respectively.

To evaluate prediction precision and bias, a random effect model was applied:

$$y'_{ij} = \mu + t_i + u_{ij}, t_i \sim N(0, \sigma_t^2),\ u_{ij} \sim N(0, \sigma_u^2) \qquad (2)$$

, to the test data set. y'_{ij} , represents the difference between conformation/fat cover set by classifier *j* for individual *i* (y_{ij}) and the associated predicted value ($\hat{y}_i$), based on the fitted model of Eq. 1. μ represents the bias of the prediction of model (1). If the model is biased towards low predictions (y_{ij} greater than $\hat{y}_i$), μ is larger than 0 and vice versa.

The total random error in model (2) for carcase *i* graded by classifier *j* is given by $t_i + u_{ij}$. t_i represents the random classification error for carcase *i* caused by the prediction given by model (1), and has corresponding variance σ_t^2. u_{ij} is the random error caused by classifier *j* for carcase *i*, and has corresponding variance σ_u^2. The intraclass correlation coefficient, corr(y_{ij},y_{ik}) = $\rho = \sigma_t^2 \ / \ \sigma_t^2 + \ \sigma_u^2$ (Dean et al., 2017) is the proportion of total misclassification due to lack of prediction precision from model (1).

The European Commission has rules for approval of automated classification methods of bovine carcases in annex IV of supplementing Regulation (EU) No 1308/2013 (European Commission, 2017). These include calculating the bias, slope of regression of median grading by classifier and automated grading method, and points based on the difference between the automated classification method and the human classifier grades. Bias must be within ±0.30 and ±0.60 for conformation and fat cover, respectively. The slope of regression must be between 1 ± 0.15 and 1 ± 0.30 for conformation and fat cover, respectively. The automated method must achieve at least 60% of maximum points for a test set of at least n = 600, with at least n = 5 independent classifiers. Though this study did not meet the requirements for sample size and number of independent classifiers, the bias, slope and points were calculated to give an indication of how well the objective method could potentially do.

Paper II

The log-odds of the six product categories (HVC, meat5, meat14, meat21, fat, and bone and waste) and the percentage forepart were calculated for the creation of a multivariate response vector of length seven.

Four predictor variables were decided a priori; conformation, fat cover, carcase weight and category. Four additional variables were considered: breed group, and the second-degree polynomials of, and interaction between, conformation and fat. To determine the inclusion or exclusion of the polynomials and the interaction an approximate F-test based on Wilks' lambda was conducted (Mardia et al., 1979). The Wilks lambda test showed significant contribution from all three variables. The polynomials and interaction were then added to the model along with the four initial predictor variables. An additional approximate F-test based on Wilks' lambda was conducted in order to test for the effect of breed group on yield. The effect was significant. However, breed group was omitted in the model as a predictor a priori, as it is viewed as a confounding factor and correlates with the other initial predictor variables. This led to a multi-response linear regression model with independent variables conformation, fat cover, their respective polynomials, the interaction between conformation and fat cover, carcase weight and category. The residuals from this model were used to create an unbiased estimator for the error variance matrix. A correlation matrix based on the estimator was also reported in the results in order to simplify the interpretation.

Regression parameter estimates with corresponding P-values based on t-tests were reported for all regression parameters. In order to simplify the understanding of the effect of a change in each of the independent variables, estimates for a standard individual were calculated. A standard individual was considered a young bull with the mean values for conformation, fat group, and weight. The estimates were transformed from the log-odds to the percentage scale for easy interpretation. The estimated effects of increasing conformation, fat group and weight by one unit, based on this standard animal, were calculated and reported. Furthermore, the effect of changing category was reported applying the same principle. To evaluate the effect of breed group a residual analysis was conducted with the residuals from the linear model as response variables and breed group as predictor variable.

Paper III

The most frequent allele of each SNP in a sample of PB individuals (n = 4014) was identified. SNPs were coded numerically, based on if the genotype was homozygous for the most frequent allele (0), heterozygous (1), or homozygous for the least frequent allele (2). In addition to the SNP data from PB pigs, SNP data from 1013 crossbred finishing pigs were collected. Using the genome data from the PB pigs, several statistical classification methods were utilised to predict the proportion of each PB in the CB data set. As the CB pedigree goes as far back as grandparents, each CB finisher pig then received a prediction with four letters (i.e., LW DD for a CB pig with Landrace, Large White, and Duroc grandparents).

The SNP data from the PBs were used with an identical by descent pattern based on the genetic map from Tortereau et al. (2012) to simulate SNP data for individuals in the F2 generation. For each individual in the F2 generation SNP data from four PB individuals, corresponding to four grandparents, were used to simulate SNP data for the F2 individual according to the breed combination of the F2 individual.

Four statistical classification methods were tested with the intention of using the classification methods to predict breed combinations in the real CB data set. The methods were Random Forest, ADMIXTURE, Partial Least Squares Regression, and PLS-QDA. Each of the methods were fitted using two data sets, the first excluded Pietrain, and the latter included all five PBs. This resulted in eight fitted methods in total. Each of the eight methods were used to classify CB individuals. This resulted in sixteen combinations of method (”four classification methods” x ”two training data sets” x ”two test data sets”), where 7000 individuals were classified per combination of method. For the sake of brevity, PLS-QDA is the only method that will be described, as it is the only method that was used to create the breed combination data used in Paper IV. For more detailed information on the other methods please see Paper IV.

PLS is a supervised dimension reduction method similar to principal component analysis (PCA), but differs from PCA in that the principal components are not created to maximise the variance of the predictor variables, but rather maximise the covariance between the response **Y** (n x q) and the predictor variables **X** (n x p), i.e., the SNPs. The response **Y** was a multivariate response matrix where each row was an indicator vector, i.e., zeros or ones, for the PB in the training data set. The predictor variables (**X**) were used to create a score matrix **T** = **XP**, where **P** (p x m) is the loading matrix that provides weights that are multiplied with each SNP. m = q -1 was used as the number of components with the reasoning that each component would separate two breeds, thus q – 1 components are required for q breeds.

The PLS score vector was used in calculating the posterior probability for different classes (CB) using:

$$P(K|t) = \frac{f_k(t)\pi_k}{\sum_{l=1}^{n_{comb}+1} f_l(t)\pi_l}$$

where $f_k(t)$ was the class-conditional density of t, PLS-scores, for an individual from class K. The class conditional density $f_k(t)$ was assumed to be multivariate normal, with mean/center as a linear combinations of means associated with the actual PB's. The covariance matrices for CB's were constructed based on general properties for the multivariate normal distribution, but additional variance was added due to variation in the actual proportions of inherited DNA from each grandparent. π_k was the prior probability of class K. π_k was an informative prior equal to the proportion of pig litters of crossbreed k among all pig litters in Norway in 2021 (Langaker et al., 2021). In addition to all the possible classes, a class of “unknown breed” was added leading to a total of $n_{comb} + 1$ classes.

Paper IV

Lean meat-, fat-, bone and waste percentage and four PCPs were calculated and utilized as univariate responses in a mixed effect regression model, resulting in seven yield variables. For each individual carcase there were two observations per response variable, one per side of the carcase. The mixed model was fitted to each of the response variables separately, resulting in seven disparate fitted models per individual:

$$
\begin{gathered}
y_{ijkl} = \boldsymbol{x_i^t \beta} + (t_i^{ID} + t_i^{UB} x_i^{UB}) + \tau_k^{side} + \tau_l^{cat} + t_j^{OP} + e_{ijkl}, \\
e_{ijkl} \sim N(0, \sigma^2),\ t_{\mathrm{j}}^{OP} \sim N(0, \sigma_{OP}^2),\ t_{\mathrm{i}}^{ID} \sim N(0, \sigma_{ID}^2), t_i^{UB} \sim N(0, \sigma_{UB}^2)
\end{gathered}
$$

i = 1, … ,3018 (ID)
j = 1, … ,10 (OP)
k = 1, 2 (side)
l = 1, 2, 3 (cat)

y_{ijkl} is the response for side *k* from carcase *i*, of category *l*, processed by operator *j*. The vector term $\boldsymbol{x_i^t \beta}$ contains the fixed effects LMP, carcase weight, the interaction between LMP and carcase weight, as well as the intercept for individual carcase *i*. For individuals with breed data, the breed combination for individual *i*, as fractions of breeds Duroc, Hampshire, Landrace and Large White, as statistically classified in Paper III, is added as fixed predictors. τ_k^{side} and τ_l^{cat} are fixed effects representing the systematic difference between the left and right side of the carcase, and the gender/castration category, respectively. t_i^{ID} is a random intercept term representing the individual carcase and has a corresponding random slope term in $t_i^{UB} x_i^{UB}$. The random intercept is interpreted as the random effect of the individual, whereas the random error, e_{ijkl}, represents the random effect of carcase side after the effect of individual is accounted for. A natural interpretation of $t_i^{ID} + e_{ijkl}$ would be the total unexplained error for side *k* and individual *i*. Every half carcase was processed by an operator *j*, and t_j^{OP} is the associated random effect of operator *j*.

Three coefficient of determination (R^2) values were calculated. Two were according to the method for linear mixed models (LMM) from Nakagawa et al., (2013; 2017). The interpretation of the marginal R^2 value is the variance explained by the fixed effects in the model, while the interpretation of the conditional R^2 value is the variance explained by the fixed and random effects. An additional coefficient of determination, denoted limiting $\mathrm{R}^2_{\mathrm{LMM(lim)}}$, for the individual in question was calculated, which can be interpreted as the maximum part of the variance at individual level, which might be explained by the fixed effects LMP, weight, category and breed together with the random effect operator.

The adjusted intraclass correlation coefficient for LMMs as defined in Nakagawa et al., (2010; 2017) was calculated. The adjusted ICC is the proportion of random variance accounted for by the between group variance while omitting other non-zero variance from other random variables (Nakagawa et al., 2017), such as confounding factors, in this case, operator variance. An additional ICC was calculated which included estimated operator variance, who's interpretation is the amount of total random error explained by the operator.

5.3 [Results and discussion]

Classification of livestock carcases is the system that influences the most control over the monetary transaction between abattoirs and producers. As such any changes to that system will have far reaching consequences. The importance of classification being objective and as accurate a proxy for potential carcase value as possible should therefore not be understated.

5.3.1 Objective- compared to human classification

The first aim of study was to examine precision and bias between objective and human classification. Paper I was the only paper to compare human classification to objective classification and as such is the only paper discussed in this section. The objective method functioned well enough to reasonably replace human classifiers for conformation. The results in support of the objective method functioning well were the R^2 values (0.92 for females and 0.93 for males) for the full model with conformation as the response, based on the training data set. For the test set, the results that supported the performance of the objective method compared to human classifiers were the low estimated bias ($\mu = 0.1056$), high correlation between predicted and observed conformation values (0.97), and the ICC ($\rho = 0.7428$) between the estimated prediction model variance ($\hat{\sigma}_t^2$) and the estimated human classifier variance ($\hat{\sigma}_u^2$). The ICC was >0.5, indicating that the prediction model was responsible for a larger part of the residual variance than the human classifiers, but the reason for this was the extremely low estimated human classifier variance (0.0985) compared to the estimated prediction model variance (0.2843). The high correlation between observed and predicted conformation grades were even slightly higher than comparative correlations based on VIA methods (Borggaard et al., 1996, Allen & Finnerty, 2000, Craigie et al., 2012, Wnęk et al., 2017). The bias showed that the model significantly underpredicted conformation compared to human classifiers, but the bias was low enough that a change from human classification to classification from the prediction model would not result in a large practical difference. In addition, the bias was low enough to be under the threshold for bias required by the European Commission for approval of automated methods.

The parameter estimates for EC requirements (European Commission, 2017) for approval of automated methods based on the test set also indicated that the predictive model for conformation functioned well. The estimated bias (0.12) was below the required threshold, the percentage points (69.3%) were above the required points and the slope (1.01) was close to one. The objective method in Paper I was even favourable in some aspects to the VIA method of classification by the EC approved BCC-3 (Larsen & Christensen, 2019); bias was lower (0.2 for BCC-3) and the slope was closer to one (1.09 for BCC-3). The BCC-3 did have noticeably higher percentage points (76.3%), and also functioned better for fat cover grading. Specifically, fat cover percentage points were much higher (54% vs. 74.3%), and slope was closer to 1 (0.78 vs. 0.91), though bias for Paper I was closer to 0 (-0.21 vs. 0.48).

The results of Paper I indicated that human classifiers are better at fat cover grading than the objective method. This was most apparent in the ICC for the full model with fat cover as the response (0.92). Unlike the ICC for conformation, the ICC for fat cover was not primarily due to low estimated human classifier variance ($\hat{\sigma}_u^2$). There was a 167% increase (0.0985 to 0.1643) in estimated human classifier variance ($\hat{\sigma}_u^2$) from conformation to fat cover, but a corresponding 664% increase (0.2843 to 1.8884) in estimated prediction model variance ($\hat{\sigma}_t^2$). Thus, it was the increase in estimated prediction model variance that was primarily the reason for the increased ICC for fat cover. The Pearson correlation between human classifier grading and predicted grading was also considerably lower for fat cover (0.61) compared to conformation (0.97). The lacking precision for fat cover grading indicates that the prediction variables in the model did not contain information that specifically relates to variation in fat, and that additional variables that can describe that variation is needed. This is potentially the reason for VIA being better suited for fat cover grading, as images of the carcase allow for a method of differentiation between meat and fat that the variables in our prediction model do not.

5.3.2 Classification and yield

The second aim of the thesis was to examine the relationship between classification and yield. Both Paper II and IV included an analysis of the relationship between classification and yield variables and will therefore be discussed in tandem. For both bovine and porcine carcases, yield was shown to be explained well by their respective classification grades, weight and gender/castration category. This was supported by the R^2 values for HVC (0.646), fat (0.639), and bone and waste (0.783) for bovine carcases, and the marginal R^2 values for lean meat (0.659) and fat (0.673), for porcine carcases. The R^2 values for bovine carcases were similar, though generally slightly lower, than previous literature (Drennan et al., 2008, Jørgenvåg et al., 2009, Conroy et al., 2010, Oliver et al., 2010), as was the case for porcine carcases compared to their respective previous literature (Bohrer et al., 2023, Goenaga, Lloveras & Améndola, 2008). Drennan et al. (2008) did have a higher R^2 value for “meat” and “fat” for bulls, but lower R^2 values for HVC and bone for bulls, and meat, fat, bone and HVC for Heifers. The R^2 values comparable to meat would be the R^2 related to the trimmings in the thesis. Meat in Drennan et al. (2008) included all trimmings, whereas Paper II separated trimmings by fat percentage, which is most likely the reason for the lower R^2 values in this thesis. This also applies to Oliver et al. (2010), who achieved a R^2 value of 0.93 for meat, but similar to Drennan et al. (2008) did not separate trimmings. Though both Bohrer et al. (2023) and Goenaga et al. (2008) achieved higher R^2 values for LMP, Goenaga et al. (2008) did not specify if the values were calculated for a training or test set, and the Bohrer et al. (2023) value was only higher for the test set (0.75), while the training set R^2 value (0.63) was slightly lower than in Paper II (0.659). Gender/castration category generally had a significant effect for both bovine and porcine carcases for all yield variables, though the effect was smaller than their respective classification grades.

The ability of classification to explain variation in lean meat-, fat- and bone and waste percentage is important, but to improve the ability of classification to function as a proxy for the potential value of the carcase, it is essential to include additional yield variables relevant to potential value that are specific to the genus of the carcase in question. Common for both bovine and porcine carcases is that the novel yield variables in Paper II and Paper IV were poorly explained by their respective statistical models, and consequently less well explained than HVC/lean meat-, fat-, and bone and waste percentage. For bovine carcases, the R^2 values for the trimmings ranged from 0.486 to 0.515. For porcine carcases, the marginal R^2 values ranged from 0.049 to 0.151 for PCPs. Though both trimmings and PCPs had lower R^2 values compared to their respective HVC/lean meat-, fat- and bone and waste percentages, it is important to note that the reasons for this are different. For bovine carcases, there is an overlap in the meat that is the basis for trimmings of 14% and 21% fat. This means that more of one product category reduces the amount of the other within the same carcase, which was illustrated by a strong negative estimated error correlation (-0.70) between meat14 and meat21. That is obviously not the case for PCPs, as it is not possible to create more of one PCP by simply categorizing weight from a different PCP as weight from the former.

5.3.3 Classification and confounding factors

The last aim of study was to examine the relationship between classification and confounding variables. Two confounding variables were considered: operator and breed. The effect of operator as a confounding variable on yield was negligible (Paper IV). Operator variance was small, both for lean meat-, fat- and bone and waste percentage (0.099 to 0.132), and for the PCPs (0.025 to 0.110). The $\mathrm{ICC}_{\mathrm{LMM(OP)}}$ for both lean meat-, fat-, and bone and waste (0.020 to 0.128) and the PCPs were small (0.027 to 0.129). This can be interpreted as operators systematically conducting the processing very similarly, all other variables considered constant. The proficiency of the operators can also be seen when compared to Nissen et al. (2006). The operator (referred to as butcher by Nissen et al.) standard deviation for LMP in this thesis (0.31) being less than half of the comparative LMP standard deviation (0.68) in Nissen et al. (2006). This could potentially be due to the operators in the Nissen et al. (2006) study being from different countries, whereas the operators in this thesis are not only from the same country but also work within the same facility and process carcases specifically to collect data for research. The response variables that had the highest estimated operator standard deviation in Paper IV were bone and waste- (0.36), loin- (0.30) and belly (0.33) percentage. This differed somewhat from Nissen et al. (2006) who found the highest standard deviation in the shoulder (0.15) and belly (0.13). The previously mentioned differences between the operators in Paper IV, and the butchers in Nissen et al. (2006) is mostly likely why loin had higher operator standard deviation compared to shoulder in Nissen et al. (2006). The reason for higher estimated standard deviation values in Paper IV is potentially due to Nissen et al. (2006) standard deviations being calculated for primal cut weights, as opposed to percentages.

Based on the results in this thesis, breed in general seems to be an important trait in regard to classification. Paper I showed that breed is especially important when predicting fat cover. The difference in estimated error variance between the full and reduced model for fat cover (0.45 for males and 1.15 for females) was larger than between the full and reduced model for conformation (0.25 for males and 0.12 for females). The reduction in precision from the full to the reduced model for conformation was small enough that not including breed information in prediction of conformation grades is feasible. Paper III did not involve carcase classification or yield, but was essential in facilitating the inclusion of breed data, and subsequent analyses in Paper IV. Paper II and IV showed breed has an effect on variation in yield outside of what can be attributed to classification. For both porcine and bovine carcases, breed had a significant effect on all yield variables. For bovine carcases, this means that some breeds are going to get paid more or less than what the value of the carcase is based on its yield. Finisher pigs in Norway are crossbreeds, and as such the same clear delineation cannot be applied as with bovine carcases, it is rather a case of degrees of difference; a finisher pig with more Duroc genetics will have higher expected lean than a finisher pig with more Z-line genetics, and therefore, all others variables considered equal, the more Duroc genetics a finisher pig has, the larger the discrepancy between the value of their yield and the actual price paid for the carcase. Breed had a significant effect on both the PCPs and lean-, fat- and bone and waste percentages, but generally had a lower effect on PCPs than the tissue percentages. Despite this, due to LMP and weight having a greater effect on lean- and fat percentages than PCPs, the effect of breed is arguably more important to PCP variation than tissue percentage variation. This is evident when looking at estimated unknown breed variance ($\widehat{\sigma}^2_{\mathrm{UB}}$) compared to fixed effect variance in Paper IV for tissue percentages compared to PCPs. For PCPs, the additional variance added based on not knowing breed ranged between 30% to 125%, whereas for lean- and fat percentage the additional variance added was only 6,4% and 6,5%, respectively.

5.4 [Identified gaps for future study / Implications of the study]

This thesis has shown that classification fulfils its function of being objective, and being a proxy for potential value of the carcase, but that there are several areas of improvement. A clear area of improvement based on the results of this thesis is the improvement of fat cover grade prediction based on the objective method in Paper I. There is a need for extracting additional information from the carcase that is relevant to fat grading to improve the precision of fat cover grading. Potential methods include information from near-infrared (NIR) spectroscopy (Chapman et al., 2020) or ultrasonography (Teixeira et al., 2006). Both these methods give information about subcutaneous tissue, which is lacking from the current objective method used for classification of bovine carcases today. NIR spectroscopy has an added benefit over ultrasonography in that it does not require active involvement of a human, whereas ultrasonography requires active involvement of humans. Better fat cover classification could potentially also reduce the difference in precision when breed information is not included in the prediction model.

Paper IV illustrated that PCPs are poorly explained by porcine classification variables, and an alternative method is needed for online estimation of PCPs. Janiszewski et al. (2018) leveraged information from the Autofom I to show high precision in predicting PCPs. The sample used in Janiszewski et al. (2018) was fairly small (n = 168), from a single producer, only including carcases of the same crossbreed combination. This means potential variance in yield that comes from breed or environmental differences does not apply to carcases in the sample. As data from Autofom is available in Norway as well, a potential future goal could be prediction of PCPs based on Autofom information applied on a larger, more varied sample than in Janiszewski et al. (2018).

5.5 [Conclusion]

Model prediction for conformation for bovine carcases functions sufficiently well to replace human classifiers, but there is need for further research to improve fat cover grade prediction. Classification does explain variation in yield fairly well, especially for LMY, but if classification is to be a better proxy for the potential value of a carcase, additional relevant yield variables and a method of predicting them are a necessity. Factors other than classification grades affect yield, but the degree to which they do so depends on the yield variable in question. Breed does affect yield, and future research should focus on how to account for the variation that comes from breed without the inclusion of breed information. The work included in this thesis is of importance to the field of classification because it, in conjunction with its associated papers, provides a detailed analysis of classification and variables important to classification, whose results can be of benefit to breeding companies, abattoirs and producers. The thesis also illuminates specific areas of improvement that would result in classification functioning as a better proxy for potential carcase value, and consequently an even fairer price per carcase.

6 References

Allen, P. & Finnerty, N. (2000). Objective beef carcass classification. A report of a trial of three VIA classification systems. *The Department of Agriculture, Food and Rural Development, and the National Food Centre, Teagasc. Dublin.* 15.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *J. Stat. Softw. 67*, 1-48

Bergen, R., Crews, D. H., Miller, J. S., & McKinnon, J. J. (2003). Predicting lean meat yield in beef cattle using ultrasonic muscle depth and width measurements. *Canadian journal of animal science*, *83*(3), 429-434. doi:10.4141/A02-093

Bohrer, B. M., Wang, Y., Dorleku, J. B., Campbell, C. P., & Mandell, I. B. (2023). An update of the predicted lean yield equation for the Destron PG-100 optical grading probe. *Journal of Animal Science*, *101*, skad199.

Borggaard, C., Madsen, N. T., & Thodberg, H. H. (1996). In-line image analysis in the slaughter industry, illustrated by beef carcass classification. *Meat Science*, *43*, 151-163. doi:10.1016/0309-1740(96)00062-9

Brøndum, J., Egebo, M., Agerskov, C., & Busk, H. (1998). On-line pork carcass grading with the Autofom ultrasound system. *Journal of Animal Science*, *76*(7), 1859-1868.

Chapman, J., Elbourne, A., Truong, V. K., & Cozzolino, D. (2020). Shining light into meat – a review on the recent advances in in vivo and carcass applications of near infrared spectroscopy. *International Journal of Food Science & Technology*, *55*(3), 935–941. doi: 10.1111/ijfs.14367

Choi, J. S., Kwon, K. M., Lee, Y. K., Joeng, J. U., Lee, K. O., Jin, S. K., Choi, Y.I. & Lee, J. J. (2018). Application of AutoFom III equipment for prediction of primal and commercial cut weight of Korean pig carcasses. *Asian-Australasian journal of animal sciences, 31*(10), 1670.

Conroy, S. B., Drennan, M. J., McGee, M., Keane, M. G., Kenny, D. A., & Berry, D. P. (2010). Predicting beef carcass meat, fat and bone proportions from carcass conformation and fat scores or hindquarter dissection. *Animal*, *4*(2), 234-241. doi:10.1017/S1751731109991121.

Craigie, C. R., Navajas, E. A., Purchas, R. W., Maltin, C. A., Bünger, L., Hoskin, S. O., ... & Roehe, R. (2012). A review of the development and use of video image analysis (VIA) for beef carcass evaluation as an alternative to the current EUROP system and other subjective systems. *Meat science*, *92*(4), 307-318. doi:10.1016/j.meatsci.2012.05.028

Dean, A., & Voss, D. (Eds.). (1999). *Design and analysis of experiments.* New York, NY: Springer New York. doi: 10.1007/978-3-319-52250-0

Delgado-Pando, G., Allen, P., Troy, D. J., & McDonnell, C. K. (2021). Objective carcass measurement technologies: Latest developments and future trends. *Trends in Food Science & Technology*, *111*, 771-782. doi:10.1016/j.tifs.2020.12.016.

Dorleku, J. B., Wormsbecher, L., Christensen, M., Campbell, C. P., Mandell, I. B., & Bohrer, B. M. (2023). Comparison of an advanced automated ultrasonic scanner

(AutoFom III) and a handheld optical probe (Destron PG-100) to determine lean yield in pork carcasses. *Journal of Animal Science*, *101*, skad058.

Drennan, M. J., McGee, M., & Keane, M. G. (2008). The value of muscular and skeletal scores in the live animal and carcass classification scores as indicators of carcass composition in cattle. *Animal*, *2*(5), 752-760. doi:10.1017/S1751731108001754.

Engel, B., Lambooij, E., Buist, W. G., & Vereijken, P. (2012). Lean meat prediction with HGP, CGM and CSB-Image-Meater, with prediction accuracy evaluated for different proportions of gilts, boars and castrated boars in the pig population. *Meat science*, *90*(2), 338-344.

European Commission. (2013). *Regulation (EU) No 1308/2013 of the European Parliament and of the Council of 17 December 2013 establishing a common organisation of the markets in agricultural products and repealing Council Regulations (EEC) No 922/72, (EEC) No 234/79, (EC) No 1037/2001 and (EC) No 1234/2007.* Retrieved 06 June 2023 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013R1308#d2046e32-801-1.

European Commission. (2017). *Commission Delegated Regulation (EU) 2017/1182 of 20 April 2017 supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards the Union scales for the classification of beef, pig and sheep carcasses and as regards the reporting of market prices of certain categories of carcasses and live animals.* Retrieved 08 June 2023 from https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32017R1182.

Font i Furnols, M., Gispert, M. (2009). Comparison of different devices for predicting the lean meat percentage of pig carcasses. *Meat science, 83*(3), 443-446.

Gangsei, L. E., Bjerke, F., Røe, M., & Alvseike, O. (2018). Monitoring lean meat percentage predictions from optical grading probes by a commercial cutting pattern. *Meat science*, *137*, 98-105.

Goenaga, P., Lloveras, M. R., & Améndola, C. (2008). Prediction of lean meat content in pork carcasses using the Hennessy Grading Probe and the Fat-O-Meater in Argentina. *Meat science*, *79*(3), 611-613.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.

Heggli, A., Gangsei, L. E., Røe, M., Alvseike, O., & Vinje, H. (2021). Objective carcass grading for bovine animals based on carcass length. *Acta Agriculturae Scandinavica, Section A—Animal Science*, *70*(2), 113-121. doi:10.1080/09064702.2021.1906940.

Janiszewski, P., Borzuta, K., Lisiak, D., Grześkowiak, E., & Stanisławski, D. (2018). Prediction of primal cuts by using an automatic ultrasonic device as a new method for estimating a pig-carcass slaughter and commercial value. *Animal Production Science*, *59*(6), 1183-1189.

Jørgenvåg, I. A., Røe, M., Malmo, T., Horsen, P. M., & Alvseike, O. (2009). Prediction of yield from carcass grading by EUROP and video image analysis. In *Proceedings of the 55th International Congress of Meat Science and Technology. Copenhagen, Denmark* (pp. 18-19). Retrieved 26 June 2022 from https://digicomst.ie/wp-content/uploads/2020/05/2009 03 04.pdf.

Kempster, A. J., Cuthbertson, A., & Harrington, G. (1982). The relationship between conformation and the yield and distribution of lean meat in the carcasses of British pigs, cattle and sheep: a review. *Meat Science*, *6*(1), 37-53. doi:10.1016/0309-1740(82)90049-3.

Kurkilahti, M., Appelberg, M., Hesthagen, T., & Rask, M. (2002). Effect of fish shape on gillnet selectivity: a study with Fulton's condition factor. *Fisheries research*, *54*(2), 153-170.

Langaker, M., Haug, D. J., Moen, Ø., Kjølseth, I.-M., and Kongsrud, S. (2021). *Ingris - årsstatistikk 2020* (in Norwegian). Available at: https://www.animalia.no/globalassets/ingris—dokumenter/arsstatistikk-2020-007—endelig.pdf.

Larsen, J., & Christensen, H. (2019). *Report on test results from BCC-3 in Holsted, Denmark 8-11 April 2019. Authorization test of BCC-3.* Copenhagen, Denmark: The Danish Classification Board for Pork, Beef and Sheep.

Marcoux, M., Bernier, J. F., & Pomar, C. (2003). Estimation of Canadian and European lean yields and composition of pig carcasses by dual-energy X-ray absorptiometry. *Meat Science*, *63*(3), 359-365.

Marcoux, M., Pomar, C., Faucitano, L., & Brodeur, C. (2007). The relationship between different pork carcass lean yield definitions and the market carcass value. *Meat Science*, *75*(1), 94-102.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press, London, England.

Mattilsynet. (2022). *Rapportering til Husdyrregisteret - Storfe | Mattilsynet.* Accessed 11 November 2023, available at: https://www.mattilsynet.no/dyr/produksjonsdyr/storfe/rapport-til-husdyrregisteret (In Norwegian)

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews*, *85*(4), 935-956.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution*, *4*(2), 133-142.

Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*(134), 20170213.

Nisbet, H., Lambe, N., Miller, G., Doeschl-Wilson, A., Barclay, D., Wheaton, A., & Duthie, C. A. (2024). Using in-abattoir 3-dimensional measurements from images of beef carcasses for the prediction of EUROP classification grade and carcass weight. *Meat Science*, *209*, 109391.

Nissen, P. M., Busk, H., Oksama, M., Seynaeve, M., Gispert, M., Walstra, P., Hansson, I. & Olsen, E. (2006). The estimated accuracy of the EU reference dissection method for pig carcass classification. *Meat Science, 73*(1), 22-28.

Nortura, (2023). *Angusavtale*. (In Norwegian) Retrieved 26 October 2023 from https://medlem.nortura.no/storfe/angusavtale/.

Oliver, A., Mendizabal, J. A., Ripoll, G., Albertí, P., & Purroy, A. (2010). Predicting meat yields and commercial meat cuts from carcasses of young bulls of Spanish breeds by the SEUROP method and an image analysis system. *Meat science*, *84*(4), 628-633. doi:10.1016/j.meatsci.2009.10.022.

Olsen, E. V., Christensen, L. B., & Nielsen, D. B. (2017). A review of computed tomography and manual dissection for calibration of devices for pig carcass classification-Evaluation of uncertainty. *Meat science*, *123*, 35-44.

Polkinghorne, R. J., & Thompson, J. M. (2010). Meat standards and grading: A world view. *Meat science*, *86*(1), 227-235.

Riordan, J. (2012). *Introduction to combinatorial analysis.* Courier Corporation. 6-7.

RStudio Team. (2020). RStudio (1.2.5033). RStudio: Integrated Development for R. http://www.rstudio.com

Røe, M. (2019). *Go'Mørning 0219 - Klassifiseringssystemet for storfe justeres fra nyttår*. Accessed 13 November 2023, available at: https://www.animalia.no/contentassets/119187080f9e422e819e301c583e8a51/gm-2-2019-web-v2.pdf (In Norwegian)

Røe, M. (2021). *301P Klassifisering av gris.* (In Norwegian) Retrieved 06 June 2023 from https://www.animalia.no/no/kjott--egg/klassifisering/klassifiseringshandboka/301p-klassifisering-av-gris-16/.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society: Series B (Methodological)*, *36*(2), 111-133.

Teixeira, A., Matos, S., Rodrigues, S., Delfa, R., & Cadavez, V. (2006). In vivo estimation of lamb carcass composition by real-time ultrasonography. *Meat Science*, 74(2), 289–295. doi: 10.1016/j.meatsci.2006.03.023

Tortereau, F., Servin, B., Frantz, L., Megens, H. J., Milan, D., Rohrer, G., & Groenen, M. A. (2012). A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. *BMC genomics*, *13*(1), 1-12.

UNECE Standards. (2016). *Bovine meat and carcass and cuts.* Retrieved 13 November 2023 from https://unece.org/fileadmin/DAM/trade/agr/standard/meat/e/Bovine_326Rev2E_2016.pdf

Walstra, P., & Merkus, G. S. (1996). *Procedure for assessment of the lean meat percentage as a consequence of the new EU reference dissection method in pig carcass classification: based on discussion in the EU Management Committee on Pig Meat and based on discussions with dissection experts during a meeting on May 18-19, 1994 at Zeist, NL* (No. 96.014). ID-DLO.

Wnęk, K., Gołębiewski, M., & Przysucha, T. (2017). Validation of the first objective evaluation system for beef carcasses. *Canadian Journal of Animal Science*, *98*(1), 53-60. doi: 10.1139/cjas-2016-0241

7 Papers I-IV

Paper I

ACTA AGRICULTURAE SCANDINAVICA, SECTION A — ANIMAL SCIENCE
2021, VOL. 70, NO. 2, 113–121
https://doi.org/10.1080/09064702.2021.1906940

Taylor & Francis
Taylor & Francis Group

Objective carcass grading for bovine animals based on carcass length

Andrew Heggli[a,b], Lars Erik Gangsei[a,b], Morten Røe[a], Ole Alvseike[a] and Hilde Vinje[b]

[a]Animalia, Oslo, Norway; [b]KBM, NMBU, Ås, Norway

ABSTRACT
The aim of the study was to evaluate performance for beef carcass grading using a novel semi-automatic method compared to human classifiers. The novelty was measuring carcass length as a predictor. The grading is given as conformation and fat cover as defined by the EUROP classification system. A training set was used to fit the model with predictors based on weight, age, breed and sex, in addition to length. Prediction performance was evaluated for a test set including carcasses graded by Norwegian classifiers, and a separate test set for carcasses graded by international classifiers. The precision for conformation was high (Pearson correlation ≥ 0.94) for both test sets, but the precision for fat cover was lower (Pearson correlation range 0.30–0.91). High correlation for conformation, together with low bias estimates, provides indication that the objective method is equipped to replace the previous human classifier system for conformation in Norway.

ARTICLE HISTORY
Received 8 December 2020
Accepted 16 March 2021

KEYWORDS
EUROP; beef; bovine animals; prediction; fat; conformation; carcass grading

Introduction

The proportion and content of lean meat, fat and bone differs in carcasses of bovine animals (cattle). Morphology, i.e. contours, are used as proxies and carcasses are graded based on this morphology. The classification strongly affects weight-based price setting of the carcass, and thus the financial settlement between livestock producer and abattoir. Classification also provides information to slaughterhouses that may be used to optimize carcass processing, and contributes with valuable feedback for livestock management. Finally, aggregated data from classification might constitute important feedback to cattle breeding companies. It is therefore important that classification is fair and consistent across producers and slaughterhouses, and across time and borders, both national and international.

In Europe, classification is conducted in accordance with the EUROP grid method (European Commission, 2013, 2017) which grades carcass conformation from E (Excellent) to P (Poor) based on muscle volume, and carcass fat cover from 1 (very thin fat layer) to 5 (very thick fat layer). For both conformation and fat cover, each grade can have a −/0/+ option, and for conformation the regulations open for including a sixth major class, S (Superior). In Norway, the −/0/+ option is used, but not the S class for conformation. Consequently, there are 15 groups for conformation and 15 groups for fat cover.

The most common classification method is to use certified classifiers who visually judge each carcass and assign grades for conformation and fat cover based on muscle volume and fat deposits. This method of classification is subjective as humans are solely responsible for classifying each carcass.

The potential for inconsistency of the classifiers is a widely recognized shortcoming of visual classification, although hard evidence for this is scarce (Craigie et al., 2012). Thus, objective grading is preferable. In addition, such indubitable processes would minimize grounds for contest between parties and countries and create incentives for continued trade between and within European countries. Objective methods are often automatic or semi-automatic, which might also reduce costs associated with carcass grading considerably.

Grading systems based on video image analysis (VIA) have been present since the 1980s for bovine carcasses (Craigie et al., 2012). The most prevalent systems in Europe are the VBS 2000 system from the German manufacturer E + V GmbH, the BCC-2/ BCC-3 systems from the Danish manufacturer FrontMatec and the MAC-2/ MAC-S systems from the French manufacturer Normaclass. These systems have been evaluated in different studies (Borggaard et al., 1996; Allen & Finnerty, 2000; Craigie et al., 2012; Wnęk et al., 2017). The European

CONTACT Andrew Heggli andrew.heggli@animalia.no Animalia, Lørenveien 38, Oslo, Norway KBM, NMBU, Ås, Norway

Commission has made detailed instructions for how to authorize automated grading methods for beef carcasses in EU (European Commission, 2017). FrontMatecs BCC-3 is authorized under these regulations for use in Denmark (Larsen & Christensen, 2019).

In Norway, all carcasses from bovine animals were graded by human classifiers until January 2019, when the method for classifying EUROP conformation, but not fat cover, was replaced with a new semi-automatic method, where the innovation was measuring carcass length (Animalia, 2020). The instruments used for length measurements are cheap compared to equipment for VIA methods. The length data are used in an algorithm along with weight, breed, age and sex as predictors for carcass conformation. The carcass grading is undertaken by slaughterhouse employees who are certified and regularly controlled by experts from Animalia. The algorithm used in this study is a modified version of the algorithm used in Norwegian slaughterhouses post January 2019.

The aim of the present study is to evaluate the performance of carcass grading for bovine animals, both conformation and fat cover, based on a novel objective method using carcass weight, length, age, sex and breed information as predictors. The performance of the objective method is compared to human grading by a methodology that utilizes repeated gradings of the same carcass. This method is implementable on datasets sampled in accordance with EU regulations for the approval of automated grading methods (European Commission, 2017).

Materials and methods

Material

Data from a total of 3825 beef carcasses were collected from January 2018 to March 2020 at 24 slaughterhouses across Norway. The data used in the algorithm were sampled from three sources for all individual carcasses, (i) measurements at the slaughterhouses, (ii) classification by Animalia's expert group and (iii) the animal livestock registry (Mattilsynet, 2016), a nationwide register of all bovine animals. These data are mostly entered by livestock producers. Results and measurements from classification in all slaughterhouses were transferred to Animalia's databases daily.

From the slaughterhouses each carcass was assigned a unique ID and data were collected for, length (cm), weight of warm carcass (kg) adjusted by a factor of 0.98 and date of slaughtering. Simultaneously as the carcass was weighed by a calibrated and officially controlled scale, the length was measured by a manually operated laser (item number MTS04700412 Meats AS, Bergen, Norway) that shot a beam perpendicular to the carcass. The length was defined as the distance from cranial edge of the first thoracic spinous process to the lowest internal point of the hooks (fixed) where the carcass was hung from the achilleas tendon (Figure 1).

All carcasses in the dataset were graded by one or more Animalia expert classifiers, a group consisting of five persons, all assigned a unique ID. Animalia experts registered results for EUROP conformation (integer 1–15, where P– equals 1, E+ equals 15) and fat cover (integer 1–15, where 1– equals 1 and 5+ equals 15). For conformation separate grades were assigned separately for back-, middle- and fore-parts, then an arithmetic mean was applied. Most carcasses were graded by one Animalia expert, but some were graded by two up to all five, disparately. Finally, from the animal livestock registry, registrations of age (days), sex (male, female or castrate) and breed based on pedigree back to great-great-grandparents were collected. A total of 69 individuals were omitted from further analysis, as they were either castrates or their sex was unknown, leaving a dataset with data from a total of 3756 beef carcasses.

Figure 1. Measurement of a carcass. The measuring starting point is the cranial edge of the first thoracic spinous process (green light on the carcass) to the bottom of the internal curvature of the hook from which the carcass hangs.

In addition to the Animalia dataset based on Animalia expert classifiers, a dataset containing data from 10 carcasses, that were graded by groups of international classifiers at a time, rather than each classifier individually, was considered. These data, including data used as predictors, i.e. length, weight, breed, sex and age, were collected in conjunction with a meeting to calibrate the classifying skills of classifiers from the Nordic countries held in Målselv in Norway during September 2019.

Method

The data were divided into a training set ($n = 3622$) and a test set ($n = 124$), in addition, the international dataset from the meeting with Nordic country classifiers was used as a separate test set ($n = 10$). The training data consisted of carcasses graded by one Animalia classifier only. The international test set consisted of carcasses from the Nordic calibration meeting, and the national test set consisted of the carcasses that had been graded by two or more Animalia classifiers. This granted the benefit of testing the model's validity in regard to precision and biases.

The following linear regression model was fitted to the training data set:

$$\begin{aligned} Y = {} & \beta_0 + \beta_k x_k + \boldsymbol{\beta}_b^t \boldsymbol{x}_b + \boldsymbol{\beta}_w^t \boldsymbol{z}_w + \boldsymbol{\beta}_l^t \boldsymbol{z}_l + \boldsymbol{\beta}_a^t \boldsymbol{z}_a + e, \\ & e \sim N(0,\ \sigma_e^2), \end{aligned} \tag{1}$$

where the response, Y, denotes either the mean (back-, middle- and fore-part) conformation, or fat cover as graded by Animalia experts. The intercept term is given by β_0 and the random error, e, is associated with variance σ_e^2. The unit of Y is EUROP conformation and EUROP fat group, hence the unit of σ_e^2 is square of the corresponding measurements.

The predictor variable x_k is the relationship between weight (kg) and cubic of length (m), known as Fulton's condition factor (Kurkilahti et al., 2002). Fulton's condition factor is invariant with size of a 3D object assuming that shape and density are kept fixed. The vector $\boldsymbol{x}_b$ represents breed groups ($n = 9$) which were defined a priori, representing the 35 different breeds registered in the animal livestock registry. An overview of the distribution of breeds in the breed groups is given in Røe (2019). Breed populations vary in size, and some breeds would not have had enough individuals to be included in the analyses if treated as a single group. Thus, we chose to aggregate breeds into breed groups based on shared qualitative traits, as to represent the variance that exists across the spectrum of breeds in Norway. The elements of $\boldsymbol{x}_b$ are positive 1/16 fractions and sum to 1. The vectors $\boldsymbol{z}_w$, $\boldsymbol{z}_l$ and $\boldsymbol{z}_a$ of length 5, 5 and 7, respectively, are constructed by evaluating weight (kg), length (cm) and the natural logarithm of age (days) as squared B-spline functions using knot sequences {100,275,450}, {150,190,230} and the natural logarithm of {100,300,730,1460,3650}, respectively. All **z**'s have positive elements which sum to one. The predictors weight, length and age were included because they all have an effect on body composition, and thereby, conformation score. Weight and length knots were chosen to have equal space between each knot, and so that the minimum and maximum knots were within the minimum and maximum weight and length in the dataset. The knots for age were chosen to correspond to the thresholds between categories of cattle, for example the threshold between a young bull and a bull is 730 days of age. The model (Equation (1)) was fitted to males and females separately. In order to achieve a prediction matrix of full rank a restriction is applied so that the elements of β_b, β_w, β_l and β_a, all sum to zero.

A subset of $2^4 = 16$ models where all combinations of including or excluding the predictor variables breed ($\boldsymbol{x}_b$), weight ($\boldsymbol{z}_w$), length ($\boldsymbol{z}_l$) and age ($\boldsymbol{z}_a$) were tested.

Akaike information criterion (AIC) (Akaike, 1974) was calculated for all models in the subset and used for model selection, as the model with lowest AIC was chosen. The procedure was conducted for both conformation and fat cover as the response, and males and females separately. This resulted in four models, i.e. models for male conformation, male fat cover and female equivalents.

Curating a pedigree can be taxing in regard to economic strain and is susceptible to mistakes in regard to parentage. Therefore, omitting breed as a predictor would be beneficial, barring negative effects on the precision of the model. Thus, reduced models excluding breed, for the four combinations of response and sex, were also evaluated. The models including breed are referred to as the full model. The models omitting breed are referred to as reduced models.

In order to evaluate prediction precision and prediction bias, we applied the random effect model:

$$y'_{ij} = \mu + t_i + u_{ij},\ \ t_i \sim N(0,\ \sigma_t^2),\ \ \ u_{ij} \sim N(0,\ \sigma_u^2), \tag{2}$$

to the two test datasets. Here, y'_{ij}, represents the difference between conformation/fat cover set by classifier j for individual i (y_{ij}) and the associated predicted value ($\hat{y}_i$), based on the fitted model of Equation (1).

The main effect, μ, represents bias of the prediction algorithm. If the model is biased towards low predictions μ is larger than 0 and vice versa. A bias equal to, or close to zero is desired.

Referring to Equation (2), the total random error for individual i graded by classifier j is given by $t_i + u_{ij}$. t_i denotes the random classification error caused by the prediction model for individual i, with corresponding variance σ_t^2. A small variance is associated with high precision for the prediction algorithm. u_{ij} is the random error caused by classifier j for carcass i, with corresponding variance σ_u^2.

The intraclass correlation coefficient, $\mathrm{corr}(y_{ij}, y_{ik}) = \rho = \sigma_t^2 \ / \ \sigma_t^2 + \ \sigma_u^2$ (Dean et al., 2017) is the proportion of total misclassification explained by the prediction algorithm. If $\rho > 0.5$, then human experts are more precise than the prediction algorithm and vice versa. Parameter estimates, and confidence intervals for μ and ρ were calculated in accordance with Dean et al. (2017). For unbalanced data, i.e. the national test set where number of classifiers per carcass varies, confidence intervals regarding μ are approximate.

Detailed rules for how to authorize automated grading methods for beef carcasses are given in annex IV of supplementing Regulation (EU) No 1308/2013 (European Commission, 2017). Under these regulations, points are reduced or given based on the disparity or lack thereof between automatic grading results and median human classifier results. For a new grading method to be accepted it must achieve 60% of maximal points on a test set consisting of at least 600 individuals (European Commission, 2017). Furthermore, the bias must be within ±0.30 and ±0.60 for conformation and fat cover, respectively. Finally, the slope of regression of median grading by classifier and automated grading must be in range 1 ± 0.15 and 1 ± 0.30 for conformation and fat cover, respectively.

In order to indicate how the method would do against the EU system's formal requirements, % of maximal points, bias and slope of regression for all model combinations applied to both test sets were calculated. For these calculations, results rounded to closest integer for both classifier median and automatic grading were used.

Data availability statement

The metadata and R-code used in the present study are freely accessible at https://doi.org/10.18710/T9SXVF. Access to the dataset with data used in the statistical analyses requires approval from Animalia AS. All statistical analyses were coded and executed in the software R (RStudio Team, 2020).

Results

Table 1 contains key statistics for the training and test data sets. The most notable difference between datasets are their sizes. In addition, the international test set is balanced, i.e. all classifiers have graded all individuals, whereas the national test set is unbalanced. Furthermore, mean conformation and fat were higher for the national test set than the training set, for both males and females. For males the standard deviations were higher for the national test set compared to the training set, for females vice versa.

When fitting the models in Equation (1), three out of the four combinations for response and sex scored the smallest AIC values when all predictor variables were included. The exception was conformation for males, where the model with lowest AIC value corresponded to the full model excluding B-splines based on length as a predictor. Three out of four combinations for the reduced models, i.e. where breed was excluded a priori, scored the lowest AIC when all other predictor variables were included. The exception was the reduced model for males with fat cover as response, where B-splines based on length was excluded as a predictor.

Based on results from training data, R^2 values, i.e. the proportion of variance explained by the models, for conformation were high (0.88–0.93), with corresponding RMSE values for conformation in the range of 0.58–0.77. R^2 full model values for fat were 0.47 and 0.65 for males and females respectively, while reduced model values for fat were 0.28 and 0.49 for males and females respectively. RMSE per breed group for full models had a range of 0.30–0.42 for conformation and 1.49–2.90 for fat. RMSE per breed group for reduced models had a range of 0.39–0.92 for conformation and 1.83–9.45 for fat.

Estimated error variances ($\hat{\sigma}_e^2$) (based on Equation (1)) ranged from 0.34 to 0.59 for conformation, and from 1.25 to 3.57 for fat. The difference between full and reduced model estimated error variances, i.e. an expression of loss when breed is excluded as a predictor, was larger between fat models (0.45 for males and 1.15 for females) than between conformation models (0.25 for males and 0.12 for females). Estimated error variances for fat were lower for males than females, indicating a somewhat better model fit for males.

The correlations between predicted values and observations for the two independent test sets are visualized in Figure 2. Pearson correlations for conformation were all high, 0.97 (national) and 0.96 (international) applying the full model, and 0.94 and 0.94 with the reduced model. For fat cover the correlations were much lower for the national test set, 0.61 and 0.30 for the full and reduced models. For the international test set correlations were high (0.91 and 0.87), which might be an artefact of small sample size ($n = 10$).

Table 1. Descriptive statistics for the training and two test datasets, split into male and female subsets. The table gives the number of unique observations and unique individuals. Furthermore, mean values ± standard deviations are given for conformation, fat cover, weight, length and age.

Set	Sex	Observations (*n*)	Individuals (*n*)	Conformation (EUROP)	Fat cover (EUROP)	Weight (kg)	Length (cm)	Age (days)
Training	Male	2140	2140	6.56 ± 2.20	6.90 ± 1.54	323 ± 80	204 ± 14	541 ± 174
	Female	1482	1482	4.59 ± 2.12	7.67 ± 2.63	263 ± 72	207 ± 18	1400 ± 959
Test national	Male	187	86	7.58 ± 2.25	7.41 ± 1.71	342 ± 64	202 ± 12	523 ± 96
	Female	76	38	5.19 ± 1.94	8.14 ± 1.87	302 ± 54	211 ± 11	2149 ± 1173
Test international	Male	35	7	6.63 ± 1.26	7.83 ± 2.12	316 ± 69	203 ± 12	565 ± 57
	Female	15	3	3.33 ± 0.50	8.40 ± 2.03	277 ± 34	220 ± 10	1346 ± 465

Table 2 shows point estimates and confidence intervals based on Equation (2) when applied to the two test datasets. In general point estimates for biasedness are fairly close to zero in the national test set, absolute numbers less than 0.11 for conformation and less than 0.2 for fat. Corresponding 95% confidence intervals for prediction bias (μ) embrace 0, except for the full model with conformation as a response, indicating a significant bias for this model. Bias estimates for the international test set were consistently higher than for the national test set, and only two out of four 95% confidence intervals spanned 0.

The estimate for human classifier variance (σ_u^2) is solely a function of the observations and independent of the predictions. Consequently, these estimates are the same for both the full and reduced model. Estimated intraclass correlation coefficients (ρ) were consistently higher for the national test set than the international test set. None of the confidence intervals for ρ in the national test set spanned 0.5, while three out of four confidence intervals for ρ in the international test set embraced 0.5. Full models had consistently lower correlation estimates than their reduced counterpart, as the estimated prediction algorithm variance (σ_t^2) was lower for the full model compared to the reduced model both for conformation and fat.

Discussion

The most important results demonstrating the proficiency of the algorithm are the high correlation, and low level of bias between predicted and observed values for conformation, shown in Figure 2. This indicates that the majority of individual carcass grades will not differ drastically from what would have been given by human classifiers. The algorithm was even precise for carcasses at the furthest ends of the EUROP range, though this is a less certain result due to few observations. The precision in predicting conformation was also reflected in the high amount of variation explained by the conformation models ($R^2 \geq 0.88$), and low estimated error variances (0.34–0.59) and low RMSE per breed group (0.30–0.92), which showed that the predictor variables chosen were well suited to predict conformation.

Results indicated low bias for conformation prediction. This is essential when it comes to the fairness and longevity of the algorithm. A producer might experience lower or higher settlements for individual carcasses than if classed by a human, due to random error. However, when the bias is low, these errors will cancel each other out over time, and the total amount paid to that producer would be close to or potentially the same as if humans were still classifying.

Though bias estimates for conformation for the test sets were low (range −0.05–0.31), the confidence intervals for the full model conformation bias did not embrace 0 and indicated a significant positive bias. This implies that the full model underpredicts conformation class. Nevertheless, point estimates for bias were low (0.11 for the national test set and 0.31 for the international) and consequently, the undesired significant bias for conformation using the full model does not reflect a large practical effect.

The precision of the full model for conformation was very high, which is reflected in narrow C.I.'s for bias parameter μ, with width 0.20 and 0.51 for the national and international test sets respectively. For the reduced model precision was slightly lower, but still high, yielding marginally wider C.I.'s for bias, which for both the national and international test set embraced 0. These results strongly support that the algorithm is well equipped to replace the old system for conformation, at least in Norway. For conformation, the reduced model seems to be a suitable candidate as its precision is only marginally lower than the corresponding full model precision.

Fat cover bias estimates for the national test set for both the full and reduced models were close to zero, with C.I.s that embraced zero. However, for the international test set bias point estimates were high, and the C.I. for the reduced model did not embrace 0. Nevertheless, results show some clear patterns for fat prediction. First and foremost, prediction precision for fat is much lower than for conformation. This is reflected in lower R^2-values in the training data set and higher

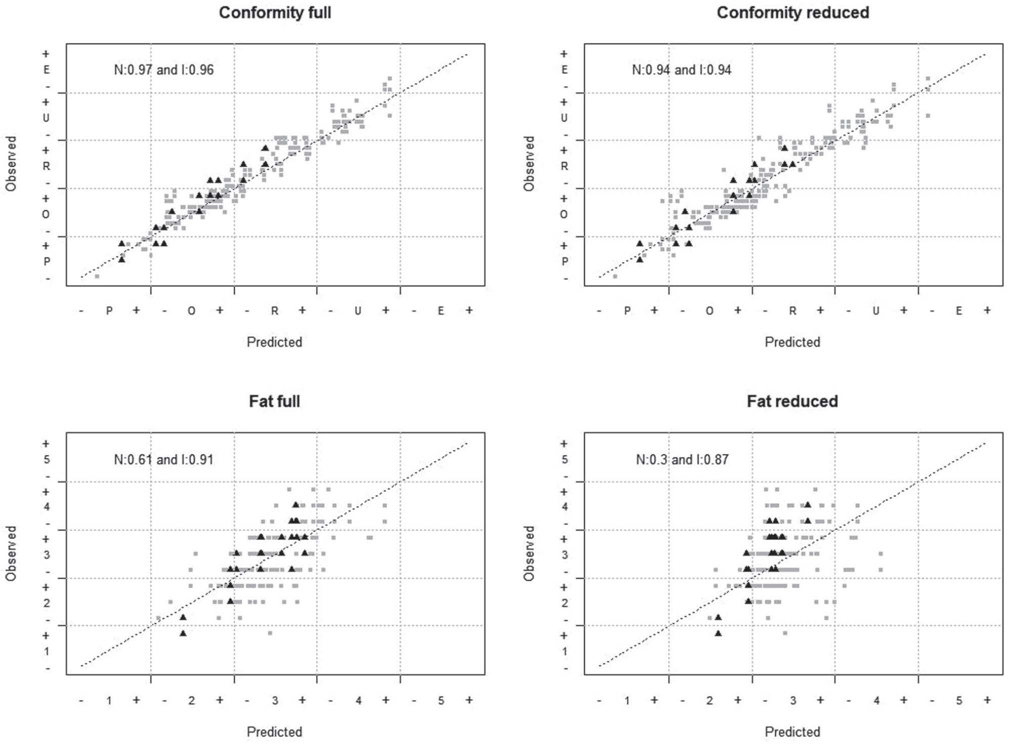


Figure 2. Predicted versus observed classifications for conformation (upper panels) and fat (lower panels) based on full (left panels) and reduced models (right panels). The dotted line represents equal predicted and observed values. Grey squares are values from the national test set and black triangles are values from the international test set. Pearson correlations for the corresponding national test set (N) and international test set (I) are given in the upper left corner within each panel.

estimated prediction error variances (σ_t^2) in the test sets. RMSE per breed group was also higher for fat than conformation, with a wide range (1.49–9.45). The wide range is due to one breed group whose RMSE (9.45) was 5.6 points above the breed group with the second highest RMSE (3.82). The highest RMSE is due to a combination of low sample size for that breed group, and the low prediction precision of the reduced model for fat.

Norwegian pedigree is highly reliable, with an error of only 3.3% when genotyping the most common breed in Norway, Norwegian Red (Larsgard et al., 2019). Therefore, including breed as a predictor in Norway gains the advantage of increased precision, while suffering little the consequence of faulty pedigrees. Breed is still an unfavourable predictor as, in addition to breed being costly to register, not all countries have detailed pedigrees available. Excluding breed would therefore make exportation of the new system more feasible. Moreover, omitting breed as a predictor is advantageous even if breed information is available, as excluding breed would negate problems associated with individuals where pedigree is not known, or where pedigree information is wrong. For these reasons, the decline in performance if breed is omitted as a predictor was evaluated through the reduced models.

Breed seems to be a more important predictor for fat cover than conformation, especially in females. This is reflected in substantially lower R^2-values for fat prediction in the training data set and higher estimated prediction error variance in the test sets when comparing males to females, and when comparing reduced to full models. This observation is not particularly surprising as some breeds are known to have unfavourable mean fat cover grading compared to their conformity grading (Hickey et al., 2007).

Bias estimates would indicate that the algorithm is possible to use for fat cover prediction as well as conformation, but Figure 2 shows that predicted values deviate from the observed considerably. Though an unbiased algorithm over time would ensure equal compensation

Table 2. Estimated population mean (μ), human classifier variance (σ_u^2), algorithm variance (σ_t^2) and intraclass correlation (ρ) based on Equation (2). Lower (C.I. 2.5%) and upper (C.I. 97.5%) confidence intervals are given for population mean and intraclass correlation. Finally, the % of maximum achievable points (1240 and 100 points for the national and international test sets respectively), bias and slope of regression line as defined in annex IV of supplementing Regulation (EU) No 1308/2013 (European Commission, 2017) is given. All estimates and confidence intervals are given for the four model combinations defined by conformation or fat cover vs. full (F) or reduced (R) models.

Data set	Parameter		Conformation (F)	Conformation (R)	Fat (F)	Fat (R)
National	μ	Estimate	0.1056	−0.0504	−0.1779	0.0057
		C.I. 2.5%	0.0035	−0.2006	−0.4268	−0.3079
		C.I. 97.5%	0.2077	0.0998	0.0710	0.3194
	σ_u^2	Estimate	0.0985	0.0985	0.1643	0.1643
	σ_t^2	Estimate	0.2843	0.6692	1.8884	3.0440
	P	Estimate	0.7428	0.8718	0.9200	0.9488
		C.I. 2.5%	0.6567	0.8242	0.8892	0.9287
		C.I. 97.5%	0.8106	0.9075	0.9427	0.9635
	EU regulations	% Points	69.3	54.8	54.0	31.4
		Bias	0.1210	−0.0645	−0.2097	0.0403
		Slope	1.0130	0.9886	0.7833	0.4300
	RMSEP	Estimate	0.6232	0.8739	1.4410	1.7843
International	μ	Estimate	0.3097	0.1654	0.4075	0.9419
		C.I. 2.5%	0.0556	−0.1879	−0.0926	0.1524
		C.I. 97.5%	0.5638	0.5187	0.9077	1.7313
	σ_u^2	Estimate	0.1900	0.1900	0.5000	0.5000
	σ_t^2	Estimate	0.1200	0.2676	0.5124	1.4258
	P	Estimate	0.3872	0.5848	0.5061	0.7404
		C.I. 2.5%	0.1222	0.3132	0.2305	0.5109
		C.I. 97.5%	0.7309	0.8447	0.8037	0.9130
	EU regulations	% Points	88.0	84.0	84.0	38.0
		Bias	0.1000	0.0000	0.4000	1.0000
		Slope	0.9012	0.8841	1.4674	1.5833
	RMSEP	Estimate	0.6246	0.6741	1.0570	1.6311

as human classifiers, a $\rho > 0.90$ (Table 2) shows that human classifiers are much better equipped to determine fat cover for now. This result was not surprising as none of the predictor variables used in this study reflect the amount of fat deposits a carcass has.

The VIA methods based on image techniques have an advantage, as they gain information about the fat cover from the image processing, and this is reflected in their correlations for fat prediction being higher than ours (Borggaard et al., 1996; Allen & Finnerty, 2000; Craigie et al., 2012; Wnęk et al., 2017). However, this was not the case for correlation between predicted and observed values for conformation; our conformation correlation for the full model for the national test set was marginally higher than all the VIA methods. Our results showed a small bias, which was also the case for the BCC-2 and the VIAscan (Allen & Finnerty, 2000), though the method of quantifying bias was different.

The objective, fully automatic, VIA methods based on image techniques are suitable for larger slaughterhouses that require faster processing, and where the purchase cost of classification equipment matters less. For smaller slaughterhouses the extra time allotted to perform the length measurement is justified as each individual carcass has high monetary value, and the purchase cost of the equipment is relatively low. Hence, for smaller slaughterhouses objective methods based on measuring carcass length are more accessible than previous VIA-based objective classification methods.

The test data sets used in this study did not meet the formal requirements for EU regulation for automatic grading regarding sample size (requirement $n = 600$), or number of independent classifiers per slaughter ($n = 5$). The results (Table 2) strongly indicate that objective grading of conformation based on carcass length would meet formal requirements for EU approval if breed is included as predictor, as bias, percentage points and slope were within the limits noted by EU regulation (European Commission, 2017). If breed is omitted as predictor results for percentage points (54.8% where 60% is the limit) indicate that approval of the method for conformation might be hard to achieve. Results for slope and bias are well inside requirements. Still, it might be possible to achieve approval, partly as an effect of less error variation associated with the human classifiers when number of classifiers are increased to five per carcass.

The good performance of the objective method based on length measurement is further supported by comparing conformation results to the BCC-3 (Larsen & Christensen, 2019); bias was lower (0.2 for BCC-3) and slope was closer to one (1.09 for BCC-3) for the method based on length measurement than BCC-3. On the other hand, the percentage points for BCC-3 at 76.3% were substantially higher than for the length method 69.3% (full model) and 54.8% (reduced model).

The good results for conformation are based on the 15 point system applied in Norway. As our range of carcass conformation spans the same as a 5 point system, which just omits the +/0/– option, it is likely that analysing our results at a 5 point scale would have little effect on conclusions, in a similar manner to the VIA systems (Allen & Finnerty, 2000). As the S class is not included in Norway, we have not taken this class into consideration. However, results indicate that the method functions well at the fringes (P and E) of the conformation scale, indicating that inclusion of S class should be achievable.

Our results strongly indicate that approval for fat cover grading is not achievable for the method based on length measurement within EU regulations, as none of the results are within the constraints for all three requirements, i.e. bias, percentage points and slope. To increase the precision for fat cover prediction an additional measuring tool or technique would be beneficial. Additional information from near-infrared (NIR) spectroscopy (Chapman et al., 2020) or ultrasonography (Teixeira et al., 2006) could potentially serve the purpose. Such methods have the potential benefit, compared to VIA and humans, of penetrating the soft tissue, in order to give better predictions for total lean/fat ratio of carcasses.

If semi-automatic methods of determining both fat group and conformation scores were shown to be precise, a natural next step would be to test if the predictor variables used to predict fat group and conformation could be used to predict monetary carcass value. The relationship between carcass value and predictor variables was outside the scope of this study as the purpose of this study was to evaluate prediction precision for conformation and fat group scores.

Castrates were excluded from statistical analyses as there is a biological distinction between the castrates and entire animals. If castrates were to be included, they should be treated as a third sex, which was unachievable in the present study due to lack of data. Furthermore, in 2020 castrates only make up 0.6% (Animalia internal data) of the total population of slaughtered cattle ($n = 295{,}002$) in Norway.

Due to the international data being sourced from a conference, the sample size of the international test set ($n = 10$) was small, but the results were included as they shed light on the important issue regarding systematic differences between classifiers in different countries.

Conclusion

The main purpose of this study was to evaluate if, and how well, a semi-automatic grading method utilizing carcass length measurements could predict EUROP classes compared to human counterparts. The results of this study show that, for conformation, the algorithm functions at a level that is sufficient to replace the older, more subjective system with human classifiers. As of today, the method is not suitable for fat cover grading. Additional measuring methods, like NIR or ultrasound are candidates as additional information sources in order to develop beef carcass grading method that is a full-fledged, semi-automatic, objective and more affordable alternative to VIA methods.

Acknowledgements

This work was funded by the Research Council of Norway under grant number 311394 and Animalia AS as a part of an industrial doctorate degree for Andrew Heggli. The authors acknowledge Nortura SA, the abattoirs, Animalia AS and the Classification board.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Research Council of Norway [grant number 311394].

ORCID

Hilde Vinje http://orcid.org/0000-0003-0318-0214

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. doi:10.1109/TAC.1974.1100705

Allen, P. & Finnerty, N. (2000). *Objective beef carcass classification*. A report of a trial of three VIA classification systems. Dublin, Ireland: The Department of Agriculture, Food and Rural Development, and the National Food Centre, Teagasc.

Animalia. (2020). Klassifiseringshåndboka. Accessed 23 November 2020, available at: https://www.animalia.no/no/kjott–egg/klassifisering/klassifiseringshandboka/ (In Norwegian).

Borggaard, C., Madsen, N. T. & Thodberg, H. H. (1996). In-line image analysis in the slaughter industry, illustrated by beef carcass classification. *Meat Science*, 43, 151–163. doi:10.1016/0309-1740(96)00062-9

Chapman, J., Elbourne, A., Truong, V. K. & Cozzolino, D. (2020). Shining light into meat – a review on the recent advances in in vivo and carcass applications of near infrared spectroscopy. *International Journal of Food Science & Technology*, 55(3), 935–941. doi:10.1111/ijfs.14367

Craigie, C. R., Navajas, E. A., Purchas, R. W., Maltin, C. A., Bünger, L., Hoskin, S. O., Ross, D. W., Morris, S. T. & Roehe, R. (2012). A review of the development and use of video image analysis (VIA) for beef carcass evaluation as an alternative to the current EUROP system and other subjective systems. *Meat Science*, 92(4), 307–318. doi:10.1016/j.meatsci.2012.05.028

Dean, A., Voss, D. & Draguljić, D. (2017). Design and Analysis of Experiments (New York: Springer International Publishing). doi:10.1007/978-3-319-52250-0.

European Commission. (2013). *Regulation (EU) No 1308/2013 of the European Parliament and of the Council of 17 December 2013 establishing a common organisation of the markets in agricultural products and repealing Council Regulations (EEC) No 922/72, (EEC) No 234/79, (EC) No 1037/*2001. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:02013R1308-20190101.

European Commission. (2017). Supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards the Union scales for the classification of beef, pig and sheep carcasses and as regards the reporting of market prices of certain categories of carcasses a. https://eur-lex.europa.eu/legal-content/SV/ALL/?uri=CELEX%3A32017R1182.

Hickey, J. M., Keane, M. G., Kenny, D. A., Cromie, A. R. & Veerkamp, R. F. (2007). Genetic parameters for EUROP carcass traits within different groups of cattle in Ireland1. *Journal of Animal Science*, 85(2), 314–3321. doi:10.2527/jas.2006-263

Kurkilahti, M., Appelberg, M., Hesthagen, T. & Rask, M. (2002). Effect of fish shape on gillnet selectivity: a study with Fulton's condition factor. *Fisheries Research*, 54(2), 153–170.

Larsen, J. & Christensen, H. (2019). *Report on test results from BCC-3 in Holsted, Denmark 8-11 April 2019. Authorization test of BCC-3*. Copenhagen, Denmark: The Danish Classification Board for Pork, Beef and Sheep.

Larsgard, A. G., Holtsmark, M. & Storlien, H. (2019). Buskap – Retting av slektskap i kukontrollen. Accessed 08. February 2021, available at: https://www.buskap.no/asset/2019/buskap-2019-07.pdf (In Norwegian).

Mattilsynet. (2016). Rapportering til Husdyrregisteret – Storfe | Mattilsynet. Accessed 23 November 2020, available at: https://www.mattilsynet.no/dyr_og_dyrehold/produksjonsdyr/merking_og_registrering_av_produksjonsdyr/rapportering_til_husdyrregisteret__storfe.4942 (In Norwegian).

Røe, M. (2019). Go'Mørning 0219 – Klassifiseringssystemet for storfe justeres fra nyttår. Accessed 08. February 2021, available at: https://www.animalia.no/contentassets/119187080f9e422e819e301c583e8a51/gm-2-2019-web-v2.pdf (In Norwegian).

RStudio Team. (2020). *RStudio* (1.2.5033). RStudio: Integrated Development for R. http://www.rstudio.com/.

Teixeira, A., Matos, S., Rodrigues, S., Delfa, R. & Cadavez, V. (2006). In vivo estimation of lamb carcass composition by real-time ultrasonography. *Meat Science*, 74(2), 289–295. doi:10.1016/j.meatsci.2006.03.023

Wnęk, K., Gołębiewski, M. & Przysucha, T. (2017). Validation of the first objective evaluation system for beef carcasses. *Canadian Journal of Animal Science*, 98(1), 53–60. doi: 10.1139/cjas-2016-0241

Paper II

Animal 17 (2023) 100854

Contents lists available at ScienceDirect

Animal

The international journal of animal biosciences

Carcase grading reflects the variation in beef yield – a multivariate method for exploring the relationship between beef yield and carcase traits

A. Heggli [a,b,*], O. Alvseike [b], F. Bjerke [b], L.E. Gangsei [a,b], J. Kongsro [b], M. Røe [b], H. Vinje [a]

[a] *Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway*
[b] *Animalia, P.O. Box 396 - Økern, NO-0513 Oslo, Norway*

ARTICLE INFO

Article history:
Received 17 February 2023
Revised 3 May 2023
Accepted 4 May 2023
Available online 16 May 2023

Keywords:
Bovine
Breed
Classification
EUROP
Lean meat yield

ABSTRACT

Beef carcases in Europe are classified as a proxy for the quantity and ratio of tissues, commonly referred to as yield. It is important that proxies accurately measure yield as they contribute to financial transactions between abattoirs and producers. The main purpose of the study was therefore to examine the ability of EUROP carcase classification to explain the variation in yield. Furthermore, the effect of breed, as a confounder, was also examined. A multivariate definition of yield separating the carcase into six product categories was utilised as a response in a linear regression analysis. The conclusion was that EUROP and carcase features explain the majority of yield variation. Breed has an effect on yield beyond what is explained by carcase features including classification. The magnitude of the breed effects varies with breed and product category.

Implications

The aim of this study was to evaluate and describe how changes in conformation, fat cover, weight and category contribute to change in yield, using a multivariate yield response that differentiated between different types of lean. In addition, the effect of breed as a confounding factor was studied. Our model explained the variance in the response, and breed was shown to contribute to residual variance independent of other predictor variables. These results could be used as basis for discussing changes in classification in Norway to improve the prediction of types of lean or to adjust for breed variations in Norway.

Introduction

The yield of bovine carcases can be defined as the ratio, quantity and distribution of tissue types a carcase consists of. It is an important description of the carcase as it affects producers, abattoirs and the market. The carcase yield is influenced by a multitude of factors, whose importance might vary between markets. The number and types of cuts, once the carcase is processed to its smallest components, are not uniform. Thus, identical carcases processed differently will have a different monetary value. In addition, the value of beef varies across markets and also varies over time (Henchion et al., 2014). Consequently, it is better to evaluate a carcase as a function of its yield rather than its monetary value in one specific market at a fixed point in time.

A classification system shall deliver a fair assessment of the ratio and quantity of meat, fat and bone within the carcase, and by proxy, the value of each carcase (Polkinghorne and Thompson, 2010; Delgado-Pando et al., 2021). Furthermore, a more detailed assessment of cuts is needed to adjust for value differences (Craigie et al., 2013) and thereby deliver a more precise assessment for producers and the meat industry. The type of assessment could depend on the market, e.g. differentiating between the value of meat for trimmings and meat for cuts in Norway, or differentiating between value of cuts using marbling assessment in the USA (Polkinghorne and Thompson, 2010). A theoretical, perfect classification system within each market would be one that can account for all the variation in yield relevant to that market, including any variation that could potentially come from other predictor variables such as weight, category or breed. In practice, a sufficiently functioning system explains the majority of the variation, cost efficiently.

The EUROP system of classification (European Commission, 2017a and 2017b) was developed by experts from member states in the European Union, based on the need of a common bovine

---

* Corresponding author at: Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway.
E-mail address: andrew.heggli@nmbu.no (A. Heggli).

https://doi.org/10.1016/j.animal.2023.100854
1751-7311/

classification system for all member states, upon request by the Organisation for European Economic Co-operation. Furthermore, the aim of the EUROP system for bovine carcases was to make a system that could describe all kinds of conformation and degrees of fatness by trained, expert human classifiers. Since EUROP was developed under these demands and not specifically for the exact yield from each carcase, the validity of the system in regard to functioning as a determining factor for yield has been the subject of study for decades (Kempster et al., 1982; Conroy et al., 2010; Oliver et al., 2010; Craigie et al., 2012), with most studies focusing on how well classification predicts yield. The definitions of yield in these studies vary slightly, but are simpler than the definition used in the present study, with one exception (Oliver et al., 2010), who used meat yield and four commercial cut quality categories as the response.

A common measurement for yield that has been used in several studies (Bergen et al., 2003; Craigie et al., 2012; Delgado-Pando et al., 2021) is a measure termed lean meat yield (**LMY**) or lean meat percentage (**LMP**), which is usually defined as g/kg of lean meat in relation to whole carcase weight. Carcase composition (Drennan et al., 2008; Conroy et al., 2010) is another measure of carcase yield that in addition to the ratio of meat, also includes the ratio of fat and bone to whole carcase weight. As information regarding fat and bone could be beneficial to abattoirs regarding the processing of a carcase, and for animal breeding goals, carcase composition is arguably a better response than just LMY.

Carcase composition describes more complexity than LMY, but it would be beneficial to add even more complexity by differentiating between high-value cuts and trimmings of different fat percentages. This would address the fact that different lean has different value. Though this would be a more precise yield measurement, it would change the response from single univariate to a multivariate response. A major objective of the present paper is to define the yield in a way which is complex enough to capture the variance of carcases smallest components, but at the same time simple enough for generalised patterns to be identified. In order to price the carcase, and to decide how it is to be further processed, actors in all markets will gather and utilise information for measurable carcase features prior to meat cutting (Polkinghorne and Thompson, 2010). This can include measurements of yield or meat qualities such as marbling scores, or both.

In the present paper, we will utilise features defined by the European Commission (European Commission, 2017a and 2017b) within the EUROP system as example for such features. These features include grading of conformation and fat group, carcase weight, age and category. EUROP scores contribute significantly to determining mercantile transactions between slaughterhouses and producers (Fisher, 2007). In Norway, an increased EUROP score increases the price paid per kg and an increased fat group score decreases the price paid per kg (Nortura AS, 2022). As it stands today, EUROP conformation might be viewed as an approximate measurement of the relationship between soft tissue (meat and fat) and bone, while fat group is the class of fat cover and reflects the relationship between fat and meat within the soft tissue. These two scores combined, form a basis for assessment of the total distribution of meat, fat and bones throughout the carcase.

In this study, a linear regression model with multivariate yield (**MVY**) response, representing high-value cuts (**HVCs**), trimmings, fat, bone & waste and percentage forepart, will be used with a focus on interpretation rather than finetuning for prediction precision. This means the focus will be on describing the relationship between carcase features/classification and the yield, rather than minimising the prediction error for the yield based on the same input. An improved understanding of these relationships will be beneficial for producers and abattoirs, but also beneficial to breeding programmes, and most importantly, understanding if the EUROP classification system actually functions as intended. The aim of study was to evaluate and describe how changes in conformation, fat cover, weight and category contribute to change in yield, i.e. the MVY. In addition, the effect of breed will be studied, in order to estimate the degree to which breed accounts for variance in the MVY as a confounding factor.

Material and methods

Data

Data from bovine animals were collected between 2011 and 2021 by Animalia's pilot plant. At the pilot plant, carcases are deboned and weights are registered daily, with the primary purpose of providing information for the Norwegian market regulation authorities, the classification system, and other stakeholders. The carcases were processed using two different commercial Norwegian cutting patterns, both of them with similarities to **UNECE** (United Nations Economic Commission for Europe) UNECE Standards (2016), but with some differences. The two Norwegian cutting patterns were similar enough to be arranged into merged product categories in one cohesive data set.

Each carcase processed at the pilot plant receives its own individual ID. For each carcase information regarding age, gender, breed, EUROP conformation and EUROP, fat group classification scores, and weight per whole carcase were also documented.

Each carcase is separated into quarters and then further processed into smaller cuts, trimmings of different fat (5, 14 and 21%), fat, bones, and waste. All cuts and trimmings were individually weighed and registered. Fat content of trimmings from the forepart and backpart of the carcase with desired fat contents of 14 and 21% were measured using the QV500, a NIRS scanner (Tomra Sorting Solutions, Asker, Norway). The original weights for trimmings of 5, 14 and 21% and fat were standardised digitally postdata collection. Either 5% trimmings or fat were added to 14 and 21% trimmings that had measured fat content lower or higher than desired, respectively.

The cuts and trimmings from both cutting patterns (see metadata and R-code freely available at https://doi.org/10.18710/TMSSJP for details) were aggregated into six disparate product categories with the objective of defining yield. The product categories were high-value cuts (**HVCs**), trimmings containing 5, 14 and 21% fat (named meat5, meat14 and meat21, respectively), fat and bone & waste. The total weights of these six product categories constituted the basis for the creation of the response variable. The smallest parts the carcases (n = 129 and n = 53 for the two cutting patterns) were deconstructed into were combined to create each product category. Product category weights, i.e. totals for the whole carcase, were set to a minimum of 0.1 kg, in order to avoid zeros in the log-odds response. Cuts that can be used without further processing, and that were considered of potential high monetary value, such as the tenderloin, were combined into the product category HVC. Trimmings with 5% fat and some cuts only defined in the most detailed cutting pattern, such as the flat iron filet, were combined into the product category meat5. Trimmings with 14% fat and 21% fat created self-titled product categories meat14 and meat21. Fat was made into its own self-titled product category. Bone and waste were combined into one product category, bone & waste.

For each carcase, registrations of conformation and fat group from EUROP classification (European Commission, 2013, 2017a and 2017b, category and carcase weight were registered online in the abattoir. EUROP conformation and fat group were registered as integers from 1 to 15 corresponding to the fifteen-point EUROP scale of five main categories each with ± variants. For conformation, this corresponds to a scale from P− (1) to E+ (15). For fat group, this corresponds to a scale from 1− (1) to 5+ (15). Weight was registered in kilograms. Category refers to a combination of age and gender, and is almost identical to categories specified by the European commission (European Commission, 2013), the categories are calf, young bull, bull, steer, heifer, young cow and cow, respectively. Details regarding categories and the minor differences to the definition given by the European commission are given in Supplementary Table S1.

Breed information was collected from two sources, the animal husbandry registry (Mattilsynet, 2022) and directly from the slaughterhouses. Data from the slaughterhouses were restricted to registration of the dominant breed. Consequently, individuals registered as a specific breed might in fact partly be crossbreeds, but for the analyses, these individuals were treated as pure breeds. Data from the animal husbandry registry were more complex, with pedigree recorded to the level of great-great-grandparents (F5). Thus, the breed distribution of each individual is registered in 16ths, and treated as such in the analyses.

The original data set consisted of 1790 individuals. Individuals with a deviation that was more than 2% between the total weight of the six product categories and carcase weight registered before processing were omitted. Furthermore, 11 individuals lacking sufficient information regarding breed and four calves, 17 bulls and eight steers were omitted from further analysis due to small sample sizes for these categories. This left a total of 1 687 individuals for the analysis. Five different breeds, Norwegian red, Hereford, Aberdeen Angus, Charolais and Simmentaler constituted their own breed group. Limousin and Blonde d'Aquitaine were merged into one breed group, and finally, one group with a combination of breeds was labelled "Other". The "Other" breed group consisted mostly of individuals that were registered as mixed breed, but with the breeds unknown, and some breeds with small sample sizes like Holstein cattle, Jersey cattle and Simmental milk, Tiroler Grauvieh, Dexter and Scottish Highland, and some Norwegian dual-purpose breeds; Sidet trønder- og nordlandsfe, Telemarkfe, Dølafe, Raukolle, Rødbroget dansk fe.

Statistical analysis

All statistical analyses were coded and executed in the software RStudio (RStudio Team, 2020). Linear regression was used to analyse the MVY outcome. The log-odds of the proportion of six product categories were used as response, together with the log-odds of the proportion made up by the forepart of the carcase, resulting in a response vector (y_i) of length seven.

The initial predictor variables were EUROP conformation and fat group, (carcase) weight, and category. An approximate F-test based on Wilks' lambda (Mardia et al., 1979) was conducted to investigate if the interaction between conformation and fat group, and their second-degree polynomials were contributing significantly to the regression model. The interaction and second-degree polynomials for fat group and category were then included in the model based on the outcome of Wilks' test. The final model then included the predictor variables EUROP conformation and fat group, as well as their significant interaction and polynomials, (carcase) weight, category, i.e.

$$y_i = \beta_0 + \beta_c x_{ci} + \beta_f x_{fi} + \beta_{cf} x_{ci} x_{fi} + \beta_{cp} x_{ci}^2 + \beta_{fp} x_{fi}^2 + \beta_w x_{wi} + \sum_{j=1}^{4} \beta_j^{cat} x_{ij}^{cat} + e_i, e_i \;\; i.i.d\, N(0, \Sigma), \quad i = 1, \cdots, 1687,$$

where i represented the individual and j represented the category. x_{ci}, x_{fi} and x_{wi} are the conformation, fat group and weight of carcase i respectively. $\beta_0, \cdots, \beta_w$ are regression parameters, each a vector of length seven, representing intercept, conformation, etc. Finally, $\beta_1^{cat}, \cdots, \beta_4^{cat}$ are each vector of length seven representing the effect of category, where the category is given by x_{ij}^{cat}, a binary variable taking value 1 if animal i has category j and else taking value 0. Category young bull was used as base category. Thus, the interpretation of β_j^{cat} is the expected change in response when the category is shifted from young bull to category j. The unexplained error for individual i, e_i, was assumed to be multivariate normally distributed with expectation 0 and covariance matrix Σ. The residuals ($\hat{e}_i$) were used to create an unbiased estimator for Σ with $\widehat{\Sigma} = \sum_i \widehat{e}_i^t \widehat{e}_i / (n - p)$ where $(n - p)$ is the degrees of freedom for the residual. The correlation matrix based on this estimate was also reported in the results in order to simplify the interpretation.

Regression parameter estimates with corresponding *P*-values based on *t*-tests were reported for all regression parameters. In order to simplify the understanding of the effect of a change in each of the independent variables, estimates for a standard individual were calculated. A standard individual was considered a young bull with the average values for conformation, fat group, and weight. The estimates were transformed from the log-odds to the percentage scale for easy interpretation. The estimated effects of increasing conformation, fat group and weight by one unit, based on this standard animal, were calculated and reported. Furthermore, the effect of changing category was reported by applying the same principle.

In addition to the prediction variables already included in the linear model, it was of interest to investigate whether breed has significant effect on yield. In order to conduct a formal test for the effect of breed on yield, an approximate F-test based on Wilks' lambda was also conducted on breed group. The effect was significant. However, breed group was omitted in the linear model as predictor a priori, as it is viewed a confounding factor and correlates with the other initial predictor variables. To evaluate the effect of breed group, without the confounding effects of the other predictors, a residual analysis was conducted with the residuals from the linear model as response variables and breed group as predictor variable. Based on the breed distribution for each animal, i.e. the data from the animal husbandry register, mean residuals were calculated for each breed group.

Results

Summary statistics for weight, conformation and fat group are presented per category in Table 1. There were large differences in sample size between the categories, with an overall unbalance towards male cattle, and young bulls specifically being overrepresented in the data. Further, summary statistics for weight, conformation and fat group are presented per breed group in Table 2. There was disparity in sample sizes between the different breed groups, with the most sampled breed group (n = 982) being the Norwegian red and the least being Simmentaler (n = 46).

The results from the approximate F-test based on Wilks' lambda showed that conformation and fat group interaction, second-order conformation and second-order fat group all were significant as predictor variables ($P < 0.001$, with approximate *F*-values at 18.09, 61.58 and 219.82, respectively). Thus, the interaction and second-order polynomials were retained in the final linear regression model.

There was also a significant effect of breed group ($P < 0.001$ with approximate *F*-value at 12.01), for the Wilks' lambda test when the other predictor variables were already included in the model. The parameter estimates for the multivariate linear regression are given in Tables 3 and 4. Table 5 shows predicted values for the percentage of a "standard individual", i.e. category young bull. All other values are in relation to this standard individual. As Table 5

Table 1
Summary statistics by category for cattle carcass weight and EUROP classification scores. Results for carcase weight, EUROP conformation and fat group are on a 15-point scale and are given as mean values ± SD. For the definition of categories, see Supplementary Table S1.

Item	Young bull	Heifer	Young cow	Cow
N	1067	111	237	272
Carcase weight (kg)	329.1 ± 58.11	232.7 ± 56.48	272.4 ± 57.85	297.7 ± 54.52
Conformation (EUROP)	6.9 ± 2.25	6.6 ± 1.84	4.3 ± 2.04	4.1 ± 1.96
Fat group (EUROP)	6.8 ± 1.5	7.7 ± 2.36	7.6 ± 2.54	7.8 ± 2.53

Table 2
Summary statistics by breed group for cattle carcass weight and EUROP classification scores. Results for carcase weight, EUROP conformation and fat group are on a 15-point scale and are given as mean values ± SD.

Item	Aberdeen Angus	Charolais	Hereford	Limousin and Blonde d'Aquitaine	Norwegian Red	Other	Simmentaler
N	54	241	68	138	982	158	46
Carcase weight (kg)	291.7 ± 86.49	351.9 ± 68.56	307.5 ± 55.58	352.4 ± 72.22	295.5 ± 52.10	293.9 ± 67.53	342.4 ± 60.19
Conformation (EUROP)	7.0 ± 1.62	8.7 ± 1.96	6.8 ± 1.75	9.4 ± 2.09	4.7 ± 1.49	6.0 ± 2.27	8.0 ± 1.45
Fat group (EUROP)	8.3 ± 2.88	7.2 ± 1.87	9.0 ± 2.45	6.6 ± 2.04	7.0 ± 1.81	6.7 ± 2.01	7.6 ± 1.78

Table 3
Regression parameter estimates from a linear regression with a multivariate yield from cattle carcases with *t*-values in brackets. Estimates with *P*-values <0.05 (*t*-test) are marked with an asterisk (*). Regression parameter estimates are estimated based on log-odds of product categories. Columns are predictor variables in the model. Rows are product categories. Conformation^2 is the quadratic version of the predictor variable Conformation, and Fat group^2 is the quadratic version of the predictor variable Fat group. C:F is the interaction between conformation and fat.

Item	Intercept	Conformation	Fat group	C:F	Conformation^2	Fat group^2
HVC	−1.7427 (−81.21)*	0.0417 (10.36)*	−0.0061 (−1.44)	−0.0010 (−2.68)*	0.0003 (1.35)	−0.0015 (−5.52)*
meat5	−2.5000 (−24.60)*	0.0697 (3.65)*	0.1227 (6.11)*	−0.0013 (−0.74)	−0.0004 (−0.36)	−0.0142 (−10.87)*
meat14	−0.4681 (−12.14)*	−0.0056 (−0.78)	0.0215 (2.82)*	0.0004 (0.56)	0.0003 (0.68)	−0.0047 (−9.44)*
meat21	−2.9922 (−46.17)*	0.0168 (1.38)	0.1381 (10.77)*	0.0021 (1.89)	−0.0037 (−4.79)*	−0.0037 (−4.44)*
Fat	−11.3203 (−28.47)*	−0.5029 (−6.73)*	1.5043 (19.13)*	0.0485 (7.23)*	−0.0013 (−0.27)	−0.0694 (−13.54)*
bone&waste	−0.6704 (−27.82)*	−0.0608 (−13.43)*	−0.0169 (−3.55)*	0.0023 (5.75)*	0.0000 (0.04)	−0.0007 (−2.25)*
forepart	−0.0265 (−1.74)	−0.0141 (−4.92)*	0.0094 (3.13)*	−0.0001 (−0.27)	−0.0002 (−1.24)	−0.0006 (−2.91)*

Abbreviations: HVCs = High-Value Cuts, Meat5 = Trimmings with 5% fat, Meat14 = Trimmings with 14% fat, Meat21 = Trimmings with 21% fat.

Table 4
Regression parameter estimates from a linear regression with a multivariate yield from cattle carcases with *t*-values in brackets. Estimates with *P*-values <0.05 (*t*-test) are marked with an asterisk (*). Regression parameter estimates are estimated based on log-odds of product categories. Columns are predictor variables in the model. Rows are product categories. σ^2 is the estimated residual variance and R^2 is the coefficient of determination per product category.

Item	Weight	Heifer	Young cow	Cow	σ^2	R^2
HVC	−0.0003 (−6.63)*	0.0898 (11.30)*	0.0981 (16.38)*	0.1082 (18.15)*	0.0045	0.646
meat5	0.0004 (1.96)	−0.0386 (−1.02)	−0.2443 (−8.62)*	−0.1241 (−4.39)*	0.0999	0.486
meat14	−0.0001 (−1.77)	−0.1479 (−10.36)*	−0.1065 (−9.90)*	−0.0387 (−3.61)*	0.0144	0.515
meat21	0.0000 (0.18)	0.1093 (4.56)*	0.0769 (4.26)*	−0.0443 (−2.46)*	0.0406	0.498
fat	0.0022 (2.83)*	0.6117 (4.16)*	0.6744 (6.08)*	0.6113 (5.53)*	1.5294	0.639
bone&waste	−0.0004 (−8.08)*	−0.0205 (−2.30)*	0.0078 (1.15)	0.0346 (5.17)*	0.0056	0.783
forepart	0.0005 (17.41)*	−0.1028 (−18.23)*	−0.1052 (−24.76)*	−0.1006 (−23.77)*	0.0022	0.583

Abbreviations: HVCs = High-Value Cuts, Meat5 = Trimmings with 5% fat, Meat14 = Trimmings with 14% fat, Meat21 = Trimmings with 21% fat.

Table 5

Predicted effects of conformation, fat group, weight, and category on expected yield. Results are based on evaluation of the linear regression model. Rows represent the product categories. The first column shows predictions for a standard individual, i.e. a young bull with corresponding mean weight, conformity and fat group. The values in the columns for conformation, fat group and weight all represent the expected change in percentage for the standard individual with an increase of one for the predictor in question. The values in the columns representing category show the expected change if the category is changed from young bull.

Item	Prediction	Conformation	Fat group	Weight	Heifer	Young cow	Cow
HVC	15.61	0.54	−0.45	0.00	1.22	1.34	1.48
meat5	14.25	0.68	−1.10	0.00	−0.46	−2.73	−1.45
meat14	35.69	0.03	−1.01	0.00	−3.32	−2.41	−0.88
meat21	10.12	−0.21	0.93	0.00	1.04	0.72	−0.40
fat	0.77	−0.14	0.98	0.00	0.64	0.73	0.64
bone & waste	22.26	−0.77	−0.19	−0.01	−0.35	0.13	0.60
forepart	51.79	−0.45	0.02	0.01	−2.57	−2.63	−2.51

Abbreviations: HVCs = High-Value Cuts, Meat5 = Trimmings with 5% fat, Meat14 = Trimmings with 14% fat, Meat21 = Trimmings with 21% fat.

values are percentage values, they are easier to interpret. Tables 3–5 should therefore be seen in tandem for a full understanding of the independent variables from the linear regression model and their significance and effect on the response.

Based on Table 3 and also reflected in 4, an increase in conformation was significantly associated with an increase in HVC and meat5, and a decrease in fat, bone & waste and percentage forepart. Higher conformation for the second-degree polynomial was significantly associated with a decrease in meat21. A higher fat group classification was significantly associated with an increase in meat5, meat14 and meat21, fat and percentage forepart, and a decrease in bone & waste.

An increase in fat group for the second-degree polynomial was significantly associated with a decrease in all product categories, indicating that the main linear effect of fat group decreases for extreme fat group registrations, i.e. both extremely lean and fat carcases, for all product categories. The interaction between conformation and fat was significantly associated with a decrease in HVC and an increase in fat and bone & waste. Weight had a significant negative effect on HVC and bone & waste, and a significantly positive effect on percentage forepart. An overview over the expected and observed proportions of the six product categories as function of EUROP conformation and fat group is visualised in Fig. 1.

Heifers, young cows and cows all had significantly positive effect on HVC and fat, and a negative effect on meat5 (heifers not significant) meat 14 and percentage forepart compared to young bulls. Heifers and young cows had significantly positive effect on meat21, while cows had significant negative effect on meat21.

Results from the residual analysis displayed in Table 6 indicate that Simmentaler and in particular Limousin and Blonde d'Aquitaine had a positive effect on HVC and meat5 compared to other breed groups. Hereford and Aberdeen Angus also had a positive effect on meat5, though not on HVC. Charolais had a positive effect on bone&waste, whereas in particular Limousin and Blonde d'Aquitaine, but also Hereford had a significantly negative effect on the percentage bone&waste. Limousin and Blonde d'Aquitaine also had significant negative effect on the percentage forepart. Norwegian red had a significantly positive, and substantial, effect estimated to be 0.83% units, on fat. Norwegian Red was the only breed group to have a significantly negative effect on HVC, and a positive effect on fat and percentage forepart. However, these estimates are small and significant mainly due to the large sample size for Norwegian red.

The residual correlation matrix, illustrated in Fig. 2, revealed that HVC had a moderately strong (−0.48) negative correlation with percentage forepart. Meat21 had moderately strong negative correlations to HVC (−0.35) and a strong negative correlation with meat14 (−0.70).

Discussion

The results confirm that EUROP classification, including conformation, fat group, weight and category, explains most of the yield variation. This is important, as weight and sex are registered in some form worldwide (Polkinghorne and Thompson, 2010) and EUROP conformation and fat group represent physical characteristics of the carcase which are interpretable also outside of Europe. The most pertinent evidence of EUROPs' ability to explain variation was the moderately high coefficient of determination estimates (>0.639) (Table 4) for HVC, fat, and bone & waste. These estimates were generally slightly lower, but not notably different, from previous literature (Drennan et al., 2008; Jørgenvåg et al., 2009, August; Conroy et al., 2010; Oliver et al., 2010) though previous literature focused on prediction. The comparatively lower coefficient of determination estimates for meat14 and meat21 (0.515 and 0.498, respectively) might be explained by an overlap in the meat they consist of, so more meat measured as one product category could lead to less of the other. This explanation is supported by the negative error correlation at –0.70 for these product categories. The inverse relationship between meat14 and meat21 is reflected in Fig. 1 where it is shown that more of the carcase consists of meat14 and less meat21 for low-fat group scores and vice versa. In other words, the plots associated with meat14 and meat21 are tilted in opposite directions.

The data used in the study were collected for industrial purposes and therefore not designed specifically for this study. If data collection could have been designed specifically for the study, a balanced distribution of breed crossed over categories would have been preferable. Additionally, one single cutting pattern following the UNECE standards (UNECE Standards, 2016) would have been used. Nevertheless, such a large data sample that reflects the population, with detailed registrations for each carcase, collected over a decade, constitutes a valuable resource for scientific analysis as conducted in the present study. The data in this study, both the yield and the classification, are based on Norwegian data, which does not take into account eating quality. In addition, eating quality is not assessed as a part of the mercantile transaction between producer and abattoirs in Norway. Consequently, eating quality was not assessed in this study.

Conformation's greatest effect was on HVC and bone & waste, while the fat group affected meat5, meat21 and especially fat more than conformation, meat14 approximately the same as conformation, and HVC and bone & waste to a lesser degree than conformation. This implies that though conformation and fat group are both important for accurately reflecting yield, they describe different aspects of the carcase. The coefficient of determination, in combination with estimated effects of conformation, fat group and weight in line with prior expectations, gives strong credence to EUROP's ability to explain yield variation.

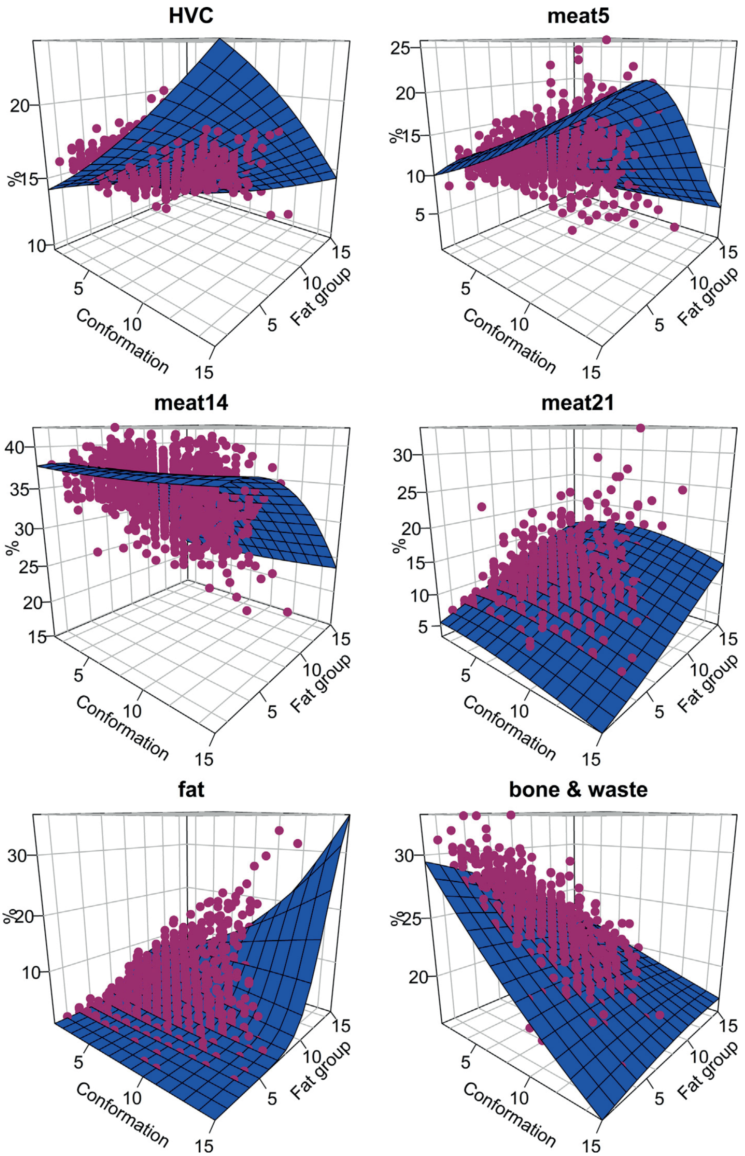


Fig. 1. Expected and observed proportions of the six product categories as function of EUROP conformation and fat group. HVC in upper left panel, meat5 in upper right panel, meat14 in middle left panel, meat21 in middle right panel, fat in lower left panel and bone & waste in lower right panel. Percentage of whole carcase on the *z*-axis, conformation on the *y*-axis, and fat group on the *x*-axis. The blue surface consists of predicted values in percentage based on a category young bull. The amaranth spheres show observed values per individual. HVCs = High-Value Cuts, Meat5 = Trimmings with 5% fat, Meat14 = Trimmings with 14% fat, Meat21 = Trimmings with 21% fat.

The estimated regression parameters (Table 3) for the fat polynomial were all negative, with the interpretation being that the relative effect on the response of a change in fat group is lower for the carcases with very low or high-fat group classification, than for carcases with fat groups close to average. The same logic can be applied to the relationship between the conformation regression parameter estimates, and its polynomial equivalent. However, even if there is significant effect of the second-order polynomial for conformation for some product categories, the overall impression, based on estimates and *t*-values (Table 3), is that most product categories scale close to linearly with conformation, which is also visible in Fig. 1.

Percentage forepart is greatly affected by weight and category (Table 4). The less weight the carcase has overall, the lower the percentage of the carcase is composed of meat and fat, and thereby the less difference between fore- and backpart due to bone making up the majority of the weight. The significant effect of weight on percentage forepart as seen by its large *t*-value (17.41) is due to

Table 6
Predicted effects of cattle breed group on expected yield. Results are based on the residual analysis. Rows represent the product categories. The values show the average percentage residual for the breed group in question. Significant effects (*P*-values <0.05, *t*-test) are marked with an asterisk (*).

Item	Norwegian red	Other	Hereford	Charolais	Aberdeen Angus	Limousin and Blonde d'Aquitaine	Simmentaler
HVC	−0.09*	0.15	0.09	−0.16*	0.00	0.87*	0.43*
meat5	0.02	0.18	1.06*	−0.08	0.59*	1.15*	0.87*
meat14	0.00	0.33	0.05	−0.04	0.33	0.17	0.05
meat21	0.23*	0.20	−0.33	0.38*	0.21	−0.45*	−0.30
fat	0.83*	0.33	−0.06	0.17	−0.10	−0.02	−0.03
bone & waste	0.04	−0.44*	0.35*	0.71*	−0.37	−0.97*	0.41
forepart	0.10*	0.24*	−0.50*	−0.13	0.01	−0.45*	−0.26

Abbreviations: HVCs = High-Value Cuts, Meat5 = Trimmings with 5% fat, Meat14 = Trimmings with 14% fat, Meat21 = Trimmings with 21% fat.

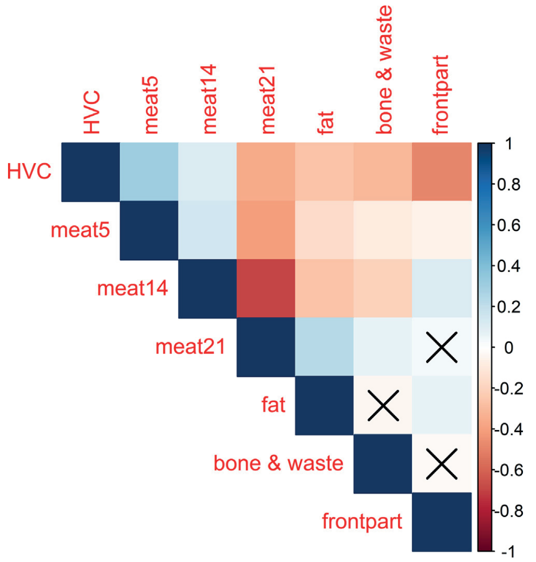


Fig. 2. Illustration of the residual correlation matrix for product categories percentage forepart from the yield of cattle carcases. Negative correlations with red, positive correlations with blue. Non-significant (5% level) correlations are shown with black "X"-s. HVCs = High-Value Cuts, Meat5 = Trimmings with 5% fat, Meat14 = Trimmings with 14% fat, Meat21 = Trimmings with 21% fat, frontpart = percentage of the carcase that is forepart.

differences in weight between forepart and backpart being exacerbated the higher the weight. The *t*-values for weight in regard to percentage forepart and in regard to bone & waste were large compared to corresponding *t*-values for other predictor variables (Tables 3 and 4). This is interpreted as weight having a greater effect on bone & waste, and percentage forepart, than predictor variables with lower *t*-values. The negative parameter estimate associated with weight for bone & waste is interpreted as heavier carcasses, having less bone and waste than can be accounted for by conformation and fat group alone. Though the value of the carcase based on EUROP classification ideally should scale with carcase composition, not weight, *t*-values associated with weight as a predictor collectively show that variation in yield is affected by carcase weight.

The effects of the female categories, i.e. heifer, young cow and cow are often considerable and significant (Table 4). One thing to take into consideration though, is that the classification variables are on a fifteen-point scale, while the category can only be changed from young bull to a single other category. The effects of category were expected as at the same EUROP classification level, different categories will still have morphological variation. The differences in ratio between product categories imply that a price difference between categories is justified.

Category had varying sample sizes. This reflects the number and the importance for the meat industry, where young bulls were sampled the most. Cattle of categories with more muscle, less fat and less bone have a larger market share, and thereby availability. Bulls fit this description as they, in general, have more muscle than cows, and young bulls specifically have relatively more muscle than older bulls, while having less fat. This can be seen in the weight of young bulls compared to the other categories (Table 1), and their mean conformation being the highest, with the corresponding standard deviation being similar to the other categories. Their mean fat group also shows that they are relatively lean compared to other categories.

The Wilks' lambda test showed that breed group was significant in addition to the variation explained by EUROP, weight and category, which is an important result as it might be concluded that breed affects yield, also after the controlling for EUROP, weight and category. A potential concern in the interpretation of the results of breed groups is that sire effects could potentially be masked as breed group effects. This is considered unlikely in this study as the time period for data collect spans a decade, and breed groups had large ($n \geq 46$) sample sizes. Another concern regarding breed groups are registrations of breeds at slaughterhouses. For the residual analysis, the individuals with breed registrations from slaughterhouses only were regarded as pure breeds, which is a simplification that might lead to minor biases in the residual analysis. However, when comparing breed group set by slaughterhouses with breed distribution registered in the animal husbandry registry for individuals with double registrations, the breed proportion in the animal husbandry registry corresponding the slaughterhouse registrations exceeded 72% for all breed groups, indicating reliable breed registrations also from the slaughterhouses.

In order to evaluate how breed affects yield, the residual analysis was needed. The most striking results are the mean residual in HVC for the Limousin and Blonde d'Aquitaine breed group at 0.87 percent, the mean residuals for bone&waste for Charolais at 0.71 percent and Limousin and Blonde d'Aquitaine breed group at −0.97 percent and finally the mean residual for fat for Norwegian red at 0.83 percent (Table 6).

When a single cattle weighs hundreds of kilos, such a seemingly small percentage difference can potentially result in a large difference in value, especially as the HVC, fat and bone&waste product categories consist of the cuts with highest (HVC) and lowest (fat and bone&waste) monetary value. Charolais seems to have an opposite effect compared to Limousin and Blonde d'Aquitaine, where the effect of Charolais is marginally less HVC and meat5, and more fat, and in particular bone & waste than expected from EUROP classification.

As breed significantly explains variation in yield as a confounding factor, an inevitable consequence is that some carcases are expected to be paid more or less than their corresponding yield value. It is beyond the scope of this study to examine how classification should potentially tackle the effect of breed. However, the results indicate that some regulation/adjustments of either classification or payment might be considered.

There were several challenges in regard to structuring the response variable. The fact that our data set included data from two different cutting patterns posed a challenge. If only carcases treated with the most detailed cutting pattern were included, the number of individuals in the data would have to be reduced by 390 individuals, who were processed under the less detailed cutting pattern. This would have been especially problematic as these individuals constituted the bulk (390 of 509) of the categories young cow and cow. Excluding these carcases would skew the balance between males and females even more. Several solutions were discussed, but in order to maintain an interpretable model and results, the solution where product components only present in the most detailed cutting pattern were changed to resemble product components also present in the least detailed cutting pattern, was chosen.

All fat group scores and majority conformation scores, i.e. for animals slaughtered before 2019 when the objective conformation classification method was implemented in Norway (Heggli et al., 2021), are set by human classifiers. As no classification system, neither automatic/objective systems nor human classifiers are flawless, some individuals will have an incorrect fat group and/or conformation grade. Heggli et al. (2021) estimated the standard deviation for fat group classified by Animalia experts to be 0.41, an estimate that must be regarded an absolute lower bound for the population of slaughterhouse classifiers. Classification errors, will in general, lead to more unexplained variance, improved classification precision will lead to an even tighter relationship between EUROP classification and the observed yield.

A source of variation that was not accounted for in the present study was the variation from different butchers. This could be partly solved by constructing the response for both sides of the carcase and use carcase ID as a random (predictor) variable in a mixed model. This was decided against as it complicated the ease of interpretation of the results, which was an important aspect of the study. Alternative methods for creating the responses, including the centred log-ratio transformation (Aitchison et al., 2000; Gloor et al., 2017), were considered, but rejected as they would inhibit or at least complicate the interpretation of the results.

Conclusion

EUROP conformation and fat group were shown to significantly explain the variation in yield, which indicates that the EUROP system is a fair system in regard to mercantile transactions, with some potential for improvements. Breed was shown to significantly explain variation in yield as a confounding factor, indicating that some breeds may be getting paid somewhat more, or somewhat less than corresponds to their yield.

Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.animal.2023.100854.

Ethics approval

Not applicable.

Data and model availability statement

The metadata and R-code used in the present study are freely accessible at https://doi.org/10.18710/TMSSJP. Access to the dataset with data used in the statistical analyses requires approval from Animalia AS.

Author ORCIDs

Andrew Heggli: https://orcid.org/0000-0002-5990-3197.
Ole Alvseike: https://orcid.org/0000-0002-1748-3470.
Frøydis Bjerke: https://orcid.org/0000-0002-5504-3636.
Lars Erik Gangsei: https://orcid.org/0000-0003-3171-544X.
Jørgen Kongsro: https://orcid.org/0000-0003-0619-9084.
Hilde Vinje: https://orcid.org/0000-0003-0318-0214.

Author contributions

Andrew Heggli: Conceptualisation, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review Editing, Visualisation.

Ole Alvseike: Conceptualisation, Methodology, Validation, Writing - Original Draft, Writing - Review Editing, Supervision, Funding acquisition.

Frøydis Bjerke: Methodology, Validation, Writing - Review Editing.

Lars Erik Gangsei: Conceptualisation, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Writing – Review Editing, Visualisation, Supervision, Project administration, Funding acquisition.

Jørgen Kongsro: Methodology, Validation, Writing - Review Editing.

Morten Røe: Conceptualisation, Methodology, Validation, Investigation, Writing - Review Editing, Supervision, Funding acquisition.

Hilde Vinje: Conceptualisation, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review Editing, Visualisation, Supervision, Project administration.

Declaration of interest

None.

Acknowledgements

The authors acknowledge Nortura SA, the abattoirs, Animalia AS and the Classification board.

Financial support statement

This work was funded by the Research Council of Norway under grant number 311394 and Animalia AS as a part of an industrial doctorate degree for Andrew Heggli.

References

Aitchison, J., Barceló-Vidal, C., Martín-Fernández, J.A., Pawlowsky-Glahn, V., 2000. Logratio Analysis and Compositional Distance. Mathematical Geology 32, 271–275. https://doi.org/10.1023/A:1007529726302.

Bergen, R., Crews, D.H., Miller Jr., S.P., McKinnon, J.J., 2003. Predicting lean meat yield in beef cattle using ultrasonic muscle depth and width measurements. Canadian Journal of Animal Science 83, 429–434. https://doi.org/10.4141/A02-093.

Conroy, S.B., Drennan, M.J., McGee, M., Keane, M.G., Kenny, D.A., Berry, D.P., 2010. Predicting beef carcass meat, fat and bone proportions from carcass conformation and fat scores or hindquarter dissection. Animal 4, 234–241. https://doi.org/10.1017/S1751731109991121.

Craigie, C.R., Navajas, E.A., Purchas, R.W., Maltin, C.A., Bünger, L., Hoskin, S.O., Ross, D.W., Morris, S.T., Roehe, R., 2012. A review of the development and use of video image analysis (VIA) for beef carcass evaluation as an alternative to the current EUROP system and other subjective systems. Meat Science 92, 307–318. https://doi.org/10.1016/j.meatsci.2012.05.028.

Craigie, C.R., Ross, D.W., Maltin, C.A., Purchas, R.W., Bünger, L., Roehe, R., Morris, S.T., 2013. The relationship between video image analysis (via), visual classification, and saleable meat yield of sirloin and fillet cuts of beef carcasses differing in breed and gender. Livestock Science 158, 169–178. https://doi.org/10.1016/j.livsci.2013.09.014.

Delgado-Pando, G., Allen, P., Troy, D.J., McDonnell, C.K., 2021. Objective carcass measurement technologies: Latest developments and future trends. Trends in Food Science & Technology 111, 771–782. https://doi.org/10.1016/j.tifs.2020.12.016.

Drennan, M.J., McGee, M., Keane, M.G., 2008. The value of muscular and skeletal scores in the live animal and carcass classification scores as indicators of carcass composition in cattle. Animal 2, 752–760. https://doi.org/10.1017/S1751731108001754.

European Commission, 2013. Regulation (EU) No 1308/2013 of the European Parliament and of the Council of 17 December 2013 establishing a common organisation of the markets in agricultural products and repealing Council Regulations (EEC) No 922/72, (EEC) No 234/79, (EC) No 1037/2001. Retrieved 26 June 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02013R1308-20211207.

European Commission, 2017a. Commission delegated regulation (eu) 2017/1182 of 20 april 2017. supplementing regulation (eu) no 1308/2013 of the european parliament and of the council as regards the union scales for the classification of beef, pig and sheep carcasses and as regards the reporting of market prices of certain categories of carcasses and live animals, annex ii. Retrieved 26 June 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32017R1182&from=SV#d1e32-92-1.

European Commission, 2017b. Commission delegated regulation (eu) 2017/1184 of 20 April 2017 laying down rules for the application of Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards the Union scales for the classification of beef, pig and sheep carcasses and as regards the reporting of market prices of certain categories of carcasses and live animals. Retrieved 19 April 2023 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R1184#d1e393-103-1.

Fisher, A., 2007. Beef carcass classification in the EU: an historical perspective. Publication-European Association for Animal Production 123, 19.

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome datasets are compositional: And this is not optional. Frontiers in Microbiology 8, 1–6. https://doi.org/10.3389/fmicb.2017.02224.

Heggli, A., Gangsei, L.E., Røe, M., Alvseike, O., Vinje, H., 2021. Objective carcass grading for bovine animals based on carcass length. Acta Agriculturae Scandinavica, Section A – Animal Science 70, 113–121. https://doi.org/10.1080/09064702.2021.1906940.

Henchion, M., McCarthy, M., Resconi, V.C., Troy, D., 2014. Meat consumption: Trends and quality matters. Meat Science 98, 561–568. https://doi.org/10.1016/j.meatsci.2014.06.007.

Jørgenvåg, I.A., Røe, M., Malmo, T., Alvseike, O., 2009. Prediction of yield from carcass grading by europ and video image analysis. Poster session presentation at 55th International Congress of Meat Science and Technology, Denmark, Retrieved 26 June 2022 from https://digicomst.ie/wp-content/uploads/2020/05/2009 03 04.pdf.

Kempster, A.J., Cuthbertson, A., Harrington, G., 1982. The relationship between conformation and the yield and distribution of lean meat in the carcasses of British pigs, cattle and sheep: A review. Meat Science 6, 37–53. https://doi.org/10.1016/0309-1740(82)90049-3.

Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. Multivariate Analysis. Academic Press, London, England.

Mattilsynet, 2022. Rapportering til Husdyrregisteret - Storfe —Mattilsynet. (In Norwegian). Retrieved 26 June 2022 from https://www.mattilsynet.no/dyr/produksjonsdyr/storfe/rapport-til-husdyrregisteret.

Nortura AS, 2022. Avregningspriser storfe - vilkaar storfe - Medlemsportal for Nortura SA. (In Norwegian). Retrieved 26 June 2022 from https://medlem.nortura.no/prislister/avregningspriser-storfe-vilkar-storfe-article16616-11969.html.

Oliver, A., Mendizabal, J.A., Ripoll, G., Albertí, P., Purroy, A., 2010. Predicting meat yields and commercial meat cuts from carcasses of young bulls of Spanish breeds by the SEUROP method and an image analysis system. Meat Science 84, 628–633. https://doi.org/10.1016/j.meatsci.2009.10.022.

Polkinghorne, R.J., Thompson, J.M., 2010. Meat standards and grading: A world view. Meat Science 86, 227–235. https://doi.org/10.1016/j.meatsci.2010.05.010.

RStudio Team, 2020. RStudio: Integrated Development Environment for R [http://www.rstudio.com/]. RStudio, PBC, Boston, MA, USA.

UNECE Standards, 2016. Bovine meat and carcass and cuts. Retrieved 15 February 2023 from https://unece.org/fileadmin/DAM/trade/agr/standard/meat/e/Bovine_326Rev2E_2016.pdf.

Paper III

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 04 December 2023
DOI 10.3389/fgene.2023.1289130

OPEN ACCESS

EDITED BY
Juan José Arranz,
University of León, Spain

REVIEWED BY
Damiano Cavallini,
University of Bologna, Italy
Pablo A. S. Fonseca,
University of León, Spain

*CORRESPONDENCE
H. Vinje,
✉ hilde.vinje@nmbu.no

RECEIVED 05 September 2023
ACCEPTED 07 November 2023
PUBLISHED 04 December 2023

CITATION
Vinje H, Brustad HK, Heggli A, Sevillano CA, Van Son M and Gangsei LE (2023), Classification of breed combinations for slaughter pigs based on genotypes—modeling DNA samples of crossbreeds as fuzzy sets from purebred founders.
Front. Genet. 14:1289130.
doi: 10.3389/fgene.2023.1289130

Classification of breed combinations for slaughter pigs based on genotypes—modeling DNA samples of crossbreeds as fuzzy sets from purebred founders

H. Vinje[1]*, H. K. Brustad[2], A. Heggli[1,3], C. A. Sevillano[4], M. Van Son[5] and L. E. Gangsei[1,3]

[1]Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway, [2]Oslo Center of Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway, [3]Animalia AS, Oslo, Norway, [4]Topigs Norsvin Research Center, Beuningen, Netherlands, [5]Norsvin SA, Hamar, Norway

In pig production, the production animals are generally three- or four-way crossbreeds. Reliable information regarding the breed of origin of slaughtered pigs is useful, even a prerequisite, for a number of purposes, e.g., evaluating potential breed effects on carcass grading. Genetic data from slaughtered pigs can easily be extracted and used for crossbreed classification. In the current study, four classification methods, namely, random forest (RF), ADMIXTURE, partial least squares regression (PLSR), and partial least squares together with quadratic discriminant analysis (PLS-QDA) were evaluated on simulated (n = 7,500) genomic data of crossbreeds. The derivation of the theory behind PLS-QDA is a major part of the current study, whereas RF and ADMIXTURE are known and well-described in the literature. Classification success (CS) rate, square loss (SL), and Kullback–Leibler (KL) divergence loss for the simulated data were used to compare methods. Overall, PLS-QDA performed best with 99%/0.0018/0.002 (CS/SL/KL) vs. 97%/0.0084/0.051, 97%/0.0087/0.0623, and 17%/0.068/0.39 for PLSR, ADMIXTURE, and RF, respectively. PLS-QDA and ADMIXTURE, as the most relevant methods, were used on a real dataset (n = 1,013) from Norway where the two largest classes contained 532 and 192 (PLS-QDA), and 531 and 193 (ADMIXTURE) individuals, respectively. These two classes were expected to be dominating *a priori*. The Bayesian nature of PLS-QDA enables inclusion of desirable features such as a separate class “unknown breed combination” and informative priors for crossbreeds, making this a preferable method for the classification of breed combination in the industry.

KEYWORDS

slaughter pigs, breed classification, crossbreeds, fuzzy classification, single-nucleotide polymorphism, partial least squares, quadratic discriminant analysis, ADMIXTURE

1 Introduction

Several meat production livestock systems rely on crossbred animals. In pig production, the production animals are generally three- or four-way crossbreeds (CBs). There are several reasons for using CB in meat production, in particular to benefit from heterosis and breed complementarity and to be flexible in creating different products for different markets

(Smith, 1964; Dickerson, 1973; Sellier, 1976). In contrast to meat production from other domestic animals, such as cattle, breed information for production pigs is not logged through the production chain, and hence, no, or at least incomplete, controls exist for the breed combination of individual slaughtered pigs.

Verification of a true CB combination is beneficial at different steps of the pig production chain. For instance, having control of CB breed origin will allow for the inclusion of CB performance from different sources after verification in the selection program of a pig breeding company. Including CB performance is desirable for achieving a sizable genetic progress for traits showing a genetic correlation between PB performance and CB performance that is lower than unity (Wientjes and Calus, 2017). Breed and crossbreeding also have a significant effect on meat quality traits (Kim et al., 2020) and the distribution of undesired mutations like halothane and Rendement Napole genes (Hamilton et al., 2000). For these reasons, verification of the breed of origin will be valuable for abattoirs and retailers. Finally, it is natural to assume that breed origin might be an unobserved nuisance factor for carcass grading; see Gangsei et al. (2018) for an elaborating discussion. If CB combinations were known for dissected carcasses, it would facilitate the evaluation of potential biases regarding the grading of different breed combinations.

The problem addressed in the current study is the classification of breed combination for individuals in the CB population based on genetic data, i.e., single-nucleotide polymorphisms (SNPs) from a 50-K SNP chip. Such genomic data have become cheaper and more accessible due to rapid developments in technology, and the number of application areas has exploded within different parts of the natural sciences, including ancestry classification tasks.

Most methods used to infer breed combinations in pig populations with genomic data were originally developed for inferring human ancestry. ADMIXTURE (Alexander et al., 2009) is one of the most popular methods used to classify individuals with an unknown ancestry into discrete populations and was developed for human populations but has been used extensively in pigs to trace commercial (Huang et al., 2014) or indigenous pig breeds (Mujibi et al., 2018; Dadousis et al., 2022; Kim et al., 2022; Yin et al., 2023). Principal component-based algorithms for determining the ancestry have also been developed, such as PCAdmix (Brisbin et al., 2012), and applied in pig populations (Schleimer et al., 2022). As pig breeding is far from human genetics, new methods are needed for better classification. An approach has been developed to assign alleles in three-way CB pigs to their PB of origin; the approach has high accuracy, but as it infers local ancestry, it is highly computationally demanding (Sevillano et al., 2016; Vandenplas et al., 2016). Another study tried using random forest for this purpose in pigs; however, the method did not accurately estimate breed composition for the breeds in question with the available markers (Chinchilla-Vargas et al., 2021). Recently, an interesting study was published showing the advantage of partial least squares regression (PLSR) and partial least square–discriminant analysis (PLS-DA) for global ancestry identification of pig breeds (Miao et al., 2023). The results showed that a wide range of breeds can be discriminated using these methods and that alternatives to human-developed methods can be beneficial for the pig industry. However, the study was restricted to the classification of PBs and treated CBs as similar to unknown breeds.

When assuming that grandparents are PB animals from a set of known PBs with known origin, i.e., breed, there will always be a possibility that CB individuals might have one or more grandparents from breeds outside the set with predefined known breeds. Ideally, classification methods should be able to identify such individuals and classify them as an "unknown breed combination." All the aforementioned methods, except PLS-DA, lack the ability to incorporate such a feature, and further elaboration is needed to achieve this objective.

The novelty of the present study is to derive the theoretical basis for partial least squares with quadratic discriminant analysis (PLS-QDA) (Boulesteix, 2004; Hastie et al., 2009) used for CB classification based on the following steps: i) PLS was used as a replacement for PCA/MDS as the primary dimension reduction method for SNP data, ii) additional variance was incorporated by modeling the proportion of DNA inherited from each grandparent as a random variable, and iii) the Bayesian nature of QDA was utilized to incorporate informative priors for CB classes and the possibility to include a class "unknown breed origin". The variance in proportion of inherited DNA was combined with breed-specific variances for PBs in order to achieve CB-specific covariance matrices for PLS components. CB-specific covariance matrices enable the use of QDA as a replacement for the more common linear discriminant analysis (LDA). An additional asset facilitated by PLS-QDA is visualization of the behavior of decision boundaries in a low (<3) dimensional space.

The overall aim of this study is to evaluate the crossbreed classification of commercial finisher pigs based on genomic data from a 50-K (Illumina) SNP chip. Two other well-known classification methods, random forest (RF) and ADMIXTURE, were compared with PLSR and PLS-QDA.

2 Materials and methods

2.1 Materials

2.1.1 Genomic data

The genotypes used in this study are data collected from the pig breeding companies Norsvin (Norway) and Topigs Norsvin (the Netherlands). Animals were genotyped using a custom GeneSeek 50-K (Illumina) SNP chip (Lincoln, NE, Unites States). Of these, 23,070 SNPs are used routinely by Topigs Norsvin and constitute the raw SNP data in the current study. Based on PB animals ($n = 4{,}014$), it was observed that from five different PBs (see details in the following section), the minimum call rate was 0.997 and minor allele frequency (MAF) was 0.045, well inside the limits used by Tusell et al. (2020) at 0.9 for call rate and 0.01 for MAF.

For each SNP, the most frequent allele in 4,014 PBs was identified. SNPs were coded into numeric vectors with zero for the homozygous genotype of the most frequent allele at the SNP in question, one for the heterozygote, and two for the homozygous genotype of the least frequent allele.

Data from five PBs ($n = 4{,}014$) are used as training data for all models. They also constitute the basis for data simulation; see the paragraphs in the following section for details. PBs are Landrace (abbreviation "L," $n = 1{,}000$), Large White ("W," n = 1,000), Duroc ("D," $n = 1{,}000$), and Hampshire ("H," $n = 14$), which are

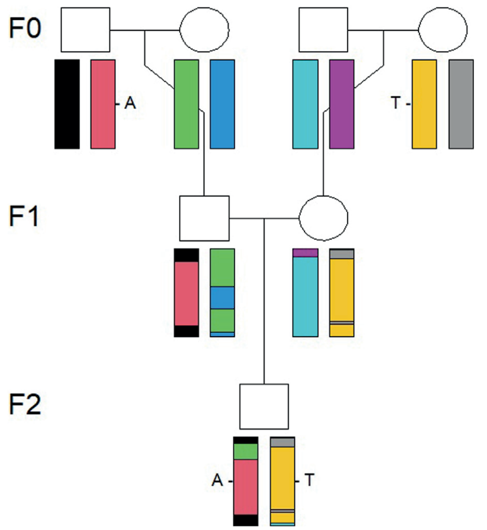


FIGURE 1
Illustration of how a chromosome pair is inherited IBD from the F0 to F2 generation. The genomes recombine from the F0 to F1 generation and then to the individual in the F2 generation, visualized by the combination of colors of the chromosomes.

dominating PBs in the Norwegian pig population. For the exploration of uncertainty measures and generalizations of the methods, a fifth breed, Pietrain ("P," n = 1,000), not present in Norway, was incorporated into the study. In the present study, the term "breed" is used extensively. In practice, the PBs described previously might be viewed as sub populations/lines primarily present in the Norwegian pig population. In addition to the SNP data from PB individuals, SNP data from 1,013 slaughter pigs with unknown breed origin were used to examine model behavior.

2.1.2 Breeds and breed combinations

The focus of this study is to classify the breed combinations of founders (F0 generation) observed in the commercial finisher pigs (F2 generation) based on genotypes from the F2 generation. It is assumed that all F0 individuals are PB.

When q PBs are present, there are q^4 (625 for $q = 5$ and 256 for $q = 4$) different unique breed permutations in the F2 generation. For example, *"LWDD"* indicates L and W as the grandfather and grandmother of the maternal line, respectively, and D as the grandfather and grandmother at the paternal line, which is the most common Norwegian finisher breed combination. The pure breeds Duroc and Hampshire are typically the paternal line for production pigs.

The q^4 unique permutations constitute a total of $n_{comb} = \binom{q+4-1}{4}$ (70 for $q = 5$ and 35 for $q = 4$) unique breed combinations when the sequence of grandparents is not taken into account. Combinations are given with letters in descending alphabetical order. For example, the combination *"DDLL"* contains the permutations {*"LLDD," "LDLD," "LDDL," "DDLL," "DLLD," "DLDL"*}, etc.

2.2 Simulation of SNP data for crossbreeds

In order to test the accuracy of classification methods, CB data with a known breed origin form a prerequisite. For the current study, such data were nonexistent, and data simulation was used to obtain relevant test datasets for the methods.

SNP data from the PBs (n = 4,014) were used as the input for the simulation. The output was combinations of the SNP data in accordance with known CB combinations. One simulated test set, TestP−, originates from PBs omitting Pietrain. The other test set, TestP+, consists of breed combinations with at least one Pietrain grandparent. For each CB combination, we performed 100 simulations, resulting in 3,500 simulations for both test sets, TestP− and TestP+.

Simulations were conducted using R packages and functions described in Vigeland (2021). The first step in the simulation procedure was to simulate an identical by descent (IBD) pattern along the genome for an individual in the F2 generation, as shown in Figure 1, based on a pedigree connecting the F0 and F2 generations. An IBD pattern shows how different parts of an individual's genome are inherited through descent from previous generations, from grandparents in the present study, using information on how the chromosomes recombine. The recombination is a stochastic process along the genome. A prerequisite for simulation of IBD patterns is a recombination map which relates the cumulative genetic distance, in centimorgans, to the cumulative physical distance, in bases, along the genome. The genetic map provided by Tortereau et al. (2012) was used as the basis for the recombination map.

The second step was to sample SNP data for the F2 individual, conditional on the simulated IBD pattern and the CB combination from the individual in question. Thus, the F0 generation was randomly selected among the 4,014 PB individuals, and then, their allele data (nucleotide bases A, C, T, and G) were transferred to the F2 individual, in accordance with the simulated IBD pattern for each SNP. Consequently, the simulated data are different combinations of the original allele data from the 4,014 PB pigs.

2.3 Evaluation of classification results

The typical goal of classification is to assign an observation to one out of a distinct set of classes. A problem arises when the goal is to classify in between such distinct classes. We will regard the CB pigs as such fuzzy sets (Zadeh, 1965), i.e., the class of CB pigs is regarded as a continuum of grades of membership in the PB classes.

Four classification methods were tested: random forest (RF), ADMIXTURE, partial least squares regression (PLSR), and partial least squares with quadratic discriminant analysis (PLS-QDA); see following sections for details. Only PB individuals were used for training the four classification methods. We applied two different training datasets, one consisting of all PB individuals and another omitting the Pietrain data, referred to as TrainP+ and TrainP−, respectively.

For the two simulated test sets, TestP+ and TestP−, breed combinations were classified using all four aforementioned methods trained on the two training sets (TrainP− and TrainP+), leading to a total of 16 ("four methods" × "two training data sets" × "two test data sets") combinations of methods, training, and test data. For each combination, 3,500 individuals were classified, i.e., 35 "CB combinations" × 100 "individuals per combination". The real data from CBs with unknown breed origin (n = 1,013), denoted "TestR", were classified using ADMIXTURE and the PLS-QDA method. The results were used to examine the effect of an informative prior distribution in the latter and whether the methods provided useful and credible results in a practical setting.

All breed combinations might be represented by a vector Δ of length q (q = 4 in TrainP− and q = 5 in TrainP+) whose elements are the proportions of grandparents from each of the q PBs, in alphabetic order, i.e., "D," "H," "L," "P," and "W." For example, an individual with breed combination $DDLW$ will have $\Delta = \frac{1}{4}[2\,0\,1\,0\,1]^t$. Consequently, the elements of Δ are quarters, which sum to 1. All four classification methods give predictions for Δ which denoted $\hat{\Delta}$. Even if the elements of Δ are quarters, the predictions are proportions, i.e., $0 \leq \hat{\delta}_{new}^{j} \leq 1$ and $\sum_j^q \widehat{\delta}_j = 1$, but the elements $(\widehat{\delta}_j)$ are not necessarily in quarters.

Two statistics are calculated for the evaluation of different methods based on simulated data where the true breed combinations (Δ) are known. The square loss for a new prediction is given by $\sum_j^q (\delta_j - \widehat{\delta}_j)^2$, and the Kullback–Leibler divergence (Kullback and Leibler, 1951) is the divergence between the two multinomial distributions for 4Δ with probability vectors Δ and $\hat{\Delta}$, respectively.

Hard classifications for RF, ADMIXTURE, and PLSR were achieved by choosing the breed combination with either the shortest Kullback–Leibler divergence or minimal square loss. For PLS-QDA, the hard prediction is the CB class with largest posterior probability.

2.4 Classification methods

2.4.1 ADMIXTURE and random forest

ADMIXTURE (Alexander et al., 2009) is an algorithm and software tool for the maximum likelihood estimation of individual ancestries, usually used for humans but also possible to apply to other species like pigs. ADMIXTURE 1.3 software (avid H. Alexander et al., 2020) was used for this analysis in a supervised mode with K-values set to 4 (for TrainP−) and 5 (for TrainP+).

RF is a widely used classification method built on the theory of tree-structured classifiers. An RF consists of a collection of K tree-structured classifiers, where K is usually a large number. In the end, all trees vote for their preferred class and RF classifies to the class with most votes (Breiman, 2001; Hastie et al., 2009). The Hampshire data were oversampled in the tree-growing process, inversely proportional to their abundance compared to other breeds (1,000/14). RF analysis was conducted via the "randomForest" package (Liaw and Wiener, 2022) in R. Only training data were used for tuning hyper parameters, with the out-of-bag (OOB) error as the performance measurement. The major hyper parameters to tune are the number of drawn candidate variables in each split (m_{try}), the number of observations drawn for each tree (sample size), node size, and number of trees (K) (Probst et al., 2019). We used $m_{try} = 151 \approx \sqrt{23070}$, number of trees $K = 100$, sample size 100, and node size 1, which gave OOB errors equal to 0 for both TrainP+ and TrainP−.

2.4.2 Partial least squares regression

Partial least squares (PLS) is a supervised method where breed information is taken into account. Wold et al. (2001) offers an overview over the fundamental principles of PLS. The basic idea of PLS regression (PLSR) is to find the multidimensional directions in the predictor variable space, i.e., the SNP ($\mathbf{X}$) that explains the maximum multidimensional variance direction in the response, i.e., the breed ($\mathbf{Y}$).

We apply a multivariate response matrix $\mathbf{Y}$ ($n \times q$) for the PLS regression. Each row in the response consists of the Δ vector for the PB in question. As all individuals in the training data are PB, all elements of $\mathbf{Y} \in \{0, 1\}$, i.e., dummy variables for the breed.

In principle, all predictor variables, i.e., SNPs, are included but assigned different weights, defined by the loading matrix $\mathbf{P}$ ($p \times m$). The score matrix $\mathbf{T} = \mathbf{XP}$, a $n \times m$ matrix, defines the relevant subspace of $\mathbf{X}$, where m is the number of relevant components.

We used $m = q - 1$ where the reasoning is that two breeds will be well separated on one axis/component, three breeds by two axis, or in general q breeds by $m = q - 1$ axis/components, where each PB should represent one node point, and one node point only, in the m-dimensional space spanned by the scores.

A frequently used method for dimension reduction is principal component analysis (PCA) (Pearson, 1901). In contradiction to PLS, PCA is an unsupervised method, not taking breed information into account when constructing the scores. For comparison of the two methods, the first four scores from PCA and PLS are compared and evaluated against the prerequisite that each PB should represent one node point, and one node point only, in the space spanned by the scores.

PLSR predictions might yield results whose elements are larger than 1 or smaller than 0. These elements were truncated to $1 - 10^{-10}$ and 10^{-10}, respectively, for the evaluation of Kullback–Leibler divergence.

The R-package "pls" (Liland et al., 2021) was used for fitting PLS and PCA models. The response ($\mathbf{Y}$ as described previously), predictors ($\mathbf{X}$ as described previously), and number of components ($q = 4$ for TrainP+ and $q = 3$ for TrainP−) were the data/parameters used as inputs for fitting the PLS and PCA models.

2.4.3 Partial least squares with quadratic discriminant analysis

The principles of classification and discriminant analysis (DA) are given in Hastie et al. (2009). The goal is to find a posterior probability for different classes (CBs):

$$P(K|\mathbf{x}) = \frac{\mathrm{f_k}(\mathbf{x})\pi_{\mathrm{k}}}{\sum_{l=1}^{\mathrm{n_{comb}}+1} \mathrm{f}_l(\mathbf{x})\pi_{\mathrm{l}}}, \tag{1}$$

where $f_k(\mathbf{x})$ is the class-conditional density of $\mathbf{x}$ (observed SNPs), assuming that the SNP sample is from an individual of class K, and π_k is the prior probability of class K.

For PLS-DA, $\mathbf{x}$ is replaced with $\mathbf{t}$, i.e., the PLS score vector. Furthermore, a common assumption is to assume that $f_k(\mathbf{t})$ is (multivariate) normally distributed with different mean parameters (μ_k). When variance is assumed to be constant among classes, the method is known as linear discriminant analysis (LDA), which is notably applied on PLS scores (PLS-LDA) (Boulesteix, 2004). In the present study, we assume different variance parameters Σ_k for each class, which is known as quadratic discriminant analysis (QDA) (Hastie et al., 2009). Initially, this assumption is applied to the scores of the PB individuals, i.e., the PLS scores (t_j) for PB individuals are assumed to be multivariate normal:

$$\mathbf{t}_j \sim \mathrm{N_m}\left(\mu_j, \Sigma_\mathrm{j}\right), \mathrm{j} = 1, \ldots, \mathrm{q}, \qquad (2)$$

where q is the number of PBs.

For the problem in the present study, only data from PB animals are used for training, and we lack observations, i.e., score vectors, for all CB classes. In order to implement CB classes, we need to find the class-conditional densities ($f_k(\mathbf{x})$) for CB classes, without having the realization of score vectors for these classes. In addition, we included a class "unknown," i.e., an unknown breed combination, leading to $n_{comb} + 1$ possible classes. The inclusion of the "unknown" class is possible for PLS-QDA, due to its Bayesian nature, where possible CB combinations are defined *a priori*, in contradiction to RF and PLSR.

A natural assumption is to assume that the scores of CB animals are distributed as linear combinations in accordance with the proportion of inherited DNA from the PB F0 generation. Let θ, a vector of length q, represent the proportion of DNA material in an F2 individual inherited from grandparents of different F0 PB individuals. Then, $0 \leq \theta_j \leq 1$, $j = 1, \ldots, q$, and $\sum_j^q \theta_j = 1$. Under the assumption that θ is known for a CB individual, it is natural to model $\mathbf{t} = \sum_j^q \theta_j \mathbf{t}^j$, where $\mathbf{t}^j$ is the score associated with PB class j. Using standard proprieties of the normal distribution, we have

$$\mathbf{t} \mid \theta \sim \mathrm{N_m}\left(\sum_{\mathrm{j}=1}^{\mathrm{q}} \theta_\mathrm{j}\mu_\mathrm{j}, \sum_{\mathrm{j}=1}^{\mathrm{q}} \theta_\mathrm{j}^2 \Sigma_\mathrm{j}\right). \qquad (3)$$

The proportion of DNA inherited from each grandparent is not exactly equal to a quarter. Thus, θ might be viewed as a random variable with $E(\theta) = \Delta$ and defined variance $V(\theta)$. By applying the law of total expectation (Adam's law) and variance (Eve's law), we find that

$$\begin{aligned} E(\mathbf{t}) &= E_\theta\left[\sum_{j=1}^{q} \theta_j \mu_j\right] = \sum_{j=1}^{q} \delta_j \mu_j = \left[\mu_1 \ldots \mu_q\right]\Delta = \mu^+\Delta, \\ V(\mathbf{t}) &= E_\theta\left[\sum_{j=1}^{q} \theta_j^2 \Sigma_j\right] + V_\theta\left[\mu^+\theta\right] = \sum_{j=1}^{q}\left(V\left(\theta_{jj}\right) + \delta_j^2\right)\Sigma_j + \mu^+ V(\theta)\mu^{+t}, \end{aligned} \qquad (4)$$

where $V(\theta_{jj})$ is the jth diagonal element of $V(\theta_l)$ and the columns of the $m \times q$ matrix μ^+ are given by the q expectation vectors (μ_j) for PBs. We will assume, even if it is an approximation, that the marginal distribution for $\mathbf{t}$ is multivariate normal with expectation and variance as given in Eq. 4, i.e.,

$$\mathbf{t} \sim \mathrm{N_m}\left(\mu^+\Delta, \sum_{\mathrm{j}=1}^{\mathrm{q}}\left(\mathrm{V}\left(\theta_\mathrm{jj}\right) + \delta_\mathrm{j}^2\right)\Sigma_\mathrm{j} + \mu^+\mathrm{V}(\theta)\mu^{+t}\right). \qquad (5)$$

Ordinary least squares (OLS) estimates are used for μ_j and Σ_j, i.e., sample means and sample covariance for the PB PLS score vectors from the training datasets. In addition, $V(\theta)$ needs to be estimated. This was done by assuming that θ is Dirichlet distributed with the concentration parameter $\alpha_0\Delta$; consequently, $E(\theta) = \Delta$, and $V(\theta) = \frac{1}{\alpha_0+1}\left(d(\Delta) - \Delta\Delta^t\right)$, where $d(\Delta)$ is the diagonal matrix with Δ on the main diagonal. Then, the only unknown parameter is α_0, which was estimated by the method of moments on simulated data. A total of 1,000 simulations of θ, applying Δ as a vector of quarters, was conducted by IBD simulation (Vigeland, 2021); see Section 2.2 for details. The diagonal elements of the simulated variance have expected values $3(16\alpha_0 + 16)^{-1}$ leading to $\widehat{\alpha_0} = 3(16\widehat{\mathrm{Var}\theta})^{-1} - 1$, where $\widehat{\mathrm{Var}\theta}$ is the mean diagonal element of the empirical variance based on simulated data, which is affected by several factors, including the genetic map (Tortereau et al., 2012).

We assumed that the class-conditional density of $\mathbf{t}$ for the unknown breed ($f_{'unknown'}(\mathbf{t})$) was uniform over the $q - 1$-dimensional space spanned by the range of PB score vectors.

We use two different prior distributions, i.e., π_k in Eq. 1, a "flat prior" for different breed-combinations, i.e., $\pi_1 = \ldots = \pi_{n_{comb}+1} = \frac{1}{n_{comb}+1}$ and an informative prior where π_k is set equal to the proportion of pig litters of crossbreed k among all pig litters in Norway in 2021 (Langaker et al., 2021). The PLS-QDA soft prediction is given by $\hat{\Delta} = \sum_{k=1}^{n_{comb}} \Delta_k P(K|x)$, where $P(K|\mathbf{x})$ is the posterior probability for class K, and Δ_k is the associated breed proportion vector; see Eq. 1. For the class unknown, we used $\Delta = 0_q$.

PLS-QDA models were fitted and evaluated in RStudio (Posit team, 2023) with custom functions, where the package "mvtnorm" (Genz et al., 2023) was used extensively. Codes are available at GitHub (Gangsei et al., 2023), a repository which also contains codes for replicating results, tables, and figures in the present study. For a more extensive exploration of the classification results presented in this article, an R-Shiny app has been made available (Gangsei, 2023).

3 Results

3.1 Comparing PCA and PLS

For visualization of the data, both PCA and PLS were conducted on the TrainP+ data. The results are shown in Figure 2. For both PCA and PLS, first and second components both split the three breeds Duroc, Landrace, and Pietrain in a similar way. Component 3 manages to separate Large White from the other breed classes for both PCA and PLS approaches. The difference lies in the last, small breed ($n = 14$) Hampshire. The fourth component from the PCA mainly spans within variation of Pietrain and Hampshire, i.e., the breed with a small sample size cluster from the other breeds but does not represent its own node point in the four-dimensional space spanned out by the first four PCA components. In contradiction, the fourth component from PLS manages to distinguish this pig breed with its own node point even with a small sample size. As CBs are regarded fuzzy sets, each PB should represent node points in the $m = q - 1$ dimensional space spanned by the scores in order to prevent equal center points (μ) for different breed combinations. PCA fails to incorporate this prerequisite for the unbalanced dataset, and hence, PCA is not included in further analysis. The total variance explained (R^2) by the first four components in TrainP+ was 34.4% and 34.9% of the X-matrix for PLS and PCA, respectively. For both PLS and PCA, $R^2 > 99.1\%$ for all PBs except Hampshire, which had $R^2 = 90.6\%$ and 5.0% for PLS and PCA, respectively. For TrainP−, the cumulative R^2 values for three components were 32.1% and 32.5% (for X with PLS and PCA), $> 99.2\%$ for all PBs except Hampshire with $R^2 = 91.1\%$ and 5.9% for PLS and PCA respectively.

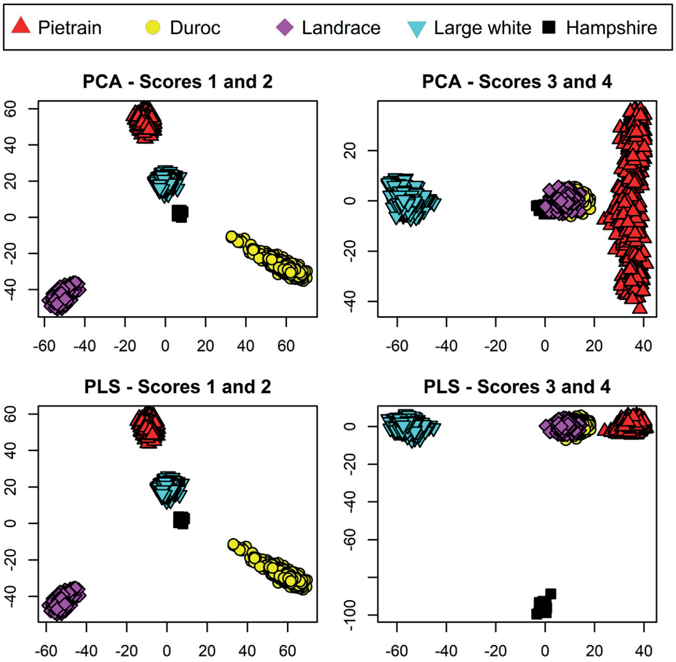


FIGURE 2
Score plots for pure breed animals based on the four primary scores using principal component analysis (PCA) displayed in the upper panels and partial least squares (PLS) displayed in the lower panels. Pietrain pigs are represented in red, Duroc in yellow, Landrace in magenta, Large White in cyan, and Hampshire in black.

3.2 Classification of simulated data

3.2.1 Comparing methods

Table 1 displays an overview of the classification results based on the simulated data, for soft and hard classification with both the Kullback–Leibler divergence and the square loss. In general, PLS-QDA performed best as it managed to correctly classify 6,901 and 6,892 (KL-dist and Sq-loss, respectively), or 99% average, out of the total of 7,000 individuals when trained with Pietrain (TrainP+) and tested for all combinations (TestP+ and TestP−). Similar results for ADMIXTURE and PLSR are 6,744 and 6,826 (97% average), and 6,768 and 6,840 (97% average), respectively. RF, by far, performed worst as it only correctly classified 458 and 2073 (17% average) individuals.

Soft classification results for different methods are visualized in Figure 3. The figure shows results based on all breed combinations except Hampshire in the first column. All methods have best precision for PBs, i.e., PB proportion for breed j ($\delta_j = 1$), or when PB is not present at all, i.e., $\delta_j = 0$. For δ_j, at 0.25, 0.5, and 0.75, the classification precision decreases with increasing δ_j for PLS-QDA, indicating that PB proportions of 0.75 are most poorly classified with PLS-QDA. Hampshire results are of particular interest as only 14 individuals were present in the training data compared to 1,000 individuals for the four other breeds. Even if RF performs poorest overall, it is more noticeable for Hampshire than the other breeds as the Hampshire proportions are heavily underestimated by RF. To some extent, this is also the case for PLSR, while ADMIXTURE and PLS-QDA seem to yield unbiased estimates also for Hampshire proportions.

3.2.2 Effect of omitting breed from training data

All methods performed well when trained and tested on their respective “alike” datasets, as well as when trained with Pietrain and tested without, as shown in the second row in Table 1.

Contrary to prior expectations, there do not appear to be large differences in the classification precision for TestP− depending on whether Pietrain is included in the training data (TrainP+) or not (TrainP−). When evaluated as the proportion of correct classified individuals, the results are marginally better with TrainP+ compared to TrainP−, in particular, for ADMIXTURE and PLSR. However, the average Kullback–Leibler divergence and square loss are, in general, marginally smaller with TrainP− compared to TrainP+. Consequently, including Pietrain in the training data does not seem to impair the model’s classification ability, even for CBs without Pietrain.

3.2.3 Pietrain regarded as an unknown breed

To get an understanding of how inclusion of unknown breeds, i.e., breeds not present in the training data, in CB combinations

TABLE 1 Evaluation of prediction errors for soft predictions as mean ± standard deviations and hard predictions given as numbers of correct classifications as the total number and % (*n* = 3,500 for TestP+ and TestP−) based on Kullback–Leibler divergences and quadratic errors, crossed over the two training (TrainP+ and TrainP−) sets and test sets (TestP+ and TestP−).

Combination	Method	RF	ADMIXTURE	PLSR	PLSQDA
TrainP+ TestP+	KL-dist (soft)	0.31 ± 0.2	0.072 ± 0.088	0.038 ± 0.038	0.0015 ± 0.013
	KL-dist (hard)	300 (8.6%)	3,326 (95%)	3,438 (98%)	3,458 (99%)
	Sq-loss (soft)	0.061 ± 0.033	0.01 ± 0.01	0.0084 ± 0.0093	0.0016 ± 0.012
	Sq-loss (hard)	1,208 (35%)	3,404 (97%)	3,424 (98%)	3,455 (99%)
TrainP+ TestP−	KL-dist (soft)	0.46 ± 0.24	0.053 ± 0.063	0.064 ± 0.08	0.0024 ± 0.016
	KL-dist (hard)	158 (4.5%)	3,418 (98%)	3,330 (95%)	3,443 (98%)
	Sq-loss (soft)	0.075 ± 0.041	0.0073 ± 0.0092	0.0084 ± 0.0096	0.002 ± 0.015
	Sq-loss (hard)	865 (25%)	3,422 (98%)	3,416 (98%)	3,437 (98%)
TrainP−TestP+	KL-dist (soft)	12 ± 16	12 ± 16	12 ± 16	9.6 ± 16
	KL-dist (hard)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	Sq-loss (soft)	0.29 ± 0.26	0.28 ± 0.29	0.26 ± 0.27	0.37 ± 0.24
	Sq-loss (hard)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
TrainP−TestP−	KL-dist (soft)	0.3 ± 0.18	0.048 ± 0.065	0.056 ± 0.081	0.0024 ± 0.017
	KL-dist (hard)	595 (17%)	3,400 (97%)	3,208 (92%)	3,445 (98%)
	Sq-loss (soft)	0.05 ± 0.035	0.0077 ± 0.0098	0.0085 ± 0.0098	0.0025 ± 0.021
	Sq-loss (hard)	1,627 (46%)	3,408 (97%)	3,406 (97%)	3,434 (98%)

affects the classification results for PLS-QDA, results for the combination TrainP− TestP+ were evaluated and are presented in Table 2. For this combination, Pietrain might be regarded as an "unknown breed" and, ideally, all individuals in TestP+ should be classified as "unknown" for the model based on TrainP+. In Table 2, "P" is still the abbreviation for Pietrain, but the combinations are generalized, i.e., "XXXX" = "DDDD," "HHHH," "LLLL," "WWWW" (PB), "XXXY" = "DDDH," . . ., "WWWL," i.e., three grandparents of the same breed, etc. Most of the CB blends with Pietrain are recognized and classified as an "unknown" breed. PB Pietrain ("PPPP"), almost PB Pietrain ("PPPX"), and the CB of one Pietrain together with three of the four other classes ("PXYW") are more likely to be classified as a mix of all the other four PB combinations ("XYLK") with 100%, 46%, and 32% classified as "XYLK," respectively. A likely explanation is that the center point for "DHLW" is close to origo, with a large associated generalized variance.

3.3 General PLS-QDA results

The Bayesian method PLS-QDA, which might be regarded as a refinement of PLSR, performs best and also has more flexibility in its modeling, e.g., it can easily be implemented with an unknown breed combination. The support for classifying to "unknown" can be adjusted by changing the associated prior probability $\pi_{Unknown}$. Changing $\pi_{Unknown}$ has some similarities to changing the significance level for hypothesis testing. If the support for an unknown breed is lowered, i.e., $\pi_{Unknown}$ is decreased, fewer individuals will be classified as "unknown breed," i.e., the probability of classification to a "real CB class" is increased, with the disadvantage that the probability of incorrect classification is increased. The analogy to hypothesis testing is that the higher significance level increases the probability of rejecting the null hypothesis but also increases the risk of doing a type I error. Consequently, if it is of huge importance to avoid incorrect classifications, $\pi_{Unknown}$ should be increased and vice versa.

Another feature which is unique for PLS-QDA is the possibility to use an informative prior. The effect of the informative prior on classification results for the real data, TestR, with PLS-QDA, trained on TrainP+, is shown in Table 3. The effect of the informative prior is conspicuous and as expected *a priori*. The number of individuals classified to the two dominating CB classes i.e., "DDLW" and "HHLW," i.e., typically F1 commercial ("LW") maternal line and Duroc ("DD") or Hampshire ("HH") paternal lines, increases, in particular, at the expense of the number classified as "unknowns." PLS-QDA with an informative prior and ADMIXTURE yields close to similar results for the real data, with one exception, the "unknown group," which is natural since classifying to "unknown" is not a feature in ADMIXTURE.

Estimates for μ_j and Σ_j in Eq. 2 are PB-specific means and covariances based on the PLS scores. The PLS scores for PBs used for these calculations are shown in the two lower panels in Figure 2. μ_j and Σ_j for CBs are linear combinations of μ_j's and Σ_j's for PBs, as shown in Eq. 5.

A crucial success factor for PLS-QDA is the incorporation of additional variance in CB covariance due to the stochastic nature of the proportion of DNA inherited from grandparents. Table 4 shows average matrix determinants at the log scale for the theoretical covariance matrices, as given in Eqs 4, 5. The size of the

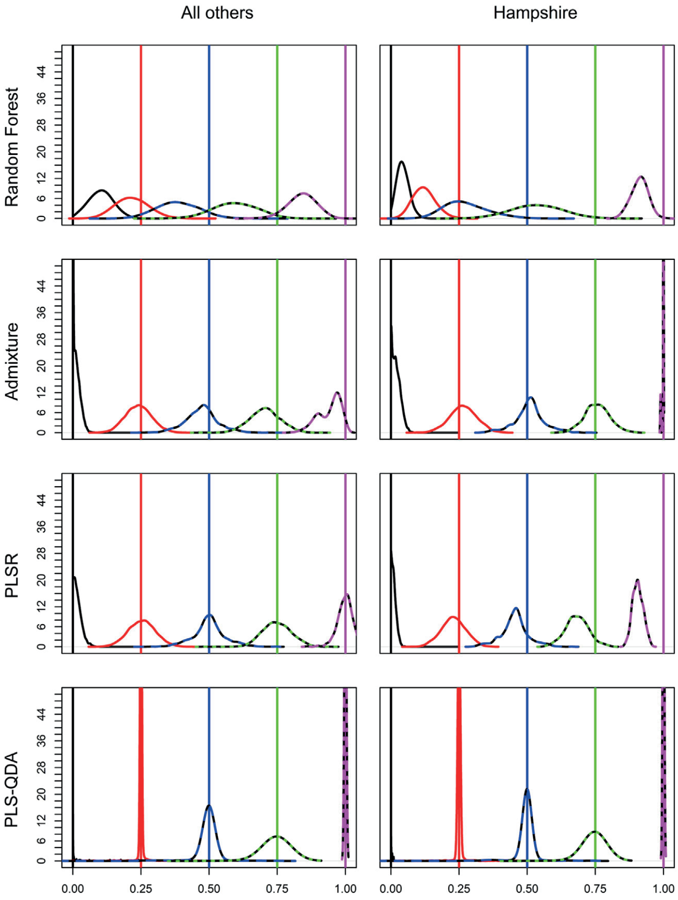


FIGURE 3
Densities for elements in the soft classifications ($\hat{\Delta}$). The left column represents results obtained from breeds Duroc, Landrace, Pietrain, and Large White, and the right column represents results from Hampshire. The rows represent the four different methods, i.e., RF, ADMIXTURE, PLSR, and PLS-QDA. The curves show empirical densities for $\hat{\delta}_j$ for δ_j = 0/4, 1/4, . . ., 4/4, with colors black, red, blue, green, and magenta, respectively. Densities are based on results from models trained on the training set TrainP + and applied to both simulated test sets, i.e., TestP+ and TestP–. Vertical lines represent the true proportions, i.e., δ_j.

determinant of the covariance matrix is referred to as the generalized variance by Wilks (1932) and might be viewed as a scalar describing the size of the covariance matrix in question. The generalized variance increases with an increasing number of PBs in the CB combination. For comparison, the log scale determinants of empirical variances from predicted scores, i.e., $\mathbf{t}^{sim} = \mathbf{X}^{sim}\mathbf{P}$, where $\mathbf{X}^{sim}$ is the matrix of simulated SNPs, are shown in the same table. The results show determinants of the same size for both training sets, except for PBs where the variance based on empirical scores is smaller.

The behavior of classification results, center points (μ), and associated covariance (Σ) for CBs of different complexities and different settings for the informative prior and support for unknown breed ($\pi_{Unknown}$) can be explored in the R-Shiny app (Gangsei, 2023).

4 Discussion

The overall aim of this study was to evaluate crossbreed classification of commercial finisher pigs based on genomic data from a 50-K (Illumina) SNP chip, with the four different methods, namely, RF, ADMIXTURE, PLSR, and PLS-QDA. The novelty was to implement PLS-QDA as an alternative method with several beneficial features to analyze the genomic SNP data.

PLS was used as an alternative dimension reduction method to PCA due to its additional features. Subsequent theoretical deductions led to the extended method, PLS-QDA. For comparison, two methods not built on dimension reduction were also executed: the model-based ADMIXTURE, which is a well-functioning software application for ancestry classification, and the

TABLE 2 Proportions (in %) of the predicted breed combination, i.e., maximum posteriori probabilities with PLS-QDA, for the model trained without Pietrain (TrainP−) vs. true breed combinations for the simulated test set with Pietrain (TestP+). Breed combinations are generalized, i.e., "XXXX" = {*"DDDD," "HHHH," "LLLL," "PPPP," "WWWW"*} (PB), "XXXY" = {*"DDDH"*, . . ., *"ZZZL"*}, i.e., three grandparents of the same breed, etc.

		Predicted combination (%)						
		Unknown	XXXX	XXXY	XXYL	XXYY	XYLK	n
True combinations	PPPP	0	0	0	0	0	100	100
	PPPX	54	0	0	0	0	46	400
	PPXX	100	0	0	0	0	0	400
	PPXY	80	0	0	0	0	20	600
	PXXX	66	28	6	0	0	0	400
	PXXY	68	0	17	8	6	0	1,200
	PXYL	41	0	0	27	0	32	400

TABLE 3 Predicted breed combinations for the test set containing real data based on models trained on all data (TrainP+). Predictions as posterior maximums applying the informative and flat prior to the PLS-QDA method and hard predictions based on the square loss for ADMIXTURE. The last columns show differences in the total number classified to different CBs for PLS-QDA with and without informative priors and ADMIXTURE. The group "other combinations" contains sums for breed combinations with fewer than 10 predictions for any of the methods.

	Number of breed combinations			Difference between methods		
	Flat	Informative	ADMIXTURE	Inf. vs. Adm	Adm. vs. Flat	Inf. vs. Flat
DDLW	497	532	531	1	34	35
HHLW	169	192	193	−1	24	23
Unknown	118	70	0	70	−118	−48
LLWW	41	44	46	−2	5	3
DDLL	43	43	53	−10	10	0
LLLL	38	39	41	−2	3	1
LLLW	23	24	22	2	−1	1
HHLL	17	16	36	−20	19	−1
DLPW	16	0	0	0	−16	−16
LWWW	10	0	11	−11	1	−10
HLLW	0	0	14	−14	14	0
Other combinations	41	53	66	−13	25	12

well-known classification method RF. The classification results, for all methods on the two simulated test sets with and without an additional breed, show that PLS-QDA had the highest accuracy and that PLSR and ADMIXTURE are both methods that meet prior expectations to classification accuracy.

It was observed, naturally, that all methods have best precision for the classification of individual elements, δ_j in Δ, when the element is associated with a PB, i.e., $\delta_j = 1$, or when the PB is not present at all, i.e., $\delta_j = 0$. For PLS-QDA, it was observed that the classification precision for δ_j was decreased at 0.25, 0.5, and 0.75, indicating that PB proportions at 0.75 are most poorly classified with PLS-QDA. A similar pattern is not evident for ADMIXTURE; however, the results for PB proportions at 0.75 are approximately equal for PLS-QDA and ADMIXTURE.

PLS-QDA has some advantages compared to ADMIXTURE and PLSR. First, it performs best when tested on the simulated data, even if this is by small margins, but more important is its ability to incorporate an unknown breed combination. The results presented in Table 2 show that the PLS-QDA method is capable of classifying CBs with Pietrain grandparents as "unknowns" to a large degree. The exception is PB Pietrain, almost PB Pietrain ("PPPX"), and the CB of one Pietrain together with three of the four other classes ("PXYW"), where many individuals were classified as a mixture of the four other PBs. Consequently, classification results with four different PBs should be interpreted with caution as it may be an unknown purebred not seen in the training set. Some misclassifications of CBs with other levels of Pietrain grandparents also occur; however, in general, the method

TABLE 4 Mean ± standard deviation for covariance matrix determinants using a logarithmic scale, crossed over generalized breed combinations i.e., "XXXX" = {*"DDDD," "HHHH," "LLLL," "PPPP," "WWWW"*} (PB), "XXXY" = {*"DDDH,"* . . ., *"ZZZL"*}, i.e., three grandparents of the same breed, etc., and if Pietrain is included in the analysis or not. The column "mixed normals" is based on Σ in the likelihood function, and the column "Simulated data" is based on empiric covariances from simulated data.

	Generalized breed combination	Mixed normals	Simulated data
P+	XXXX ($n = 5$)	5 ± 1.3	2.4 ± 0.63
	XXXY ($n = 20$)	5.8 ± 0.72	5.4 ± 0.48
	XXYL ($n = 30$)	7.1 ± 0.38	7.8 ± 0.45
	XXYY ($n = 10$)	5.5 ± 0.45	6 ± 0.36
	XYLK ($n = 5$)	9.1 ± 0.25	9.7 ± 0.26
P−	XXXX ($n = 4$)	4 ± 1.2	2 ± 0.51
	XXXY ($n = 12$)	5.3 ± 0.62	4.8 ± 0.48
	XXYL ($n = 12$)	7 ± 0.3	7.3 ± 0.19
	XXYY ($n = 6$)	5.2 ± 0.45	5.5 ± 0.39
	XYLK ($n = 1$)	9.3 ± NA	8.7 ± NA

performs reasonably well for these combinations, in particular for individuals with 50% Pietrain grandparents.

Another advantage PLS-QDA has in comparison to the other methods is the possibility to use different prior distributions for the CB populations. These priors might change in time and space, for instance, in other target populations, e.g., countries with other dominating breed combinations. By assigning high prior weight to the unknown breed group, more individuals will be classified as unknowns at the expense of the known CBs. As argued in the results, the interpretation of the prior weight for an unknown breed has similarities to the interpretation of the significance level in hypothesis tests, i.e., a higher significance level/lower prior for unknown breed not only leads to increased strength for classifying a known CB/reject the null hypothesis, but also an increased possibility of misclassification/type I errors. This is a desirable feature enabled by the Bayesian nature of QDA and, to the best of our knowledge, a novelty in classification of crossbreed pigs.

A main result is that, at least for unbalanced data, PLS is preferable over PCA as PLS fulfills the prerequisite of assigning one node point in the $q-1$-dimensional space to each PB, which is, thus, a necessity for classification purposes and, in addition, is a considerable advantage for visualization. This is illustrated in Figure 2 where PLS assigns one PB to each node point in the $q-1$-dimensional space defined by the scores. This is a prerequisite for the PLS and PLS-QDA methods as CBs are considered blends in a space where the PBs represent the extremes. For purely practical purposes, this might have been solved by having a more balanced dataset, i.e., the same number of Hampshires as for the other breeds. However, the insight has significance since new breeds might fairly easily be added to the model without the need for observations from a large number of individuals when PLS is used.

It should be recognized that PLS, at least for genomic data such as the data used in the present study, fulfills the need for dimension reduction, with better results than PCA in the sense of more dense and dispersed clusters of PBs for the first m PCA/PLS components. As a key finding of the present study, we highlight that PLS might be used as a complementary method for dimension reduction of SNP data under the assumption that a "supervising" feature in the present study "breed," is available.

An important feature of a general method is the stability when exposed to new or unknown breeds that are not included in the PB F0 generations. The results show that both PLS-QDA and PLSR, as well as ADMIXTURE, are flexible in the sense that new PBs might be added to the training data without substantial loss with respect to classification accuracy, even if some of the included PBs are not present in the target population. As strongly anticipated, classification performance for breeds not included in training data was poor. However, the results showed marginal deterioration when the method was trained with the new breed, TrainP+, and classified without, TestP−. Hence, the disadvantage of training on a variety of PBs is small even if the possible crossbreed combinations are well known *a priori*. Consequently, for all methods, other PBs might be included in the training data, with small or even negligible loss of classification precision in populations where one or more of PBs is not present. The possibility of adding new PBs to the model without the need for a large sample size for PBs in question is a highly desirable feature for PLSR, PLS-QDA, and ADMIXTURE.

RF performed, by far, the poorest for classification of the simulated data, in particular for CBs, including Hampshire. This is in line with prior assumptions as Hampshire was hugely underrepresented in the training data. RF is built on tree prediction, where a considerable number of trees collectively favor the class with the highest probability. With few observations in the training datasets, Hampshire will most likely not be included in the training of all tree models, and therefore, RF will give a skewed result and suffer when presented with a small class in the test data. RF performance depends on the tuning of hyper parameters. In the present study, hyper parameters were tuned to values giving OOB errors at 0 in the training set and at the same time yielded small computational cost. It is not unlikely that RF performance could be improved more by extensive tuning of the hyper parameters based on the model's performance on the simulated test data. However, as models for all methods were fitted based on training data only, the same principle should be applied to RF. The results show that for the present study, RF is a sub-optimal classification method, most likely due to the unbalanced data structure. This could also be adjusted in favor of RF if operated with more balanced data, but as argued previously, it is beneficial with methods that perform well on unbalanced data and balanced for generalization purposes. The three other methods, at least PLS-QDA and ADMIXTURE, seem to be robust against the unbalanced training dataset and without extensive hyper parameter tuning.

Two different measures, namely, square loss and Kullback–Leibler divergence, were both used for two purposes: comparing the accuracy of soft predictions between methods, and transformation of soft predictions to hard predictions for RF, ADMIXTURE, and PLSR. Formally, the transformations from soft to hard predictions based on the square loss and Kullback–Leibler divergence are just discriminant classification performed on the soft predictions, $\hat{\mathbf{\Delta}}$, with a flat prior for CBs. The square loss corresponds to LDA, and Kullback–Leibler is an alternative discriminant function. The results from both methods are quite similar but differ marginally. We view it as a strength that the evaluation of the results seems to be affected to a negligible degree by the choice of distance measurement.

The purely practical applications for the models included in the present study are limited to breed and breed combinations for the five PBs included, i.e., Duroc, Hampshire, Landrace, Large White, and Pietrain. However, through the results and principles, we show that both ADMIXTURE and PLSR/PLS-QDA are methods where other PBs might be fairly easily included, even when PB data for new breeds are scarce, which are key findings in the study. Another limitation to the study is the unbalanced training set, containing only 14 pigs of the breed Hampshire. This has been seen as an opportunity to evaluate the methods in a more realistic setting than a balanced dataset would provide. Therefore, it has been kept this way intentionally instead of pruning the data by, for instance, taking out Hampshire as PB.

In the present study, breed combinations, not breed permutations, were used as classifying units. Variation in the percentage of DNA material inherited from PB animals is affected by different breed permutations under the same breed combination. By only considering combinations, potential information associated with different permutations might be lost. For example, consider the combination *"LLWW"* consisting of the six permutations *"LLWW"*, *"LWLW"*, *"LWWL"*, *"WLLW"*, *"WLWL"*, and *"WWLL"*. When calculating mean (μ_j) and covariance (Σ_j) for CBs (see Eq. 5), the assumption is that the proportion of DNA inherited from grandparents, i.e., θ, was Dirichlet distribution with the concentration parameter $\alpha_0\Delta$, leading to $V(\theta) = \frac{1}{\alpha_0+1}\left(d(\Delta) - \Delta\Delta^t\right)$, where Δ represents the breed combinations in the F0 generation. This assumption seems reasonable for all permutations; however, for the two permutations *"LLWW"* (F1 commercial maternal line) and *"WWLL"*, both with two PB individuals in the F1 generation, we know that the proportion of DNA inherited from the two PBs in question is 50% exactly, which is not the case for the other four permutations. Consequently, for "*LLWW*" and "*WWLL*", $\theta = \Delta = [0\ 0\ ½\ ½\ 0]^t$ (i.e., zero variance for θ). For the four other permutations, it is natural to assume non-zero variance for θ. The consequence, referring to Eq. 5, is that covariance, Σ_j, associated with permutations *"LLWW"* and *"WWLL"* should be smaller than the other permutations, as all elements including $V(\theta)$ in Eq. 5 should be excluded for these permutations. This information might be possible to utilize in order to, at least to some extent, distinguish different breed permutations under the same breed combination. However, the strength of classifying different permutations is likely to be low as the means, i.e., μ_j in Eq. 5 are unaffected by $V(\theta)$. The effect of permutation clustering within combinations is easy to observe for real data classified as breed combination *"HHLZ"* in the 3D Shiny app (Gangsei, 2023). These individuals cluster inside their associated limiting spheres. From prior information, it is overwhelmingly likely that the only permutation existing within this combination is the crossing of the Hampshire paternal and TN70 maternal lines, which also highlights that permutations might be identified by an informative prior.

The software program used for the simulation of data in the study was developed with a primary area of application for the human genome, in particular kinship analyses and forensic genetics. Due to the genetic map provided by Tortereau et al. (2012), it was possible to apply the software application to the pig genome in a realistic manner. The genetic map is averaged over sex and four different breed combinations ("pedigrees") containing PBs Large White, Meishan, Yorkshire, Berkshire, Duroc, and Landrace. The recombination rates varied between breed combinations and sexes (Tortereau et al., 2012). Consequently, the use of an average genetic map in the present study is an approximation. However, the effect of variations in the genetic map is assumed to be of minor importance as it will only have limited effects on the parameter α_0 scaling the variance of the proportion of the inherited genomic material from the four grandparents (θ). Higher recombination rates would yield larger values for α_0 and lower variance for θ. In the present study, α_0 was kept constant at its estimated value at 73.58. A possible topic for future research is to evaluate the effects of changing this value and thereby the covariance matrices for CBs.

A challenge with the SNP data is that they only contain information regarding the two nucleobases that are present at each SNP but no information regarding whether the nucleobases originate from the paternal or maternal line. For homozygote SNPs, this data structure causes no problems. For heterozygote SNPs, the two nucleobases were randomly assigned to the maternal or paternal chromosomes of the F0 generation when assigning the nucleobases to a simulated IBD chromosome structure.

For the simulation study, breed permutations were drawn randomly within each breed combination. A topic for future studies might be to design simulations for different breed permutations and apply a classification model for permutations based on theoretically different variances, in order to classify permutations within the same combination. Such studies would, to the best of our understanding, be of more theoretical than practical interest.

The simulation study provides SNP simulations for CBs and behaves as a credible realization for SNP data in real CB individuals. This is supported, although not proved, by the fact that simulated data are distributed in accordance with the model, both regarding expected values and variances. Consequently, it is reasonable to assume that the evaluation of methods based on the simulated results, to a great extent, describes the real precision and reliability for different methods and breed combinations. To explore how PLS-QDA and ADMIXTURE behave when applied to a real example, the trained PLS-QDA and ADMIXTURE models were tested on real data, TestR. The distribution of CB classes was in accordance with prior knowledge, i.e., the dominating CBs were *"DDLW"* and *"HHLW"*, even when using the flat prior for PLS-QDA. The flat vs. informative prior results for the PLS-QDA method appear to be a textbook example of how an informative prior might be utilized in a Bayesian setting. The inclusion of the informative prior has a substantial effect by allocating more individuals to breed combinations known to be dominating *a priori*, at the expense of the "unknown" class and CBs known to be rare *a priori*. Still, the informative prior does not totally dominate the classification results. ADMIXTURE classifies closer to PLS-QDA with informative priors, which may indicate that the method is adequate in adjusting for actual populations. The prior information can neither be added nor changed. This result again advocates for ADMIXTURE as a reliable method for classification. It could be interesting to see how the two different methods, PLS-QDA with informative priors and ADMIXTURE, behave on real data from other real situations with other CB combinations.

Some of the real data are classified to CBs containing Pietrain, even if Pietrain should not be present in the Norwegian pig population. The CBs with Pietrain are *"DLPW"* and *"HLPW"*, i.e., four breed combinations. Inclusion of some genetics of Pietrain origin cannot be totally ruled out in Norway; however, for the last 15–20 years, the policy of breeding companies operating in Norway has been to avoid using Pietrain genetics. From the simulation results, we observed that PB individuals from different breeds were not part of the training set and were generally classified as four-breed combinations. Consequently, the four-breed classification results should be interpreted with care as they might, in fact, be PBs or close to PB individuals, from breeds not included in the training dataset. In Norway, at least the Mangalica breed is present and, in fact, a possible candidate for these classifications. A natural development of the work presented in this study would be to incorporate Mangalica as a new PB in the training data.

The results in this study can beneficially be used for generalization to other problems in several ways. The simulation tool showed an excellent generalization from humans to pigs and can be generalized to other breeds/populations/countries or to other species with genomic data available in the form presented in the current study. Prior knowledge of recombination rates, i.e., the study of Tortereau et al. (2012), was essential for the present study, both in the simulation and in order to estimate $V(\theta)$ and thereby Σ. If similar information regarding recombination rates is available, the methods described in the present study might be transferred to similar problems for other species, assuming that genomic data are available.

Another interesting topic, which falls outside the scope of this study, is to consider other responses than breeds. For instance, a feature such as color could be treated in a similar way, where some colors are viewed as references, i.e., the counterpart to PBs in the current study and other color combinations as blends, i.e., the counterpart to CBs.

ADMIXTURE and RF were tested as possible candidate methods. Other candidates could also have been included, for instance, different classification methods that deal better with unbalanced data. Although RF failed as a real candidate, ADMIXTURE performed well for both simulated data and real data. Thus, the result of this study confirms ADMIXTURE's suitability as a standard software program for classifying genetic origins, not only for human ancestry. Kim et al. (2022) indicate how ADMIXTURE, in combination with PCA, behaves nicely and provides useful information for both classification and visualization in a pig population.

Partial least squares with linear discriminant analysis (PLS-DA) has recently been shown to perform well on other problems with similar SNP data (Miao et al., 2023). The derivation of PLS-QDA for CBs was initiated and conducted prior to the publication by Miao et al. (2023). However, PLS-QDA might be viewed as an elaboration of PLS-DA utilized in Miao et al. (2023) in the sense that i) PLS-QDA was applied to CB classification in contrast to PLS-DA used for PB classification only, and ii) the derivation of CB-specific covariance matrices is a prerequisite for QDA and novel to the present study. In particular, a research topic for further analysis could be to apply PLS-QDA and simulation of CBs to the data used in Miao et al. (2023) where the number of PBs was much higher than that in the present study ($n = 91$).

The main focus of the study was to evaluate the PLS-based methods, in particular to derive equations for the expected values μ's and covariances used in the likelihood functions for CBs. Another important objective was to show that PLS extends and improves classification in a more robust way for unbalanced data and when faced with unknown breed combinations, which is a reality when working with real data from slaughterhouses.

5 Conclusion

In the present study, it has been shown that PLS-QDA, PLSR, and ADMIXTURE are well suited methods for the crossbreed classification of pigs based on genomic data from a 50-K (Illumina) SNP chip from purebred grandparents. ADMIXTURE is a well-proven method that is suited for ancestry classification tasks with genetic SNP data. It originates from kinship in humans but proved to work nicely and was easy to transfer to pigs. The method of the main focus in the present paper, PLS-QDA, has some advantages compared to the other methods. It has the highest classification accuracy, which supports the inclusion of an "unknown breed combination" class and an informative prior. Finally, it facilitates informative visualization in 3D format. Accurate CB classification has important applications, in particular, related to research and development topics in the pig industry, including breeding progress, carcass grading, meat yield, and quality. Another important contribution from the current study is to incorporate the stochasticity in the proportion of inherited DNA from ancestors as a feature utilized for PLS-QDA as an extension of PLS-DA.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the studies involving animals in accordance with the local legislation and institutional requirements because data used in the study were from two sources: 1. SNP data from purebred animals were already collected and stored at Topigs Norsvin, and they were not collected specifically for the present study. 2. The SNP data for crossbreeds from Animalia AS were taken from commercial pork-meat/carcasses, i.e., animals already slaughtered. Written informed consent was not obtained from the owners for the participation of their animals in this study because the SNP data from both Topigs Norsvin and Animalia AS are from animals/carcasses owned by themselves.

Author contributions

HV: conceptualization, formal analysis, methodology, visualization, writing–original draft, and writing–review and editing. HB: data curation, methodology, visualization, and writing–review and editing. AH: writing–review and editing. CS: formal analysis, writing–review and editing, and methodology. MS: conceptualization and writing–review and editing. LG: conceptualization, formal analysis, methodology, project administration, software, visualization, writing–original draft, and writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. AH was funded by The Research Council of Norway, grant number 311394. All other authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to thank BioBank AS (Hamar) for sample handling and DNA extraction, the Centre for Integrative Genetics (CIGENE) at the Norwegian University of Life Sciences, for providing laboratory facilities, and Hanne Hamland at Norsvin for performing genotyping. They would also like to thank Norsvin and Topigs Norsvin for providing genotypes from purebred animals.

Conflict of interest

Authors AH and LG were employed by the company Animalia AS. Author MS was employed by the company Norsvin SA.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alexander, A. H., Shringarpure, S. S., Novembre, J., and Lange, K. (2020). *ADMIXTURE: fast ancestry estimation. ADMIXTURE version 1.3.*

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr.094052.109

Boulesteix, A.-L. (2004). PLS dimension reduction for classification with microarray data. *Stat. Appl. Genet. Mol. Biol.* 3, Article33. doi:10.2202/1544-6115.1075

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

Brisbin, A., Bryc, K., Byrnes, J., Zakharia, F., Omberg, L., Degenhardt, J., et al. (2012). PCAdmix: principal components-based assignment of ancestry along each chromosome in individuals with admixed ancestry from two or more populations. *Hum. Biol.* 84, 343–364. doi:10.3378/027.084.0401

Chinchilla-Vargas, J., Bertolini, F., Stalder, K., Steibel, J., and Rothschild, M. (2021). Estimating breed composition for pigs: a case study focused on mangalitsa pigs and two methods. *Livest. Sci.* 244, 104398. doi:10.1016/j.livsci.2021.104398

Dadousis, C., Muñoz, M., Ovilo, C., Fabbri, M. C., Araújo, J. P., Bovo, S., et al. (2022). Admixture and breed traceability in European indigenous pig breeds and wild boar using genome-wide SNP data. *Sci. Rep.* 12, 7346. doi:10.1038/s41598-022-10698-8

Dickerson, G. E. (1973). Inbreeding and heterosis in animals. *J. Animal Sci.* 1973, 54–77. doi:10.1093/ansci/1973.Symposium.54

Gangsei, L. E. (2023). *Crossbreed classification of pigs.* Available at: https://larserikgangsei.shinyapps.io/appdir/.

Gangsei, L. E., Bjerke, F., Røe, M., and Alvseike, O. (2018). Monitoring lean meat percentage predictions from optical grading probes by a commercial cutting pattern. *Meat Sci.* 137, 98–105. doi:10.1016/j.meatsci.2017.11.010

Gangsei, L. E., Brustad, H., Heggli, A., Sevillano, C. A., van Son, M., and Vinje, H. (2023). *CrossBreedClassificationPublic.* Available at: https://github.com/LarsErikGangsei/CrossBreedPredictionPublic.

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., et al. (2023). *mvtnorm: multivariate Normal and t Distributions. R package version 1.2-3.* Available at: https://cran.r-project.org/web/packages/mvtnorm/mvtnorm.pdf.

Hamilton, D. N., Ellis, M., Miller, K. D., McKeith, F. K., and Parrett, D. F. (2000). The effect of the Halothane and Rendement Napole genes on carcass and meat quality characteristics of pigs. *J. Animal Sci.* 78, 2862–2867. doi:10.2527/2000.78112862x

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*, 2. Berlin, Germany: Springer.

Huang, Y., Bates, R., Ernst, C., Fix, J., and Steibel, J. (2014). Estimation of us yorkshire breed composition using genomic data. *J. Animal Sci.* 92, 1395–1404. doi:10.2527/jas.2013-6907

Kim, J. A., Cho, E. S., Jeong, Y. D., Choi, Y. H., Kim, Y. S., woo Choi, J., et al. (2020). The effects of breed and gender on meat quality of duroc, pietrain, and their crossbred. *J. animal Sci. Technol.* 62, 409–419. doi:10.5187/jast.2020.62.3.409

Kim, Y.-M., Seong, H.-S., Kim, Y.-S., Hong, J.-K., Sa, S.-J., Lee, J., et al. (2022). Genome-wide assessment of a Korean composite pig breed, woori-heukdon. *Front. Genet.* 13, 779152. doi:10.3389/fgene.2022.779152

Kullback, S., and Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Statistics* 22, 79–86. doi:10.1214/aoms/1177729694

Langaker, M., Haug, D. J., Moen, Ø., Kjølseth, I.-M., and Kongsrud, S. (2021). *Ingris - årsstatistikk 2020 (in Norwegian).* Available at: https://www.animalia.no/globalassets/ingris—dokumenter/arsstatistikk-2020-007—endelig.pdf.

Liaw, A., and Wiener, M. (2022). *randomForest:classification and regression by randomforest. R package version 4.7-1.1.*

Liland, K. H., Mevik, B.-H., Wehrens, R., and Hiemstra, P. (2021). *Pls: partial least squares and principal component regression. R package version 2.8-0*. Available at: https://CRAN.R-project.org/package=pls.

Miao, J., Chen, Z., Zhang, Z., Wang, Z., Wang, Q., Zhang, Z., et al. (2023). A web tool for the global identification of pig breeds. *Genet. Sel. Evol.* 55, 18–12. doi:10.1186/s12711-023-00788-0

Mujibi, F. D., Okoth, E., Cheruiyot, E. K., Onzere, C., Bishop, R. P., Fèvre, E. M., et al. (2018). Genetic diversity, breed composition and admixture of kenyan domestic pigs. *PLoS One* 13, e0190080. doi:10.1371/journal.pone.0190080

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philosophical Mag. J. Sci.* 2, 559–572. doi:10.1080/14786440109462720

Posit team (2023). *RStudio: integrated development environment for R*. Boston, MA: Posit Software, PBC.

Probst, P., Wright, M. N., and Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip. Rev. data Min. Knowl. Discov.* 9, e1301. doi:10.1002/widm.1301

Schleimer, A., Richart, L., Drygala, F., Casabianca, F., Maestrini, O., Weigand, H., et al. (2022). Introgressive hybridisation between domestic pigs (sus scrofa domesticus) and endemic Corsican wild boars (s. s. meridionalis): effects of human-mediated interventions. *Heredity* 128, 279–290. doi:10.1038/s41437-022-00517-1

Sellier, P. (1976). The basis of crossbreeding in pigs; a review. *Livest. Prod. Sci.* 3, 203–226. doi:10.1016/0301-6226(76)90016-6

Sevillano, C. A., Vandenplas, J., Bastiaansen, J. W. M., and Calus, M. P. L. (2016). Empirical determination of breed-of-origin of alleles in three-breed cross pigs. *Genet. Sel. Evol.* 48, 55–12. doi:10.1186/s12711-016-0234-9

Smith, C. (1964). The use of specialised sire and dam lines in selection for meat production. *Animal Sci.* 6, 337–344. doi:10.1017/S0003356100022133

Tortereau, F., Servin, B., Frantz, L., Megens, H.-J., Milan, D., Rohrer, G., et al. (2012). A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. *BMC genomics* 13, 586–612. doi:10.1186/1471-2164-13-586

Tusell, L., Bergsma, R., Gilbert, H., Gianola, D., and Piles, M. (2020). Machine learning prediction of crossbred pig feed efficiency and growth rate from single nucleotide polymorphisms. *Front. Genet.* 11, 567818. doi:10.3389/fgene.2020.567818

Vandenplas, J., Calus, M. P. L., Sevillano, C. A., Windig, J. J., and Bastiaansen, J. W. M. (2016). Assigning breed origin to alleles in crossbred animals. *Genet. Sel. Evol.* 48, 61–22. doi:10.1186/s12711-016-0240-y

Vigeland, M. D. (2021). *Pedigree analysis in R*. United States: Academic Press. doi:10.1016/c2020-0-01956-0

Wientjes, Y. C. J., and Calus, M. P. L. (2017). Board invited review: the purebred-crossbred correlation in pigs: a review of theory, estimates, and implications. *J. Animal Sci.* 95, 3467–3478. doi:10.2527/jas.2017.1669

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* 24, 471–494. doi:10.1093/biomet/24.3-4.471

Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemom. intelligent laboratory Syst.* 58, 109–130. doi:10.1016/S0169-7439(01)00155-1

Yin, S., Li, Z., Yang, F., Guo, H., Zhao, Q., Zhang, Y., et al. (2023). A comprehensive genomic analysis of Chinese indigenous ningxiang pigs: genomic breed compositions, runs of homozygosity, and beyond. *Int. J. Mol. Sci.* 24, 14550. doi:10.3390/ijms241914550

Zadeh, L. A. (1965). Fuzzy sets. *Inf. Control* 8, 338–353. doi:10.1016/s0019-9958(65)90241-x

Paper IV

Evaluation of the relationship between porcine carcase classification, breed and yield of tissue types and ratios of primal cuts

A. *Heggli*[a,b], *O. Alvseike*[b], *F. Bjerke*[b], *H. Vinje*[a], *J. Kongsro*[b], *M. Røe*[b], *L.E. Gangsei*[a,b]

[a] Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway

[b] Animalia, P.O. Box 396 - Økern, NO-0513 Oslo, Norway

Abstract

Porcine carcase classification in Europe is determined by the lean meat percentage of each carcase. Carcases that receive the same classification score can still have morphological variation that affects the potential value of the carcase, such as the difference in the ratios between primal cuts. In addition, finishing pigs are often crossbreeds, and breed combination could potentially affect yield variation outside of the variation explained by classification. The main aim of the study was to examine the ability of EUROP classification (i.e., lean meat percentage, LMP) from optical probes to explain several yield variables including lean meat percentage, fat percentage, bone percentage, and ratio between primal cuts, while taking into account factors such as carcase weight, breed combination, operator, and carcase side. A secondary aim was to examine if and how yield is affected by gender/castration category. Results showed that lean meat- and fat percentage were highly negatively correlated and well explained by LMP from optical probes with R^2 (marginal) at 66% and 67% respectively. Bone and waste percentage (marginal R^2 at

29%) and ratios between primal cuts were only to a small degree explained by LMP and weight, with marginal R^2 in the range 5-15%. Furthermore, results showed that breed combination had substantial, though not large effect on ratios between primal cuts. We conclude that information regarding LMP and weight alone is far from sufficient to predict ratios between primal yields in a meaningful way.

Keywords: pigs, grading, LMY, EUROP, crossbreed

Introduction

Yield can be defined as the ratio, quantity and distribution of tissue types within an animal carcase. Carcases are usually classified as an indirect method of yield measurement. Finisher pigs in Europe are classified (European Commission, 2013) based on one criterion, lean meat percentage (LMP), which is defined as the proportion of the total weight of the carcase that is made up of lean meat weight. LMP is a determining factor in price setting, thus it is important that it accurately reflects the yield the abattoirs get from the carcase. Consequently, LMP and yield are important to producers, abattoirs, breeding companies and the meat industry as a whole.

Legislation (European Commission, 2017) requires that instruments used for estimation of LMP are quality assured through total or partial dissection (Walstra & Merkus, 1996) or computed tomography (CT) (European Commission, 2017; Olsen et al., 2017). LMP is determined by optical probe (Engel et al., 2012; Gangsei et al., 2018) or ultrasound instruments (Font i Furnols & Gilbert, 2009). In Norway, the Hennessy Grading Probe 7 (HGP7) (Hennessy Grading Systems, Auckland, New Zealand. Web page: hennessy-technology.com) is the optical probe used for classifying pigs (Gangsei et al., 2018; Røe, 2021) and the Autofom III (FrontMatec, Kolding, Denmark. Web page: frontmatec.com) is the ultrasound alternative (Brøndum et al., 1998; Font i Furnols & Gilbert, 2009; Røe, 2021). Slaughterhouses in all countries that are subject to the EUROP regulations will have an LMP prediction available for all pork carcases. This prediction should in principle be independent of country and instrument used for grading. Consequently, LMP is a natural predictor to use for carcase yield.

Lean meat-, fat- and bone percentage are common measurements of yield for both bovine and porcine carcases, and are used as proxies for the potential value of a carcase. For porcine carcases in particular, primal cuts (Choi et al., 2018; Marcoux et al., 2003) are common commercial commodities with differing prices. The value of the carcase is especially affected by the ratio between primal cuts (Marcoux et al., 2007), and the value of the primal cuts varies over time, and by season. For this reason, the ability to predict weight of primal cuts has been studied (Choi et al., 2018; Dorleku et al., 2023; Lisiak et al., 2014), but the prediction of percentage primal cuts (Janiszewski et al., 2018) seems to be sparse, though it is important as it contributes to the value of the carcase.

As the monetary value of both primal cuts and pig carcases as a whole can vary over time and across markets, it is better to primarily examine the relationship between carcass grading and yield measurements, as this enables easier comparison between studies. Monetary value at a specific time in a specific market has limited use, and can potentially be calculated additionally or separately. It would be beneficial to include percentage primal cuts as yield measurements in addition to lean meat-, fat- and bone percentage when examining the relationship between classification and yield. The more yield variation relevant to the potential value of the carcase that can be explained, the better, as this would be an indicator that the classification is a better judgement of potential value of the carcase, regardless of market.

In the present study, yield is defined by seven variables, each registered per half of the carcase based on weight data from pig carcases. Three of the yield variables are the lean meat-, fat-, and bone and waste percentage. Waste is a small percentage of each carcase and similarly to bone, it is of little value, and is therefore combined with bone percentage. In addition, primal cut percentages (forepart, loin, belly and hindpart) are used as yield variables.

Animalia's (The Norwegian Meat and Poultry Research Centre) pilot plant processes carcases daily on assignment from the Norwegian market regulation authorities for the purpose of providing information to the regulation authorities, the classification system and other stakeholders. The processing is done using a commercial cutting pattern (CCP). Gangsei et al. (2018) showed that LMP calculated based on CCP is highly correlated to the LMP from manual dissection (MD) ($\rho = 0.88$) and CT ($\rho = 0.91$) specified by the European Commission (2017). It is important to note that in any reference method for calculating LMP, MD included, there are sources of inaccuracy (Nissen et al., 2006; Olsen et al, 2017).

In addition to LMP, carcase weight and gender/castration category (immunologically castrated pigs (I.C.s), gilts, surgically castrated pigs) are potentially important predictor variables for yield as they are both commonly registered variables and are known within the industry to have an effect on yield. It is also important to study how variation in yield is explained by breed, as classification ideally should function the same regardless of breed. Finishing pigs in Norway are mainly crossbreeds whose breed are not registered during production. Recently, in Vinje et al., (2023), genetic data was used to predict the breed combination for individual crossbred finishing pigs in Norway. The use of breed information based on genetic data to determine breed combination of crossbred finishing pigs in studies regarding yield is a novel contribution to the study of porcine yield and should be of benefit to pig breeding companies as well as abattoirs.

The aim of this study is to evaluate how well lean, fat, and bone and waste yield (%), as well as primal cut yield (%), might be explained by LMP, while taking into account carcase weight, carcase side, the effect of different breed combinations and operators (butchers). A secondary aim is to study how the yield is affected by gender/castration category, i.e., between finisher pigs that were immunocastrated, gilts or surgically castrated.

Material and methods

Data

Data from 3669 finishing pigs were collected from Animalia's pilot plant between 2012 and 2023. The processing of and data collection from pig, lamb, cattle and reindeer carcases is done at the pilot plant (Gangsei et al., 2018; Heggli et al., 2023) for the primary purpose of providing information for the Norwegian market regulation authorities, the classification system, and other stakeholders, and as such the data from the processed finisher pig carcases in this study were not collected specifically for this study. Each carcase was slaughtered and split into left and right halves at the abattoir. Each abattoir registered the LMP of each carcase, either by use of the optical probe Hennessy Grading Probe 7 (HGP7), the Autofom III using ultrasound, or both. In addition to LMP, the abattoirs registered carcase number, slaughter date, category and warm carcase weight, which was adjusted by a two percent reduction, which is the standard adjustment in Norway to approximate cold carcase weight. Once the pilot plant received each carcase, each side was registered and weighed before being cut using a commercial cutting pattern (CCP) as detailed in Gangsei et al., (2018), with one exception. Most carcases had the entire tenderloin removed from one side of the carcase (n = 1972), some from both sides of the carcase (n = 220), and a few did not have the tenderloin removed before sectioning (n = 16). The forepart, middle section and hindpart were each separated into progressively smaller parts. In every step of the process, all parts and trimmings were registered and weighed, until the smallest part weight (SPW) was obtained.

The smallest parts were divided into seven categories; (i) cuts and trimmings that were considered lean meat, (ii) fat, (iii) bone, (iv) cuts and trimmings including fat that were scanned to measure their fat percentage, (v) cuts with meat and fat that were not scanned, (vi) cuts with meat, fat and bone that were not scanned, and (vii) waste. Fat percentage measurements were done using QV500, a NIR scanner (Tomra Sorting Solutions, Asker, Norway). The trimmings from the forepart were scanned in three groups, two groups from the upper and one group from the lower forepart. The middle section trimmings were all scanned together to ensure enough trimmings to get an accurate measurement, and the same applied to trimmings from the hindpart. The weight of one side of the carcase minus the feet and jowl was compared to the sum of the SPW registrations as part of the data screening process. The weight of the side of the carcase minus the feet and jowl with be referred to as side half carcase weight. Any half carcase side that had an absolute difference of more than two percent between the sum of SPWs and corresponding side half carcase weight was excluded from further analysis. SPWs were aggregated into four categories (the forepart, loin, belly and hindpart) to calculate the weight of each primal cut. The primal cut weights were used to calculate the percentage each primal cut constituted of its corresponding half carcase, i.e., primal cut weight divided by side half carcase weight. This variable will be referred to as primal cut percentage (PCP).

The SPWs were also aggregated into total lean weight, fat weight and "bone and waste" weight per half carcase. The weight of cuts and trimmings that were scanned each had their registered weight multiplied by the measured fat percentage to calculate the fat content, or multiplied by one minus the fat percentage to calculate the lean content. The lean weight of these cuts and trimmings was added to the weight of other SPWs that were lean meat to calculate the total lean weight of the half carcase. The fat weight of the cuts and trimmings was added to the weight of other SPWs that were fat to calculate the total fat weight of the half carcase. The weight of cuts that consisted of meat and fat, or meat, fat and bone, but not scanned, were separated into meat, fat and bone weight by the same method as stated in Gangsei et al., (2018) and added to the total lean meat, fat, and bone and waste weight of the half carcase, respectively. These aggregated weights were then half carcase lean-, fat-, and bone and waste weight.

The lean-, fat-, and bone and waste weights were each divided by the side half carcase weight to calculate the percentage of the half carcase that consisted of lean, fat and "bone and waste". Any individuals that had no registration or a weight of zero for any of the aggregated variables, or no category information were removed from the data set. This resulted in a post cleaning data set with a sample of 5235 half carcases from 3018 individuals.

Finisher pigs are generally three- or four-way crossbreds, but the breed information for production pigs is not registered through the production chain. To include breed as a potential associated variable in the model, genetic DNA data were used to statistically classify the breed combinations following the PLS-QDA method described in (Vinje et al., 2023) from 693 individuals (1302 half carcases) where the genotyping data was available.

Statistical analysis

In order to explore the relationship between yield in its seven categories (i)-(vii) and predictors, seven mixed effect regression models were fitted, with one yield category at a time as response variable. LMP, weight, the interaction between LMP and weight, gender/castration category, breed and side were used as fixed predictors. LMP and weight were centered, whereas their interaction was standardised to zero mean and unit variance. Operator, individual and breed status were added as random prediction factors in the model. The mixed effect regression model includes two observations for each response for each individual, one for each carcase side. In addition, carcase side (right or left) is added as a fixed effect to measure the systematic difference between sides of the carcase. The individual carcase will be added as a random effect to separate unexplained error variation based on the individual and variation within individual that is due to discrepancies between the left and right side of the carcase.

The mixed model is formalised in Equation (1) below.

$$y_{ijkl} = \boldsymbol{x_i^t \beta} + \left(t_i^{ID} + t_i^{UB} x_i^{UB}\right) + \tau_k^{side} + \tau_l^{cat} + t_j^{OP} + e_{ijkl}, \qquad (1)$$

$$e_{ijkl} \sim N(0, \sigma^2),\ t_{\mathrm{j}}^{OP} \sim N(0, \sigma_{OP}^2),\ t_{\mathrm{i}}^{ID} \sim N(0, \sigma_{ID}^2),\ t_i^{UB} \sim N\left(0, \sigma_{UB}^2\right)$$

i = 1, … ,3018 (ID)

j = 1, … ,10 (OP)

k = 1, 2 (side)

l = 1, 2, 3 (cat)

y_{ijkl} is the response for carcase side *k* from individual *i*, of category *l*, processed by operator *j*. y_{ijkl} was one of the seven variables; lean-, fat-, bone and waste percentage, or one of the four PCPs. The vector term $\mathbf{x_i^t \beta}$ contains fixed effects for the specific individual *i*, i.e., the intercept, LMP, carcase weight and the interaction between carcase weight and LMP. The intercept might be interpreted as the average for the response in question for an individual with no breed data, as LMP and weight are centred, and their interaction standardised. In addition, the breed combination for individual *i*, as fractions of breeds Duroc, Hampshire, Landrace and Large White, see Vinje et al. (2023) for details, is added as fixed predictors with corresponding regression parameters.

For each individual a random intercept, t_i^{ID}, and a random slope, $t_i^{UB} x_i^{UB}$, term is added. The interpretation of the random intercept is the random effect of the individual, whereas the random error (e_{ijkl}) is the effect of the two different carcase sides after the effect of individual is accounted for. The random slope is additional variance added to the individual proportional to the degree of unknown breed x_i^{UB}, which is the proportion of unknown breed combination for individual *i*. For individuals categorized in accordance with Vinje et al. (2023), x_i^{UB} $(0 < x_i^{UB} < 1)$ is the posterior proportion of category “unknown breed”, whereas for the 2325 individuals without breed information, x_i^{UB} = 1.

The parameter τ_k^{side} is a fixed effect representing the systematic effect of carcase side defined so that $\tau_{left}^{side} = -\tau_{right}^{side}$. Hence, the interpretation of $2\tau_{left}^{side}$ is the change in the response, y_{ijkl} , when left side is compared to the right side of the carcase. The same principle was applied to categories, i.e., $\tau_{castrate}^{cat} + \tau_{gilts}^{cat} + \tau_{im.vac}^{cat} = 0.$

Each half carcase was processed by one or several operators out of the ten potential operators. Operator was treated as a factor and the operator that processed the highest number of SPWs for a particular half carcase was registered as the operator for that half carcase. The interpretation of t_j^{OP} is thereby the random effect of operator *j* on the response y_{ijkl}.

Three coefficient of determination (R^2) values were calculated. Two were according to the method for linear mixed models (LMM) from Nakagawa et al., (2013; 2017). The marginal R^2 value was calculated by

$$R_{LMM(m)}^2 = \frac{\hat{\sigma}_F^2}{\hat{\sigma}_F^2 + \hat{\sigma}_{OP}^2 + \hat{\sigma}_{ID}^2 + \overline{x_{ub}^2}\hat{\sigma}_{UB}^2 + \hat{\sigma}^2}$$

and the conditional R^2 value was calculated by

$$\mathrm{R}_{\mathrm{LMM(c)}}^2 = \frac{\hat{\sigma}_{\mathrm{F}}^2 + \hat{\sigma}_{\mathrm{OP}}^2 + \hat{\sigma}_{ID}^2 + \overline{x_{ub}^2}\hat{\sigma}_{UB}^2}{\hat{\sigma}_{\mathrm{F}}^2 + \hat{\sigma}_{\mathrm{OP}}^2 + \hat{\sigma}_{ID}^2 + \overline{x_{ub}^2}\hat{\sigma}_{UB}^2 + \hat{\sigma}^2}$$

with $\hat{\sigma}_{\mathrm{F}}^2$ defined as

$$\hat{\sigma}_F^2 = var(\hat{\beta}_{LMP} x_i^{LMP} + \hat{\beta}_w x_i^w + \hat{\beta}_{LMPw} x_i^{LMP} x_i^w + \hat{\tau}_k^{side} + \hat{\tau}_l^{cat})$$

The interpretation of the marginal R^2 value is the variance explained by the fixed effects in the model, while the interpretation of the conditional R^2 value is the variance explained by the fixed and random effects.

A natural interpretation of $t_i^{ID} + e_{ijkl}$ would be the unexplained error for side *k* and individual *i*. Due to the model assumption of independence between t_i^{ID} and e_{ijkl} the sum $t_i^{ID} + e_{ijkl}$ is normal with mean 0 and variance $\sigma_{ID}^2 + \sigma^2$. The decomposition of the total unexplained error into one part by individual (t_i^{ID}) and one by side (e_{ijkl}) leads to an interpretation of e_{ijkl} as the random error that differs between sides for individual j and t_i^{ID} as the random error for the individual that does not differ between sides from individual j.

Even if breed is known, and there is no random effect of operator, it will be impossible to achieve more precise predictions than allowed for by $\sigma_{ID}^2 + \sigma^2$.

Furthermore, in the interest of describing the average over the two sides, i.e., the average response for the individual in question, the prediction precision is limited by σ_{ID}^2. Consequently, an additional coefficient of determination (R^2), denoted limiting $\mathrm{R}^2_{\mathrm{LMM(lim)}}$, for the individual in question was calculated by

$$R^2_{LMM(lim)} = \frac{\hat{\sigma}_F^2 + \hat{\sigma}_{OP}^2 + \overline{x_{ub}^2}\hat{\sigma}_{UB}^2}{\hat{\sigma}_F^2 + \hat{\sigma}_{OP}^2 + \overline{x_{ub}^2}\hat{\sigma}_{UB}^2 + \hat{\sigma}_{ID}^2}$$

, which might be interpreted as the maximum part of the variance at individual level, which might be explained by the fixed effects LMP, weight, category and unknown breed together with the random effect operator.

The adjusted intraclass correlation coefficient for LMMs as defined in Nakagawa et al., (2010; 2017) was calculated by

$$ICC_{LMM(adj)} = \frac{\hat{\sigma}_{ID}^2}{\hat{\sigma}_{ID}^2 + \hat{\sigma}^2}$$

The adjusted ICC is the proportion of random variance accounted for by the between group variance while omitting other non-zero variance from other random variables (Nakagawa et al., 2017), such as confounding factors, in this case, operator variance. Between group variance in the present study is between individual variance, and thus the adjusted ICC gives the random variance accounted for by the between individual variance and gives an estimate as to how much of the random variance is explained by the individual as opposed to the sides of the carcase.

An additional ICC was calculated which included estimated operator variance, who's interpretation is the amount of total random error explained by the operator and was calculated by:

$$ICC_{LMM(OP)} = \frac{\hat{\sigma}_{\mathrm{OP}}^2}{\hat{\sigma}_{\mathrm{OP}}^2 + \hat{\sigma}_{ID}^2 + \hat{\sigma}^2}$$

All statistical analyses were coded and performed in Rstudio software (Rstudio Team, 2020). Estimates are reported for fixed effect regression parameters, i.e., $\beta's$ and $\tau's,$ and approximate p-values from (Student-)t tests applying Satterthwaite's method (Giesbrecht & Burns, 1985; Hrong-Tai Fai & Cornelius, 1996) via the R-package "lmerTest" (Kuznetsova, Brockhoff & Christensen, 2017). Furthermore, sum of squares from regression associated with LMP, weight, their interaction, side, category and breed and associated approximate p-values from F(isher)- test with Satterthwaite's method are reported. Sums of Squares is calculated by the so-called ANOVA type I method in order to distinguish the attitude for different factors. An adjusted sum of squares for breed, with adjustment factor $\frac{n}{(n-\sum x_i^{UB})}$, in order to compensate for individuals without breed information, and thereby comparable to the other factors without missing data, is also reported.

Results

Summary statistics are presented in Table 1 for all categories. Castrates and gilts had similar sample sizes, while I.C.s had a smaller sample size. All categories had similar mean weight, though I.C.s had a larger SD than the other categories. Measured LMP was higher for, and similar between, gilts and I.C.s, and lower for castrates. Primal cut weights were similar between categories, with the largest difference being between forepart weight for gilts and I.C.s.

Table 1

Summary statistics by category with mean values ±SD for LMP measured by HGP7, weight per carcase measured at the abattoirs, and percentages of Lean, Fat and Bone, and percentages of the four primal cuts.

Category	Castrate	Gilt	I.C.
N	1494	1373	151
LMP	58.7 ±3.1	60.8 ±3.0	60.4 ±2.9
Weight	81.2 ±7.7	80 ±7.8	80.1 ±11.3
Lean	62.5 ±3.7	65.1 ±3.7	65.7 ±3.3
Fat	20.6 ±4.2	17.7 ±4.3	15.8 ±4.2
Bone & Waste	16.8 ±1.1	17.2 ±1.2	18.5 ±1.6
Forepart	31.2 ±0.9	30.9 ±0.9	32 ±1.2
Loin	19.1 ±0.9	19.1 ±0.9	18.4 ±1
Belly	17 ±0.9	16.8 ±0.9	16.7 ±1.1
Hindpart	32.8 ±1.0	33.2 ±0.9	32.8 ±1.2

[1] I.C.: Immunologically castrated pigs.

Table 2 shows variance component estimates for random variables ID, operator and the error, as well as $\mathrm{ICC}_{\mathrm{LMM(adj)}}$, $\mathrm{ICC}_{\mathrm{LMM(OP)}}$, $\mathrm{R}^2_{\mathrm{LMM(m)}}$, and $\mathrm{R}^2_{\mathrm{LMM(c)}}$. $\mathrm{R}^2_{\mathrm{LMM(m)}}$ was moderately high, 66 % and 67 % respectively, for lean- and fat percentage, but lower, 29%, for bone and waste percentage. The same pattern applied to the $\mathrm{R}^2_{\mathrm{LMM(c)}}$, at 96.8%, 97.6% and 83.4% for lean-, fat- and bone and waste percentages respectively. Thus, the increase from $\mathrm{R}^2_{\mathrm{LMM(m)}}$ to $\mathrm{R}^2_{\mathrm{LMM(c)}}$ was larger for bone and waste percentage than lean- and fat percentage. The large values for $\mathrm{R}^2_{\mathrm{LMM(c)}}$ means that results from the two sides are highly correlated for these responses, which is illustrated in the two upper panels in Figure 1.

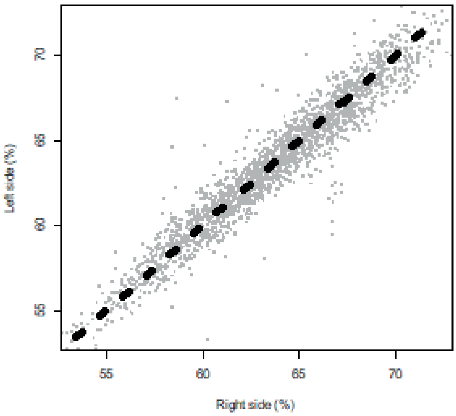
Table 2

Variance component estimates for the fixed effects (σ_F^2), unknown breed (σ_{UB}^2), individual (σ_{ID}^2), operator (σ_{OP}^2) and the error (σ^2), the estimates for the adjusted intraclass correlation coefficient ($ICC_{LMM(adj)}$) and intraclass correlation coefficient for operators ($ICC_{LMM(OP)}$), and the marginal ($\mathrm{R}^2_{\mathrm{LMM(m)}}$), the limited ($\mathrm{R}^2_{\mathrm{LMM(l)}}$), and conditional ($\mathrm{R}^2_{\mathrm{LMM}(c)}$) coefficient of determination for all seven response variables based on model (1).

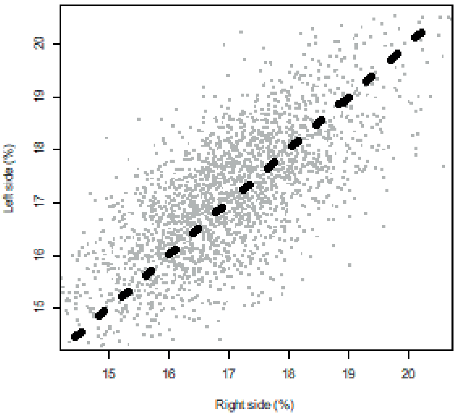
Response variable	$\hat{\sigma}_F^2$	$\hat{\sigma}_{\mathrm{UB}}^2$	$\hat{\sigma}_{\mathrm{ID}}^2$	$\hat{\sigma}_{\mathrm{OP}}^2$	$\hat{\sigma}^2$	$\mathrm{ICC}_{\mathrm{LMM(adj)}}$	$\mathrm{ICC}_{\mathrm{LMM(OP)}}$	$\mathrm{R}^2_{\mathrm{LMM(m)}}$	$\mathrm{R}^2_{\mathrm{LMM(l)}}$	$\mathrm{R}^2_{\mathrm{LMM(c)}}$
Tissue Percentages										
Lean	10.547	0.675	4.333	0.099	0.514	0.894	0.020	0.659	0.720	0.968
Fat	14.171	0.921	5.537	0.132	0.517	0.915	0.021	0.673	0.731	0.976
Bone and Waste	0.479	0.257	0.599	0.129	0.279	0.682	0.128	0.285	0.573	0.834
Primal Cut Percentages										
Forepart	0.113	0.048	0.544	0.025	0.355	0.605	0.027	0.105	0.244	0.670
Loin	0.049	0.061	0.384	0.091	0.421	0.477	0.102	0.049	0.327	0.576
Belly	0.158	0.047	0.376	0.110	0.363	0.508	0.129	0.151	0.447	0.652
Hindpart	0.170	0.072	0.519	0.030	0.404	0.562	0.031	0.144	0.329	0.657

For PCPs $\mathrm{R}^2_{\mathrm{LMM(m)}}$ were low (in the range from 5% to 15%), especially for the forepart and hindpart. For PCPs $\mathrm{ICC}_{\mathrm{LMM(adj)}}$ takes values around 50% (range 48% to 61%) as opposed to $\mathrm{ICC}_{\mathrm{LMM(adj)}}$ values for lean- and fat percentages at around 95%, which is well reflected in Figure 1. The interpretation is that for PCPs the variance between sides in one individual is of approximately the same magnitude as the unexplained variance explained by the individuals for PCPs. As a consequence of the relatively large variances associated with side, $\mathrm{R}^2_{\mathrm{LMM(c)}}$, for PCPs were at relatively low levels, in the range 58% to 67%, whereas the difference between $\mathrm{R}^2_{\mathrm{LMM(m)}}$ and $\mathrm{R}^2_{\mathrm{LMM(lim)}}$ is more perspicuous for PCPs than for lean- and fat percentages. Even though $\mathrm{R}^2_{\mathrm{LMM(lim)}}$ are at substantially higher levels than $\mathrm{R}^2_{\mathrm{LMM(m)}}$ for PCPs they are still at low levels, range 24 to 33%, except for belly percentage where $\mathrm{R}^2_{\mathrm{LMM(lim)}}$ is a little higher at 45%.

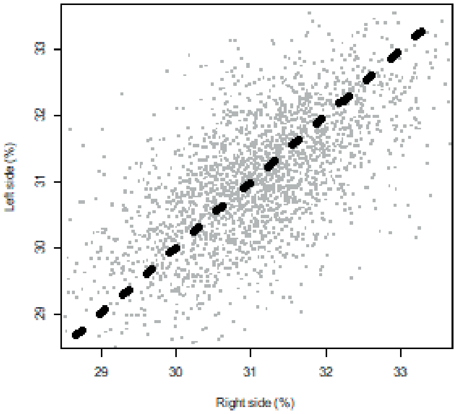
LeanProsHC (Pearson = 0.96)



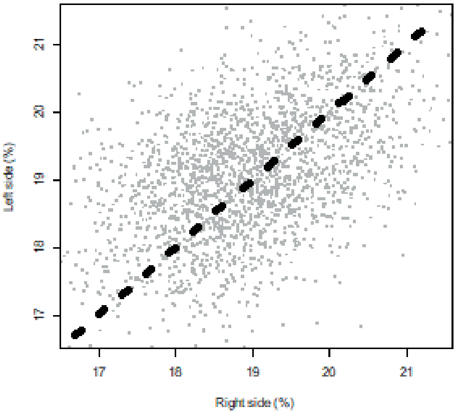
BWProsHC (Pearson = 0.72)



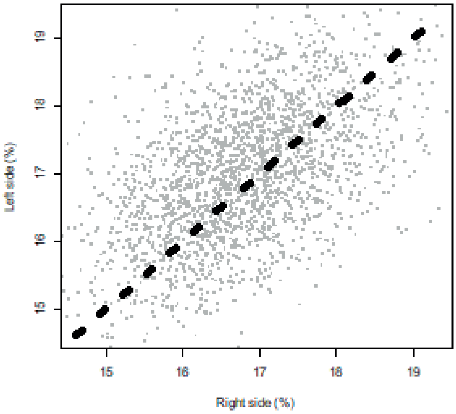
ShoulderProsHC (Pearson = 0.64)



LoinProsHC (Pearson = 0.48)



BellyProsHC (Pearson = 0.55)



HamProsHC (Pearson = 0.62)

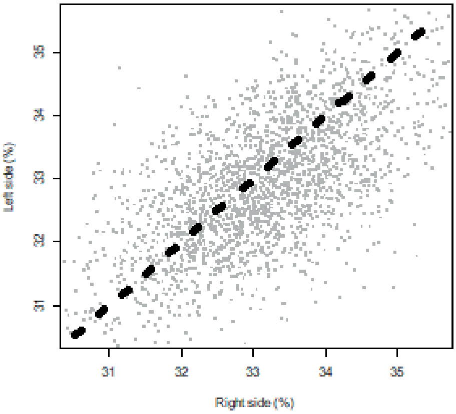


Fig. 1. Scatterplot of registration from right side vs. left side of carcases for the six responses “LeanProsHC”, “BWProsHC”, “ShoulderProsHC”, “LoinProsHC”, “BellyProsHC” and “HamProsHC”. “LeanProsHC” is the lean percentage per half carcase calculated from deconstruction of the carcase, and “BWProsHC” is corresponding the bone and waste percentage. “FatProsHC” which is highly correlated to “LMPProsHC” is omitted. “ShoulderProsHC”, “LoinProsHC”, “BellyProsHC” and “HamProsHC” are percentage primal cuts per half carcase. Black dotted lines show the one-to-one relationship.

$\mathrm{ICC}_{\mathrm{LMM(OP)}}$ was generally low, in particular for lean-, fat-, shoulder- and ham percentages ($\mathrm{ICC}_{\mathrm{LMM(OP)}}$<3%). For bone and waste, loin- and belly percentages $\mathrm{ICC}_{\mathrm{LMM(OP)}}$s were still small, but marginally larger, in the range 10% to 13%.

Regression parameter estimates and their respective p-values are shown in Table 3a and Table 3b. The main effects of measured LMP and weight had a significant, negative effect on fat percentage, and a significant, positive effect on lean- and bone and waste percentage. The interaction between LMP and weight is significant negative for lean- and bone and waste percentages, and significant positive for fat percentage.

Most parameter estimates for LMP, weight and their interaction were non-significant for PCPs. LMP had a significant, but small, negative effect on belly percentage, and a slightly higher, significant, positive effect on hindpart percentage, i.e., fatter pigs tend to have more belly and less hindpart.

There is a significant effect of category for all response variables. If all other predictors are held constant I.C.s will on average be leaner than gilts which in turn will be leaner than castrates. The difference for lean percentage between I.C.s and castrates is approximately 1% unit, where the associated counterparts is that I.C.s have approximately 1.5% units less fat and 0.8% units more bone and waste than castrates with gilts in a middle position.

For PCPs the differences are largest between gilts and I.C.s with castrates in a middle position. If all other predictors are held constant I.C.s will on average have approximately 1.2% more shoulder, 0.5 % less loin, 0.3% less belly and 0.4% less hindpart compared to gilts. Castrates are in the middle position, but are closer to gilts than I.C.s.

There is considerable effect of breed combinations. If all other predictors are held constant the expected lean-, and bone and waste percentages will increase when proportion of Duroc increases, and decrease with proportion of Large White (Z-line). The effect on fat percentages is the opposite. The effects of Hampshire proportions are small and not significant, whereas Landrace proportion has a positive effect on expected lean percentage and negative effect on expected bone and waste percentage. If these estimates were to be applied on purebreds, effects had been considerable with expected differences between Duroc and Z-line approximately 3%, 5%, and 2% units for lean-, fat- and bone and waste percentages respectively.

For the PCPs the effect of Duroc is positive on expected shoulder and hind part percentages and negative on belly- and loin percentages. Landrace is the contrast of

Duroc with higher expected percentages for loin and belly and lower for shoulder- and hind part percentages. Hampshire has a significant negative effect on belly and positive effect on hind part percentages. Finally, large white (Z-line) has a positive effect on expected percentage of loin and negative effect of expected hind part percentage. It is important to note (Table 2) that for PCPs the estimates for $\hat{\sigma}_{UB}^2$, i.e., the extra variance added to individuals with unknown breed, is considerable, proportions from 30% (belly) to 125% (loin) compared to $\hat{\sigma}_F^2$, the estimated variance explained by the fixed factors. Consequently, for the PCPs, “breed” is an important part of $\hat{\sigma}_F^2$ for individuals with known breed, and only small parts of the variance are explained by measured LMP and weight.

Table 3a

Regression parameter estimates and p-values for seven response variables based on the same linear mixed model (1) that was fitted seven times for each univariate response. Columns are fixed predictor variables in the model, rows are response variables. LMP:Weight is the interaction between LMP and weight. Gender/castration category and side are modeled so the sum of the parameter estimates is 0.

Response variable	Intercept	LMP	Weight	LMP:Weight	Castrates	Gilts	Left Side
Tissue Percentages							
Lean	64.004	1.294	0.252	-2.337	-0.715	-0.152	-0.082
	(0.000)	(0.000)	(0.003)	(0.001)	(0.000)	(0.059)	(0.000)
Fat	18.375	-1.542	-0.310	3.272	1.254	0.544	0.053
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Bone and Waste	17.622	0.241	0.051	-0.882	-0.536	-0.385	0.019
	(0.000)	(0.000)	(0.162)	(0.004)	(0.000)	(0.000)	(0.016)
Primal Cut Percentages							
Forepart	31.368	0.029	-0.002	-0.005	-0.133	-0.533	-0.060
	(0.000)	(0.514)	(0.595)	(0.985)	(0.000)	(0.000)	(0.000)
Loin	18.884	-0.043	0.000	0.142	0.096	0.212	0.063
	(0.000)	(0.298)	(0.989)	(0.587)	(0.002)	(0.000)	(0.000)
Belly	16.771	-0.144	-0.022	0.307	0.077	0.095	0.119
	(0.000)	(0.000)	(0.463)	(0.220)	(0.010)	(0.001)	(0.000)
Hindpart	32.973	0.159	0.041	-0.457	-0.042	0.226	-0.118
	(0.000)	(0.000)	(0.228)	(0.109)	(0.209)	(0.000)	(0.000)

Table 3b

Regression parameter estimates and p-values for seven response variables based on the same linear mixed models (1) that was fitted seven times for each univariate response. Columns are fixed predictor variables in the model, rows are response variables. Each breed is a fraction of the breed combination for each individual.

Response variable	Duroc	Hampshire	Landrace	Z-line
Tissue Percentages				
Lean	1.298 (0.000)	-0.346 (0.498)	1.660 (0.000)	-2.150 (0.000)
Fat	-2.023 (0.000)	0.469 (0.414)	-0.897 (0.031)	3.164 (0.000)
Bone and Waste	0.714 (0.000)	-0.133 (0.519)	-0.770 (0.000)	-1.014 (0.000)
Primal Cut Percentages				
Forepart	1.003 (0.000)	-0.025 (0.902)	-0.835 (0.000)	-0.219 (0.356)
Loin	-0.523 (0.000)	-0.101 (0.585)	0.542 (0.000)	0.661 (0.002)
Belly	-1.060 (0.000)	-0.501 (0.005)	0.530 (0.000)	0.418 (0.045)
Hindpart	0.579 (0.000)	0.637 (0.002)	-0.230 (0.116)	-0.862 (0.000)

Table 4

Sum of Squares (ANOVA type I) associated with different regression factors with approximate p- value based on F-test in parentheses. Columns represent factors in the model, rows are response variables. The last column gives only an adjusted Sum of Squares for breed in order to adjust for individuals with missing data.

Response factors	LMP	Weight	LMP:Weight	Gender/castration Category	Side	Breed	Breed Adjusted
Tissue Percentages							
Lean	2998.01 (0.000)	10.00 (0.000)	3.89 (0.006)	38.16 (0.000)	34.06 (0.000)	27.77 (0.000)	121.73
Fat	3079.52 (0.000)	87.77 (0.000)	8.32 (0.000)	99.22 (0.000)	14.25 (0.000)	26.36 (0.000)	115.58
Bone and Waste	170.59 (0.000)	159.92 (0.000)	4.37 (0.000)	78.46 (0.000)	2.04 (0.007)	25.91 (0.000)	113.61
Primal Cut Percentages							
Forepart	3.290 (0.002)	33.225 (0.000)	0.097 (0.601)	93.062 (0.000)	17.294 (0.000)	30.464 (0.000)	133.555
Loin	8.805 (0.000)	36.673 (0.000)	0.617 (0.226)	21.289 (0.000)	17.653 (0.000)	19.494 (0.000)	85.459
Belly	184.532 (0.000)	23.769 (0.000)	0.152 (0.518)	3.244 (0.012)	69.421 (0.000)	33.304 (0.000)	146.005
Hindpart	176.614 (0.000)	14.904 (0.000)	0.583 (0.229)	26.394 (0.000)	64.560 (0.000)	9.898 (0.000)	43.394

Discussion

The primary aim of the study was to evaluate how the yield, i.e., lean, fat and bone and waste percentages, as well as primal cut percentages, could be explained by LMP and weight, also considering side of the carcase, gender/castration category, and breed combination. The carcase characteristics used in the models, including LMP, weight, and category are important as they are commonly registered in several countries worldwide (Delgado-Pando et al., 2021) and therefore allow understanding and comparisons outside the confines of Europe and its classification system.

The results showed that yield was explained almost perfectly for lean ($\mathrm{R}^2_{\mathrm{LMM(c)}}$ = 0.968), fat ($\mathrm{R}^2_{\mathrm{LMM(c)}}$ = 0.976) and bone and waste ($\mathrm{R}^2_{\mathrm{LMM(c)}}$ = 0.834) based on the conditional coefficients of determination (Table 2). The marginal R^2 values (Table 2), i.e., the total variance explained by the fixed variables in the models, showed a poor (PCPs and bone and waste) to moderately high (lean- and fat percentage) ability to explain total variance depending on the yield variable in question. Lean and fat percentage were the highest (>0.65), while bone and waste, and the PCPs were low (<0.285). Previous studies have achieved higher R^2 values for lean percentage (Bohrer et al., 2023; Goenaga, Lloveras & Améndola, 2008). However, prediction variables in these studies were not LMP given by the probes as in this study, but rather fat depth and meat depth variables based on measurements by the probes. Goenaga achieved an R^2 of 0.79, but does not specify if this is for a calibration/training set or a validation/test set. Bohrer achieved an R^2 of 0.75 when calculated for the validation/test set, but had a lower R^2 (0.63) for the calibration/training set than in the current study. The R^2 for the calibration/training set is arguably more comparable to the marginal R^2 value calculated in the current study, as this study did not use a validation/test set.

As the fixed effect variance not only includes variance from LMP and carcase weight, but also variance from side, category and breed combination, it is important to understand how much of the fixed effect variance, and consequently the marginal R^2, is affected by side, gender/castration category and breed combination. Though the effect of side on lean and fat percentage was significant, compared to LMP, weight and their interaction its effect was small, with the largest estimated systematic difference between the left and right side of the carcase being 0.164 (Table 3a). This was also reflected in the high correlation between the two sides of the carcase (Figure 1), and comparatively low sum of squares (Table 4).

Both gender/castration category and breed have significant effects (Table 3a & 3b), which is in line with Gangsei et al., (2018) and Engel et al., (2012). Though both castrates and gilts appear to have lower lean percentage than I.C.s, the effect was not significant for gilts. Both castrates and gilts seemed to have significantly more fat than I.C.s. The lack of significant difference between I.C.s and gilts for lean percentage could be due to I.C.s having a lower sample size than the two other gender/castration categories. Kress et al. (2020) had a larger sample size and similarly did not find any significant difference between gilts and I.C.s, though their study also had fewer I.C.s than gilts. The differences between I.C.s and castrates in this study are congruent with the findings between I.C.s and castrates in meta-analyses by Batorek et al. (2012) and Nautrup et al. (2018). Though the effect of gender/castration category on lean- and fat percentage is significant (Table 3a), the effect is small compared to the combined effect of LMP, weight and their interaction. This is especially clear when comparing the sum of squares between gender/castration category and LMP (Table 4).

The effect of breed on lean- and fat percentage (Table 3b) was significant for Duroc, Landrace and Z-line, but small compared to the combined effect of LMP, weight and their interaction (Table 3a). This was also reflected when comparing the sum of squares (Table 4) between breed and LMP, even when the sum of squares for breed was adjusted for the individuals with missing breed data. Though parameter estimates for Duroc, Landrace, and Z-line are comparable to the parameter estimates for LMP, weight and their interaction, the scale of the variables differ, and thus LMP, weight and their interaction have a greater effect. Based on results discussed, it is likely that the majority of variance in the fixed effect variance for lean and fat percentage (Table 2) is due to LMP, weight and their interaction, and consequently, the majority of the proportion of total variance explained as illustrated by the marginal R^2 for lean and fat percentage is due to LMP, weight and their interaction.

The primary aim of porcine carcase grading is differentiating between carcases of varying levels of leanness, not prediction of PCPs. Nevertheless, PCPs are an important factor for evaluation of the yield. If ordinary EUROP carcase grading is sufficient for precise PCP prediction it would have been a valuable asset. As of today, PCP prediction is a feature offered by for instance the Autofom carcase grading system (Choi et al., 2018; Janiszewski et al., 2018). Based on results from the current study, LMP and weight alone have small to negligible effect for explaining variances in primal cut proportions. This is especially apparent for the differences in sum of squares between breed and LMP and weight (Table 4) for the forepart and loin. It is also apparent when comparing the parameter estimates between LMP and weight (Table 3a) and breed (Table 3b) for PCPs.

For primal cuts there is considerable variance between left and right sides, with Pearson correlations as low as between 0.48 and 0.64 between sides (Figure 1). The relatively low correlations between left and right sides might partly be a result of real differences, but most likely differences occur due to inaccuracies either when carcasses are split in two at the slaughterhouses, or in the cutting process. Consequently, it might be argued that $\mathrm{R}^2_{\mathrm{LMM(lim)}}$ in particular is the most important measurement for evaluating primal percentages as this measurement omits the variance between the two halves of a carcase.

The $\mathrm{R}^2_{\mathrm{LMM(m)}}$ values for PCPs in the range 5% (loin) to 15% (belly) show that information from carcase grading alone, i.e., predicted LMP and carcase weight, contains almost no information for precise prediction of PCPs. Even though $\mathrm{R}^2_{\mathrm{LMM(m)}}$ was highest for the belly, prior expectations were that the belly percentage to a larger degree could be predicted from LMP, as leaner carcases are known to have lower belly PCP than fatter carcases (Kongsro et al., 2017, Pulkrábek et al., 2006). The $\mathrm{R}^2_{\mathrm{LMM(lim)}}$ values for PCPs were in the range 24% (forepart) to 48% (belly). This is an improvement compared to $\mathrm{R}^2_{\mathrm{LMM(m)}}$, but $\mathrm{R}^2_{\mathrm{LMM(lim)}}$ shows that even on an individual level, PCPs are difficult to predict without further information.

Regarding the effect of sex, I.C.s. differed considerably from gilts and castrates, where the shoulder part in particular was larger at the expense of the other cuts. A larger shoulder/forepart is a typical characteristic of male pigs (Lei et al., 2023), as tends to be the case with male mammals in general (McPherson & Chenoweth, 2012). Consequently, this characteristic might be viewed as I.C.s preserving more of the typical male physiology than castrates, which are the middle category compared to gilts. I.C.s having a larger shoulder than castrates is consistent with previous literature (Nautrup et al., 2018). Documentation of difference in shoulder size for I.C.s compared to gilts is sparce, with Kress et al. (2020) showing no significant difference. In the present study, the difference between I.C.s and castrates/gilts is substantial for shoulder PCP, with estimated effect at 1.2% units. Consequently, the yield not only on individual level, but also on population level would be heavily affected if surgical castration is to be replaced with immunological castration. The total effect of this change would have to be considered with other aspects including, but not limited to, intramuscular fat content, growth rate, and concentration of androstenone and skatole (Nautrup et al., 2018).

The results from the current study does not support the findings in Ko et al. (2023) who reported larger belly yield from castrates than gilts in the Korean pig population, which is very comparable to the Norwegian population in the sense that the crossbreed with Duroc as the paternal line and crossing of Landrace and Yorkshire as the maternal line is the dominating finisher pig. The reason for this discrepancy is unclear. One explanation is that Ko et al. (2023) used PCP from Autofom III as responses on an immense data set (350 179 pigs), whereas the current study is based on a smaller dataset (3018 individuals), and the responses, PCPs, are based on manual butchering.

The effects of breeds are considerable, however, the interpretations should be conducted with some care. Firstly, only a fraction of the pigs have predictions for their breed combinations, predictions which fortunately are subject to little uncertainty (Vinje et al., 2023). The interpretation of breed effects is as if the individuals in question are purebred animals, which are only rare exceptions, most of the individuals are crossbreds. Breed has an effect on both tissue percentages and PCPs. The most striking effect of Yorkshire Z-line and Duroc was on tissue percentages. High proportion of Duroc genetics yields more lean, and bone and waste, and less fat than predicted by LMP grading and weight alone, whereas the opposite is true for carcases with high proportion of Yorkshire Z-line. It might be argued that the breed information should be part of the regression equation for LMP, at least LMP based on HGP7. However, this is unfeasible, as reliable data for breed combination of individual pigs is not present at the slaughter line.

For the PCPs the typical paternal breeds, Duroc and Hampshire, seem to have larger shoulder and hindpart, and less loin and belly, than accounted for by LMP and weight alone. It is worth noting that this might be explained by the length of the back, or number of vertebras, which differs among individuals and breeds, with Duroc on average having a lower total number of vertebras (28.72) than Landrace (29.78) (Van Son, et.al, 2019).

The effects of operator were mainly small, with a small exception for the tissue category bone and waste with $\mathrm{ICC}_{\mathrm{LMM(OP)}}$ at 0.128, and for PCPs from loin and belly with $\mathrm{ICC}_{\mathrm{LMM(OP)}}$ at 0.102 and 0.129, respectively. The interpretation of operator effect is that some of the operators systematically tend to register lower/ higher values for the response in question than other operators, all other factors assumed fixed. In particular for the loin/belly PCPs is natural to assume that different operators have slightly different opinions on where to conduct the saw cut between loin and belly, a cut which is not anatomically well defined. The $\mathrm{ICC}_{\mathrm{LMM(OP)}}$ is the measurement of random effect of operators, not the precision of operators. Consequently, the variance between sides, represented by the error variance, σ^2, might still to some extent be an effect of operator inaccuracies.

The current study does not aim at evaluating Autofom performance, but to evaluate to which extent LMP from the HGP7 and weight alone or in combination with gender/castration category and breed might explain variance in primal cut percentages. The scientific literature documenting Autofom performance for PCP prediction is scarce, however Janiszewski et al. (2018) documented that Autofom to a great extent is capable of predicting PCPs. In Janiszewski et al. (2018) a slightly different definition of primal cuts is applied, with the shoulder split in two, “shoulder” and “neck” respectively. Nevertheless, they report R^2 at 0.42 (neck), 0.68 (shoulder), 0.81 (belly), 0.82 (ham/ hindpart) and 0.93 (loin) for PCPs. Results in Janiszewski et al. (2018) are from one breed combination and the same producer, where the same environment and feeding management were applied to all pigs. Hence, variation due to breed, and potential other environmental factors are removed from carcases used in Janiszewski et al. (2018). However, the results are obtained only by evaluating halve carcases, i.e., the left sides. Thus, it is hard to directly compare results from Janiszewski et al. (2018) to R^2 statistics reported in the current study, but it is obvious that Autofom adds substantial information for PCP prediction.

The Autofom data for Norwegian pigs are scarce, and LMP from Autofom and HGP7 and lean percentage are all highly correlated with Pearson correlations in the range 0.799 to 0.815. However, in future studies these data (n = 216), and if possible, even more data should be evaluated with suitable methods like Partial Least Squares (PLS) regression to evaluate Autofom prediction precision for PCPs. Another topic of interest is to evaluate if the breed effects on tissue percentages are, if not removed, at least reduced.

Conclusion

EUROP classification in the form of lean meat percentage from the HGP7 explains yield to varying degrees. Most importantly, lean meat percentage and fat percentage based on processed carcases were explained to a moderately high degree, though some of the variance can be attributed to carcase weight, gender/castration category, breed, operator and carcase side. Bone and waste percentage, as well as all PCPs, were poorly explained by lean meat percentage from the HGP7.

Gender/castration category generally had a significant effect on all yield variables, though the effect was greater on lean-, fat-, and bone and waste percentage, than on the PCPs.

References

Batorek, N., Čandek-Potokar, M., Bonneau, M., & Van Milgen, J., 2012. Meta-analysis of the effect of immunocastration on production performance, reproductive organs and boar taint compounds in pigs. Animal, 6(8), 1330-1338.

Bohrer, B. M., Wang, Y., Dorleku, J. B., Campbell, C. P., & Mandell, I. B., 2023. An update of the predicted lean yield equation for the Destron PG-100 optical grading probe. Journal of Animal Science, 101, skad199.

Brøndum, J., Egebo, M., Agerskov, C., & Busk, H., 1998. On-line pork carcass grading with the Autofom ultrasound system. Journal of Animal Science, 76(7), 1859-1868.

Choi, J. S., Kwon, K. M., Lee, Y. K., Joeng, J. U., Lee, K. O., Jin, S. K., Choi, Y.I. & Lee, J. J., 2018. Application of AutoFom III equipment for prediction of primal and commercial cut weight of Korean pig carcasses. Asian-Australasian journal of animal sciences, 31(10), 1670.

Delgado-Pando, G., Allen, P., Troy, D. J., & McDonnell, C. K., 2021. Objective carcass measurement technologies: Latest developments and future trends. Trends in Food Science & Technology 111, 771–782. doi:10.1016/j.tifs.2020.12.016.

Dorleku, J. B., Wormsbecher, L., Christensen, M., Campbell, C. P., Mandell, I. B., & Bohrer, B. M., 2023. Comparison of an advanced automated ultrasonic scanner (AutoFom III) and a handheld optical probe (Destron PG-100) to determine lean yield in pork carcasses. Journal of Animal Science, 101, skad058.

Engel, B., Lambooij, E., Buist, W. G., & Vereijken, P., 2012. Lean meat prediction with HGP, CGM and CSB-Image-Meater, with prediction accuracy evaluated for different proportions of gilts, boars and castrated boars in the pig population. Meat science, 90(2), 338-344.

European Commission, 2013. Regulation (EU) No 1308/2013 of the European Parliament and of the Council of 17 December 2013 establishing a common organisation of the markets in agricultural products and repealing Council Regulations (EEC) No 922/72, (EEC) No 234/79, (EC) No 1037/2001 and (EC) No 1234/2007. Retrieved 06 June 2023 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013R1308#d2046e32-801-1.

European Commission, 2017. Commission Delegated Regulation (EU) 2017/1182 of 20 April 2017 supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards the Union scales for the classification of beef, pig and sheep carcasses and as regards the reporting of market prices of certain categories of carcasses and live animals. Retrieved 08 June 2023 from https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32017R1182.

Font i Furnols, M., Gispert, M., 2009. Comparison of different devices for predicting the lean meat percentage of pig carcasses. Meat science, 83(3), 443-446.

Gangsei, L. E., Bjerke, F., Røe, M., & Alvseike, O., 2018. Monitoring lean meat percentage predictions from optical grading probes by a commercial cutting pattern. Meat science, 137, 98-105.

Giesbrecht, F. G., & Burns, J. C., 1985. Two-stage analysis based on a mixed model: large-sample asymptotic theory and small-sample simulation results. Biometrics, 477-486.

Goenaga, P., Lloveras, M. R., & Améndola, C., 2008. Prediction of lean meat content in pork carcasses using the Hennessy Grading Probe and the Fat-O-Meater in Argentina. Meat science, 79(3), 611-613.

Heggli, A., Alvseike, O., Bjerke, F., Gangsei, L. E., Kongsro, J., Røe, M., & Vinje, H., 2023. Carcase grading reflects the variation in beef yield–a multivariate method for exploring the relationship between beef yield and carcase traits. animal, 17(6), 100854.

Hrong-Tai Fai, A., & Cornelius, P. L., 1996. Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments. Journal of statistical computation and simulation, 54(4), 363-378.

Hviid, M., Erbou, S. G. H., & Olsen, E. V., 2011. The content of meat–Is there a difference between left and right sides of pig carcasses?. Proc. of the 57th ICoMST, Gand, Belgique, P045.

Janiszewski, P., Borzuta, K., Lisiak, D., Grześkowiak, E., & Stanisławski, D., 2018. Prediction of primal cuts by using an automatic ultrasonic device as a new method for estimating a pig-carcass slaughter and commercial value. Animal Production Science, 59(6), 1183-1189.

Ko, E., Park, Y., Park, K., Woo, C., Kim, J., Kim, K., & Choi, J., 2023. Comparison of pork belly characteristics and weights of primal cuts between gilt and barrow of Landrace× Yorkshire× Duroc pigs measured by AutoFomIII. Journal of Animal Science and Technology, 65(2), 412.

Kongsro, J., Gangsei, L. E., Karlsson-Drangsholt, T. M., & Grindflek, E., 2017. Genetic parameters of in vivo primal cuts and body composition (PigAtlas) in pigs measured by computed tomography (CT). Translational Animal Science, 1(4), 599-606.

Kress, K., Hartung, J., Jasny, J., Stefanski, V., & Weiler, U., 2020. Carcass characteristics and primal pork cuts of gilts, boars, immunocastrates and barrows using AutoFOM III data of a commercial abattoir. Animals, 10(10), 1912.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H., 2017. lmerTest package: tests in linear mixed effects models. Journal of statistical software, 82, 1-26.

Lei, X. I. E., QIN, J. T., Lin, R. A. O., CUI, D. S., Xi, T. A. N. G., XIAO, S. J., ... & HUANG, L. S., 2023. Effects of carcass weight, sex and breed composition on meat cuts and carcass trait in finishing pigs. Journal of Integrative Agriculture, 22(5), 1489-1501.

Lisiak, D., Duziński, K., Janiszewski, P., Borzuta, K., & Knecht, D., 2014. A new simple method for estimating the pork carcass mass of primal cuts and lean meat content of the carcass. Animal Production Science, 55(8), 1044-1050.

Marcoux, M., Bernier, J. F., & Pomar, C., 2003. Estimation of Canadian and European lean yields and composition of pig carcasses by dual-energy X-ray absorptiometry. Meat Science, 63(3), 359-365.

Marcoux, M., Pomar, C., Faucitano, L., Brodeur, C., 2007. The relationship between different pork carcass lean yield definitions and the market carcass value. Meat Science, 75(1), 94-102.

McPherson, F. J., & Chenoweth, P. J., 2012. Mammalian sexual dimorphism. Animal reproduction science, 131(3-4), 109-122.

Nakagawa, S., & Schielzeth, H., 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biological Reviews, 85(4), 935-956.

Nakagawa, S., Schielzeth, H., & O'Hara, R. B., 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol Evol, 4, 133–142.

Nakagawa, S., Johnson, P. C., & Schielzeth, H., 2017. The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of the Royal Society Interface, 14(134), 20170213.

Nautrup, B. P., Van Vlaenderen, I., Aldaz, A., & Mah, C. K., 2018. The effect of immunization against gonadotropin-releasing factor on growth performance, carcass characteristics and boar taint relevant to pig producers and the pork packing industry: A meta-analysis. Research in Veterinary Science, 119, 182-195.

Nissen, P. M., Busk, H., Oksama, M., Seynaeve, M., Gispert, M., Walstra, P., Hansson, I. & Olsen, E., 2006. The estimated accuracy of the EU reference dissection method for pig carcass classification. Meat Science, 73(1), 22-28.

Olsen, E. V., Christensen, L. B., & Nielsen, D. B., 2017. A review of computed tomography and manual dissection for calibration of devices for pig carcass classification-Evaluation of uncertainty. Meat science, 123, 35-44.

Pulkrábek, J., Pavlík, J., Valis, L., & Vítek, M., 2006. Pig carcass quality in relation to carcass lean meat proportion. Czech Journal of Animal Science, 51(1), 18.

RStudio Team, 2020. RStudio: Integrated Development Environment for R. RStudio, PBC. Boston, MA. http://www.rstudio.com/.

Røe, M., 2021. 301P Klassifisering av gris. (In Norwegian) Retrieved 06 June 2023 from https://www.animalia.no/no/kjott--egg/klassifisering/klassifiseringshandboka/301p-klassifisering-av-gris-16/.

Van Son, M., Lopes, M. S., Martell, H. J., Derks, M. F., Gangsei, L. E., Kongsro, J., Wass, M. N., Grindflek, E. H., & Harlizius, B., 2019. A QTL for number of teats shows breed specific effects on number of vertebrae in pigs: bridging the gap between molecular and quantitative genetics. Frontiers in genetics, 272.

Vinje, H., Brustad, K. H., Heggli, A., Sevillano, C., Van Son, M., Gangsei, L.E., 2023. Classification of breed combinations for slaughter pigs based on genotypes -Modelling DNA samples of crossbreeds as fuzzy sets from purebred founders. Frontiers in Genetics, 14.

Walstra, P., & Merkus, G. S., 1996. Procedure for assessment of the lean meat percentage as a consequence of the new EU reference dissection method in pig carcass classification: based on discussion in the EU Management Committee on Pig Meat and based on discussions with dissection experts during a meeting on May 18-19, 1994 at Zeist, NL (No. 96.014). ID-DLO.

ISBN: 978-82-575-2131-8
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

119578 / MEDIEHUSET-ANDVORD.NO