

DISSERTATION

ASSESSING THE PREDICTIVE VALUE OF DAIRY FACIAL BIOMETRICS FOR
MEASURES OF PRODUCTIVITY, HEALTH, AND SOCIAL DOMINANCE

Submitted by

Catherine McVey

Department of Animal Sciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2018

Master's Committee:

Advisor: Pablo Pinedo

Temple Grandin

Bailey Fosdick

Fiona Maunsell

Copyright by Catherine McVey 2018

All Rights Reserved

ABSTRACT

ASSESSING THE PREDICTIVE VALUE OF DAIRY FACIAL BIOMETRICS FOR MEASURES OF PRODUCTIVITY, HEALTH, AND SOCIAL DOMINANCE

The purpose of this thesis was to identify and characterize robust correlations between variations in bovine facial morphology and measures related to productivity, longevity, and social temperament in dairy cattle. In humans, use of facial features as indicators of health and personality dates back several thousand years in both Eastern and Western cultures. Historic records date similar techniques back at least two centuries in animals, and it is still practiced by prominent modern horse trainers. While research in humans has largely focused on the predictive potential of singular facial traits for targeted personality traits and health risks, recent research has underscored the value of comprehensive assessments of facial morphology in the prediction of more complex outcome measures such as diagnosis of autism spectrum disorders. Research in animals has similarly focused on targeted facial traits, but results of the fox farm experiment, a foundational study in the field of behavioral genetics, suggests that a more holistic analysis of facial morphology could correlate to a broader range of traits related to temperament, reproduction, and health.

The first chapter of this thesis details the process of developing image analysis algorithms capable of comprehensively quantifying subtle variations in bovine facial structures. Here a novel geometric approach was developed to produce more intuitive measures of facial shape. The statistical properties of these geometric biometrics were then compared to those of simple normalized linear facial measurements to determine which measurement system was better-suited

to subsequent inclusion in statistical models. This was done by acquiring bilateral images of lactating Holstein dairy cows at the feed bunk over a series of three subsequent days. Images were annotated with anatomical reference points and pixel coordinates extracted using the image processing tools in MatLab programming environment. This process was repeated in two separate annotation replicates, from which two sets of geometric and normalized length measures were calculated. Subsequent analyses of between-photo error terms revealed geometric biometrics to be slightly more resistant to variations in image resolution, particularly for smaller facial traits. Nested mix models were used to quantify sources of variance related to cow, bilateral asymmetry, between-day error, and within-photo error. Analysis of these results indicated that geometric biometrics demonstrated a slight advantaged over normalized length measures with respect to measurement repeatability, particularly for larger facial structures. Finally, geometric biometrics demonstrated lower levels of correlation in error between metrics as compared to normalized length measures, a common simplifying assumption for many standard statistical models.

The second chapter explores correlations between facial biometrics and measures of genomic merit for productivity, fertility, and health. Images were generated from a convenience sample of 594 mature milk cows from a fully genotyped purebred Holstein herd. One lateral image, either from the left or the right side of the face, was acquired from each cow while moving through the parlor and sorting stocks according to their normal farm routine. Annotation of these images with anatomical reference points was performed in two replicated, with the resulting 60 biometric valued computes averaged over replicated to reduce measurement error. These biometrics were then combined with genomic estimates for standard structure traits as candidate predictor variables. A total of 23 response traits were considered comprising both the standard Holstein genomic panel and the Zoetis Clarified health panel. Three statistical models, optimized using a

standard cross-validation scheme and validated with a fully blinded hold-out set, were used to explore correlations between biometrics values and these response values: LASSO, penalized smoothing spline, and boosted regression tree. Results indicated that, while biometric did not provide reproducible improvements in predictive performance over structure traits, a significant number of biometric terms were included in several response models, particularly those related to calving ease and still births. Further, several biometrics were retained multiple independent response models, indicating they might be indicators for more broadly adaptive traits. Finally, results of the spline and regression tree models yielded some evidence for significant nonlinearity and interaction effects, suggesting that the relationship between facial biometrics and genomic merit may be more complex than a simple linear model.

Finally, chapter 3 explored relationships between facial biometrics and estimates of social dominance. Daily milk order data was collected over a 150 day observation period for a closed herd of 203 organic milking cows – the same animals photographed for analyses in chapter 1. Exploratory data analysis revealed milk order to be dynamic over this time range, and PCA visualizations indicated a significant shift in milk order midway through the observation period when cows were granted access to pasture. Rank order was thus calculated separately for pen and pasture environments using 31 and 50 days of milk order records respectively, which in turn boasted complete records for 186 and 182 cows respectively. Weighted adjacency matrices were generated from pen and pasture data, where an incidence of a directed dyad was defined as one cow entering the milking parlor directly ahead of another. These adjacency matrices were augmented with information from indirect social interactions quantified via a percolation algorithm of length 3 through the network using the *Perc* package (Fujii *et al* 2016). Augmented adjacency matrices were then converted to a beta random field, to which an annealing algorithm

was then applied to generate an optimized linear rank order. Rank estimates generated from pen and pasture data proved surprising uncorrelated ($R^2=0.004$). A highly significant correlation was found between pen rank and bilateral estimate of Nostril Position Angle, a trait traditionally associated with dominance in horses, and also a significant predictor of production traits in Chapter 2. There was also some evidence that biometrics calculated from the right side of the face offered a slight advantage in predicting pen rank, despite the inherent increase in measurement error. Finally, the unaugmented adjacency matrices was used to calculate the assortativity of biometric values within the network. Eye length proportion demonstrated significant negative assortativity within both the pen and pasture networks. Additionally, the overwhelming majority of biometrics demonstrated negative assortativity values, which while not individually significant, may indicate an overall preference of cows for a more heterogenous social structure.

ACKNOWLEDGEMENTS

I would like to thank the National Science Foundation Graduate Research Fellowship Program for providing the funds to support my studies and this research. Additionally, I would like to thank Dr. Juan Velez and Aurora Organic Dairy personnel, as well as Don Bennink and North Florida Holsteins team for providing access to their animals and records to make this project possible. We are also grateful to Dr. Francisco Peñagaricano for his assistance in providing genotypic data for the North Florida Holstein herd.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	vi
LIST OF TABLES.....	viii
LIST OF FIGURES.....	x
Review of Literature.....	1
Introduction.....	1
Early History.....	2
Modern Research In Humans.....	6
Research in Livestock.....	14
Final Remarks.....	21
REFERENCES.....	22
CHAPTER 1 – GEOMETRIC BIOMETRICS AS A ROBUST APPROACH TO THE QUANTIFICATION OF LIVESTOCK PHENOTYPE.....	27
Introduction.....	27
Materials and Methods.....	42
Results – Eye Biometrics.....	49
Results – Muzzle Biometrics.....	61
Results – Topline Biometrics.....	74
Results – Forehead Biometrics.....	86
Discussion.....	103
Conclusions.....	105
REFERENCES.....	107
CHAPTER 2 – FACIAL BIOMETRICS AS PREDICTORS OF GENOMIC MERIT.....	109
Introduction.....	109
Materials and Methods.....	110
Results.....	117
Discussion.....	126
Conclusions.....	129
Implications.....	129
REFERENCES.....	130
CHAPTER 3 – FACIAL BIOMETRICS AS PREDICTORS OF SOCIAL RANK.....	131
Introduction.....	131
Materials and Methods.....	132
Results.....	140
Discussion.....	151
Conclusions.....	155
Implications.....	156
REFERENCES.....	157
EXECUTIVE SUMMARY.....	159
APPENDIX.....	162

LIST OF TABLES

TABLE 1 - Proportion of total variability in eye biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps).....	53
TABLE 2 - Proportion of total variability in eye biometrics attributed to error in image acquisition (variance between days/photos).....	54
TABLE 3 - Repeatability of Geometric Eye Biometrics from a Single Coordinate Extraction...	55
TABLE 4 - Repeatability of Geometric Eye Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction.....	56
TABLE 5 - Pairwise Correlation between Geometric Eye Biometrics Coordinate System 6.....	59
TABLE 6 - Proportion of total variability in muzzle biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps).....	66
TABLE 7 - Proportion of total variability in muzzle biometrics attributed to error in image acquisition (variance between days/photos).....	67
TABLE 8 - Repeatability of Geometric Muzzle Biometrics from a Single Landmark Coordinate Extraction.....	68
TABLE 9 - Repeatability of Geometric Eye Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction.....	69
TABLE 10 - Pairwise Correlation between Geometric Muzzle Biometrics.....	71
TABLE 11 - Proportion of total variability in topline biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps).....	79
TABLE 12 - Proportion of total variability in topline biometrics attributed to error in image acquisition (variance between days/photos).....	80
TABLE 13 - Repeatability of Geometric Topline Biometrics from a Single Landmark Coordinate Extraction.....	81
TABLE 14 - Repeatability of Geometric Topline Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction.....	82
TABLE 15 - Pairwise Correlation between Geometric Topline Biometrics.....	83
TABLE 16 - Proportion of total variability in topline biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps) – Part A.....	91
TABLE 17 - Proportion of total variability in topline biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps) – Part B.....	92
TABLE 18 - Proportion of total variability in forehead and jaw biometrics attributed to error in image acquisition (variance between days/photos) – Part A.....	93
TABLE 19 - Proportion of total variability in forehead and jaw biometrics attributed to error in image acquisition (variance between days/photos) – Part B.....	94
TABLE 20 - Repeatability of Geometric Topline Biometrics from a Single Coordinate Extraction – Part A.....	95
TABLE 21 - Repeatability of Geometric Topline Biometrics from a Single Coordinate Extraction – Part B.....	96
TABLE 22 - Repeatability of Geometric Forehead and Jaw Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction – Part A.....	97
TABLE 23 - Repeatability of Geometric Forehead and Jaw Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction – Part B.....	98

TABLE 24 - Pairwise Correlation between Geometric Forehead and Jaw Biometrics.....	100
TABLE 25 - LASSO Model Performance on Training and Validation Data.....	118
TABLE 26 - Count of Influential Variables Across Response Variables for LASSO Models...	119
TABLE 27 - Significance of Spline Components Across Genomic Response Models.....	122
TABLE 28 - Performance of Boosted Regression Models.....	124
TABLE 29 - Count of Influential Variables Across Response Variables for Regression Tree Models.....	125
TABLE 30 - Kendal Tau Estimates for correlation between rank order and facial biometrics...	146
TABLE 31 - Performance of boosted regression tree models against pen and pasture rank.....	147
TABLE 32 - Performance of boosted regression tree models against estimates of vertex betweenness.....	149
TABLE 33 - Assortativity of facial biometrics within the pen and pasture networks.....	150

LIST OF FIGURES

FIGURE 1 - Basic Anatomy of a Digital Image (Singh 2012).....	29
FIGURE 2 - Impact of Out-of-Plane Variations in Face Angle on Coordinate Locations.....	33
FIGURE 3 - Illustration of Elimination of Scale Effect by Division Operator.....	34
FIGURE 4 - When horizontal eye landmarks are constant, error in point selection leads to triangular relationship between edges.....	36
FIGURE 5 - Example of Geometric Biometric using Orthogonal Projects.....	38
FIGURE 6 - Orthogonalization of Error Component.....	40
FIGURE 7 - Example of an Interpolated Landmark Point (red).....	41
FIGURE 8 - Frame-to-Face Ratio.....	45
FIGURE 9 - Overall Face Angle.....	46
FIGURE 10 - Overall Face Angle.....	47
FIGURE 11 - Above – landmark points of the eye; below – anatomical lines of the eye.....	49
FIGURE 12 - Distribution of 3 rd & 4 th moments for normalized length & geometric eye biometrics.....	50
FIGURE 13 - Comparison of Between and Within-Day Repeatability of Normalized and Geometric Eye Biometric Measures.....	52
FIGURE 14 - Comparing level of correlation between pairwise combinations of normalized length and geometric eye biometrics.....	58
FIGURE 15 - Comparing levels of correlation in error terms between normalized length and geometric eye biometrics.....	60
FIGURE 16 - Above – Proportion of error attributed to variations in image scale and rotation; Below – Proportion of error attributed to variations in camera position.....	61
FIGURE 17 - Landmark points of the muzzle and geometric biometrics of the muzzle.....	62
FIGURE 18 - Distribution of 3 rd & 4 th moments for normalized length & geometric muzzle biometrics.....	63
FIGURE 19 - Comparison of Between and Within-Day Repeatability of Normalized and Geometric Muzzle Biometric Measures.....	64
FIGURE 20 - Comparing level of correlation between pairwise combinations of normalized length and geometric muzzle biometrics.....	70
FIGURE 21 - Comparing levels of correlation in error terms between normalized length and geometric muzzle biometrics.....	71
FIGURE 22 - Above – Proportion of error attributed to variations in image scale and rotation; Below – Proportion of error attributed to variations in camera position.....	73
FIGURE 23 - Relationship between changes in Nostril Height Proportion and changes in Face Angle and distance between camera and cow.....	74
FIGURE 24 - Landmark points of the topline and geometric biometrics of the topline.....	75
FIGURE 25 - Distribution of 3 rd & 4 th moments for normalized length & geometric topline biometrics.....	76
FIGURE 26 - Comparison of Between and Within-Day Repeatability of Normalized and Geometric Topline Biometric Measures.....	77
FIGURE 27 - Comparing level of correlation between pairwise combinations of normalized length and geometric topline biometrics.....	84

FIGURE 28 - Comparing levels of correlation in error terms between normalized length and geometric topline biometrics.....	85
FIGURE 29 - Above – Proportion of error attributed to variations in image scale and rotation; Below – Proportion of error attributed to variations in camera position.....	86
FIGURE 30 - Landmark points and geometric biometrics of the forehead and jaw.....	87
FIGURE 31 - Distribution of 3 rd and 4 th moments for normalized length and geometric forehead and jaw biometrics.....	88
FIGURE 32 - Comparison of Between and Within-Day Repeatability of Normalized and Geometric Forehead and Jaw Biometric Measures.....	89
FIGURE 33 - Comparing level of correlation between pairwise combinations of normalized length and geometric forehead and jaw biometrics.....	101
FIGURE 34 - Comparing levels of correlation in error terms between normalized length and geometric forehead and jaw biometrics.....	101
FIGURE 35 - Above – Proportion of error attributed to variations in image scale and rotation; Below – Proportion of error attributed to variations in camera position.....	103
FIGURE 36 - Distribution of cows ages at start of trial.....	138
FIGURE 37 - Examples of quantile plots of milk order over time by individual cow demonstrating both highly variable (left) but potentially cohesive (right) patterns in milk order.....	140
FIGURE 38 - Visualizations of dimensionality results from PC analysis by cow.....	141
FIGURE 39 - Results of PCA to approximate and visualize overall herd structure.....	142
FIGURE 40 - Visualization of the first three principal components embedding milking observations.....	143
FIGURE 41 - Heatmap visualization of pair-wise dominance probabilities ordered by descending rank order for pen data (left) and pasture data (right).....	144
FIGURE 42 - Comparison variable importance values from boosted regression tree results....	147
FIGURE 43 - Relationship between estimated rank and vertex betweenness.....	149

REVIEW OF LITERATURE

Introduction

The agricultural sciences are unique amongst the scientific fields in a number of ways – their dedication to extension, the diversity of technical skill sets encompassed, but perhaps the most romantic distinction is their depth of history. Most scientific disciplines are relatively modern phenomena, seldom dating back more than a few centuries. Their seminal papers have publication dates and their founders have names. But the pursuit of better means of raising livestock reaches back millennia. A complete record of agricultural knowledge would not only predate print, but writing itself, and the names of the field’s earliest innovators - the earliest cultivators of wheat, the first herdsman and horse tamers - have long been forgotten. Thus, innovation in agriculture is unique in that modern problems are not only solved by novel scientific discovery, but also creative repurposing of ideas that may extend backwards in time many generations. The ancient Irish understood the value of feeding seaweed as cattle fodder (Patterson 1994; Walling 2014) long before recent advances in technology revealed its promise in curbing livestock methane emissions (Kinley *et al* 2016; Machado *et al* 2014). And ancient Bedouin horse breeders, some of the earliest chroniclers of livestock bloodlines, realized the value of keeping meticulous pedigree records (Asil Araber 2007) millennia before Henderson and his contemporaries began to lay the mathematical foundations for systematic pedigree analysis in livestock selection (Gianola & Rosa 2015).

While the past offers agriculturalists a rich and practically endless source of inspiration for new research, it also presents unique challenges in the pursuit of objective scientific results. Techniques developed over years of trial and error as opposed to rigorous experimentation, whether that be through lack of inclination, or perhaps more formidably, a lack of technological

means, are invariably susceptible to human bias. And the longer such an idea stays in this limbo between discovery and scientific verification, the more opportunity is presented for biological phenomena to be swayed by the tides of culture, politics, and religion. Such is the case for the prospective biological relationship between facial morphology and individual variation in temperament and health.

Early History

Physiognomy, defined most broadly as the practice of discerning character in the face or form of the body, can be traced back in Western consciousness to the ancient Greeks (Cox 2003). Much of this early work seems focused on physiognomy as a tool for medical diagnostics (McKeown 2016). But philosophical musings on the subject, which focused on the face as a means of discerning the character of a man, take a decidedly more moralistic position. An oft recounted story in early texts is of the physiognomic reading of Socrates, wherein Zopyrus attributed all manner of vice to the shape of his face and neck. Instead of rejecting these claims, however, Socrates confirmed the veracity of this reading, and admonished to his students that it was his dedication to logic that prevented him from being governed by this predisposition towards such moral failings (Hoyland 2006; Hunfeld 2008).

As western medicine and the science of human anatomy advanced in the ensuing centuries, the connections between facial morphology and health appear to have faded, leaving the focus mainly on assessment of character, but that too seemed to wane in popularity. Giambattista della Porta's 1586 publication on physiognomy *De humana physiognomonia libri IIII* has more the feel of a coffee table novelty than a rigorous academic text, with a number of striking illustrations contrasting the facial traits of man to various animal forms and their corresponding behavioral

archetypes. Thomas Hill's 1556 text *A brief and most pleasant epitomye of the whole art of phisiognomie*, on the other hand, features stronger religious undertones. Though largely a recounting of work from various early Greek physiognomists, his text positions physiognomy as a means of discerning the inherent sinful nature of man, and extends Socrates expositions on the redeeming qualities of logic to the importance of divine salvation (Hunfeld 2008).

Academic interests in physiognomy began to rekindle in the 19th century with the popularization of the natural sciences and Darwin's theories of evolution. Darwin touches briefly on some of these ideas in his book *The Expression of the Emotions in Man and Animals*, but largely defers to the expertise of his contemporary Sir Charles Bell, and praises his 1806 publication *The Anatomy and Philosophy of Expression*. While this work places greater emphasis on objective measurement of facial features, with illustrations that emphasize the importance of angle, proportion, and underlying anatomical features, there also appear clear influences from phrenology, a more modern concept relating the size and shape of the head to cognitive functions. Bell's text, while perhaps one of the most scientific explorations of physiognomy to that date, also illustrates a decided shift in the underlying rhetoric surrounding this proposed biological phenomenon. Where earlier works had emphasized the redemption of character flaws through awareness, physiognomy post-evolution theory take a decidedly more deterministic tone. By the early 1900's, examination of facial features had become a largely pseudo-scientific spinoff of early criminology, and included work by such prominent figures as state Supreme Court justices (Cox 2003) and even Sir Francis Galton, an early innovator in the area of fingerprint identification (New Zealand Police Museum 2000). These ideas subsequently fed into increasingly popular but largely racially motivated assertions that criminality was a heritable trait (New Zealand Police Museum 2000; Wolfgang 1961), an idea which in turn drove the rhetoric behind turn-of-the-

century eugenics and forced-sterilization programs in numerous Western countries (Curators of the University of Missouri 2012), and ultimately the mass genocide of the Nazi Regime (*Boaz 2012*). As these ideas are as abhorrent as they were unscientific, no further description will be afforded to them here. Suffice to say that the notable dearth of scientific work in the area of physiognomy among western academics throughout the latter half of the 20th century was not without clear motive.

A review of the early history of face reading would not be complete, however, without also an appraisal of eastern traditions. Modern practitioners tout that eastern physiognomy, or Mien Shiang (“face reading”), dates back to early Taoist philosophies. Guiguzi, or the Ghost Valley Scholar (481-221 BC) is cited as offering the earliest written references to physiognomy (George 2014; Kohn 1986). The earliest text fully dedicated to the art of body divination, however, is not found until the *Shenxiang Quanian* (“Complete Guide to Spirit Physiognomy”) in the 10th century, though historic records indicate that physiognomists held prominent places in the Chinese court before this (Kohn 1986). Given that many such texts appear to have been lost, and even fewer have been translated to Western languages (Kohn 1986; McCarthy 2007), it is difficult to determine how closely Western practitioners abide to original teachings, or to what degree their work may be influenced by Western ideas, but most teach the same core themes.

Whereas the medical dimension of Western physiognomy was lost over time, these diagnostic properties are central to eastern face reading. This difference may have been driven by the cultural distinction that, in the East, dictums of modesty traditionally prevented physicians from touching females patients, forcing them to rely more heavily on visual diagnostics. Specific regions of the face are attributed to the five phases - wood, fire, earth, metal, water - found throughout traditional Chinese medicine. These elemental factors are in turn attributed to

groupings of organs or larger organ systems (Bridges 2012; Haner 2008; McCarthy 2007). These elemental regions largely coincide with the facial markers of energy meridians (Bridges 2012). Wrinkles, discoloration of the skin, or disfiguration of underlying cartilage are all signs of energy imbalances within an elemental region, which will over time manifest in health problems in the corresponding organs. Asymmetrical irregularities were considered particularly telling as, depending on the individual's gender, one side of the face was attributed to interactions with the outer world and social dealings, whereas the other side of the face was reflective of the individual's inner world and spiritual dealings (Bridges 2012; Haner 2008). Finally, some areas of the face were thought to correspond with specific stages of development, allowing the physician to assess an individual's medical history as far back as infancy to distinguish between chronic and acute stressors (Bridges 2012; Haner 2008).

Once a diagnosis of an energy imbalance was made, facial indicators were also referenced to direct treatment. The shape of the forehead and philtrum, where the Du and Ren channels converged, was said to be indicative of an individual's Qi, or life force (Bridges 2012; McCarthy 2007). Markings in these regions indicated that an individual was drawing excessively on their energy reserves, and that major life changes were needed to avoid major illness. The relative size of the upper, middle, and lower portions of the face, divided at the level of the eyebrows and nostrils along the central meridian, were indicative of how an individual responded to outside stimuli, and could be used to bring their decision-making processes more in line with their natural preferences (Bridges 2012; Haner 2008). Within each elemental region of the face, distinctive shapes of finer facial features were prescribed to quite detailed dimensions of personality, which together culminated into a broader elemental personality type. Such classes of personality could be positive or negatively nurtured, where the latter could lead to predictable health complications

(Bridges 2012; Haner 2008; McCarthy 2007). While the tenets of Eastern face reading are by no means scientifically verified, their emphasis on the diagnostic potential of the face, particularly the use of facial shape as indicators of both past and future health complications, arguably makes this historic system of facial inference more approachable to rigorous and objective study.

Modern Research In Humans

In more recent decades, a handful of targeted facial traits and their relationships to health and personality have captured the interest of modern scientists. One such line of work is the presence of diagonal creases across the earlobe as a potential indicator of coronary heart disease. This relationship was first proposed in a paper in the *New England Journal of Medicine* in 1973 (Frank 1973). In the following decades, it remained a contentious conclusion, as reliable diagnosis of artery disease could be difficult, and age served as a significant confounder. A large cohort study by Shmilovich *et al* (2012) offers perhaps the most complete examination of this phenomena to date. In their fully blinded study of 430 mixed-ethnicity patients, the presence of a diagonal earlobe crease (DELC) was established by consensus by two independent observers, and the presence and severity of coronary artery disease (CAD) was quantified using CT angiography results analyzed by two study-blinded experts in medical imaging using both a 0-4 scale for presence of disease in main arteries and the American Heart Association's 15 segment coronary artery tree model. Presence and severity of CAD were coded as binary responses, and analyzed using multivariate logistic regression analysis. After adjusting for a range of confounding factors - gender, diagnosis of diabetes mellitus, history of smoking, family history of premature CAD, symptoms of chest pain, presence of hypertension, presence of hyperlipidemia - diagonal earlobe creases remained significantly correlated to the presence, extent, and severity of CAD.

A number of physiological mechanisms underlying a correlation between heart disease and DELC have been proposed, though none have been fully validated. Perhaps the most straight forward is that wrinkling in the ear is simply an indicator of vascular disease, resulting in skin atrophy as the underlying connective tissue matrix is starved of nutrients and begins to break down, though this theory does not explain why the ear specifically appears such a good indicator (Evrengül *et al* 2004). It has also been suggested that tissues of the myocardium and ear lobe are generated from the same genetically originated end-arterioles, and thus may be commonly influenced by genetic factors or mutual biochemical pathways (Evrengül *et al* 2004). Some early work suggested a link between the atherosclerotic C3-F gene and increased levels of B27, though the statistical rigor of these results seem somewhat dubious, and confirmation of these results using modern genomic techniques has not been pursued (Kristensen 1980). A more recent pilot study in Japan (Higuchi 2009) compared telomere length, a proposed indicator for biological aging of the cardiovascular system, in male patients with and without bilateral earlobe creases that were match by age and risk factors for metabolic syndrome (glucose intolerance, hypertension, dyslipidemia, and visceral fat accumulation). They found that the telomere length in the peripheral blood cells of men with ear lobe creases were significantly shorter compared to men without creases. These results suggest that earlobe creases might be an outward indicator of oxidative stress and inflammation.

Another facial feature that has recently received targeted interest is face width-to-height ratio (fWHR) as an indicator of aggression. This idea can be traced back to two early studies that determined fWHR to be a sexually dimorphic trait. The first, a study of an ontogenetic series of 121 skulls from a modern native South African population (68 male, 53 female), regressed facial measurements against age to reveal that, while measures of facial height did not differ significantly

between genders, growth curves for bizygomatic width diverged at puberty between male and female skulls. Further, they determined that the resulting sex difference in face width-to-height ratios could not be fully accounted for by ontogenetic scaling, and suggested that this size-independent facial variant could be a target for mate selection as a physical indicator of ‘hormonal markers’ (Weston *et al* 2007). The second study by Carre and McCormick (2008) corroborated a statistically significant difference in fWHR between sexes in a convenience sample of 88 North American college students (37 male, 51 female) of mixed ethnicity, where facial metrics were extracted from digital images of live subjects with high inter-rater measure reliability ($r > 0.9$). More recent research, however, has failed to confirm the presence of sexual dimorphism of fWHR for larger samples of 2D and 3D images where ethnicity and age were more tightly controlled, suggesting these earlier results could simply reflect sampling bias (Lefevre *et al* 2012). In fact, 3D images revealed a statistically significant trend in the opposite direction of the original body of research, with women demonstrating larger fWHR than men, though this trend also became insignificant when Body Mass Index (BMI) was incorporated into the model, which may reflect important differences in measurement of fleshy traits between sexes.

Though the sexually dimorphic nature of this trait remains contested, this has not prevented researchers from exploring correlations between fWHR and a number of masculine personality traits. In the seminal paper, Carre and McCormick (2008) reported significant correlations to two measures of aggressions. In the first, 88 undergraduates participated in a modified Point Subtraction Aggression Paradigm (PASP). While under the impression that they were competing against another student for a monetary prize, and not in reality a pre-scheduled computer program, students were able to earn points by utilizing one of three buttons: one that added points to their score, one that removed points from their opponent’s score, and one that protected their own points

from their opponent. Use of button two was tracked as an indicator of reactive aggression (Cherek 1981; Gerra *et al* 2007), and subsequent regression models revealed fWHR to be a significant predictor of PASP results for males ($p = 0.02$) but not for females ($p = 0.27$). Additionally, Carre and McCormick generated from freely available web sources two data sets of hockey players that participated in the Canadian University (21 players) and professional (127 players) leagues that consisted of both front-facing facial photographs and total penalty minutes incurred during the 2007-2008 season. They found that facial width-to-height ratio explained 29% of the variability in total penalty minutes among college players ($p = 0.01$) and 9% of the variability in total penalty minutes among professional players ($p = 0.005$). More recent research, however, that utilized a larger sample of players ($n = 518$) from all 30 NHL teams and accounted for variability in player size reported much lower magnitudes of correlation between fWHR and total career penalty minutes adjusted for total games played (Deaner *et al* 2012), which suggests confounding factors could be at play in Carre and McCormick's secondary results.

Subsequent analyses have further explored the relationship between face width-to-height ratio measures of aggression, as well as a broader range of personality traits. Carre, McCormick and Mondloch (2009) presented 42 undergraduate volunteers (32 women, 15 men) with a randomized sequence of images comprised of 24 clean-shaven Caucasian college-age males and asked them to predict their aggressive reactivity using a 7-point scale. They determined that observer estimates of aggression were significantly and positively correlated to the face width-to-height ratio displayed in the corresponding photo ($p < 0.001$). Subsequently, observer estimates of aggression correlated positively and significantly with PASP aggression scores of the corresponding photo subjects ($p < 0.001$). They determined that together these results suggest that fWHR might qualify as an honest signal of aggressive behavior. Haselhuhn and Wong (2012)

found evidence of significant correlations between face width to height ratio and use of deceptive negotiation strategies, as determined by applying the Bullard House negotiation exercise to a class of 192 masters students in business administration, but this effect only proved significant for male participants ($p = 0.01$). A significant correlation between fWHR and self-reported feelings of power was also reported, though fWHR still retained a marginally significant relationship to deceptive behaviors ($p = 0.06$) when regressed with power. Stirrat, Stulp, and Pollet (2012) mined skeletal morphology metrics from US forensics databases to reveal a significant relationship between fWHR and risk of dying by contact violence in men ($p = 0.012$), though here increased risk of homicide was actually associated with males with narrower facial features. Lewis, Lefevre, and Bates (2012) were even able to discern a significant correlation between achievement drive and fWHR ($p < 0.01$) from historic images of 29 US presidents. But perhaps the most intriguing positive result for fWHR comes from fMRI studies performed by Carre, Murphey, and Hariri (2013). Working off the theory that variations in fWHR were primarily driven by individual variation in pubertal testosterone levels (Verdonck et al 1999), and given that animal models suggest pubertal testosterone influences development of neural structures in the brain, the brain function of 64 healthy adults (28 men) were tracked while presented with a randomized sequence of shapes and emotional faces. Results indicated that right amygdala activity in response to aggressive faces was significantly correlated with self-reported scores for physical aggression, but only for men with high fWHR, which suggests that fWHR might serve as a physical indicator for variations in development that can have a persistent modulating influence on behavioral responses.

While the majority of work on fWHR has demonstrated significant correlations to a range of behavioral metrics, the magnitude of such associations consistently appear quite small (Haselhuhn *et al* 2015). This may be attributable to the fact that most historic systems for face

reading emphasize a more holistic assessment of a broader range of facial traits than is seen in either body of research exploring face width-to-height ratios or diagonal earlobe creases. Additionally, neither of these lines of research acknowledges a synergistic relationship between behavior and health found in the traditional Eastern teachings. A smaller but promising body of research in humans that perhaps aligns more closely to historic Eastern face reading techniques is the use of facial morphology in the study of Autism Spectrum Disorders (ASD).

Face reading in the diagnosis of ASD can in fact be traced back as far as one of the field's original founders, Leo Kanner, who in endeavoring to emphasize the oft overlooked intelligence of his patients, would frequently point out the physiognomic merits of their facial shape (Cohmer 2014). In the following decades, epidemiological and clinical studies proposed a number of facial phenotypes that might serve as a visual indicator for autism spectrum traits. Rodier *et al* (1997) proposed an autistic facial phenotype comprised of reduced inter-pupillary distance, ptosis, or a drooping upper eyelid, strabismus, or eye misalignment, lop ears, and hypotonia in the lower face. Hammon *et al* (2008), on the other hand, suggested instead that autism spectrum disorders could be characterized by greater levels of facial asymmetry in both the affected individual and closely related relatives. But modern research into the potential biological underpinnings of autism and related disorders have underscored the complex genetic, epigenetic, and developmental relationships between the tissues that form the forebrain and face, suggesting that the relationship between facial phenotypes and clinical diagnoses of ASD may not be so straightforward a relationship as initially supposed (Aldridge *et al* 2011).

In 2011 Aldridge *et al* made the critical leap from simple anthropometric studies of targeted facial traits, to a high-dimensional statistical learning methodology based around the analysis of high-quality 3D images. The goal was to determine if facial biometrics could be used to distinguish

between typically developing (TD) and ASD children. Images were collected from 105 Caucasian boys (64 ASD, 41 TD) between the ages of 8 and 12. The 3D facial images were annotated with 17 anatomical landmark points, as defined by Farkas (1994), by two separate observers. The Euclidean distances between all unique pairwise combinations of points was computed to produce a set of 136 candidate predictor variables for each boy, and globally normalized to adjust for image resolution. A nonparametric bootstrapping approach was used to generate confidence intervals for the range of each facial biometrics within the control and treatment groups, revealing 39 of the 136 to be statistically distinguishable between treatment groups. Application of an unsupervised principal coordinate analysis algorithm subsequently showed modest separation with quite a bit of overlap between groups, but two subgroups of ASD boys were observed that appeared distinct from the main cluster. Closer assessment of subgroup traits revealed some distinct trends in clinical parameters. The first subgroup, characterized by reduced distances in the nasion, inner canthus, and glabella regions and increased distances in the mouth and chin regions, showed the severest forms of autism, with reduced performance on cognitive tests and higher levels of regression. Subgroup two, on the other hand, characterized by reduced distances in inferior nasion, philtrum, and lateral mouth area and increased distances in the upper face region, demonstrated a higher composition of Asperger diagnoses and marginally improved verbal scores.

In 2015 Obafemi-Ajayi *et al* performed a follow-up analysis on this image database, augmented with 11 additional ASD boys and two additional landmark coordinates, with the goal of identifying clinical subgroups of ASD individuals using facial biometrics. Geodesic, as opposed to Euclidean, distances between points were calculated in an attempt to better capture variation in soft tissue features. Optimal clustering was achieved using a k-means clustering algorithm of size $k = 3$, as determined by optimal scores on both the Davies-Bouldin and Calinski-Harabasz

clustering indices when compared to results from expectation maximization, self-organizing maps, and partitioning around medoid algorithms. Optimality of these clustering results were subsequently confirmed by training three types of classifiers - support vector machines, feed-forward multilayer perceptron, and random forests - on the full set of geodesic measures and comparing classification results using measures of prediction. Finally, feature selection of the most predictive geodesic distances was performed via consensus of three model reduction techniques: parallel scatter search, best first search, and linear forward selection. The result was a subset of 31 geodesic lengths that yielded either equivalent or superior classification results as compared to the full model. Of these, 12 metrics were determined to be significantly different between all three clusters using both ANOVA analysis and paired t-tests at the 0.05 significance level. In comparing these results to clinical tests, clusters 1 and 3 showed significant overlap with data from typically developing boys held out of the training data, but cluster 2 proved well distinguished from these controls. This cluster consisted primarily of boys diagnosed with ASD (79%) with the lowest incidence of Asperger Syndrome of any of the three clusters, along with the severest social and verbal regression scores, indicating that cluster 2 also represented some of the severest forms of ASD. Results of this more statistically sophisticated analysis thus mirror fairly closely the results of the original preliminary study: that severe ASD is physically distinguishable from typically developing boys and that those with mild ASD by facial phenotypes characterized by wider mouths and decreased facial height along the midline. Overall these results suggest that holistic assessments of facial structures using provided by modern computational tools and statistical techniques can produce robust predictive models for behavioral traits with medical implications.

Research in Livestock

While use of facial morphology as an indicator of livestock behavior and health may not have as rich a history as with humans, our domesticated partners certainly were not excluded from such practices. Perhaps one of the earliest written accounts prescribing personality attributes to features of the animal face comes from Major Roger Upton in his now classic memoir “Gleanings from the Desert of Arabia” (1881). In his thorough description of the physical type of the Keheilan, or genuine Arabian horse, Major Upton utilizes no less than 3 of the 13 the pages he devotes to the topic to define the distinctive characteristics of a well-bred Arabian’s facial structure. One excerpt from this account underlies the close ties placed between facial morphology and personality in such horses:

“Such a head is often supposed to denote a violent temper. It is the type, however, of the head of the Arabian horse, and is, we thought, more marked and to be seen more frequently among the Anazah tribes than elsewhere. Every Arabian horse may be said to have a high temper of some extent, but it is balanced or controlled by the power of the large and well-developed cerebrum. The head I have described of horses we have seen denotes the highest order of qualities - intelligence, energy, and unconquerable courage. It is almost human in its expression of nobility, dignity, and sagacity. Other horses have much fire, but it is but too often the habitual and only expression, not called forth by occasion and controlled at other times by higher organs; indeed, a spirit of the highest order is characteristic of the Arabian. With regard to the great development of the upper part of the head and the fineness of the muzzle, I have seen instances of the former measuring nearly two and a half to one; witness a measurement of thirty-seven inches over the forehead and under the jaws, taken in a line horizontal to the bone, against one of fifteen inches, or perhaps a line over, round the muzzle above the nostrils, and of perhaps just over thirty-seven inches around the forehead, and sixteen inches, or just under, round the muzzle; there may be examples of even greater difference.”

Unfortunately, it is difficult to determine if this account reflects authentic Bedouin breeding traditions, or simply a superposition of European equestrian attitudes prevalent in that period, as by his own accounts Upton was not exposed to breeding records that he deemed authentic. Such practices certainly seemed prevalent amongst Western horsemen at the turn of the century, often receiving passing mentions in popular horse training manuals like Professor Beery’s *Mail in Horse*

Training Course (Beery 1908). Perhaps the most thorough delineation of the association between equine facial morphology and personality comes from the more modern TTouch and TTeam systems (Tellington-Jones 1995). Developed by award winning horsewoman Linda Tellington-Jones and purportedly based on eastern gypsy traditions taught to her grandfather while training race horses for Czar Nicholas II in turn of the century Russia, this system assigns personality traits to both bony and cartilaginous features of the equine face, as well as the number and relative distribution of facial hair whorls. For some facial traits associated with stronger personality types, warnings of potential training and health complications are offered. Overall, the system emphasizes a more holistic and comprehensive approach to face reading as a means of better tailoring the training and management of a horse to their innate nature - a philosophy that seems to fall closely in line with the teachings of traditional Chinese face reading.

Unfortunately, very little of this antiquated knowledge has been subjected to the rigors of the modern scientific method. Perhaps the singular exception in the collective body of research in livestock management is a series of studies relating to facial hair whorl position in cattle. In the first academic report to suggest a connection between facial whorls and temperament, Tanner *et al* (1994) reported a relationship amongst dairy cattle between hair whorl position and laterality in the milking parlor. In a follow-up study, Grandin *et al* (1995) subsequently found a significant association between height of hair whorls, relative to position of the eyes, and ordinal measures of calm temperament in range-bred beef cattle. In this study, 1500 feedlot cattle were observed while undergoing routine management procedures in a squeeze chute. One observer scored the reaction of the cattle to restraint on a four-point scale, and a second observer scored their behavior on a three-point scale as they left the chute. They found that cattle with hair whorls above the eye were significantly more agitated when restrained ($P < 0.001$) and were also more excitable when exiting

the chute ($p < 0.01$), a trend that was consistent for both *Bos taurus* breeds and *Bos taurus* x *Bos indicus* crossbreds. In a subsequent study of over 1600 range-raised cattle utilizing a similar behavioral scoring system, Lanier *et al* (2001) found cattle with high hair whorls to be significantly more excitable in the auction ring ($p = 0.01$), and also that cattle with lateral displacement of hair whorls showed a greater amount of variability in behavior scores.

To expand upon these results Randle (1997) collected a broader range of temperament assessments on a group of 57 well-handled *Bos taurus* type beef cattle. The only significant association found with hair whorl position was for response to an unfamiliar human, with responses to novel objects, familiar humans, and performance on cognitive tests showing no correlation to whorl position, suggesting that this morphological indicator might only be deterministic of a very narrow range of conditional behaviors. In a separate study designed to confirm the robustness of results found in range-raised animals amongst more routinely handled cattle populations, Oloms and Turner (2008) repeated the methodology reported in the original study presented by Grandin, but here also collected continuous measures for whorl position, behavior in the chute, and flight speed leaving the chute. Using only 76 animals of various *Bos taurus* breeds, they confirmed the significant association found between ordinal measures of whorl position and behavior in the chute reported by Grandin *et al* (1995), but not with ordinal measures of flight speed. Additionally, they found a borderline significant linear association ($p = 0.056$) between whorl position, as expressed as a ratio normalized by overall face length, and ethogram data aggregated using principal component analysis to produce an overall measure of movement in the squeeze chute. They did not, however, find any significant correlations to measures of performance such as average daily gain (ADG).

Correlations have also been found between facial whorl morphology and several measures of fertility. In a study utilizing data from breeding soundness exams of 150 Angus bulls, Meola *et al* (2003) found that bulls with whorls that formed normal spirals, as opposed to those that presented as abnormal lines, had a significantly higher percentage normal spermatozoa ($p < 0.05$), and also that a significantly higher proportion of these bulls met the $>70\%$ normal spermatozoa cutoff. No significant associations were found, however, to measures of sperm motility or to scrotal circumference. A follow-up study by Evans *et al* (2005), however, failed to find any significant correlations between facial whorl morphology and measures of semen quality among Holstein AI bulls. These results may indicate the presence of breed-specific relationships between hair whorls and reproductive performance; alternatively, these results may simply reflect a sampling bias towards an inherently more uniform population with respect to fertility, as the majority of the animals utilized in this study were older and proven bulls, an explanation supported by the observation that this study population of Holstein studs also showed less variability in whorl morphology than reported previously among younger Angus animals.

Looking beyond research focused on livestock, however, there is additional evidence of a relationship between facial morphology and temperament in animal models. Perhaps the most compelling evidence comes from a foundational study in the field of behavioral genetics: the fox farm experiment. Over more than a 50-year period, geneticists at the Russian Institute of Cytology and Genetics selectively bred Russian silver foxes (*Vulpes vulpes*) based exclusively on behavioral measures for temperament traits related to tamability. After decades of intensive selection pressure, this population of domesticated foxes demonstrated a range of morphological and behavioral changes that closely mirrored traits seen across a range of domesticated species. Whereas control populations were highly fearful of humans, often exhibiting aggressive behavior

in their attempts to evade physical contact, the majority of pups born to the domesticated line of foxes actively sought out human contact, whimpering to attract attention and even fighting their littermates for the favor of their handlers. Researchers identified significant changes in developmental plasma corticosteroid levels of these highly social foxes, which resulted in an imprinting window that opened several days and closed several weeks later than their wild-type counterparts and closely resembled developmental landmarks seen in domesticated dogs. Beyond the neonatal stages of development, researchers also identified reductions in the activity of the adrenal glands of domesticated foxes, resulting in major reductions in baseline corticosteroid levels in the blood. They also observed significant increases in serotonin levels present in the brain of domesticated foxes, as well as associated enzymes and metabolites.

While changes observed in the behavioral traits under selection were impressive in both magnitude and rate, changes observed in physiological traits not placed under direct selection pressure were even more surprising. Within 10 generations, piebald coat patterns rarely found in wild populations were observed, first as star patterns on the face, and later so extreme that they mirrored the markings of modern border collies. Floppy ears and curled tails subsequently emerged in this domesticated population, followed by shorter legs and changes to the face that were so significant that underbites and overbites became notably more prevalent. Changes were even seen in the reproductive system. Domesticated foxes reached sexual maturity on average a month earlier than the standard farm fox, and demonstrated a significantly lengthened breeding season, with some females even breeding out of season - a feat fur farmers had previously failed to achieve in decades of concerted effort. To explain this broad suite of physiological changes, researchers proposed that, through strong selection pressure on behavioral traits, they had indirectly targeted genes exerting high-level control over early development, particularly those related to hormonal

control. By altering ontogenesis, they had in turn indirectly altered development on a broad suit of traits, a biological mechanism which might explain the consistent set of morphological and physiological changes seen across a range of temporally and geographically distinct domestication events (Trut 1999). Thus, this research not only suggests a genetic basis for the connection between facial morphology and behavior, but also underscores fundamental biological connection between facial morphology and a range of physiological traits.

The fox farm experiment is also not the only line of research to identify physiological indicators of subtle variations in developmental processes among both domestic and wild species. Academic interest in anatomical symmetry and the developmental processes reinforcing this highly conserved biological trend date back as far as Darwin (Palmer 1996). In 1962, however, researchers became interested in measures of physiological asymmetry as a practical and objective indicator of developmental stressors (Van Valen, 1962). The biological preface underlying this experimental technique was relatively simple: while the exact physiological drivers may not be fully understood, symmetry was clearly the developmental ideal for most mammalian features, and thus an animal should put as much energy as they had available towards developmental processes reinforcing structural symmetry. If, however, an animal were systematically stressed during development, less energy would be available to reinforce structural symmetry, and the chances of random divergences from symmetry would become more likely (Palmer 1996). Thus, when measures of bilateral traits are analyzed amongst a cohort of animals, developmental stressors should be detectable as significantly higher levels of variance in directional asymmetry (Graham *et al* 1993; Leary and Allendorf 1989; Palmer 2001).

Fluctuating asymmetry has been used to explore a number of biological stressors. In one of the field's seminal studies, Sumner and Huestis (1921) noticed, contrary to Mendelian principles, that crosses of highly inbred strains of laboratory mice produced greater levels of asymmetry in the F2 generation than in the F1 parents. They subsequently determined that structural asymmetry could be used to compare levels of genetic stress - inbreeding, hybridization, chromosomal abnormalities, mildly deleterious recessive genes - between populations. Fluctuating facial asymmetry has also been used extensively by ecologists and applied ethologists to compare levels of environmental stress in a number of animal species ranging from aquatic species (Ottaviano and Scapini 2010; Clarke 1993), to reptiles (Vervust *et al* 2008; Lazic *et al* 2013), to macaques (Newell-Morris *et al* 1989; Hallgrimsson 1999; Willmore *et al* 2007), and even poultry (Eriksen 2003; Yang 1998). This work not only underscores the link between variability in developmental processes and a wide range of physiological traits, but also the scope of genetic, epigenetic, and environmental influences that collectively drive such associations (Parsons 1990).

Final Thoughts

While research exploring the relationship between facial morphology and facets of behavior, reproduction, and health may be scattered, consistent results for a number of traits have been reported across a range of species. While the collective results of such studies frequently prove statistically significant, the predictive potential of individual morphological traits may be limited, suggesting that a holistic approach of analysis of facial morphology is needed. Direct physical metric of a range of morphological characteristics, however, present researchers with a number of practical restrictions to experimental investigation, particularly in the case of large and often difficult to handle livestock species. Thus, the principal goal of this thesis was to create and

validate computational tools to holistically quantify the facial morphology of livestock species that would prove both sufficiently robust in highly variable farm environments while also minimizing stressed placed on the animals themselves. To lay the groundwork for future research, this thesis also began to assess the performance of facial biometric in the prediction of a range of health, performance, and behavioral traits among both conventionally and organically managed dairy cattle.

REFERENCES

- Asil Club. "Asil Araber VI: The Noble Arabian Horses" 2007.
- Beery, Jesse. "Professor Beery's Mail Course in Horsemanship - Lesson Two: Disposition and Subjection." 1908. Reprint. 2010. CreateSpace Independent Publishing Platform.
- Boaz, R.E. "In Search of Aryan Blood: Serology in Interwar and National Socialist Germany." 2012. Central European University Press. Budapest. 162-163.
- Bridges, Lillian. "Face Reading in Chinese Medicine: Second Edition" 2012. Elsevier Inc.
- Carre, J.M., McCormick, C.M. "In your face: facial metrics predict aggressive behavior in the laboratory and in varsity and professional hockey players." 2008. Proceedings of the Royal Society B-Biological Sciences. 275:2651-2656.
- Carre, J.M., Murphy, K.R., Hariri, A.R. "What Lies Beneath the Face of Aggression." 2013. SCAN. 8: 224 - 229.
- Carre, J.M., McCormick, C.M., Mondloch, C.J. "Facial Structure is a Reliable Cue of Aggressive Behavior." 2009. Psychological Science. 20:10:1194-1198.
- Cherek, D.R. "Effects of smoking different doses of nicotine on human aggressive behaviour". 1981. Psychopharmacology 75: 339-345.
- Clarke, G.M. "Fluctuating facial asymmetry of invertebrate populations as a biological indicator of environmental quality." 1993. 82. 207-211.
- Cox, J.A. "Beginning Face Reading." 2003. Southstar Publishing.
- Curators of the University of Missouri. "Controlling Heredity: The American Eugenics Crusade". 2012. Online. <https://library.missouri.edu/exhibits/eugenics/origins.htm>
- Deaner, R.O., Goetz, S.M.M., Shattuck, K. Schnotala, T. "Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players." 2012. Journal of Research in Personality. 46: 235-238.
- Eriksen, M.S., Haug, A., Torjesen, P.A., Bakken, M. "Prenatal exposure to corticosterone impairs embryonic development and increases fluctuating asymmetry in chickens." 2003. British Poultry Science. 44:5. 690-697.
- Evans, R.D., Grandin, T., DeJarnette, J.M., Deesing, M., Garrick, D.j. "Phenotypic relationships between hair whorl characteristics and spermatozoal attributes in holstein bulls." 2005. Animal Reproduction Science. 85. 95-103.

- Evrengül H, Dursunog D, Kaftan A, Zoghi M, Tanriverdi H, Zungur M, Kiliç M. “Bilateral diagonal earlobe crease and coronary artery disease: a significant association.” 2004. *Dermatology* 209:271–275.
- Frank S. “Aural sign of coronary-artery disease.” *N Engl J Med* 1973; 289:327–328.
- Farkas LG: Anthropometry of the head and face in clinical practice. In *Anthropometry of the Head and Face*. 2 edition. 1994. Edited by: Farkas LG. New York: Raven Press. 71-77.
- George, Shrin V. “An Explanatory Study on Perceptual Processes in Fortune Telling”. 2014. Thesis. Sree Sankaracharya University of Sanskrit. Online.
<http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/79619>
- Gerra, G., Zaimovic, A., Raggi, M.A., Moi, G., Branchi, B., Moroni, M., Brambilla, F. “Experimentally induced aggressiveness in heroin-dependent patients treated with buprenorphine: comparison of patients receiving methadone and healthy subjects”. 2007. *Psychiatry Res.* 149: 201–203.
- Gianola, D., Rosa, G.J.M. “One Hundred Years of Statistical Developments in Animal Breeding.” 2015. *Annual Review.* 31. 1. 19-56.
- Graham, J.H., Freeman, D.C., Emlen, J.M., “Developmental stability: a sensitive indicator of populations under stress.” 1993. In “*Environmental Toxicology and Risk Assessment.*” Philadelphia (PA): American Society for Testing and Mate
- Grandin, T., Deesing, M.J., Struthers, J.J., Swinker, A.M. “Cattle with hair whorl patterns above the eyes are behaviorally agitated during restraint.” 1995. *Applied Animal Behaviour Science.* 46. 117-123.
- Hallgrimsson, Benedikt. “Ontogenetic Patterning of Skeletal Fluctuating Asymmetry in Rhesus Macaques and Humans: Evolutionary and Developmental Implications.” 1998. *International Journal of Primatology.* 20:1. 121-151.
- Hammond P., Forster-Gibson C., Chudley A.E., Allanson J.E., Hutton T.J., Farrell S.A., McKenzie J., Holden J.J., Lewis M.E. “Face-brain asymmetry in autism spectrum disorders.” 2008. *Molecular Psychiatry.* 13:614-623.
- Haner, Jean. “The Wisdom of your Face.” 2008. Hay House Inc. US.
- Haselhuhn, M.P., Ormiston, M.E., Wong, E.M. “Men’s Facial Width-to-Height Ratio Predicts Aggression: A Meta-Analysis.” *PLOS One.* 2015.
- Haselhuhn, M.P., Wong, E.M. “Bad to the Bone: facial structure predicts unethical behavior.” 2012. *Proceedings of the Royal Society B.* 279: 571-576.

- Higuchi, Y.; Maeda, T; Guan, JZ; Oyama, J; Sugano, M; Makino, N. "Diagonal Earlobe Crease are Associated With Shorter Telomere in Male Japanese Patients With Metabolic Syndrome." 2009. *Circ J.* 73: 274 – 279
- Hoyland, R.G. "Polemon's Encounter with Hippocrates and the status of Islamic physiognomy." 2006. *JSAI* 32.
- Hunfeld, C. "The "Science of the Countenance": Full-Bodied Physiognomy and the Cosmography of the Self in Seventeenth-Century England." 2008. Dissertation. Dalhousie University.
- Kinley, R.D., de Nys, R., Vucko, M.J., Machado, L., Tomkins, R.W. "The red macroalgae *Asparagopsis taxiformis* is a potent natural antimethanogenic that reduces methane production during in vitro fermentation with rumen fluid." 2016. *Animal Production Science.* 56. 282-289.
- Kohn, Livia. "A Textbook of Physiognomy: The Tradition of the "Shenxiang quanbian" 1986. *Asian Folklore Studies.* 45. 2. 227-258.
- Kristensen, BO. "Ear-lobe crease and vascular complications in essential hypertension". *Lancet* 1980. 265.
- Lafevre, C.E., Lewis, G.J., Bates, T.C., Dzhelyova, M., Coetzee, V., Deary, I.J., Perrett, D.I. "No evidence for sexual dimorphism of face width-to-height ratio in four large adult samples." 2012. *Evolution and Human Behavior.* 33:623-627.
- Lanier, J.L., Grandin, T., Green, R., Avery, D., McGee, K. "A note on hair whorl position and cattle temperament in the auction ring." 2001. *Appl Animal Behaviour Sci.* 73:93-101.
- Lazic, M.M., Kaliontzopoulou, A., Carretero, M.A. Crnobrnja-Isailovic, J. "Lizards from Urban Areas are More Asymmetric: Using Fluctuating Asymmetry to Evaluate Environmental Disturbance." 2013. *PLOS ONE.* 8:12. 1-9.
- Leary, R.F., Allendorf, F.W. "Fluctuating asymmetry as an indicator of stress in conservation biology." 1989. *Trends in Ecology and Evolution.* 4. 214-217.
- Lewis, G.J., Lefevre, C.E., Bates, T.C. "Facial width-to-height ratio predicts achievement drive in US presidents." 2012. *Personality and Individual Differences.* 52: 855-857.
- Machado, L., Magnusson, M., Paul, N.A., de Nys, R., Tomkins, N. "Effects of Marine and Freshwater Macroalgae on In Vitro Total Gas and Methane Production". 2014. *PLOS ONE.* 9-1.
- McCarthy, Patrician. "The Face Reader". 2007. Penguin Books. England.
- McKeown, J.C., Smith, J. "The Hippocrates Code: Unraveling the Ancient Mysteries of Modern Medical Terminology" 2016. Hackett Publishing Company

- Meola, M.G., Grandin, T., Burns, P., Deesing, M. "Hair whorl patterns on the bovine forehead may be related to breeding soundness measures." 2004. *Theriogenology*. 62. 450-457.
- Newell-Morris, L.L, Fahrenbruch, C.E., Sackett, G.P. "Prenatal psychological stress, dermatoglyphic asymmetry, and pregnancy outcome in pigtailed macaque." 1989. *Biol Neonate*. 56. 61-75.
- Obafemi-Ajayi, T., Miles, J.H., Takahashi, T.N., Qi, W., Aldridge, K., Zhang, M., Xin, S.Q., He, Y., Duan, Y. "Facial Structure Analysis Separates Autism Spectrum Disorders into Meaningful Clinical Subgroups." 2015. *Journal of Autism Developmental Disorders*. 45: 1302:1317.
- Ottaviano, O., Scapini, F. "Can fluctuating asymmetry in *Talitrus saltator* populations be used as an indicator of stress on sandy beach environments." 2010. *Oceanologia*. 52:2. 259-280.
- Palmer, A.R, Strobeck, C. "Fluctuating Asymmetry Revisited." 2001. In: *Developmental Instability: Causes and Consequences*. 2003. Online.
<http://www.biology.ualberta.ca/palmer.hp/pubs/03BookChapt/PalmerStrobeckChapt.pdf>
- Parsons, P.A. "Fluctuating Facial Asymmetry: an epigenetic measure of stress." 1990. *Biological Reviews*. 65:2. 131-145.
- Patterson, Nerys. "Cattle Lords and Clansmen: The Social Structure of Early Ireland". 1994. University of Notre Dame, USA.
- Randle, H.D. "Facial hair whorl position and temperament in cattle." 1998. *Applied Animal Behaviour Science*. 1998. 56. 139-147.
- Rodier PM, Bryson SE, Welch JP "Minor malformations and physical measurements in autism: data from Nova Scotia." 1997. *Teratology* 55:319-325.
- Shmilovich H, Cheng VY, Rajani R, Dey D, Tamarappoo BK, Nakazato R, Smith TW, Otaki Y, Nakanishi R, Gransar H, Paz W, Pimental RT, Hayes SW, Friedman JD, Tomson LEJ, Berman DS. Relation of diagonal ear lobe crease to the presence, extent, and severity of coronary artery disease determined by coronary computed tomography angiography. *Am J Cardiol* 2012;109:1283–1287
- Stirrat, M. Stulp, G., Pollet, T.V. "Male facial width is associated with death by contact violence: narrow faced males are more likely to die by contact violence." 2012. *Evolution and Human Health*. 33: 551:556
- Sumner, J.L., Huestis, R.R. "Bilateral Asymmetry and its relation to certain problems in genetics." 1921. *Genetics*. 6. 445-485.
- Tanner, M., Grandin, T. Cattell, M. Deesing, M. "The relationship between facial hair whorls and milking parlor side preference." 1994. *Journal of Animal Science*. 72:1. 207. Abstract.

- Tellington-Jones, Linda. "Getting in TTouch: Understanding and Influence Your Horse's Personality." 1995. Trafalgar Square Publishing: North Pomfret, Vermont.
- The New Zealand Police Museum. "Physiognomy, Photography and the criminal look." 2000. Online.
<http://www.police.govt.nz/about-us/history/museum/exhibitions/suspicious-looking-19th-century-mug-shots/physiognomy-photography-criminal-look>
- Trut, L.N. "Early Canid Domestication: The Fox Farm Experiment." 1999. *Scientific American*. 87. 160-169.
- Olmos, G., Turner, S.P. "The relationship between temperament during routine tasks, weight gain and facial whorl position in frequently handled beef cattle." 2008. *Applied Animal Behaviour Science*. 115. 25-36.
- Upton, Roger D. "Gleanings from the Desert of Arabia." 1881. London. Reprinted. 1985. Olms Presse: Zurich.
- U.S. National Library of Medicine. "The Horse: A Mirror of Man: Parallels in Early Human and Horse History". 2011. Online.
<https://www.nlm.nih.gov/exhibition/horse/physiognomy.html>
- VanValen, L. "A study of fluctuating asymmetry." 1962. *Evolution*. 16. 125-142.
- Verdonck, A., Gaethofs, M. Carels, C., de Zegher, F. "Effect of low-dose testosterone treatment on craniofacial growth of young boys with delayed puberty." 1999. *European Journal of Orthodontics*. 21: 137 - 143.
- Vervust, B., Dongen, S.V., Grbac, I., Van Damme, R., "Fluctuating Asymmetry, Physiological Performance, and Stress in Island Populations of Italian Wall Lizard" 2008. *Journal of Herpetology*. 42:2. 367-377.
- Walling, Philip. 2014. "Counting Sheep: A Celebration of the Pastoral Heritage of Britain." Profile Books LTD. London, UK.
- Weston, E.M., Friday, A.E., Lio, P. "Biometric Evidence that Sexual Selection has Shaped the Hominem Face." 2007. *PLOS One*. 8.
- Willmore, K.E., Klingenberg, C.P., Hallgrímsson, B. "The relationship between fluctuating asymmetry and environmental variance in rhesus macaques." 2005. 59:4. 898:909
- Wolfgang, M.E. "Pioneers in Criminology: Cesare Lombros" 1961. *Journal of Criminal Law and Criminology*. 52:4. 361-391.
- Yang, A. "Bilateral Asymmetry in Chickens of Different Genetic Backgrounds." 1998. Thesis.

CHAPTER 1 - GEOMETRIC BIOMETRICS AS A ROBUST APPROACH TO THE QUANTIFICATION OF LIVESTOCK PHENOTYPES

Introduction

In most standard linear models, model specification includes the simplifying assumption that covariates are known without error (Kutner et al 2005). For some simple categorical variables, like gender or herd of origin, this is perhaps true, excluding errors in data entry (Meyer 1997). Measurements that seek to capture more complex and dynamic sources of information, however, can seldom be represented without error. In this case, a thorough understanding of the nature of the error associated with a measurement system is necessary to fully evaluate the appropriateness of such simplifying assumptions, and where necessary, make accommodations and adjustments in the development of a robust model. At the outset of a chapter focused entirely on metric validation and characterization of error, it therefore seems prudent to briefly reflect more abstractly on the process of measurement of complex features.

When approaching the task of extracting anatomical information from a digital image, it is essential that one not conceptualize a cow simply as a solid object of finite dimensions existing in three-dimensional space. Instead, conceive of an image of a cow as existing in a high dimensional space, sometimes colloquially referred to in the natural sciences as a hyperspace, built from a composite of information of numerous types (Hurlbert 1981; Ojiem et al 2006; Van Heel 1984). Part of the high dimensional space in which such an image lives will capture information about the physical attributes of an animal, but many other dimensions will capture extraneous information classified in this application as noise - age, coat length, coat color, cleanliness, emotional state, position, light exposure, shadow exposure, background, *etc.* In developing a novel measurement system, the goal is to extract the maximum amount of information of a desired type,

the signal, from information deemed extraneous, the noise (Measurement Systems Analysis Work Group 2010). This is effectively done through a series of data compressions steps. Careful consideration must be given to the assumptions made at each stage of dimensionality reduction to account for potential sources of error introduced by the compression technique selected (Kirby 2001). Finally, full characterization of a measurement system requires a thorough characterization of the resulting metrics to ensure they demonstrate traits amenable to standard methods of statistical inference and predictive modeling (Kutner et al 2005).

The first major compression of information utilized by this measurement system occurred in reducing the temporal dimension down to a single time point. It was assumed that boney structural features of a mature cow's head would not change significantly over a time window spanning only two weeks, excluding any obvious physical injuries. It should be stated that this is fully an assumption, as it does not appear that this issue has been addressed in the existing body of published research, and validation through an extended longitudinal study of the boney structures of the bovine face has been left to future research. In the process of obtaining facial photos, it was observed that this assumption may have been violated for select boney and cartilaginous structures obscured by a significant amount of soft tissue, as variations in facial expressions might obscure measurement of facial structure on a much finer time scale. This observation was explored indirectly as part of the larger metric validation procedure.

The second and perhaps most significant compression of information utilized by this measurement system was exclusion of pixel exposure information so that facial structures were represented only by the distances between landmark structural points. An image is typically represented by an $m \times n \times 3$ matrix, where each pixel index contains, depending on the format, a real number value reflecting the exposure level or hue intensity at that position in the captured

scene. Even for a camera with moderate pixel resolution, this represents a massive amount of information contained in many thousand pixel values. Photos are often compressed to a grayscale $m \times n \times 3$ matrix for image analysis purposes, but even for modestly large image databases, standard pixel-based analysis techniques, like Eigenface analysis, can quickly become computationally intractable (Kirby 2010). This would impose computational constraints that could limit the applicability of such a system, particularly if it were to be implemented at a breed-level.

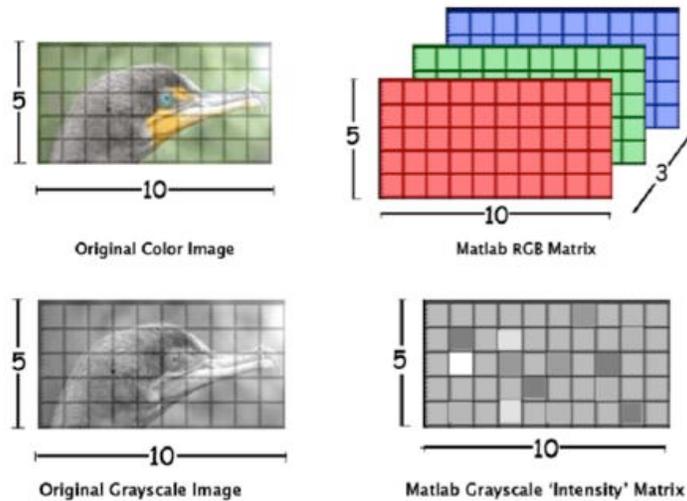


Figure 1: Basic Anatomy of a Digital Image (Singh 2012)

Perhaps of greater concern, however, is the susceptibility of pixel-based image analysis techniques to extraneous noise. At its core, eigen face analysis is just a form of principal components analysis, where in this case each pixel index represents a variable. Given a large number of images captured under very controlled conditions, eigen face analysis can be an effective means of dimension reduction, eliminating noise and redundancy in a set of training images to yield a much smaller set of basis images that concentrate the information of the signal. Like standard principal components analysis, however, eigen face algorithms are greedy

algorithms, creating at each step a basis image which captures the maximum amount of variation found in the original training set possible. Subsequently, eigen face algorithms offer very little control with regards to how information is partitioned within basis images (Kirby 2001). This makes them a poor choice for isolating a specific class of features, such as facial anatomy, when conditions do not allow for tight control of other factors influencing image quality. Fluctuations in image exposure alone, with all other factors held constant, typically account for at least 10 dimensions in the resulting image space (Beveridge et al 2009). They also cannot discern between changes in the intended subject of the image (foreground) and random artifacts in the background, which can be difficult to control in a field setting. Additionally, in animals, eigen face algorithms seem more influenced by changes in coat pattern than overall facial morphology, a major concern for application to Holstein populations (Caiafa et al 2005). Neural network-based image analysis techniques, which can be thought of as a non-linear extension of principal components based eigen-techniques, are a newer approach, and subsequently not as well defined (Kirby 2001). However, deep learning algorithms, likely by virtue of their multiple differentiable layers, seem more adept at paring complex components of images down into their simpler components, making them perhaps a better algorithmic candidate for extraction of facial phenotypes from farm quality images (RSPI Vision 2017). Unfortunately, robust networks require large and diverse image data sets to train, making them difficult to implement for applications without existing databases.

Given these constraints, it was determined that a face mesh approach was more appropriate for this domain of application. With this approach, key anatomical landmarks of the face are determined *a priori*. All images in a database are subsequently annotated with these landmark points, and their coordinate location within the pixel matrix of each image recorded and used in subsequent analyses. Large image databases have been used to train fully automated algorithms

for landmark point extraction for applications in humans, but such work has not been pursued for animal populations, which again imposes practical constraints on the scale at which this technique can be imposed. Here again, however, deep learning algorithms have shown promise in this area, and new research indicates that learning algorithms for landmark point extraction trained on larger human databases may be effectively adapted to livestock features with much smaller reference data bases when strategic constraints are applied (Rashid 2017). For the purposes of this largely exploratory study, it was deemed sufficient to simply extract landmark points manually. By using MatLab's GINPUT tool to interactively select a predetermined series of key anatomical reference points on the face, and storing their coordinate locations within the pixel matrix, extraneous information related coat pattern and features of the farm environment, like variable lighting exposure and changing background content, were effectively excluded. It should be noted, however, that in applying this compression, a significant amount of structural information was inevitably lost as well, with only structural points that had been identified as descriptive *a priori* being retained for further analysis. This could serve as a source of bias, if certain regions of the face or types of structural variations were not adequately described by the landmark points defined. Additionally, physical selection of these anatomical points within the image was not without error, requiring targeted analysis to determine the magnitude and systematic nature of this source of measurement error.

The third major compression of information came from reducing facial structures from 3 to 2 physical dimensions. It would be possible to represent structural features of the face via a 3-dimensional image, and thereby capture all dimensions of facial shape, up to the resolution of the camera and accuracy of the stitching algorithm (Aldridge et al 2011; Obafemi-Ajayi et al 2014). To do so, however, would be an expensive and time-consuming endeavor, requiring specialized

equipment and greater restraint of the animal. Thus, this compression decision was driven predominantly by practical concerns, as it was deemed that loss of information was outweighed by gains in accessibility realized by developing this measurement system around the specifications of any standard quality digital camera. In projecting a 3D object onto a 2D plane, however, several sources of error are introduced. The first and most important is angle of the object relative the plane of the camera. Significant variations in angle related to depth can effectively distort the resulting image as it is projected onto the plane, effectively warping the relative distances between facial structures. This is a major concern, as it not only distorts the perceptions of facial shape, but because of the underlying geometry, tends to do so in a systematic way. In other words, errors from this source are not necessarily random, and tend to be correlated, which breaks the assumptions of many statistical models (Kutner et al 2005). This source of error was addressed procedurally by attempting to reduce variation in camera angle as much as is possible on a farm working with large and at times disagreeable animals. Side profile images were obtained parallel to the surface of the cheek. This was partially achieved by attempting to center the image on the eye, and then aligning as closely as possible the ridges of the eye orbitals on either side of the forehead. Front profile images were obtained parallel to plane of the forehead. This was achieved by attempting to equalize the distance between either eye and the center of the forehead on either side of the face, and then seeking to obtain an image where the nose appeared as long as possible.

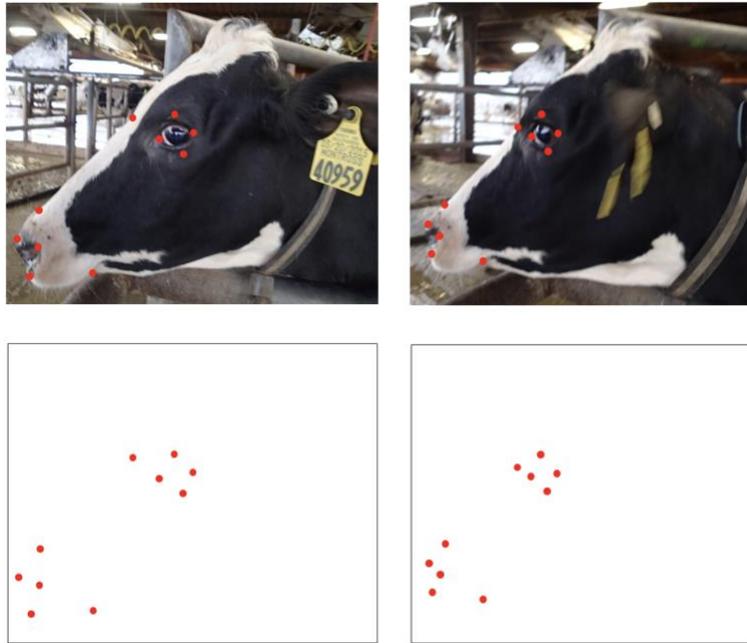
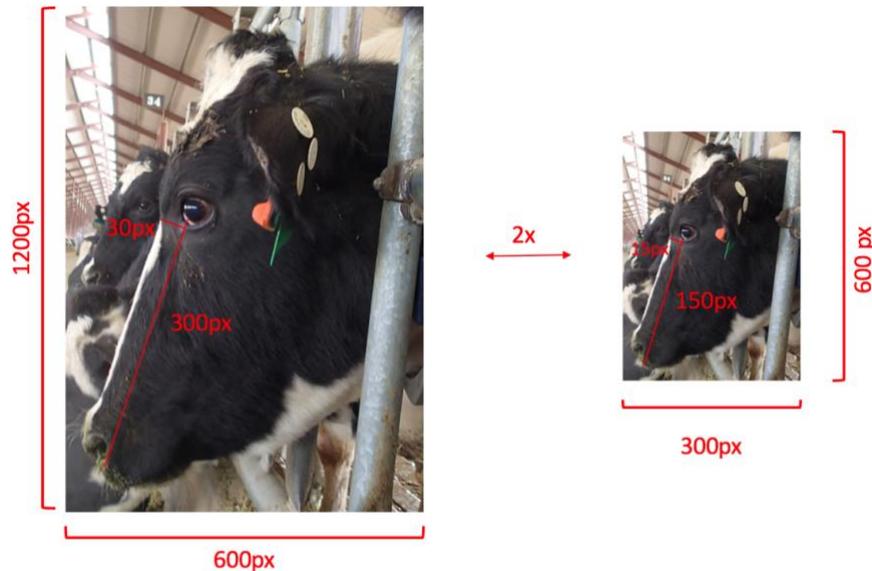


Figure 2: Impact of Out-of-Plane Variations in Face Angle on Coordinate Locations

The second major source of error introduced by this compression came from variations in relative position of the camera to the cow. Varying distances between the camera and object changes the proportion of the frame dedicated to the cow's face. This in turn changes the number of pixels dedicated to capturing structural features of the cow's face (*i.e.* image resolution). If raw pixel distances between anatomical points were used, changes in image resolution would become a major source of error due to differences in scaling. This issue is often addressed by scaling the image to a known reference length of an object with the frame of the image, but this solution was deemed impractical on a working farm environment. Attempting to place a reference object in the frame near the cow so that it would be in a plane equidistant to the camera with the cow would not only significantly increase the amount of time required to obtain an image, but also increase stress experienced by the animal and put the handlers in a more exposed position. Instead this issue was addressed by developing biometrics that either reported angles or distances as proportions.

Computation of angles between traits are of course geometrically dependent only on their relative, not absolute, distances. Similarly, by using proportions to report relative distance measures, the scaling factor of the image was effectively "divided out". Thus, this measurement system should be inherently robust to changes in image resolution that result from variable distances between camera and cow, as well as any variations in specs of the camera used or degree of zoom applied. Practically speaking, this greatly simplified the process of acquiring images of the cows, and allowed greater focused to be placed on reducing variations in image angle.



$$\frac{300}{30} = \frac{2(150)}{2(15)} = \frac{150}{15} = 10$$

Figure 3: Illustration of Elimination of Scale Effect by Division Operator

One exception to this assumption of distance invariance was that, when the photograph was taken extremely close to the cow's face, as frequently happened when photos were acquired in the feed bunk, there seemed to be a significant interaction between position and angle. Put simply, when quite close to the cow, aligning the camera using the eye structural reference point

created the correct 90-degree angle for the central part of the face, but still left a significant angle between the camera and distal parts of the face, namely the nose. Auxiliary image measurements were used in an attempt to correct for this potential source of error in such images.

The fourth and final source of compression comes from converting the 2D coordinate vectors representing the locations of key structural points into 1D descriptive measures that could be used as covariates in predictive models. Previous studies have frequently accomplished this by simply taking the Euclidean distance between all pairwise combinations of anatomical points, globally normalizing by the sum of all lengths to correct for differences in image resolution, and then reducing the number of candidate variables by using a multivariate compression technique such as principal components (Cole et al 2016; Aldrige et al 2001), or else using data clustering techniques designed for high dimensional input (Obafemi-Ajayi et al 2014). While this procedure is quite simple to apply, it has two key drawbacks. The first and most significant is that the resulting distance measures are directly geometrically related, resulting in complex correlations structures. A slight change in the relative position of one anatomical point would be reflected in slight changes in all pairwise distances of which that point is a member. When points change their relative positions due to underlying face shape, associated Euclidean distance terms will change as well, but so many of these points would change simultaneously that it becomes difficult, if not impossible, to discern the nature of this geometric shift just from direct appraisal of the data.

Principal component analysis is a means to concentrate this redundant information, but in doing so assumptions of linearity are necessary. When a number of facial features shift simultaneously, their cumulative effects on individual pairwise distances may not necessarily be additive, which could potentially lead to inflation of the parameter space or misleading reparameterizations (Kirby 2001; Johnson & Richard 2007). When the relative position of points

change due to error in point selection, as opposed actual changes in facial shape, this error is also subject to geometric constraints between pairwise combinations of points, potentially leading to correlation in the error structures. Most correlation-based multivariate techniques, including principal component analysis, require the assumption that error terms are uncorrelated. When this is not in fact true, correlation in error is mathematically interpreted as correlation in the signal. As a result, application of these dimension reduction techniques lead to concentration of both signal and error simultaneously (Johnson & Richard 2007), which is at best inefficient but also a potential source of bias in downstream analysis.

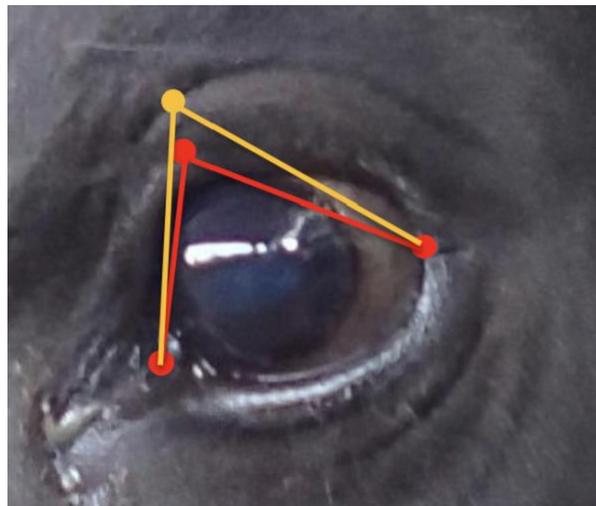


Figure 4: When horizontal eye landmarks are constant, error in point selection leads to triangular relationship between edges

The second drawback of this technique is that principal component analysis, while an effective means of dimension reduction, is limited in terms of its descriptive ability. For exceedingly high dimensional input, it is difficult to determine from the orthogonal bases vectors what information is captured in each new transformed variable. In other words, it might be possible to determine from the relative scale of orthogonal basis values that a given dimension is dedicated

largely to describing variations in eye structure, but it would be difficult if not impossible to determine what this would relate to in terms of the underlying structural variability without a means of effectively regenerating the face (Nielsen et al 2011). This makes any subsequent models built using reparameterized variables difficult to interpret. While this is perhaps sufficient for purely predictive models, it makes it difficult if not impossible to assess the biological appropriateness of such results. Further, as principal component analysis is not a model-based technique, it is not generally considered readily extrapolatable to novel data sets, which makes it more appropriate for descriptive studies as opposed to predictive modeling (Johnson & Richard 2007).

In an effort to overcome these drawbacks, a geometric approach to biometric extraction was developed. This approach had two key goals. The first was to minimize correlation between resulting biometrics, attempting to isolate specific changes in shape using targeted geometric relationships on the front end of the algorithm to create independence between measurements, as opposed to applying an indiscriminate orthogonalization technique like PCA on the back-end. This was done in two ways. The first was that, as opposed to normalizing pixel distances between points using the sum of all pairwise distances, it divided by distances between nearby points that were selected to produce more intuitive interpretations of shape. For example, instead of describing the height of the eye as a proportion of overall face size using the sum of distances, which would in turn be influenced by many other unrelated anatomical factors like jowl depth or nose length, it was compared directly to the length of the eye, or to the depth. The second means of achieving this goal was to make use of projection lengths over simple Euclidean distances. The coordinate locations of many key anatomical points were frequently observed to be influenced by multiple independent variations in facial shape. By projecting such a point onto a number of carefully

selected reference slopes from nearby facial features, the effects of these independent shape variants could be more effectively broken up into distinct distance measures to isolate their independent effects. For example, the location of the highest point of the eye is influenced by two variants in eye shape: how tall the eye is, and how angular the top of the eye is (*i.e.* how far forward is the highest point). Simple Euclidean distances would capture both effects at the same time. By instead relying on projections, the angularity of the eye is captured by projecting the highest point of the eye onto the horizontal plane of the eye (Eye Height Point Proportion - EHPP), and the height of the eye is captured by projecting the highest point of the eye onto the plane perpendicular to the horizontal reference plane of the eye (Eye Height Proportion - EHP).

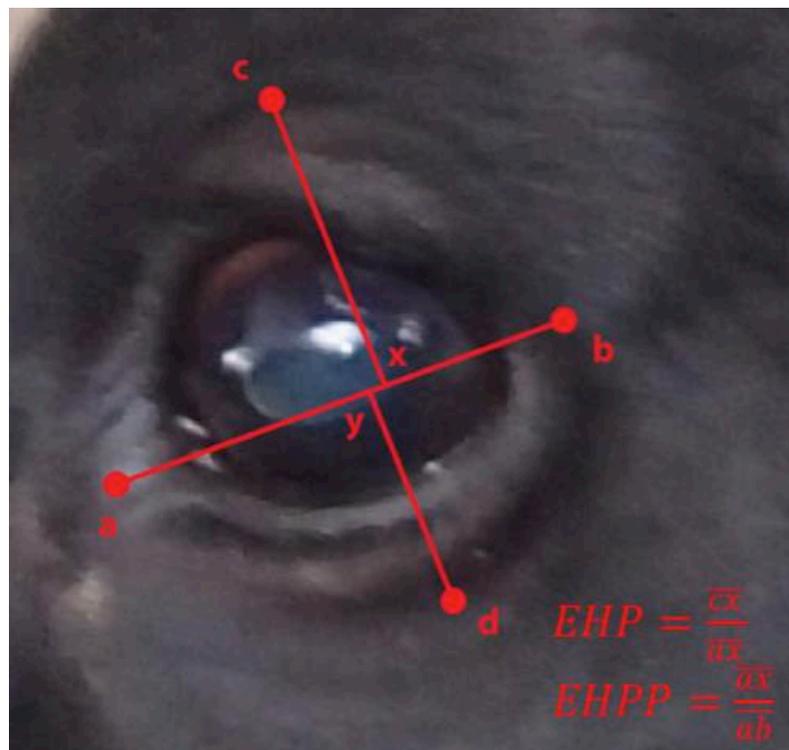


Figure 5: Example of Geometric Biometric using Orthogonal Projects

The second goal pursued with a geometric approach to biometric extractions was to reduce the impact of errors in point selection. This was done in several ways. First, by attempting to define facial shapes using targeted comparisons of length measures, as opposed to Euclidean distances between all pairwise combinations of anatomical reference points, and by relying on a local as opposed to global normalization scheme, any error incurred in the selection of the coordinate location of a given anatomical reference point was effectively isolated to only a targeted handful of metrics, and not amplified across the broader set of metrics. Put more simply, if the coordinate location of the highest point of the eye was selected poorly in a given picture, that error was only seen in a subset of the eye biometrics, and had no impact on biometrics extracted from the nose, topline, or forehead. This characteristic was enhanced by relying predominantly on projection lengths. Just as the projections were used to break down distances into the distinct influences of shape, they also effectively orthogonalized the components of error. This was particularly helpful for traits where coordinate selection was perhaps clear in one direction but less easy to distinguish for another. For example, take highest point of the eye. For very rounded eyes, multiple coordinate selections might return points with very similar vertical distances horizontal plane of the eye, but a great deal of variability in the horizontal distance. This error in point selection would in turn be isolated only to metrics that relied on the horizontal component of this point location, and have virtually no influence on metrics that rely only on the vertical distance, whereas for simple Euclidean distance this error would influence any pairwise combination that involved this point.



Figure 7: Example of an Interpolated Landmark Point (red)

Full derivations of the 104 candidate geometric biometrics developed to fully describe the shape of the bovine face can be found in Appendix A. To assess the efficacy of this novel approach to biometric extraction from digital images, geometric biometrics will be compared to standard normalized length measures within each region of the face to determine which strategy demonstrates more robustness to measurement error while minimizing correlations between metrics without use of dimension reduction techniques. Final estimates of repeatability will then be used to select which candidate biometrics demonstrate sufficient robustness to warrant farther study in predictive models of dairy productivity and longevity.

Materials and Methods

To assess the relative performance of geometric and normalized length biometrics in the quantification of bovine facial structures, an image database was compiled consisting of 108 mature Holstein dairy cows over a two week period. Cows were photographed while locked in the feed bunk of their home pen while standard herd checks were being performed by farm staff after the first morning milking. Cows were never required to stand locked for longer than 1.5 hours for the purposes of image acquisition, per IACUC approved protocol. The procedural target was to photograph each cow on three separate days, acquiring each day side profile images for both the left and right side of the face. Of the 108 cows recorded, 74 demonstrated complete image profiles across three separate days. The remaining cows represent either a failure to follow up on the third day, or more commonly, an individual image was discarded farther down the analysis pipeline due to failure to meet quality control standards. In total, 551 images were deemed suitable for analysis. Images were analyzed using the image analysis toolbox in the MatLab programming environment. Images were read in as true color pixel matrices at a consistent screen ratio. Coordinate locations of key anatomical points were extracted using the GInput tool in two separate replicates of the point extraction process in order to distinguish between measurement errors stemming from point selection and measurement error related to variability between photos. Custom MatLab scripts were then used to extract from this coordinate information geometric biometrics computed within four separate facial sub regions: eye, muzzle, topline, forehead/jowl/overall face shape. Additionally, to allow for more direct comparisons, normalized length measures were computed and normalized using subset of anatomical landmarks points within each distinct facial region.

Once metric data was extracted from raw images, subsequent analysis of measurement system performance was completed in the R programming environment (R Core Team 2016). To allow for a more thorough examination of biometrics performance, metric validation procedures were applied independently in each of the facial subregions, and will subsequently be summarized by anatomical region in this report. First, the normality of geometric and normalized length biometrics was assessed to determine if a method showed an advantage in terms of production of extreme values. This was done by computing standard skew and kurtosis (tail thickness) measures for each metric using the moments package (Lukasz *et al* 2015), and visually comparing the overall performance of each measurement system by aggregating the results of individual metrics using a simple histogram.

Next the error structure and repeatability of each geometric biometric was assessed via a nested mixed model approach using the lme4 package (Bates et al 2015). Side of the face was nested within cow, to account for any potential variability due to facial asymmetry. The three separate day observations were subsequently nested within side of the face to quantify error due to variations between photos. Finally, the 1102 individual coordinate replicates were nested within each cows-by-side-by-day cross to estimate the error due to uncertainty in point selection. Variance components extracted from the mixed model for each individual biometric were used to estimate three metrics: proportion of variability attributed to between-image error (*i.e.* error due to image acquisition), proportion of variability attributed to within-image error (*i.e.* error in image annotation), and proportion of variance attributed to the cow by side interaction (*i.e.* repeatability of the biometric). Additionally, the bootMer functionality within the lme4 package was used to estimate a 95% bootstrap confidence interval for each of these descriptive statistics. This methodology was also replicated with the response being the average of each metric over the two

coordinate selection replicates in order to estimate the measurement repeatability achieved by concentrating signal using a simple averaging technique. Results for the individual geometric biometrics are fully reported. To compare the robustness of geometric biometrics to that of normalized length measures, point estimates for between-day and within-day measurement repeatability were visually compared using a histogram.

Next, with the distribution and robustness of individual metrics having been appraised, between-metric characteristics within a given metric system were assessed. First, in order to eliminate artificial redundancy in both measurement systems due to the presence of alternative versions of some anatomical landmarks, results of the repeatability analysis were used to select the optimal landmark set within each coordinate system. Normalized distances calculated from dropped landmark points were dropped from the set of candidate biometrics, and only the best performing version of a given metric derivation was retained within the geometric biometric candidate set. After the data set had been culled, correlations between individual biometrics within each measurement system were calculated. To account for the hierarchical structure of the data and subsequently its variance structures, the `statsBy` function within the `psych` package was used to calculate correlations at the cow by side interaction level (Revelle 2017). Correlations matrices for the geometric biometrics are reported. To compare the overall level of redundancy between measurement systems, pairwise correlation values were aggregated and visually compared using histograms. Next, the same methodology was used to compare the levels of correlation in the error structures of each measurement system, using the error values extracted from the nested mixed model at the cow by side interaction level.

Finally, by leveraging some assumptions about the structure of error within this data set, it is possible to directly explore correlations between error structure and several dimensions of image quality. By assuming that a boney and cartilaginous facial structures should remain constant for a given cow on a specific side of her face, then any variations observed in a measurement between days must be measurement error. Thus, by finding the difference between all pairwise combinations of metrics within a cow by side cross, it is possible to orthogonalize the data to isolate measurement error. If changes in a metric between any two pairs of images occur systematically with changes in any auxiliary measurements not related to facial structure, this suggests algorithmic techniques underlying a given metric system may not be robust to corresponding changes in image quality. These measures of image attributes included:

1) Frame-to-Face Ratio = Pixel area occupied by a cow's head relative to overall size of the frame. Pixel area attributed to the cow's head is approximated by the pixel area occupied by the polygon fitted using the polyarea function in MatLab's image processing toolbox. Frame size is simply frame length x frame height. This metrics is meant to capture measurement error attributed to inadequacies in a given biometric's scale invariance



$$FFR = \frac{A(L, E_{lower}, T_{poll}, T_{int})}{FrameHeight \times FrameWidth}$$

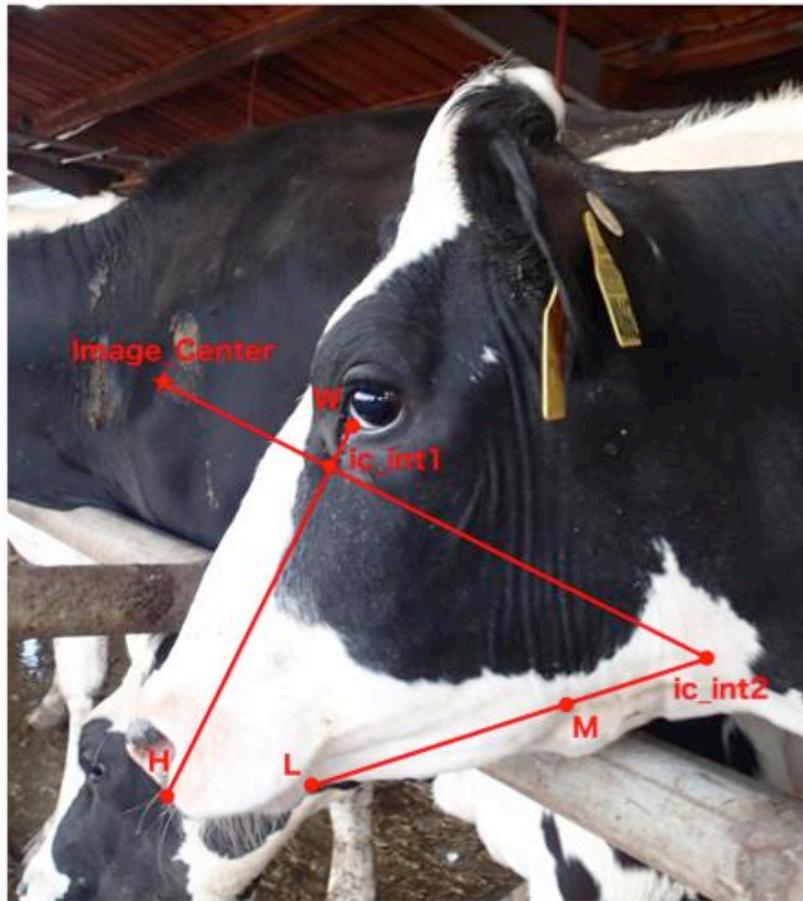
Figure 8: Frame-to-Face Ratio

2) Overall Face Angle = Angle of the overall plane of the face \overline{HW} relative to the horizontal plane of the image frame. This metric is meant to capture measurement error attributed to inadequacies in a given biometric's resistance to in-plane variations in face angle.



Figure 9: Overall Face Angle

3) Eye-Center Displacement Proportion - Image Center is defined as the coordinate location of the centroid of the original uncropped image. Eye-Center Displacement Proportions capture the vertical and horizontal displacement of the center of the image from the rostral-most point of the eye relative to the overall plane of the face. These metrics are meant to capture measurement error attributed to inadequacies in a given biometric's resistance to variations in angle position relative to the facial trait being measured.



$$ECO_{vert} = CF * \frac{\|ic_{cent}, ic_{int1}\|}{\|HW\|}$$

$$ECO_{hort} = CF * \frac{\|W, ic_{int1}\|}{\|HW\|}$$

$$CF = \frac{\|ic_{cent}, ic_{int2}\| - \|ic_{int1}, ic_{int2}\|}{\| \|ic_{cent}, ic_{int2}\| - \|ic_{int1}, ic_{int2}\| \|}$$

$$CF = \frac{\|ic_{int1}, H\| - \|H, W\|}{\| \|ic_{int1}, H\| - \|H, W\| \|}$$

Figure 10: Overall Face Angle

Changes in each observed biometric for each pair-wise combination of days (Δ biometric) were then regressed against changes in measures of image quality across days (Δ attribute) using the stepAIC function, an automated backwards selection procedure using AIC as model adequacy criterion found in the MASS package in R (Venables & Ripley 2002). Two separate models were fit to explore the resistance of geometric biometrics and normalized lengths to changes in image

quality. The first model explored resistance to variations in image quality that were deemed uncontrollable in a farm environment: image scaling (Frame-to-Face Ratio), and in-plane variations in face angle (Overall Face Angle). The full model used to initialize the backward selection procedure was:

$$\begin{aligned} \Delta\text{biometric} = & \Delta\text{Frame-to-Face Ratio} + (\Delta\text{Frame-to-Face Ratio})^2 + \Delta\text{Overall Face Angle} \\ & + \Delta\text{Frame-to-Face Ratio} \times \Delta\text{Overall Face Angle} \\ & + (\Delta\text{Frame-to-Face Ratio})^2 \times \Delta\text{Overall Face Angle} \end{aligned}$$

The second model explored resistance to variations in image quality related to camera position relative to the traits being measured. The image acquisition protocol sought to control these factors by aligning the center to the camera to the eye so that the eye orbitals were as closely aligned as possible - an attempt to control for out-of-plane variations in face angle. This model sought to validate the effectiveness of this protocol, and more specifically, to explore any inadequacies due to interactions between camera position and proximity of the camera to the face due to potential fish-eye effects.

$$\begin{aligned} \Delta\text{biometric} = & \Delta\text{Eye-Center Displacement}_{\text{Vertical}} + \Delta\text{Eye-Center Displacement}_{\text{Horizontal}} \\ & + (\Delta\text{Eye-Center Displacement}_{\text{Vertical}})^2 + (\Delta\text{Eye-Center Displacement}_{\text{Horizontal}})^2 \\ & + \Delta\text{Frame-to-Face Ratio} \times \Delta\text{Eye-Center Displacement}_{\text{Vertical}} \\ & + (\Delta\text{Frame-to-Face Ratio})^2 \times \Delta\text{Eye - Center Displacement}_{\text{Vertical}} \\ & + \Delta\text{Frame-to-Face Ratio} \times \Delta\text{Eye-Center Displacement}_{\text{Horizontal}} \\ & + (\Delta\text{Frame-to-Face Ratio})^2 \times \Delta\text{Eye - Center Displacement}_{\text{Horizontal}} \end{aligned}$$

Reduced models were optimized to each candidate biometrics from the full models defined above, and the resulting R^2 values, which can be roughly interpreted as the proportion of between-day measurement error attributed to changes in image attributes, were then compared across the two measurement systems to assess advantages in robustness to image quality.

Results - Eye Biometrics

Eye biometrics are formulated uniquely amongst the subregions of the face. The shape of the eye is defined by a total of eight landmark points – a, b, c, d, e, f, g, h . For each landmark point, there are three alternative definitions – along the line of the inner eye lid, along the line of the outer eyelid, and along the ridge of the eye orbital. From these three alternative landmark point definitions, a total of six combinations of landmark points were developed. A total of 21 candidate geometric biometrics were derived to describe the shape of the eye, with each of these metrics being calculated using all six coordinate systems (see Figure 11).

Eye Depth Proportion – Full Length
 Eye Depth Proportion – Front Eye Length
 Eye Depth Point Proportion
 Eye Height Proportion – Full Length
 Eye Height Proportion – Front Eye Length
 Eye Height Point Proportion
 Eye Width-to-Height Ratio
 Eye Displacement Proportion
 Eye Length Proportion – Combined Height
 Eye Length Proportion – Length
 Eye Roundness Proportion – Upper Front – Poly
 Eye Roundness Proportion – Upper Back – Poly
 Eye Roundness Proportion – Lower Back – Poly
 Eye Roundness Proportion – Lower Front – Poly
 Eye Roundness Proportion – Upper Front – Linear
 Eye Roundness Proportion – Upper Back – Linear
 Eye Roundness Proportion – Lower Back – Linear
 Eye Roundness Proportion – Lower Front – Linear
 Eye Roundness Proportion – Total – Poly
 Eye Roundness Ratio – Top-to-Bottom – Poly
 Eye Roundness Ratio – Front-to-Back – Poly

$V1 = \{aa, ba, ca, da, ea, fa, ga, ha\} \rightarrow \{a, b, c, d, e, f, g, h\}$
 $V2 = \{ab, bb, cb, db, eb, fb, gb, hb\} \rightarrow \{a, b, c, d, e, f, g, h\}$
 $V3 = \{ac, bc, cc, dc, ec, fc, gc, hc\} \rightarrow \{a, b, c, d, e, f, g, h\}$
 $V4 = \{aa, ba, cb, db, eb, fb, gb, hb\} \rightarrow \{a, b, c, d, e, f, g, h\}$
 $V5 = \{ab, bb, cc, dc, ec, fc, gc, hc\} \rightarrow \{a, b, c, d, e, f, g, h\}$

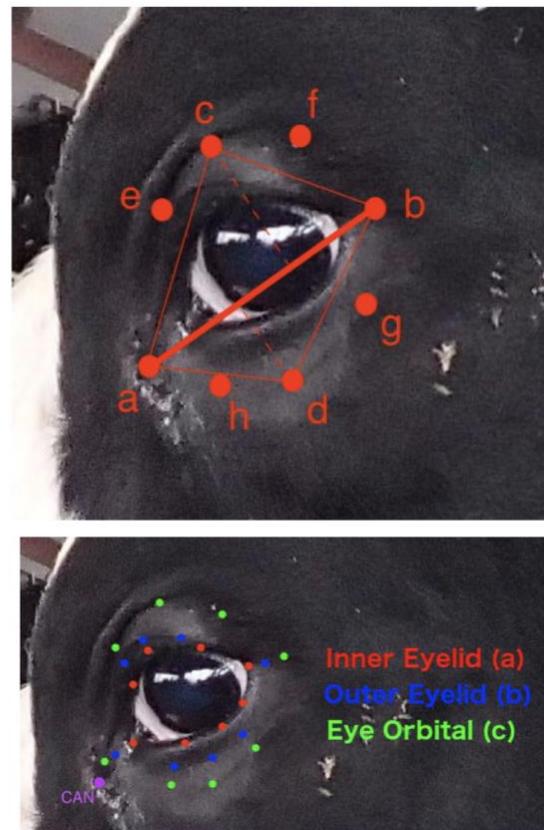


Figure 11: Above – landmark points of the eye; below – anatomical lines of the eye (further Appendix A)

First, an appraisal of the global behavior of the higher moments of biometrics in each measurement system (see Figure 12). Normalized length measures were roughly as likely to be positively as negatively skewed, being roughly centered around zero, with some negative outliers. Geometric biometrics had a consistent tendency to be positively skewed, but with perhaps slightly fewer measures demonstrating extreme skew values. In comparing kurtosis measures, geometric biometrics showed perhaps marginally less tendency for thick tails, with slightly more density near zero than for the normalized length measures, and also fewer extreme values.

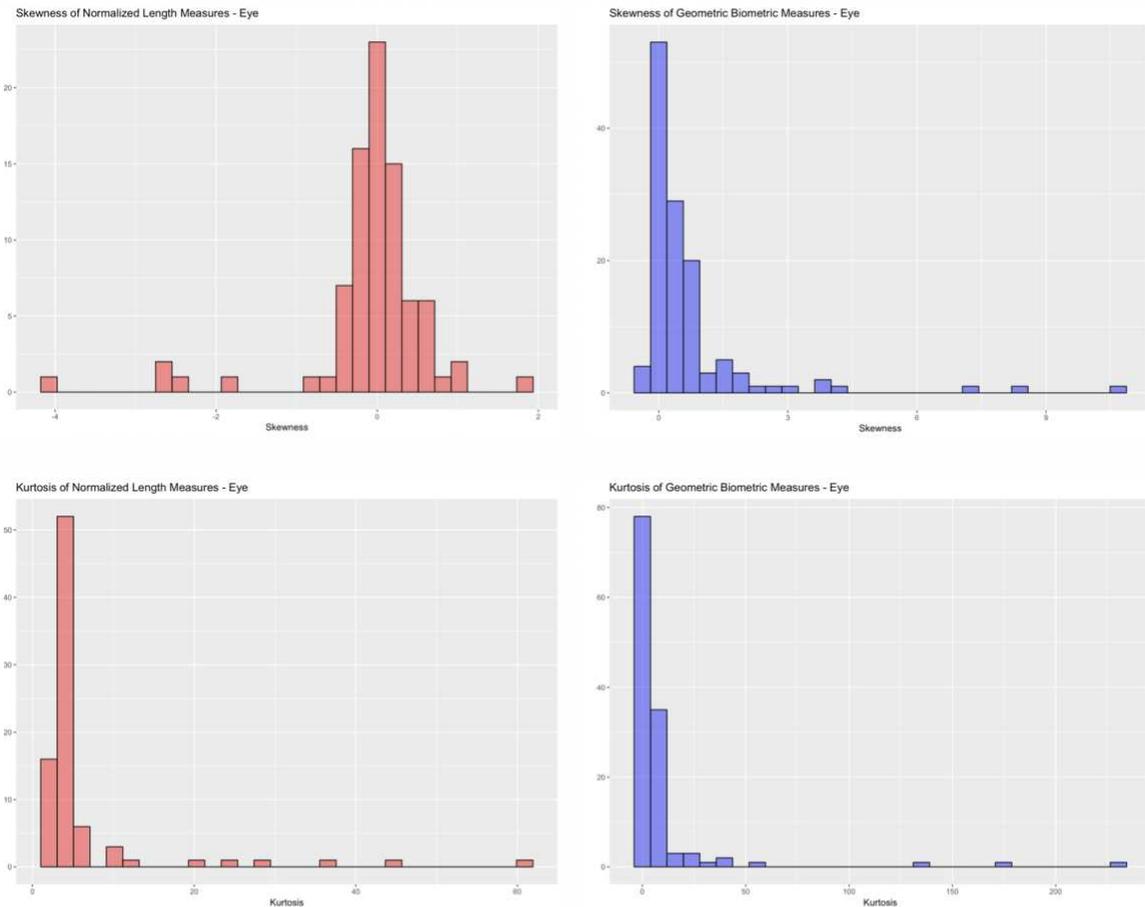


Figure 12: Distribution of 3rd & 4th moments for normalized length & geometric eye biometrics

Second, a comparison of the performance of normalized and geometric biometrics in terms of repeatability (see Figure 13). One comparison is the repeatability of these two measurement systems within-photo, which corresponds to errors in points selection, which here are quite similar. The overall average within-photo repeatability of normalized length measures was 0.32, with landmark points from the outer eyelid showing the best performance with a mean repeatability of 0.39 and landmark points from the ridge of the eye orbital demonstrating the worst performance with a mean repeatability of 0.27. The overall average within-photo repeatability of geometric biometrics was marginally better at 0.36, with the worst performance coming from coordinate version 2 with a mean repeatability of 0.27, and the best performance coming from coordinate version 6 with a mean repeatability of 0.42. Thus, for eye biometrics, geometric biometrics demonstrate a slight advantage in terms of resistance to error in landmark point selection.

Of perhaps greater practical importance is the comparative performance of these measurements systems in terms of repeatability between-photos, which can be thought of as their overall measurement repeatability (see Figure 13). Again, the two distributions appear quite similar, but the geometric distribution shows a slightly more desirable range, with considerably more metrics displaying between-day repeatability's above 0.5. The overall average repeatability for geometric biometrics is 0.45, with coordinate version 2 again showing the worst performance with an average repeatability of 0.36, and coordinate version 2 showing the best performance with mean repeatability of 0.54. In contrast, the overall average repeatability of normalized length biometrics 0.43, with the coordinate version 3 demonstrating the worst performance with an average repeatability 0.38, but the optimal average repeatability from coordinate version 3 being on 0.48. Thus, geometric biometrics again demonstrate a slight advantage over normalized length metrics in terms of overall metric repeatability

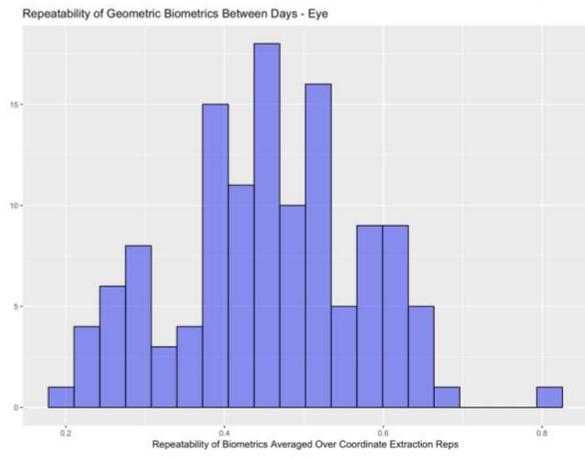
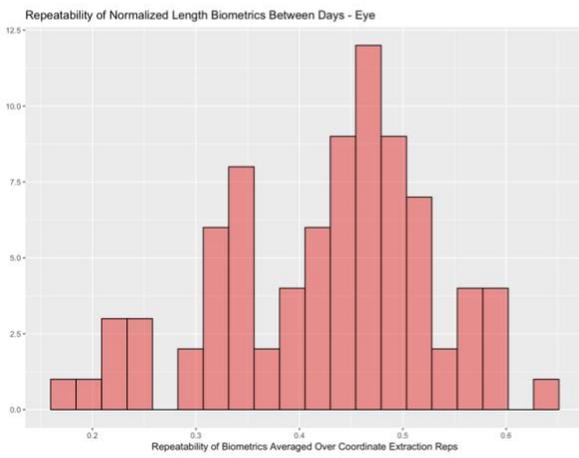
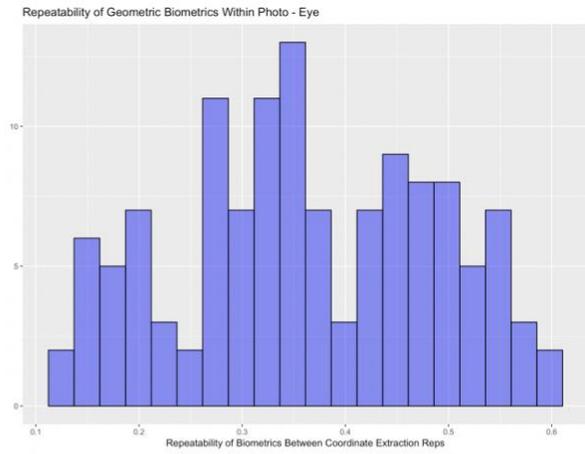
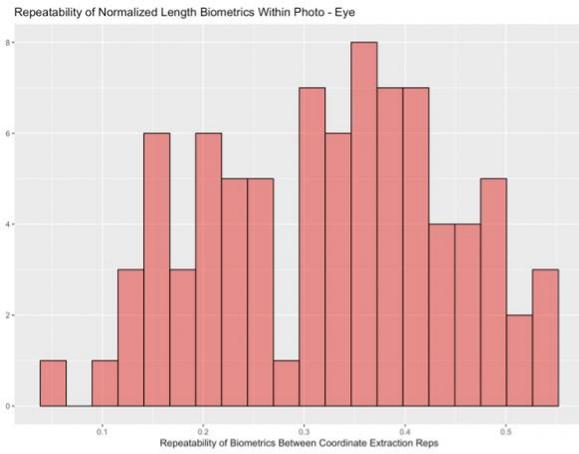


Figure 13: Comparison of Between and Within-Day Repeatability of Normalized and Geometric Eye Biometric Measures

Table 1: Proportion of total variability in eye biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps)

	Coord 1	Coord 2	Coord 3	Coord 4	Coord 5	Coord 6
Eye Depth Proportion – Full Length	0.300 [0.24, 0.35]	0.352 [0.29, 0.41]	0.426 [0.36, 0.49]	0.289 [0.23, 0.34]	0.340 [0.28, 0.39]	0.279 [0.23, 0.32]
Eye Depth Proportion – Front Eye Length	0.228 [0.18, 0.27]	0.451 [0.38, 0.52]	0.426 [0.36, 0.49]	0.479 [0.41, 0.55]	0.355 [0.30, 0.41]	0.344 [0.29, 0.39]
Eye Depth Point Proportion	0.360 [0.29, 0.42]	0.579 [0.50, 0.65]	0.273 [0.22, 0.32]	0.566 [0.49, 0.64]	0.241 [0.19, 0.28]	0.230 [0.19, 0.27]
Eye Height Proportion – Full Length	0.209 [0.17, 0.25]	0.265 [0.35, 0.52]	0.371 [0.15, 0.28]	0.203 [0.16, 0.24]	0.285 [0.23, 0.33]	0.196 [0.16, 23]
Eye Height Proportion – Front Eye Length	0.268 [0.21, 0.31]	0.338 [0.28, 0.39]	0.396 [0.33, 0.46]	0.242 [0.19, 0.28]	0.285 [0.23, 0.33]	0.248 [0.20, 0.29]
Eye Height Point Proportion	0.163 [0.13, 0.19]	0.295 [0.24, 0.35]	0.297 [0.24, 0.35]	0.217 [0.17, 0.16]	0.230 [0.18, 0.27]	0.238 [0.19, 0.28]
Eye Width-to-Height Ratio	0.144 [0.11, 0.17]	0.229 [0.18, 0.27]	0.382 [0.32, 0.44]	0.156 [0.12, 0.18]	0.254 [0.21, 0.27]	0.164 [0.13, 0.19]
Eye Displacement Proportion	0.250 [0.20, 0.29]	0.399 [0.33, 0.45]	0.214 [0.17, 0.25]	0.406 [0.34, 0.46]	0.229 [0.18, 0.26]	0.262 [0.21, 0.30]
Eye Length Proportion – Combined Height	0.394 [0.33, 0.46]	0.563 [0.48, 0.64]	0.310 [0.25, 0.36]	0.564 [0.48, 0.64]	0.235 [0.19, 0.28]	0.229 [0.18, 0.27]
Eye Length Proportion - Length	0.351 [0.28, 0.41]	0.583 [0.50, 0.66]	0.334 [0.27, 0.39]	0.573 [0.49, 0.65]	0.259 [0.21, 0.31]	0.267 [0.21, 0.31]
Eye Roundness Proportion – Upper Front – Poly	0.756 [0.67, 0.84]	0.563 [0.49, 0.64]	0.597 [0.51, 0.68]	0.396 [0.33, 0.46]	0.534 [0.45, 0.61]	0.380 [0.31, 0.44]
Eye Roundness Proportion – Upper Back – Poly	0.550 [0.47, 0.62]	0.701 [0.62, 0.78]	0.455 [0.38, 0.52]	0.583 [0.51, 0.66]	0.330 [0.27, 0.38]	0.248 [0.20, 0.29]
Eye Roundness Proportion – Lower Back – Poly	0.463 [0.39, 0.53]	0.742 [0.66, 0.82]	0.557 [0.48, 0.63]	0.650 [0.56, 0.73]	0.451 [0.38, 0.51]	0.395 [0.34, 0.46]
Eye Roundness Proportion – Lower Front – Poly	0.615 [0.53, 0.69]	0.791 [0.72, 0.88]	0.553 [0.47, 0.63]	0.715 [0.63, 0.79]	0.563 [0.48, 0.64]	0.468 [0.40, 0.53]
Eye Roundness Proportion – Upper Front – Linear	0.760 [0.68, 0.84]	0.559 [0.49, 0.63]	0.590 [0.51, 0.67]	0.556 [0.48, 0.63]	0.552 [0.47, 0.63]	0.525 [0.45, 0.59]
Eye Roundness Proportion – Upper Back – Linear	0.511 [0.43, 0.58]	0.709 [0.63, 0.79]	0.525 [0.44, 0.60]	0.646 [0.56, 0.72]	0.478 [0.41, 0.55]	0.433 [0.37, 0.49]
Eye Roundness Proportion – Lower Back – Linear	0.515 [0.44, 0.58]	0.756 [0.67, 0.84]	0.562 [0.48, 0.63]	0.707 [0.62, 0.79]	0.508 [0.43, 0.58]	0.502 [0.43, 0.57]
Eye Roundness Proportion – Lower Front – Linear	0.437 [0.37, 0.50]	0.580 [0.50, 0.66]	0.288 [0.23, 0.33]	0.549 [0.47, 0.62]	0.231 [0.18, 0.27]	0.213 [0.17, 0.25]
Eye Roundness Proportion – Total – Poly	0.686 [0.60, 0.77]	0.715 [0.63, 0.80]	0.536 [0.46, 0.61]	0.651 [0.57, 0.73]	0.483 [0.41, 0.55]	0.467 [0.40, 0.53]
Eye Roundness Ratio – Top-to-Bottom – Poly	0.612 [0.54, 0.69]	0.729 [0.64, 0.81]	0.460 [0.39, 0.52]	0.674 [0.59, 0.76]	0.449 [0.38, 0.51]	0.451 [0.38, 0.51]
Eye Roundness Ratio – Front-to-Back – Poly	0.537 [0.46, 0.61]	0.648 [0.57, 0.72]	0.489 [0.41, 0.56]	0.672 [0.59, 0.75]	0.481 [0.41, 0.55]	0.469 [0.40, 0.53]

Table 2: Proportion of total variability in eye biometrics attributed to error in image acquisition (variance between days/photos)

	Coord 1	Coord 2	Coord 3	Coord 4	Coord 5	Coord 6
Eye Depth Proportion – Full Length	0.216 [0.16, 0.27]	0.224 [0.16, 0.28]	0.259 [0.19, 0.33]	0.221 [0.16, 0.28]	0.281 [0.21, 0.35]	0.264 [0.20, 0.32]
Eye Depth Proportion – Front Eye Length	0.182 [0.13, 0.23]	0.211 [0.14, 0.28]	0.233 [0.16, 0.30]	0.202 [0.13, 0.27]	0.263 [0.20, 0.33]	0.216 [0.15, 0.27]
Eye Depth Point Proportion	0.172 [0.12, 0.23]	0.141 [0.20, 0.35]	0.192 [0.14, 0.24]	0.149 [0.08, 0.22]	0.204 [0.15, 0.25]	0.209 [0.15, 0.26]
Eye Height Proportion – Full Length	0.324 [0.27, 0.39]	0.300 [0.35, 0.52]	0.216 [0.15, 0.28]	0.344 [0.28, 0.41]	0.274 [0.21, 0.33]	0.316 [0.25, 0.38]
Eye Height Proportion – Front Eye Length	0.257 [0.19, 0.32]	0.241 [0.18, 0.30]	0.228 [0.16, 0.29]	0.299 [0.23, 0.36]	0.278 [0.21, 0.34]	0.264 [0.20, 0.32]
Eye Height Point Proportion	0.312 [0.24, 0.38]	0.259 [0.19, 0.32]	0.218 [0.16, 0.27]	0.312 [0.24, 0.38]	0.236 [0.17, 0.29]	0.222 [0.16, 0.27]
Eye Width-to-Height Ratio	0.359 [0.28, 0.43]	0.337 [0.27, 0.40]	0.257 [0.28, 0.45]	0.356 [0.28, 0.42]	0.341 [0.27, 0.41]	0.347 [0.27, 0.41]
Eye Displacement Proportion	0.321 [0.25, 0.39]	0.247 [0.18, 0.31]	0.234 [0.18, 0.29]	0.249 [0.18, 0.32]	0.228 [0.17, 0.28]	0.228 [0.16, 0.28]
Eye Length Proportion – Combined Height	0.218 [0.16, 0.28]	0.157 [0.09, 0.23]	0.194 [0.14, 0.25]	0.144 [0.08, 0.22]	0.212 [0.15, 0.26]	0.204 [0.15, 0.25]
Eye Length Proportion - Length	0.178 [0.12, 0.23]	0.140 [0.07, 0.22]	0.203 [0.14, 0.26]	0.141 [0.07, 0.22]	0.214 [0.15, 0.27]	0.187 [0.13, 0.23]
Eye Roundness Proportion – Upper Front – Poly	0.084 [0.01, 0.17]	0.232 [0.16, 0.32]	0.114 [0.04, 0.19]	0.288 [0.23, 0.36]	0.135 [0.07, 0.21]	0.241 [0.18, 0.30]
Eye Roundness Proportion – Upper Back – Poly	0.183 [0.13, 0.26]	0.139 [0.06, 0.32]	0.181 [0.12, 0.25]	0.200 [0.13, 0.28]	0.248 [0.18, 0.31]	0.301 [0.23, 0.36]
Eye Roundness Proportion – Lower Back – Poly	0.241 [0.17, 0.31]	0.116 [0.03, 0.20]	0.122 [0.05, 0.19]	0.139 [0.07, 0.23]	0.160 [0.10, 0.22]	0.00 [0.00, 0.00]
Eye Roundness Proportion – Lower Front – Poly	0.148 [0.08, 0.23]	0.030 [0.00, 0.06]	0.166 [0.10, 0.24]	0.052 [0.00, 0.10]	0.136 [0.07, 0.21]	0.093 [0.03, 0.14]
Eye Roundness Proportion – Upper Front – Linear	0.109 [0.03, 0.20]	0.265 [0.20, 0.35]	0.118 [0.05, 0.19]	0.259 [0.19, 0.34]	0.133 [0.07, 0.21]	0.196 [0.12, 0.26]
Eye Roundness Proportion – Upper Back – Linear	0.193 [0.14, 0.27]	0.128 [0.05, 0.21]	0.147 [0.08, 0.22]	0.159 [0.19, 0.34]	0.193 [0.13, 0.26]	0.222 [0.15, 0.28]
Eye Roundness Proportion – Lower Back – Linear	0.241 [0.17, 0.32]	0.097 [0.01, 0.19]	0.118 [0.05, 0.19]	0.097 [0.02, 0.18]	0.152 [0.08, 0.22]	0.144 [0.08, 0.20]
Eye Roundness Proportion – Lower Front – Linear	0.207 [0.14, 0.27]	0.140 [0.07, 0.22]	0.177 [0.12, 0.23]	0.173 [0.10, 0.25]	0.189 [0.14, 0.24]	0.206 [0.15, 0.25]
Eye Roundness Proportion – Total – Poly	0.091 [0.03, 0.17]	0.129 [0.05, 0.22]	0.136 [0.07, 0.21]	0.144 [0.07, 0.23]	0.168 [0.10, 0.24]	0.200 [0.13, 0.26]
Eye Roundness Ratio – Top-to-Bottom – Poly	0.253 [0.19, 0.35]	0.113 [0.04, 0.20]	0.230 [0.16, 0.30]	0.142 [0.07, 0.23]	0.186 [0.12, 0.25]	0.185 [0.12, 0.24]
Eye Roundness Ratio – Front-to-Back – Poly	0.181 [0.11, 0.26]	0.149 [0.07, 0.23]	0.154 [0.09, 0.22]	0.122 [0.04, 0.20]	0.174 [0.11, 0.24]	0.193 [0.12, 0.25]

Table 3: Repeatability of Geometric Eye Biometrics from a Single Coordinate Extraction

	Coord 1	Coord 2	Coord 3	Coord 4	Coord 5	Coord 6
Eye Depth Proportion – Full Length	0.484 [0.41, 0.57]	0.424 [0.35, 0.51]	0.316 [0.24, 0.40]	0.490 [0.42, 0.58]	0.379 [0.30, 0.46]	0.458 [0.38, 0.54]
Eye Depth Proportion – Front Eye Length	0.590 [0.52, 0.67]	0.338 [0.26, 0.42]	0.341 [0.26, 0.42]	0.318 [0.24, 0.40]	0.382 [0.30, 0.46]	0.440 [0.37, 0.52]
Eye Depth Point Proportion	0.468 [0.40, 0.55]	0.279 [0.20, 0.35]	0.535 [0.47, 0.62]	0.285 [0.21, 0.36]	0.555 [0.49, 0.64]	0.560 [0.49, 0.64]
Eye Height Proportion – Full Length	0.467 [0.38, 0.55]	0.434 [0.35, 0.52]	0.413 [0.34, 0.50]	0.453 [0.36, 0.54]	0.441 [0.36, 0.53]	0.489 [0.41, 0.58]
Eye Height Proportion – Front Eye Length	0.475 [0.40, 0.57]	0.421 [0.34, 0.51]	0.376 [0.30, 0.46]	0.459 [0.43, 0.59]	0.437 [0.36, 0.53]	0.487 [0.41, 0.57]
Eye Height Point Proportion	0.525 [0.45, 0.62]	0.446 [0.34, 0.51]	0.485 [0.41, 0.57]	0.471 [0.39, 0.56]	0.534 [0.46, 0.62]	0.540 [0.47, 0.62]
Eye Width-to-Height Ratio	0.497 [0.41, 0.59]	0.434 [0.35, 0.53]	0.361 [0.28, 0.45]	0.488 [0.40, 0.58]	0.405 [0.33, 0.50]	0.488 [0.41, 0.58]
Eye Displacement Proportion	0.429 [0.35, 0.52]	0.354 [0.27, 0.44]	0.551 [0.48, 0.63]	0.345 [0.26, 0.43]	0.543 [0.47, 0.63]	0.509 [0.44, 0.59]
Eye Length Proportion – Combined Height	0.388 [0.31, 0.48]	0.280 [0.21, 0.36]	0.496 [0.42, 0.58]	0.293 [0.22, 0.37]	0.552 [0.48, 0.64]	0.567 [0.50, 0.65]
Eye Length Proportion - Length	0.471 [0.40, 0.56]	0.277 [0.21, 0.35]	0.463 [0.39, 0.55]	0.316 [0.23, 0.40]	0.527 [0.46, 0.61]	0.546 [0.48, 0.63]
Eye Roundness Proportion – Upper Front – Poly	0.160 [0.09, 0.22]	0.206 [0.13, 0.27]	0.289 [0.22, 0.36]	0.218 [0.14, 0.29]	0.331 [0.26, 0.41]	0.378 [0.30, 0.46]
Eye Roundness Proportion – Upper Back – Poly	0.267 [0.19, 0.34]	0.160 [0.09, 0.22]	0.363 [0.29, 0.44]	0.212 [0.14, 0.28]	0.422 [0.35, 0.51]	0.451 [0.37, 0.54]
Eye Roundness Proportion – Lower Back – Poly	0.296 [0.22, 0.38]	0.143 [0.08, 0.20]	0.321 [0.25, 0.39]	0.233 [0.16, 0.30]	0.389 [0.31, 0.47]	0.605 [0.55, 0.67]
Eye Roundness Proportion – Lower Front – Poly	0.238 [0.16, 0.31]	0.180 [0.12, 0.24]	0.281 [0.21, 0.36]	0.185 [0.11, 0.25]	0.301 [0.23, 0.38]	0.439 [0.37, 0.52]
Eye Roundness Proportion – Upper Front – Linear	0.131 [0.07, 0.19]	0.176 [0.10, 0.24]	0.292 [0.22, 0.37]	0.195 [0.12, 0.26]	0.315 [0.24, 0.39]	0.279 [0.20, 0.36]
Eye Roundness Proportion – Upper Back – Linear	0.296 [0.21, 0.37]	0.163 [0.09, 0.23]	0.328 [0.25, 0.40]	0.196 [0.13, 0.26]	0.330 [0.25, 0.41]	0.345 [0.27, 0.42]
Eye Roundness Proportion – Lower Back – Linear	0.245 [0.16, 0.32]	0.147 [0.08, 0.21]	0.320 [0.29, 0.39]	0.278 [0.20, 0.35]	0.341 [0.27, 0.42]	0.355 [0.28, 0.43]
Eye Roundness Proportion – Lower Front – Linear	0.356 [0.28, 0.44]	0.280 [0.20, 0.35]	0.535 [0.47, 0.61]	0.308 [0.21, 0.40]	0.581 [0.52, 0.66]	0.581 [0.52, 0.65]
Eye Roundness Proportion – Total – Poly	0.224 [0.15, 0.29]	0.156 [0.09, 0.22]	0.328 [0.25, 0.41]	0.205 [0.13, 0.27]	0.349 [0.27, 0.43]	0.332 [0.26, 0.41]
Eye Roundness Ratio – Top-to-Bottom – Poly	0.135 [0.06, 0.19]	0.159 [0.09, 0.22]	0.310 [0.23, 0.39]	0.185 [0.11, 0.25]	0.365 [0.29, 0.44]	0.365 [0.29, 0.44]
Eye Roundness Ratio – Front-to-Back – Poly	0.282 [0.21, 0.36]	0.203 [0.13, 0.27]	0.357 [0.28, 0.43]	0.206 [0.13, 0.27]	0.345 [0.27, 0.42]	0.339 [0.27, 0.42]

Table 4: Repeatability of Geometric Eye Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction

	Coord 1	Coord 2	Coord 3	Coord 4	Coord 5	Coord 6
Eye Depth Proportion – Full Length	0.569 [0.48,0.66]	0.515 [0.43, 0.62]	0.401 [0.31, 0.51]	0.573 [0.49, 0.67]	0.457 [0.36, 0.56]	0.532 [0.44, 0.63]
Eye Depth Proportion – Front Eye Length	0.666 [0.60, 0.75]	0.436 [0.34, 0.54]	0.436 [0.34, 0.53]	0.419 [0.32, 0.52]	0.464 [0.38, 0.56]	0.532 [0.45, 0.62]
Eye Depth Point Proportion	0.570 [0.49, 0.66]	0.393 [0.29, 0.50]	0.620 [0.55, 0.70]	0.398 [0.30, 0.50]	0.632 [0.55, 0.70]	0.633 [0.56, 0.70]
Eye Height Proportion – Full Length	0.522 [0.43, 0.61]	0.501 [0.41, 0.59]	0.507 [0.42, 0.61]	0.504 [0.41, 0.60]	0.514 [0.42, 0.61]	0.542 [0.45, 0.64]
Eye Height Proportion – Front Eye Length	0.549 [0.46, 0.64]	0.507 [0.42, 0.61]	0.469 [0.38, 0.57]	0.522 [0.43, 0.60]	0.509 [0.42, 0.61]	0.556 [0.47, 0.65]
Eye Height Point Proportion	0.572 [0.49, 0.67]	0.523 [0.43, 0.63]	0.569 [0.49, 0.67]	0.528 [0.44, 0.63]	0.603 [0.52, 0.70]	0.613 [0.53, 0.70]
Eye Width-to-Height Ratio	0.536 [0.45, 0.64]	0.490 [0.40, 0.59]	0.446 [0.35, 0.55]	0.529 [0.44, 0.63]	0.464 [0.37, 0.57]	0.532 [0.44, 0.63]
Eye Displacement Proportion	0.490 [0.40, 0.59]	0.442 [0.35, 0.54]	0.617 [0.54, 0.70]	0.433 [0.34, 0.54]	0.614 [0.54, 0.70]	0.586 [0.51, 0.68]
Eye Length Proportion – Combined Height	0.483 [0.39, 0.59]	0.391 [0.29, 0.49]	0.587 [0.51, 0.68]	0.407 [0.31, 0.51]	0.626 [0.55, 0.72]	0.640 [0.57, 0.73]
Eye Length Proportion - Length	0.572 [0.49, 0.67]	0.391 [0.29, 0.50]	0.556 [0.47, 0.65]	0.400 [0.30, 0.50]	0.605 [0.53, 0.70]	0.630 [0.55, 0.72]
Eye Roundness Proportion – Upper Front – Poly	0.258 [0.16, 0.36]	0.286 [0.18, 0.38]	0.412 [0.32, 0.52]	0.394 [0.30, 0.50]	0.451 [0.36, 0.56]	0.467 [0.37, 0.56]
Eye Roundness Proportion – Upper Back – Poly	0.368 [0.26, 0.46]	0.247 [0.14, 0.35]	0.471 [0.38, 0.57]	0.307 [0.20, 0.41]	0.506 [0.42, 0.60]	0.515 [0.42, 0.61]
Eye Roundness Proportion – Lower Back – Poly	0.385 [0.29, 0.49]	0.227 [0.13, 0.32]	0.445 [0.36, 0.54]	0.313 [0.21, 0.41]	0.502 [0.42, 0.59]	0.810 [0.77, 0.86]
Eye Roundness Proportion – Lower Front – Poly	0.343 [0.24, 0.45]	0.297 [0.20, 0.40]	0.388 [0.29, 0.49]	0.362 [0.26, 0.47]	0.418 [0.32, 0.52]	0.573 [0.50, 0.66]
Eye Roundness Proportion – Upper Front – Linear	0.211 [0.11, 0.31]	0.244 [0.14, 0.33]	0.414 [0.32, 0.52]	0.256 [0.15, 0.35]	0.435 [0.34, 0.54]	0.379 [0.28, 0.48]
Eye Roundness Proportion – Upper Back – Linear	0.398 [0.30, 0.49]	0.252 [0.15, 0.35]	0.445 [0.35, 0.55]	0.288 [0.19, 0.39]	0.433 [0.34, 0.53]	0.441 [0.34, 0.54]
Eye Roundness Proportion – Lower Back – Linear	0.329 [0.23, 0.43]	0.236 [0.14, 0.33]	0.445 [0.35, 0.54]	0.304 [0.20, 0.40]	0.457 [0.37, 0.55]	0.474 [0.39, 0.57]
Eye Roundness Proportion – Lower Front – Linear	0.456 [0.36, 0.56]	0.394 [0.30, 0.50]	0.625 [0.56, 0.70]	0.383 [0.29, 0.49]	0.657 [0.59, 0.73]	0.650 [0.58, 0.73]
Eye Roundness Proportion – Total – Poly	0.340 [0.24, 0.44]	0.243 [0.14, 0.34]	0.448 [0.35, 0.55]	0.303 [0.20, 0.40]	0.460 [0.37, 0.56]	0.434 [0.33, 0.54]
Eye Roundness Ratio – Top-to-Bottom – Poly	0.194 [0.09, 0.27]	0.250 [0.15, 0.34]	0.403 [0.31, 0.51]	0.278 [0.17, 0.37]	0.470 [0.38, 0.57]	0.471 [0.38, 0.57]
Eye Roundness Ratio – Front-to-Back – Poly	0.386 [0.29, 0.49]	0.300 [0.20, 0.40]	0.472 [0.39, 0.56]	0.311 [0.21, 0.41]	0.454 [0.36, 0.55]	0.442 [0.35, 0.54]

The third appraisal is a deeper exploration of the error structures underlying the geometric biometric measures of the eye. From these results, several overarching trends can be observed. The first is that, for eye biometrics, the overwhelming majority of the error comes from uncertainty in extraction of landmark coordinates (see Table 1). This makes intuitive sense, as not only are the boney and cartilaginous features of the eye obscured by a large amount of soft tissues, but being

in the interior of the face, as opposed to the outline of the head, these features also demonstrate less contrast within the image, making them difficult to see against both white and black fur. Unfortunately, this trend suggests that the only means of improving repeatability of these measures is either an improved system for coordinate identification, higher quality images (added expense), or simply increasing the number of coordinate extraction replicates (added labor). It should also be noted, however, that several eye biometrics showed a notable amount of error attributed to variability between photo. In particular, metrics related to the height and depth of the eye seemed susceptible to error between days. As the protocol for image acquisition used the eye to center the image, it seems unlikely that this error is attributable to variations in angle between the cow and camera are to blame for this error. It seems more likely that changes in facial expression (squinting, staring, glaring) are more likely to blame for this phenomenon, though further research would be needed to confirm this suspicion (see Table 2).

With respect to the repeatability results, several other important trends emerge in the data (see Table 4). The first is that coordinate system 6 clearly demonstrates superior performance. In comparing the performance of coordinate systems holistically, the inner eyelid yielded the best approximation for the horizontal line of the eye, while the ridge of the boney eye orbital yielded the best results for the vertical and diagonal points. It is also clear that, for the majority of eye metrics to show acceptable levels of repeatability, information from at least two landmark coordinate extractions are necessary. Overall, the points related to eye height and depth performed better than metrics seeking to describe roundness traits. This may reflect the difficulty in identifying the eye roundness landmarks amongst the soft tissue folds of the eye, which were quite prone to shadow and overexposure. But also, as the eye roundness landmarks are selected using reference lines, any error in selection of the horizontal and vertical landmarks of the eye would

have been indirectly amplified into the roundness landmarks. Overall, eye roundness metrics calculated using polygon area appeared more robust to measurement error as compared corresponding metrics utilizing linear projections.

Fourth, it is informative to explore the correlation structures between eye biometrics. The first such relationship is simply the direct correlation between all pairwise combinations of metrics within each measurement system. It should be noted that, for eye biometrics, pairwise correlations were calculated between metrics within coordinate systems, and aggregated over all versions. In comparing the distribution between measurement systems, the normalized length metrics demonstrate a far thicker right tail than the geometric biometrics. This indicates that geometric biometrics overall show a higher degree of independence (see Figure 14).

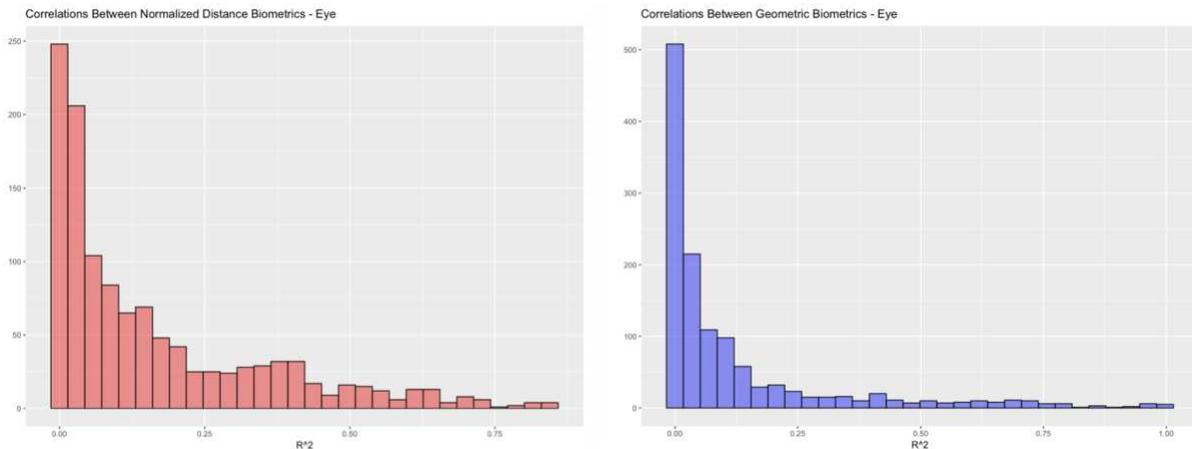


Figure 14: Comparing level of correlation between pairwise combinations of normalized length and geometric eye biometrics

Table 5: Pairwise Correlation between Geometric Eye Biometrics Coordinate System 6

	EyeDepthProportionFullLength	EyeDepthProportionFrontEyeLength	EyeDepthPointProportion	EyeHeightProportionFullLength	EyeHeightProportionFrontEyeLength	EyeHeightPointProportion	EyeWidth-to-HeightRatio	EyeDisplacementProportion	EyeLengthProportionCombinedHeight	EyeLengthProportionLength	EyeRoundnessProportionUpperFrontPoly	EyeRoundnessProportionUpperBackPoly	EyeRoundnessProportionLowerFrontPoly	EyeRoundnessProportionLowerBackPoly	EyeRoundnessProportionTotalPoly	EyeRoundnessRatioTop-to-BottomPoly	EyeRoundnessRatioFront-to-BackPoly	EyeRoundnessProportionUpperFrontLinear	EyeRoundnessProportionUpperBackLinear	EyeRoundnessProportionLowerFrontLinear	EyeRoundnessProportionLowerBackLinear
EyeDepthProportionFullLength	1.00	0.20	0.20	0.38	0.21	-0.05	0.80	0.18	-0.27	-0.01	0.25	0.09	0.23	0.11	0.43	-0.44	0.11	0.20	0.08	0.30	-0.35
EyeDepthProportionFrontEyeLength	0.20	1.00	-0.85	0.17	0.13	-0.03	0.22	-0.62	0.15	0.24	0.09	0.07	-0.32	0.77	0.24	-0.17	-0.50	0.01	0.03	-0.14	0.70
EyeDepthPointProportion	0.20	-0.85	1.00	-0.01	-0.05	0.03	0.11	0.73	-0.34	-0.33	0.02	-0.03	0.59	-0.58	-0.02	-0.06	0.64	0.08	-0.01	0.28	-0.86
EyeHeightProportionFullLength	0.38	0.17	-0.01	1.00	0.39	0.07	0.86	-0.05	-0.34	-0.07	0.21	0.28	0.07	0.18	0.46	0.38	0.00	0.05	0.02	0.10	0.21
EyeHeightProportionFrontEyeLength	0.21	0.13	-0.05	0.39	1.00	-0.88	0.37	0.56	0.60	0.78	0.65	-0.60	-0.07	0.06	0.02	0.05	-0.36	0.17	-0.38	-0.05	-0.08
EyeHeightPointProportion	-0.05	-0.03	0.03	0.07	-0.88	1.00	0.02	-0.66	-0.84	-0.90	-0.57	0.82	0.11	0.04	0.22	0.14	0.40	-0.15	0.43	0.09	0.21
EyeWidth-to-HeightRatio	0.80	0.22	0.11	0.86	0.37	0.02	1.00	0.07	-0.37	-0.05	0.28	0.23	0.17	0.18	0.54	0.00	0.06	0.14	0.05	0.23	-0.06
EyeDisplacementProportion	0.18	-0.62	0.73	-0.05	0.56	-0.66	0.07	1.00	0.32	0.36	0.40	-0.58	0.37	-0.46	-0.17	-0.14	0.21	0.16	-0.30	0.15	-0.79
EyeLengthProportionCombinedHeight	-0.27	0.15	-0.34	-0.34	0.60	-0.84	-0.37	0.32	1.00	0.95	0.32	-0.75	-0.40	-0.01	-0.39	-0.09	-0.58	0.02	-0.36	-0.20	0.06
EyeLengthProportionLength	-0.01	0.24	-0.33	-0.07	0.78	-0.90	-0.05	0.36	0.95	1.00	0.45	-0.73	-0.38	0.06	-0.23	-0.09	-0.60	0.07	-0.36	-0.15	0.04
EyeRoundnessProportionUpperFrontPoly	0.25	0.09	0.02	0.21	0.65	-0.57	0.28	0.40	0.32	0.45	1.00	-0.31	0.06	0.07	0.38	0.13	0.06	0.86	-0.14	0.06	-0.10
EyeRoundnessProportionUpperBackPoly	0.09	0.07	-0.03	0.28	-0.60	0.82	0.23	-0.58	-0.75	-0.73	-0.31	1.00	0.11	0.14	0.59	0.34	-0.13	-0.01	0.78	0.14	0.35
EyeRoundnessProportionLowerFrontPoly	0.23	-0.32	0.59	0.07	-0.07	0.11	0.17	0.37	-0.40	-0.38	0.06	0.11	1.00	-0.12	0.28	-0.22	0.41	0.14	0.06	0.48	-0.44
EyeRoundnessProportionLowerBackPoly	0.11	0.77	-0.58	0.18	0.06	0.04	0.18	-0.46	-0.01	0.06	0.07	0.14	-0.12	1.00	0.45	-0.26	-0.16	0.05	0.11	-0.02	0.63
EyeRoundnessProportionTotalPoly	0.43	0.24	-0.02	0.46	0.02	0.22	0.54	-0.17	-0.39	-0.23	0.38	0.59	0.28	0.45	1.00	0.07	0.04	0.49	0.63	0.48	0.27
EyeRoundnessRatioTop-to-BottomPoly	-0.44	-0.17	-0.06	0.38	0.05	0.14	0.00	-0.14	-0.09	-0.09	0.13	0.34	-0.22	-0.26	0.07	1.00	-0.05	0.15	0.29	-0.40	0.33
EyeRoundnessRatioFront-to-BackPoly	0.11	-0.50	0.64	0.00	-0.36	0.40	0.06	0.21	-0.58	-0.60	0.06	0.13	0.41	-0.16	0.04	-0.05	1.00	0.33	-0.16	0.00	-0.50
EyeRoundnessProportionUpperFrontLinear	0.20	0.01	0.08	0.05	0.17	-0.15	0.14	0.16	0.02	0.07	0.86	-0.01	0.14	0.05	0.49	0.15	0.33	1.00	0.06	0.13	-0.09
EyeRoundnessProportionUpperBackLinear	0.08	0.03	-0.01	0.02	-0.38	0.43	0.05	-0.30	-0.36	-0.36	-0.14	0.78	0.06	0.11	0.63	0.29	-0.16	0.06	1.00	0.13	0.27
EyeRoundnessProportionLowerFrontLinear	0.30	-0.14	0.28	0.10	-0.05	0.09	0.23	0.15	-0.20	-0.15	0.06	0.14	0.48	-0.02	0.48	-0.40	0.00	0.13	0.13	1.00	-0.22
EyeRoundnessProportionLowerBackLinear	-0.35	0.70	-0.86	0.21	-0.08	0.21	-0.06	-0.79	0.06	0.04	-0.10	0.35	-0.44	0.63	0.27	0.33	-0.50	-0.09	0.27	-0.22	1.00

Another correlation structure that carries the potential to influence performance of biometrics as model covariates is correlation amongst error terms (see Figure 15). In comparing the distribution of correlation values for pair-wise between-day error, normalized length measured again demonstrated a notably thicker right tail as compared to the geometric biometrics. Subsequently, the overall average error correlation amongst normalized length measures was 0.17, whereas geometric biometrics demonstrated an overall average correlation of only 0.11. Thus, geometric biometrics appear to have a marginal advantage in terms of error structure for eye biometrics.

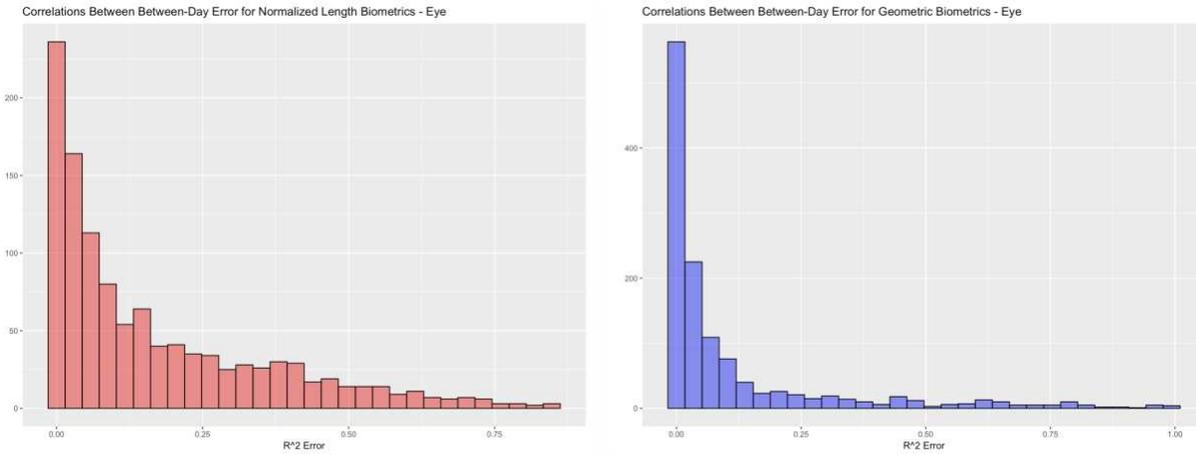


Figure 15: Comparing levels of correlation in error terms between normalized length and geometric eye biometrics

Fifth and finally, the proportion of change in biometrics between days that can be attributed to change in image quality was assessed for both measurement systems (see Figure 16). With respect to attributes related to image scale and rotation, geometric biometrics clearly out-perform normalized length measures. The average proportion of error attributed to image scale and rotation measures was only 4%, with very few geometric biometrics showing an error correlation of more than 20%. Normalized length biometrics demonstrated an average proportion of error attributed to image scale and rotation measures of 18%, with several metrics having more than 50% of their error attributed to these image quality attributes. On the other hand, both measurement systems proved quite robust to changes in image attributes related to camera position. The average proportion of variance attributed to these image attributes for normalized length and geometric biometrics was 2% and 3% respectively. While geometric biometrics demonstrated a slightly narrower range, the performance of these two measurement systems were effectively equivalent.

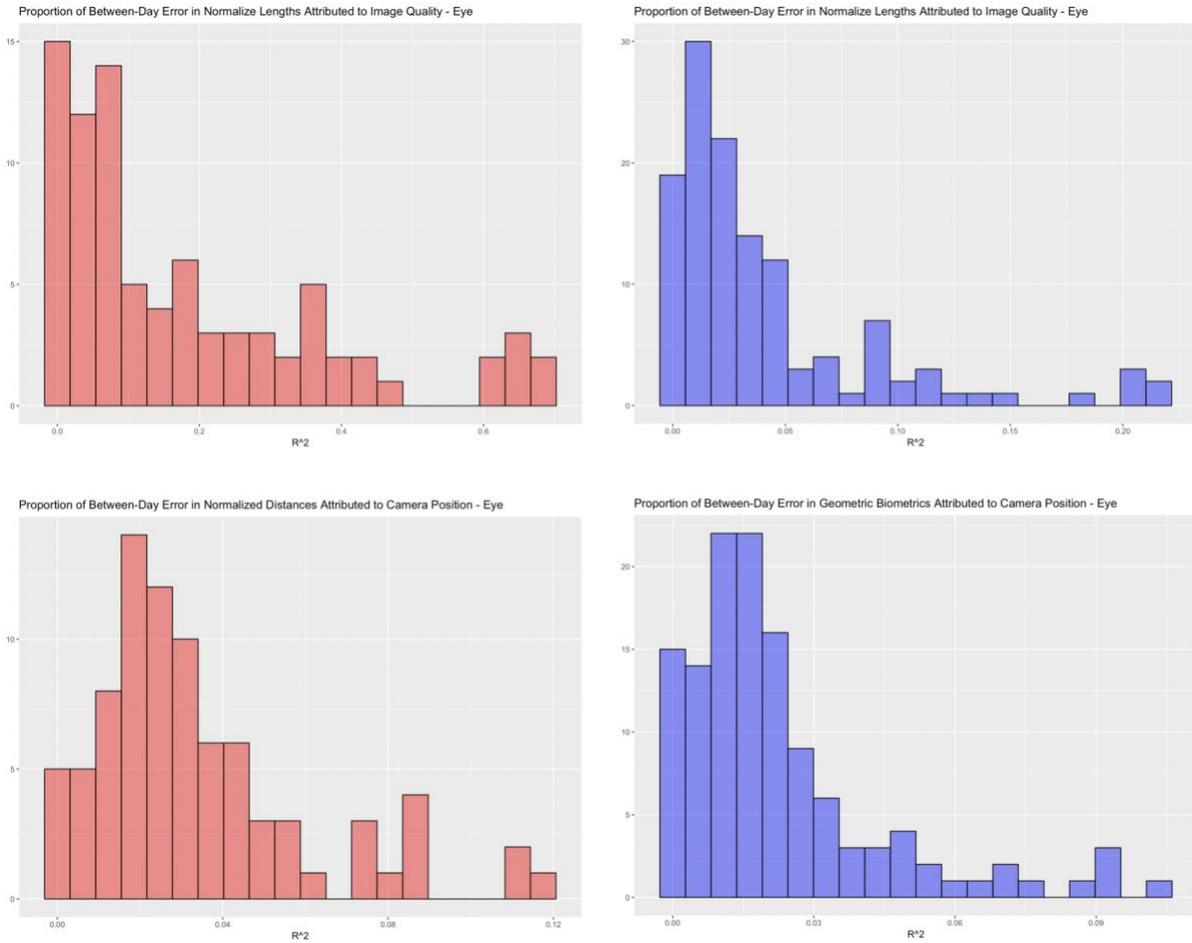


Figure 16: Above – Proportion of error attributed to variations in image scale and rotation;
 Below – Proportion of error attributed to variations in camera position

Results - Muzzle Biometrics

The shape of the muzzle is defined by a total of 12 landmark points: A_extrap, E_upper, H, I, J, K, JI, JH, KI, KH, L, Q. From these landmark points, a total of 25 candidate geometric biometrics were extracted. For each geometric biometric, alternative combinations of landmark points were used to compute between one and four alternative versions of the same geometric derivation (see Figure 17). Details of these combinations can be found in the derivations of the geometric biometrics in Appendix A.

Nostril Flare Proportion – Upper Front
 Nostril Flare Point Proportion – Upper Front
 Nostril Flare Proportion – Upper Back
 Nostril Flare Point Proportion – Upper Back
 Nostril Flare Proportion – Lower Back
 Nostril Flare Point Proportion – Lower Back
 Nostril Flare Proportion – Lower Front
 Nostril Flare Point Proportion – Lower Front
 Nostril Depth Proportion - Linear
 Nostril Depth Proportion - Area
 Nostril Depth Point Proportion
 Nostril Height Proportion - Linear
 Nostril Height Proportion - Area
 Nostril Height Point Proportion
 Nostril Position Angle
 Nostril – Muzzle Ratio - Area
 Nostril – Muzzle Ratio - Height
 Nostril – Muzzle Ratio - Length
 Mouth Eye-to-Extrap Offset - Height
 Mouth Eye-to-Extrap Offset - Length
 Upper Lip Roundness Proportion
 Upper Lip Roundness Point Proportion
 Muzzle Thickness Proportion
 Chin Thickness Proportion
 Chin-to-Lip Thickness Proportion

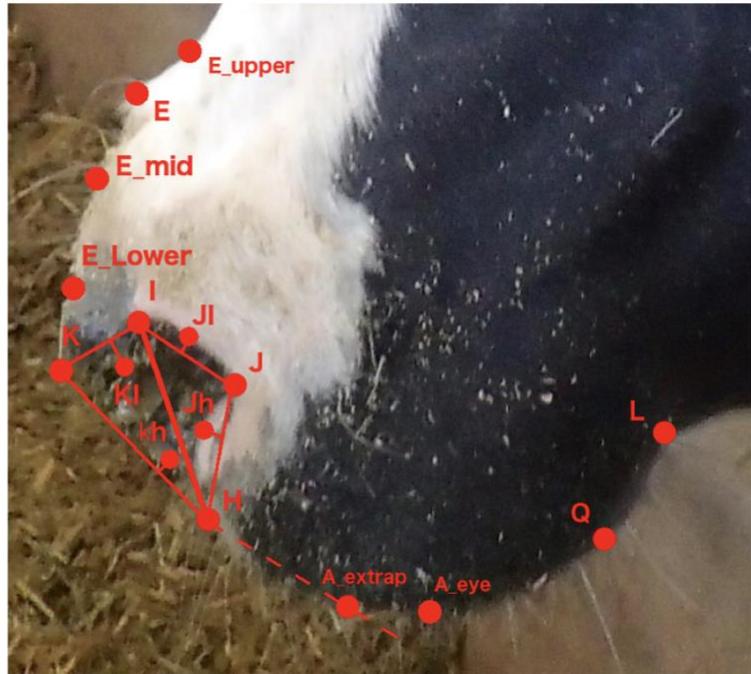


Figure 17: Landmark points and geometric biometrics of the muzzle (further Appendix A)

First is an assessment of the global behavior of the higher moments of biometrics in each measurement system (see Figure 18). With respect to measures of skew (3rd moment), normalized length values are well-centered around zero, indicating that normalized length values are as likely to be skewed towards negative as positive values. The majority of geometric biometrics are also well-centered around zero, but here there is a clear tendency to produce some extremely skewed distributions, both in the positive and negative directions. With respect to measures of kurtosis (4th moment), normalized length values appear to have a slight tendency towards producing measures with thicker tails, but again geometric biometrics show a capacity to produce measures with extreme values for tail thickness.

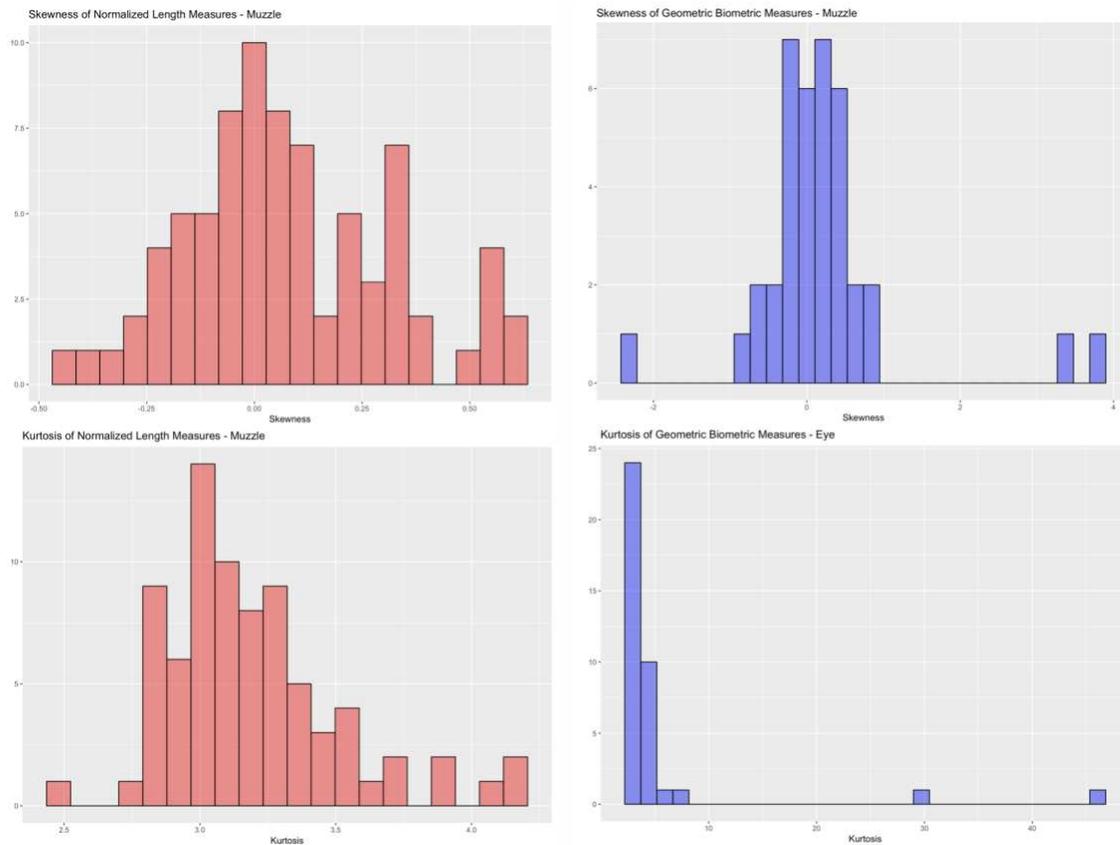


Figure 18: Distribution of 3rd & 4th moments for normalized length & geometric muzzle biometrics

Secondly, is a comparison of performance of normalized and geometric biometrics in terms of repeatability (see Figure 19). First is the repeatability of these two measurement systems within-photo, which corresponds to their robustness to errors in points selection. Here we see that while repeatability values of both measurement systems occupy roughly the same range, density for the geometric biometrics appears to be shifted slightly towards the right representing higher repeatability. With respect to overall average within-photo repeatability, normalized lengths marginally outperform geometric biometrics with a value of 0.36 to 0.34. But it should also be noted that geometric biometrics yield a relatively greater number of measurements exceeding a repeatability of 0.6 as compared to normalized length measures.

Also of concern is the comparative performance of these measurement systems in terms of between-photo repeatability, or overall measurement repeatability. For geometric biometrics, we see a nice tight clustering of repeatability values around 0.5, with very few biometrics demonstrating truly poor repeatability values (<0.3). The normalized length measures demonstrate more variable performance, with a number of low repeatability values, but also a number of repeatability values exceeding 0.6, which geometric biometrics did not achieve. The overall average repeatability for geometric biometrics was 0.39, with normalized length again performing marginally better with a value of 0.41. Thus, it appears that for muzzle biometrics normalized length demonstrate a slight advantage with respect to measure repeatability.

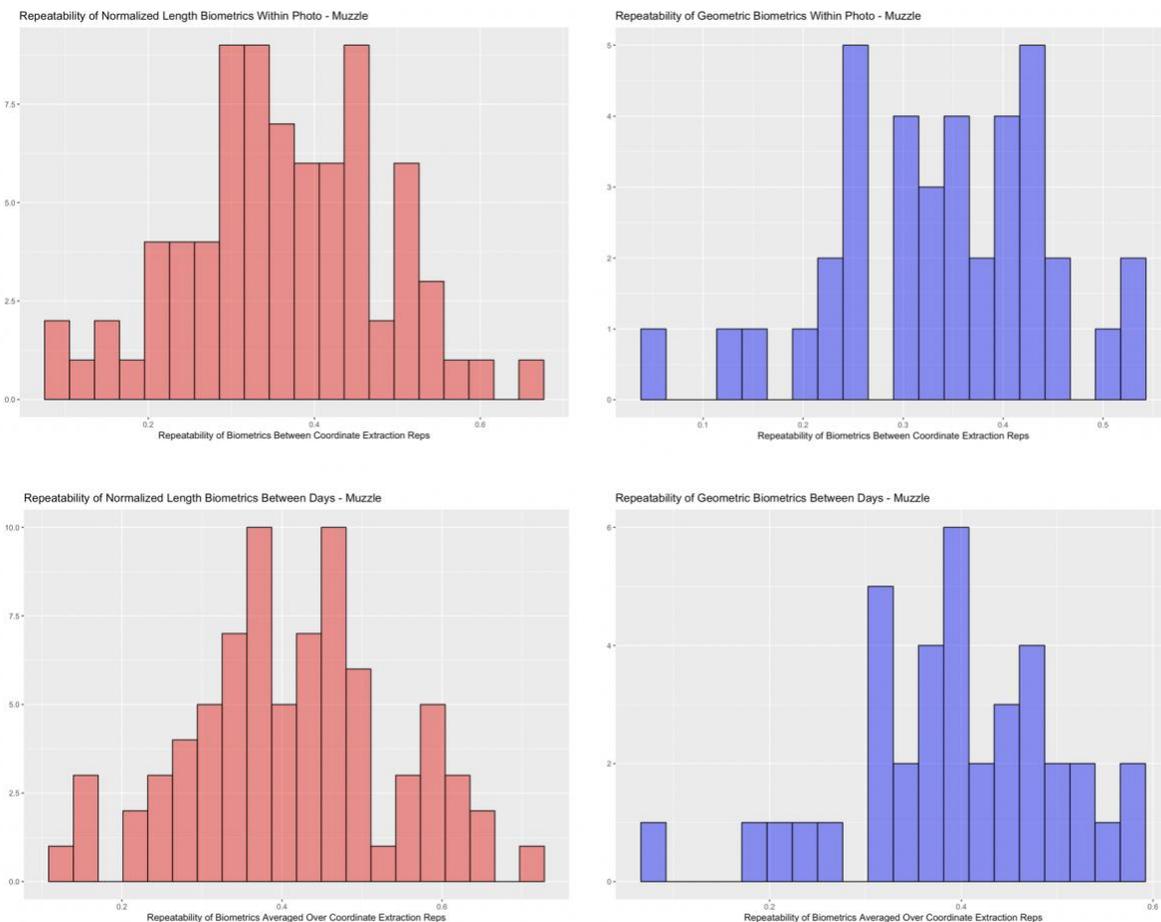


Figure 19: Comparison of Between and Within-Day Repeatability of Normalized and Geometric Muzzle Biometric Measures

Third is a deeper exploration of the error structures underlying the geometric biometric measures of the eye. From these results, several overarching trends can be observed. In stark contrast to the results returned from eye biometrics, muzzle biometrics as a whole demonstrated very little error as a result of error in landmark coordinate extraction (see Table 6 & 7). While the flare metrics calculated from the secondary nose landmarks proved difficult to select, the larger structures of the muzzle appear to be quite reliably identified. Thus, the majority of the error in muzzle biometrics are attributed to variability in the photos themselves on distinct days. As the image was aligned to the face using anatomical features near the eye, which are relatively distant to the point of the muzzle, this could indicate issues controlling for out-of-plane facial angle (ie – a fisheye effect). However, as the muzzle is also perhaps the fleshiest part of the face, it is also possible that, as with the eye, biometric values were potentially being influence by changes in the micro expressions of the cow that varied between days. This variability between images resulted in exceedingly poor repeatability of muzzle biometric measures. Only a handful of chin and mouth traits demonstrated repeatability values above 0.5, with no nostril traits proving sufficiently robust for use in predictive modeling applications (see Table 9).

Table 6: Proportion of total variability in muzzle biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps)

	Version 1	Version 2	Version 3	Version 4
Nostril Flare Proportion – Upper Front	0.645 [0.57, 0.72]			
Nostril Flare Point Proportion – Upper Front	0.404 [0.34, 0.46]			
Nostril Flare Proportion – Upper Back	0.412 [0.35, 0.47]			
Nostril Flare Point Proportion – Upper Back	0.327 [0.27, 0.38]			
Nostril Flare Proportion – Lower Back	0.461 [0.39, 0.52]			
Nostril Flare Point Proportion – Lower Back	0.740 [0.66, 0.82]			
Nostril Flare Proportion – Lower Front	0.203 [0.16, 0.24]			
Nostril Flare Point Proportion – Lower Front	0.152 [0.12, 0.18]			
Nostril Depth Proportion - Linear	0.091 [0.07, 0.11]			
Nostril Depth Proportion - Area	0.285 [0.24, 0.33]			
Nostril Depth Point Proportion	0.224 [0.19, 0.26]			
Nostril Height Proportion - Linear	0.299 [0.24, 0.35]			
Nostril Height Proportion - Area	0.244 [0.20, 0.28]			
Nostril Height Point Proportion	0.138 [0.11, 0.16]			
Nostril Position Angle	0.119 [0.09, 0.14]			
Nostril – Muzzle Ratio - Area	0.151 [0.12, 0.18]	0.137 [0.11, 0.16]		
Nostril – Muzzle Ratio - Height	0.281 [0.23, 0.32]	0.239 [0.20, 0.28]		
Nostril – Muzzle Ratio - Length	0.165 [0.13, 0.19]			
Mouth Eye-to-Extrap Offset - Height	0.298 [0.25, 0.34]			
Mouth Eye-to-Extrap Offset - Length	0.276 [0.23, 0.32]			
Upper Lip Roundness Proportion	0.170 [0.13, 0.20]	0.401 [0.33, 0.46]		
Upper Lip Roundness Point Proportion	0.411 [0.35, 0.47]	0.392 [0.33, 0.45]		
Muzzle Thickness Proportion	0.146 [0.12, 0.17]	0.256 [0.21, 0.30]		
Chin Thickness Proportion	0.189 [0.15, 0.22]	0.245 [0.20, 0.28]	0.276 [0.23, 0.32]	0.284 [0.23, 0.33]
Chin Thickness Point Proportion	0.372 [0.31, 0.43]	0.301 [0.25, 0.35]	0.590 [0.51, 0.66]	0.505 [0.43, 0.57]
Chin-to-Lip Thickness Proportion	0.171 [0.14, 0.20]	0.225 [0.18, 0.26]		

Table 7: Proportion of total variability in muzzle biometrics attributed to error in image acquisition (variance between days/photos)

	Version 1	Version 2	Version 3	Version 4
Nostril Flare Proportion – Upper Front	0.229 [0.14, 0.32]			
Nostril Flare Point Proportion – Upper Front	0.345 [0.27, 0.42]			
Nostril Flare Proportion – Upper Back	0.382 [0.31, 0.46]			
Nostril Flare Point Proportion – Upper Back	0.420 [0.35, 0.50]			
Nostril Flare Proportion – Lower Back	0.386 [0.32, 0.47]			
Nostril Flare Point Proportion – Lower Back	0.211 [0.14, 0.31]			
Nostril Flare Proportion – Lower Front	0.388 [0.31, 0.50]			
Nostril Flare Point Proportion – Lower Front	0.320 [0.25, 0.38]			
Nostril Depth Proportion - Linear	0.535 [0.44, 0.62]			
Nostril Depth Proportion - Area	0.388 [0.32, 0.46]			
Nostril Depth Point Proportion	0.557 [0.48, 0.64]			
Nostril Height Proportion - Linear	0.323 [0.25, 0.39]			
Nostril Height Proportion - Area	0.462 [0.38, 0.54]			
Nostril Height Point Proportion	0.502 [0.42, 0.58]			
Nostril Position Angle	0.479 [0.39, 0.56]			
Nostril – Muzzle Ratio - Area	0.512 [0.43, 0.59]	0.506 [0.42, 0.59]		
Nostril – Muzzle Ratio - Height	0.411 [0.34, 0.48]	0.395 [0.32, 0.47]		
Nostril – Muzzle Ratio - Length	0.484 [0.40, 0.56]			
Mouth Eye-to-Extrap Offset - Height	0.374 [0.30, 0.45]			
Mouth Eye-to-Extrap Offset - Length	0.319 [0.25, 0.39]			
Upper Lip Roundness Proportion	0.302 [0.23, 0.36]	0.176 [0.12, 0.23]		
Upper Lip Roundness Point Proportion	0.342 [0.27, 0.42]	0.359 [0.28, 0.43]		
Muzzle Thickness Proportion	0.431 [0.35, 0.51]	0.322 [0.25, 0.39]		
Chin Thickness Proportion	0.302 [0.24, 0.36]	0.303 [0.24, 0.37]	0.300 [0.23, 0.36]	0.319 [0.25, 0.39]
Chin Thickness Point Proportion	0.315 [0.25, 0.39]	0.391 [0.33, 0.47]	0.186 [0.11, 0.27]	0.248 [0.17, 0.32]
Chin-to-Lip Thickness Proportion	0.398 [0.32, 0.47]	0.328 [0.26, 0.39]		

Table 8: Repeatability of Geometric Muzzle Biometrics from a Single Landmark Coordinate Extraction

	Version 1	Version 2	Version 3	Version 4
Nostril Flare Proportion – Upper Front	0.126 [0.06, 0.19]			
Nostril Flare Point Proportion – Upper Front	0.250 [0.17, 0.33]			
Nostril Flare Proportion – Upper Back	0.206 [0.12, 0.28]			
Nostril Flare Point Proportion – Upper Back	0.253 [0.17, 0.34]			
Nostril Flare Proportion – Lower Back	0.153 [0.07, 0.22]			
Nostril Flare Point Proportion – Lower Back	0.048 [0.00, 0.08]			
Nostril Flare Proportion – Lower Front	0.409 [0.33, 0.40]			
Nostril Flare Point Proportion – Lower Front	0.528 [0.46, 0.61]			
Nostril Depth Proportion - Linear	0.375 [0.29, 0.48]			
Nostril Depth Proportion - Area	0.327 [0.24, 0.41]			
Nostril Depth Point Proportion	0.219 [0.13, 0.31]			
Nostril Height Proportion - Linear	0.378 [0.30, 0.46]			
Nostril Height Proportion - Area	0.293 [0.21, 0.39]			
Nostril Height Point Proportion	0.360 [0.27, 0.46]			
Nostril Position Angle	0.402 [0.31, 0.50]			
Nostril – Muzzle Ratio - Area	0.338 [0.25, 0.44]	0.357 [0.27, 0.45]		
Nostril – Muzzle Ratio - Height	0.308 [0.23, 0.40]	0.366 [0.28, 0.46]		
Nostril – Muzzle Ratio - Length	0.351 [0.27, 0.45]			
Mouth Eye-to-Extrap Offset - Height	0.329 [0.24, 0.42]			
Mouth Eye-to-Extrap Offset - Length	0.405 [0.33, 0.49]			
Upper Lip Roundness Proportion	0.528 [0.45, 0.61]	0.423 [0.35, 0.51]		
Upper Lip Roundness Point Proportion	0.248 [0.16, 0.33]	0.249 [0.16, 0.33]		
Muzzle Thickness Proportion	0.423 [0.34, 0.52]	0.421 [0.34, 0.52]		
Chin Thickness Proportion	0.509 [0.43, 0.59]	0.452 [0.38, 0.54]	0.424 [0.35, 0.51]	0.397 [0.32, 0.48]
Chin Thickness Point Proportion	0.313 [0.23, 0.40]	0.308 [0.22, 0.39]	0.225 [0.15, 0.30]	0.247 [0.17, 0.32]
Chin-to-Lip Thickness Proportion	0.431 [0.35, 0.52]	0.448 [0.37, 0.54]		

Table 9: Repeatability of Geometric Eye Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction

	Version 1	Version 2	Version 3	Version 4
Nostril Flare Proportion – Upper Front	0.186 [0.08, 0.28]			
Nostril Flare Point Proportion – Upper Front	0.314 [0.21, 0.41]			
Nostril Flare Proportion – Upper Back	0.260 [0.16, 0.36]			
Nostril Flare Point Proportion – Upper Back	0.303 [0.20, 0.40]			
Nostril Flare Proportion – Lower Back	0.200 [0.10, 0.28]			
Nostril Flare Point Proportion – Lower Back	0.077 [0.00, 0.14]			
Nostril Flare Proportion – Lower Front	0.455 [0.36, 0.56]			
Nostril Flare Point Proportion – Lower Front	0.572 [0.50, 0.66]			
Nostril Depth Proportion - Linear	0.392 [0.29, 0.50]			
Nostril Depth Proportion - Area	0.381 [0.29, 0.48]			
Nostril Depth Point Proportion	0.247 [0.14, 0.34]			
Nostril Height Proportion - Linear	0.444 [0.35, 0.55]			
Nostril Height Proportion - Area	0.334 [0.23, 0.44]			
Nostril Height Point Proportion	0.387 [0.29, 0.49]			
Nostril Position Angle	0.427 [0.33, 0.53]			
Nostril – Muzzle Ratio - Area	0.365 [0.27, 0.47]	0.383 [0.28, 0.49]		
Nostril – Muzzle Ratio - Height	0.358 [0.26, 0.46]	0.415 [0.32, 0.51]		
Nostril – Muzzle Ratio - Length	0.383 [0.28, 0.49]			
Mouth Eye-to-Extrap Offset - Height	0.388 [0.28, 0.49]			
Mouth Eye-to-Extrap Offset - Length	0.470 [0.38, 0.57]			
Upper Lip Roundness Proportion	0.578 [0.50, 0.67]	0.529 [0.44, 0.63]		
Upper Lip Roundness Point Proportion	0.312 [0.21, 0.41]	0.310 [0.21, 0.41]		
Muzzle Thickness Proportion	0.456 [0.36, 0.56]	0.483 [0.39, 0.59]		
Chin Thickness Proportion	0.562 [0.48, 0.65]	0.515 [0.43, 0.60]	0.492 [0.41, 0.58]	0.462 [0.38, 0.55]
Chin Thickness Point Proportion	0.385 [0.29, 0.49]	0.362 [0.26, 0.46]	0.319 [0.22, 0.42]	0.330 [0.23, 0.43]
Chin-to-Lip Thickness Proportion	0.472 [0.38, 0.57]	0.504 [0.42, 0.60]		

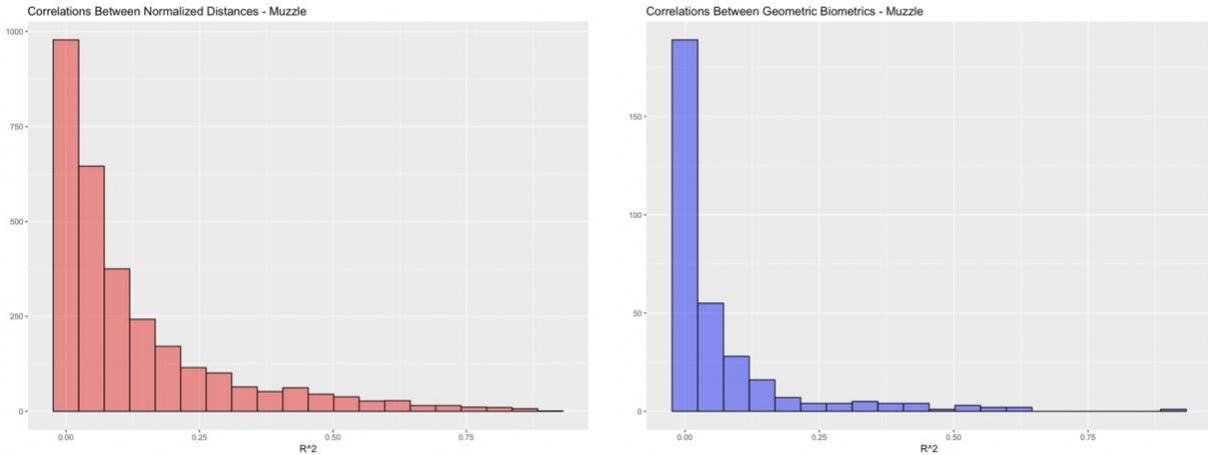


Figure 20: Comparing level of correlation between pairwise combinations of normalized length and geometric muzzle biometrics

Fourth, is an exploration of the correlation structures between muzzle biometrics. Based on the results of the repeatability analysis, only version 1 metrics were kept, and all other version dropped in order to avoid redundancy between biometrics with the same geometric derivations, which would artificially inflate measures of correlation. In comparing the distributions of pairwise correlations between observed biometrics, the density for normalized length biometrics is clearly shifted right, with a thicker upper tail (see Figure 20). This is reflected in the average pairwise correlation value, with normalized length values showing double the overall correlation level of 0.13 to 0.07 from geometric biometrics. And while correlation between the majority of biometrics is quite low for both measurement systems, a greater proportion of normalized length metrics demonstrate correlation levels above 0.5. Thus, overall, geometric biometrics demonstrate a higher degree of independence.

Table 10: Pairwise Correlation between Geometric Muzzle Biometrics

	NostrilFlareProportionUpperFront	NostrilFlarePointProportionUpperFront	NostrilFlareProportionUpperBack	NostrilFlarePointProportionUpperBack	NostrilFlareProportionLowerBack	NostrilFlarePointProportionLowerBack	NostrilFlareProportionLowerFront	NostrilFlarePointProportionLowerFront	NostrilDepthProportionLinear	NostrilDepthProportionArea	NostrilDepthPointProportion	NostrilHeightProportionLinear	NostrilHeightProportionArea	NostrilHeightPointProportion	NostrilPositionAngle	NostrilMuzzleRatioAreaV1	NostrilMuzzleRatioHeightV1	NostrilMuzzleRatioLength	MouthEye-to-ExtrapOffsetHeight	MouthEye-to-ExtrapOffsetLength	UpperLipRoundnessProportionV1	UpperLipRoundnessPointProportionV1	MuzzleThicknessProportionV1	ChinThicknessProportionV1	ChinThicknessPointProportionV1	Chin-to-LipThicknessProportionV1
NostrilFlareProportionUpperFront	1.00	0.50	-0.03	-0.09	-0.06	-0.11	-0.06	0.25	-0.18	0.15	-0.03	-0.21	0.17	0.30	-0.22	0.03	-0.04	0.28	-0.12	0.12	0.12	0.18	0.09	0.09	0.00	0.09
NostrilFlarePointProportionUpperFront	0.50	1.00	-0.09	-0.39	0.09	0.07	-0.18	0.42	-0.64	-0.06	-0.19	0.02	0.56	0.79	-0.60	0.10	-0.33	0.71	-0.15	0.13	0.11	0.18	0.15	0.14	-0.02	0.07
NostrilFlareProportionUpperBack	-0.03	-0.09	1.00	-0.22	0.03	0.06	-0.02	-0.08	-0.01	-0.07	-0.16	0.11	0.00	0.11	-0.03	0.03	0.02	0.00	0.00	0.05	-0.07	-0.07	-0.12	0.15	0.01	0.21
NostrilFlarePointProportionUpperBack	-0.09	-0.39	-0.22	1.00	0.00	-0.02	0.15	-0.12	0.57	0.30	-0.02	-0.12	-0.34	-0.52	0.47	-0.04	0.31	-0.46	-0.02	-0.02	0.20	-0.15	-0.04	-0.03	0.07	0.03
NostrilFlareProportionLowerBack	-0.06	0.09	0.03	0.00	1.00	0.28	0.04	0.10	-0.07	-0.05	-0.13	0.03	0.05	0.09	-0.09	-0.02	-0.11	0.09	0.09	-0.11	0.15	-0.06	0.03	-0.10	0.04	-0.08
NostrilFlarePointProportionLowerBack	-0.11	0.07	0.06	-0.02	0.28	1.00	-0.07	-0.04	-0.09	-0.14	-0.05	0.15	0.04	0.08	-0.10	0.04	-0.04	0.06	0.02	-0.01	-0.01	-0.14	0.02	0.01	-0.05	0.03
NostrilFlareProportionLowerFront	-0.06	-0.18	-0.02	0.15	0.04	-0.07	1.00	0.00	0.15	-0.18	0.14	-0.09	-0.40	-0.25	0.16	-0.33	-0.11	-0.33	-0.03	0.02	0.28	0.06	0.15	-0.02	-0.02	-0.17
NostrilFlarePointProportionLowerFront	0.25	0.42	-0.08	-0.12	0.10	-0.04	0.00	1.00	-0.22	-0.02	0.02	0.00	0.16	0.39	-0.18	-0.28	-0.32	0.18	-0.01	-0.02	0.14	0.20	0.12	0.05	0.07	-0.01
NostrilDepthProportionLinear	-0.18	-0.64	-0.01	0.57	-0.07	-0.09	0.15	-0.22	1.00	0.53	0.28	-0.13	-0.52	-0.71	0.65	0.13	0.37	-0.69	0.12	-0.14	0.09	-0.15	-0.13	-0.11	0.12	-0.04
NostrilDepthProportionArea	0.15	-0.06	-0.07	0.30	-0.05	-0.14	-0.18	-0.02	0.53	1.00	0.13	-0.39	0.28	-0.14	0.30	0.41	0.47	0.02	-0.03	-0.01	0.04	-0.09	-0.14	0.02	0.15	0.18
NostrilDepthPointProportion	-0.03	-0.19	-0.16	-0.02	-0.13	-0.05	0.14	0.02	0.28	0.13	1.00	-0.27	-0.29	-0.32	0.37	-0.20	0.10	-0.36	0.01	-0.01	-0.04	-0.05	0.10	-0.02	-0.03	-0.20
NostrilHeightProportionLinear	-0.21	0.02	0.11	-0.12	0.03	0.15	-0.09	0.00	-0.13	-0.39	-0.27	1.00	0.18	0.19	-0.28	0.36	-0.35	0.10	0.22	-0.17	-0.20	-0.15	-0.23	0.00	0.07	0.09
NostrilHeightProportionArea	0.17	0.56	0.00	-0.34	0.05	0.04	-0.40	0.16	-0.52	0.28	-0.29	0.18	1.00	0.66	-0.48	0.44	-0.10	0.77	-0.07	0.08	-0.15	0.03	-0.09	0.14	0.05	0.26
NostrilHeightPointProportion	0.30	0.79	0.11	-0.52	0.09	0.08	-0.25	0.39	-0.71	-0.14	-0.32	0.19	0.66	1.00	-0.73	0.20	-0.40	0.78	-0.15	0.17	-0.02	0.17	0.08	0.18	-0.06	0.16
NostrilPositionAngle	-0.22	-0.60	-0.03	0.47	-0.09	-0.10	0.16	-0.18	0.65	0.30	0.37	-0.28	-0.48	-0.73	1.00	-0.19	0.66	-0.76	0.20	-0.19	-0.06	-0.28	-0.39	-0.14	0.14	0.00
NostrilMuzzleRatioAreaV1	0.03	0.10	0.03	-0.04	-0.02	0.04	-0.33	-0.28	0.13	0.41	-0.20	0.36	0.44	0.20	-0.19	1.00	0.18	0.44	0.08	-0.06	-0.17	-0.16	-0.34	0.04	0.13	0.23
NostrilMuzzleRatioHeightV1	-0.04	-0.33	0.02	0.31	-0.11	-0.04	-0.11	-0.32	0.37	0.47	0.10	-0.35	-0.10	-0.40	0.66	0.18	1.00	-0.29	0.19	-0.12	-0.20	-0.33	-0.61	-0.05	0.30	0.45
NostrilMuzzleRatioLength	0.28	0.71	0.00	-0.46	0.09	0.06	-0.33	0.18	-0.69	0.02	-0.36	0.10	0.77	0.78	-0.76	0.44	-0.29	1.00	-0.17	0.17	0.00	0.15	0.08	0.15	-0.01	0.20
MouthEye-to-ExtrapOffsetHeight	-0.12	-0.15	0.00	-0.02	0.09	0.02	-0.03	-0.01	0.12	-0.03	0.01	0.22	-0.07	-0.15	0.20	0.08	0.19	-0.17	1.00	-0.95	-0.32	-0.41	-0.37	-0.54	0.56	-0.15
MouthEye-to-ExtrapOffsetLength	0.12	0.13	0.05	-0.02	-0.11	-0.01	0.02	-0.02	-0.14	-0.01	-0.01	-0.17	0.08	0.17	-0.19	-0.06	-0.12	0.17	-0.95	1.00	0.26	0.40	0.24	0.64	-0.59	0.30
UpperLipRoundnessProportionV1	0.12	0.11	-0.07	0.20	0.15	-0.01	0.28	0.14	0.09	0.04	-0.04	-0.20	-0.15	-0.02	-0.06	-0.17	-0.20	0.00	-0.32	0.26	1.00	0.24	0.41	0.10	-0.17	-0.16
UpperLipRoundnessPointProportionV1	0.18	0.18	-0.07	-0.15	-0.06	-0.14	0.06	0.20	-0.15	-0.09	-0.05	-0.15	0.03	0.17	-0.28	-0.16	-0.33	0.15	-0.41	0.40	0.24	1.00	0.32	0.24	-0.37	-0.02
MuzzleThicknessProportionV1	0.09	0.15	-0.12	-0.04	0.03	0.02	0.15	0.12	-0.13	-0.14	0.10	-0.23	-0.09	0.08	-0.39	-0.34	-0.61	0.08	-0.37	0.24	0.41	0.32	1.00	0.04	-0.35	-0.56
ChinThicknessProportionV1	0.09	0.14	0.15	-0.03	-0.10	0.01	-0.02	0.05	-0.11	0.02	-0.02	0.00	0.14	0.18	-0.14	0.04	-0.05	0.15	-0.54	0.64	0.10	0.24	0.04	1.00	-0.27	0.64
ChinThicknessPointProportionV1	0.00	-0.02	0.01	0.07	0.04	-0.05	-0.02	0.07	0.12	0.15	-0.03	0.07	0.05	-0.06	0.14	0.13	0.30	-0.01	0.56	-0.59	-0.17	-0.37	-0.35	-0.27	1.00	0.14
Chin-to-LipThicknessProportionV1	0.09	0.07	0.21	0.03	-0.08	0.03	-0.17	-0.01	-0.04	0.18	-0.20	0.09	0.26	0.16	0.00	0.23	0.45	0.20	-0.15	0.30	-0.16	-0.02	-0.56	0.64	0.14	1.00

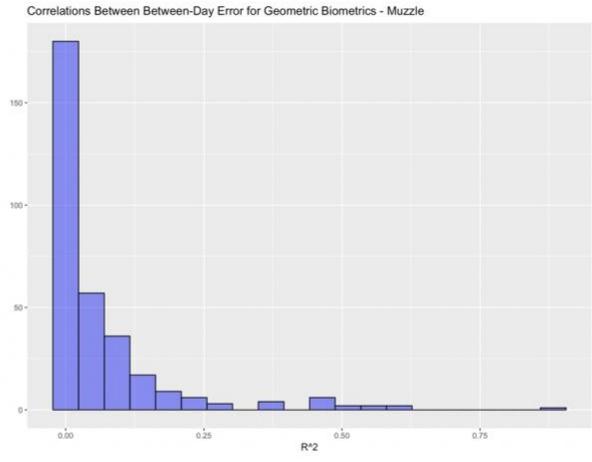
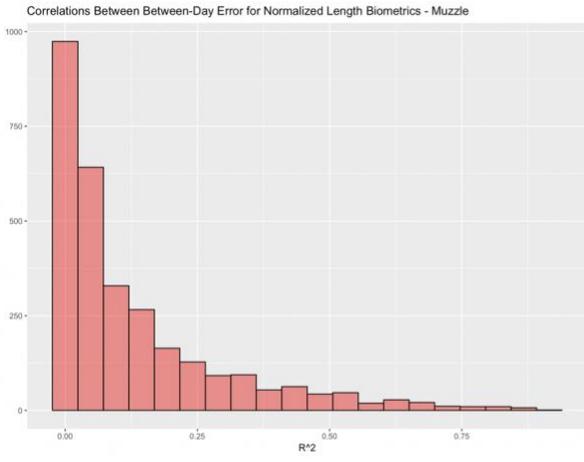


Figure 21: Comparing levels of correlation in error terms between normalized length and geometric muzzle biometrics

Pairwise correlations amongst error terms for the two measurement systems revealed similar results to correlations between observed biometrics, which is not surprising given the high level of between-day error in muzzle biometrics (see Figure 21). Again, distribution density for normalized length measures was clearly shifted right relative to geometric biometrics. The average level of pairwise correlation in error terms for normalized lengths was 0.13, as compared to 0.07 for geometric biometrics, and again normalized length measured demonstrated a higher proportion of metrics above the 0.5 correlation level. Thus, geometric biometrics also appear to have a marginal advantage in error structure for muzzle biometrics as well.

Finally, the proportion of change in biometrics between days that can be attributed to changes in image quality yielded some interesting results (see Figure 22). With respect to attributes related to image scale and rotation, both measurement systems produced negligible correlations to metric error. Both measurement systems yielded average R^2 values of only 0.025, with nearly all metrics producing correlations to changes in image attributes under 10%, indicating both measurement systems are quite robust to these variables for muzzle biometric traits. An interesting exception to this was the Nostril Height Proportion (linear derivation). Further analysis revealed that, increases in nostril height between days corresponded to an increase in nominal face angle (*i.e.* nose abducted towards body). There was also a slight tendency for increases in nostril height to correspond to smaller distances between camera and cow (see Figure 23). This could represent measurement invariance of geometric biometrics to changes in these image attributes, though this seems unlikely, given that this trend manifested in only a single biometric. An alternative explanation that seem perhaps more feasible is that, on days that a photographer was forced to stand closer to a given cow in the feed bunk, some cows became upset, pulling their head back into the head lock (ie – increasing face angle) and snorting. Similarly, both measurement systems

proved quite robust to changes in image attributes related to camera position. Normalized length metrics showed a slightly higher range than geometric biometrics, which corresponded to a higher average R^2 value of 0.05 compared to 0.02 for geometric biometrics, but these differences are likely negligible. While neither measurement system shows a clear advantage in robustness to image attributes, these results do suggest that the high between-day error rate observed in muzzle biometrics are likely attributable to changes in facial expression.

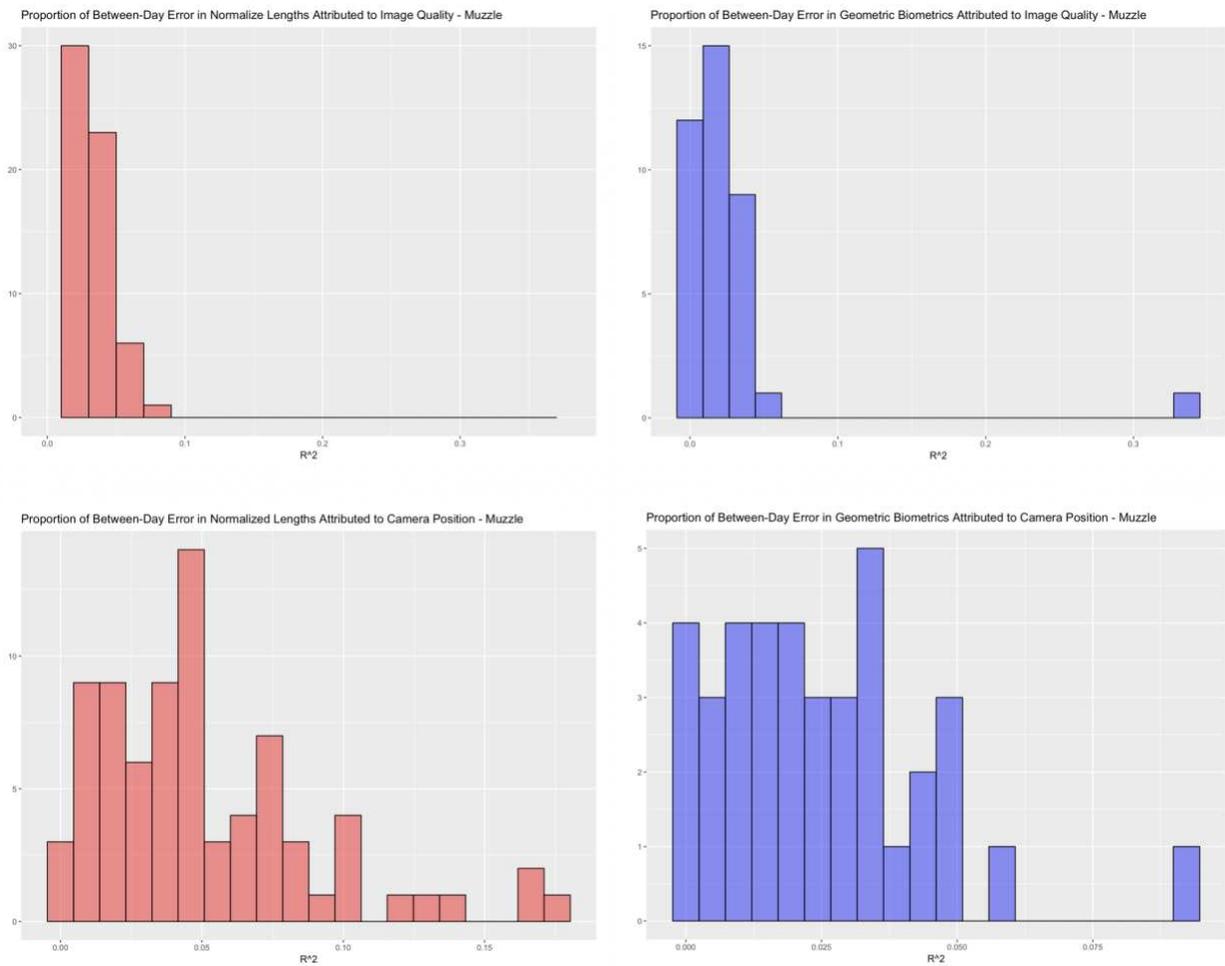


Figure 22: Above – Proportion of error attributed to variations in image scale and rotation;
Below – Proportion of error attributed to variations in camera position

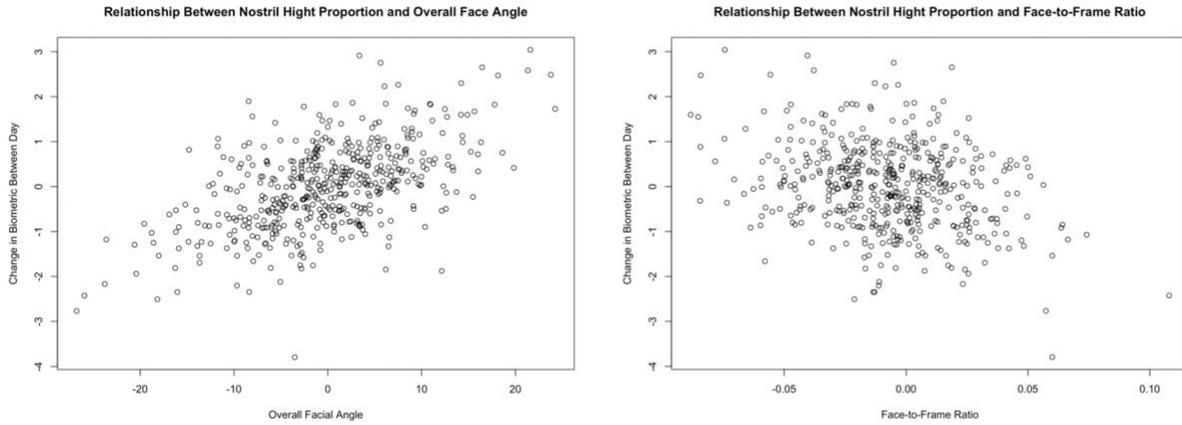


Figure 23: Relationship between changes in Nostril Height Proportion and changes in Face Angle and distance between camera and cow

Results - Topline Biometrics

The shape of the topline is defined by a total of 7 landmark points: C_extrap, D, F_extrap, E_upper, E_mid, E_lower, L_full. From these landmark points, a total of 22 candidate geometric biometrics were extracted. For each geometric biometric, alternative combinations of landmark points were used to compute between one and four alternative versions of the same geometric derivation (see Figure 24). Details of these combinations can be found in the derivations of the geometric biometrics in Appendix A. Additionally, downstream quality checks revealed that two photos – cow 33513, day 3, left side; and 31904, day 1, right side – featured corrupted landmark point extractions, likely as a result of an auto-save error to the coordinate .mat files. As the exact nature of these autosave errors was not clear, these photos were simply dropped from further analyses.

Nares Thickness Proportion
 Nares Roundness Proportion
 Nares Roundness Point Proportion
 Sinus-Midface Rounding Proportion
 Midface-Nose Rounding Proportion
 Midface Divergence Proportion
 Nose Divergence Proportion
 Nares Divergence Proportion
 Midface Inflection Proportion
 Midface Inflection Point Proportion
 Nose Inflection Proportion
 Nose Inflection Point Proportion
 Nares-Topline Length Proportion
 Nose-Topline Length Proportion
 Midface-Topline Length Proportion
 Sinus-Topline Length Proportion
 Nares-Nose Length Proportion
 Upper-Lower Topline Length Proportion
 Sinus-Midface Length Proportion
 Midface-Nose Length Proportion
 Sinus Projection Proportion

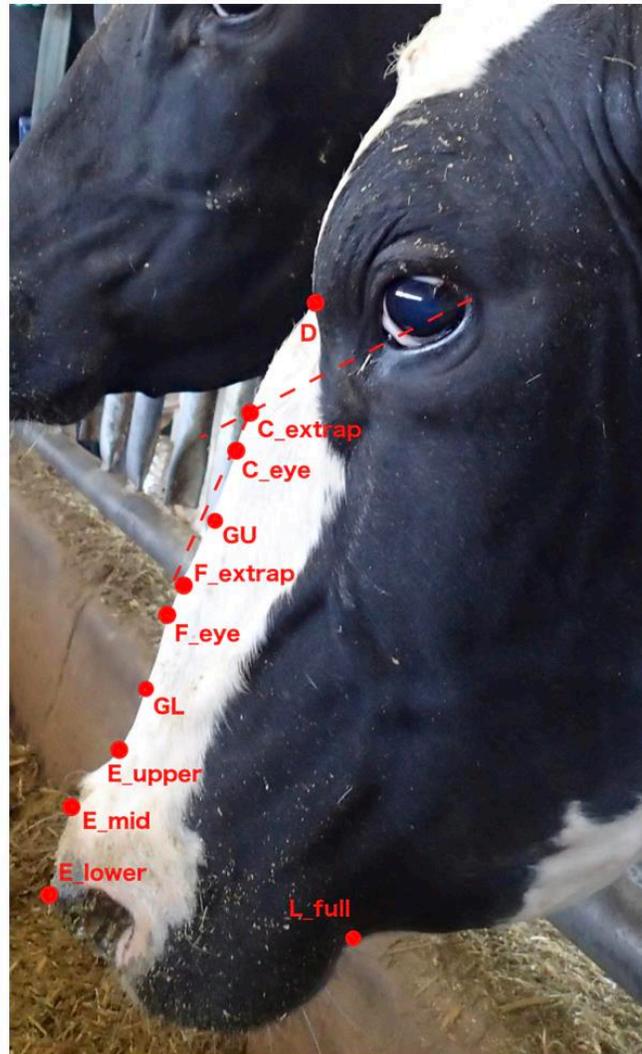


Figure 24: Landmark points and geometric biometrics of the topline (further Appendix A)

First is an assessment of the global behavior of the higher moments of biometrics in each measurement system (see Figure 25). With respect to measures of skew the majority of normalized length measures show a slight tendency to be skewed right (3rd moment), but a fair number also demonstrate strong left tails, whereas geometric biometrics show a clear tendency to positively skewed tails. With respect to measures of kurtosis (4th moment), both normalized length and geometric biometrics show some tendency to produce metric distributions with thickened tails.

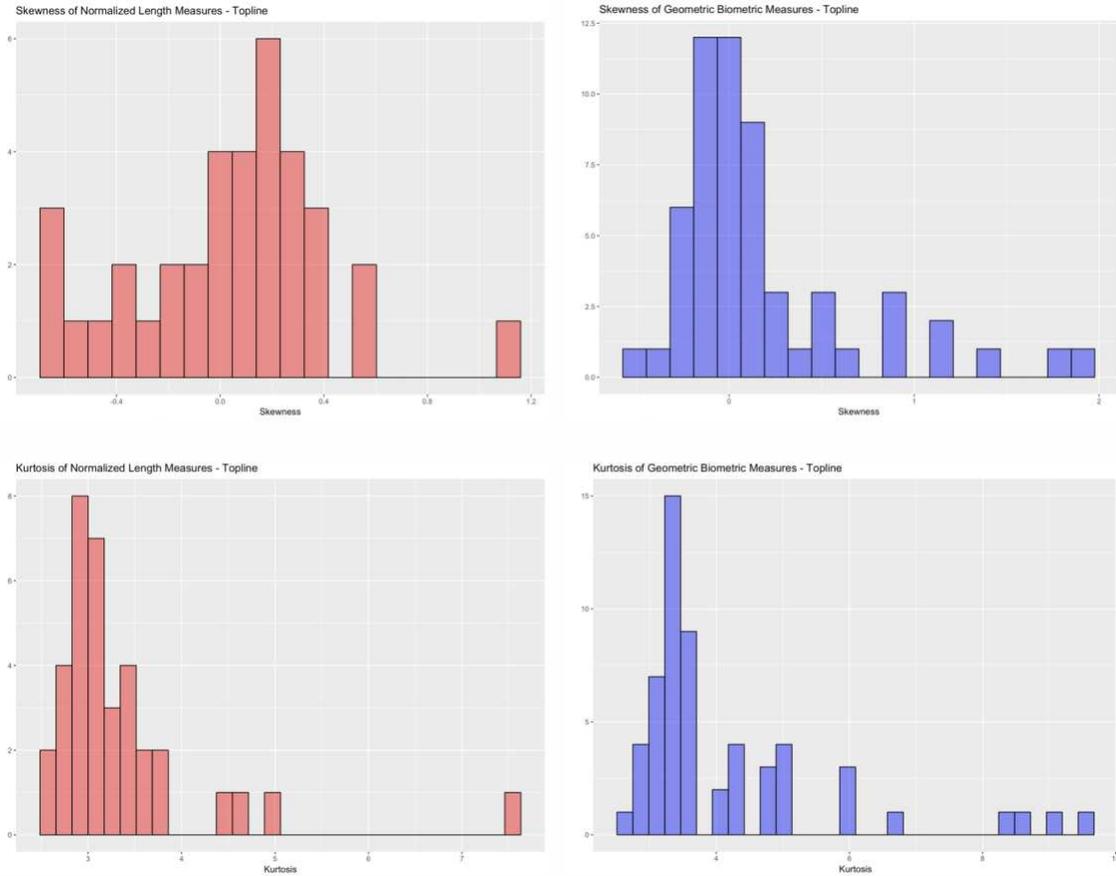


Figure 25: Distribution of 3rd & 4th moments for normalized length & geometric topline biometrics

Second is a comparison of the performance of normalized and geometric biometrics in terms of repeatability (see Figure 26). First is the repeatability of these two measurement systems within-photo, which corresponds to their robustness to errors in points selection. Here geometric biometrics clearly out-perform normalized length measures, with far more density shifted towards the upper end of the repeatability scale. The overall average repeatability for geometric biometrics was 0.43, whereas for normalized length measures it was only 0.30, and a far greater proportion of geometric biometrics exceeded the 0.5 repeatability threshold as compared to normalized lengths. Thus, in terms of robustness to errors in coordinate extraction, geometric biometrics have a clear advantage for topline biometrics.

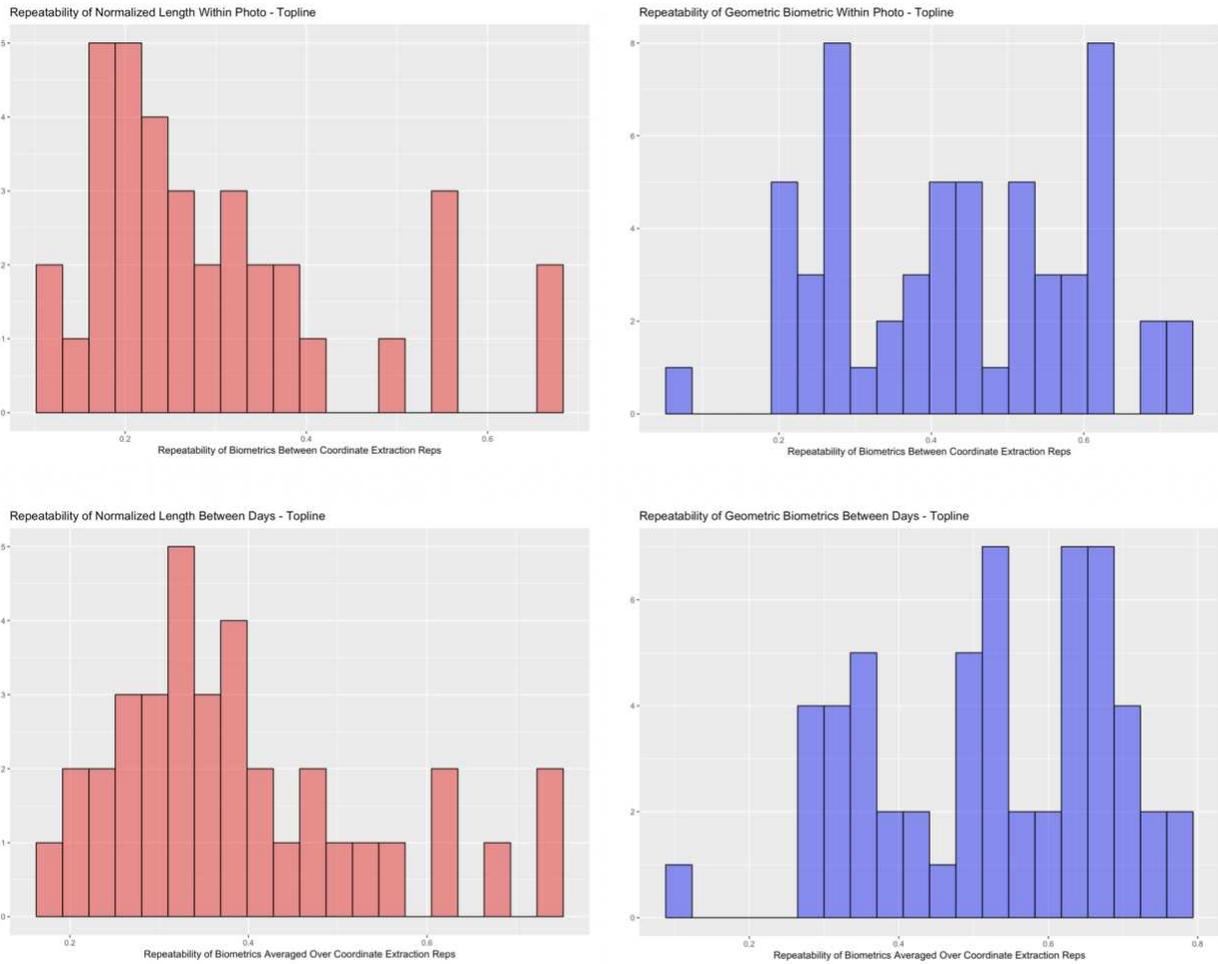


Figure 26: Comparison of Between and Within-Day Repeatability of Normalized and Geometric Topline Biometric Measures

The next concern is the comparative performance of these measurement systems in terms of between-photo repeatability, or overall measurement repeatability (see Figure 26). Here again the distribution density is clearly shifted right for geometric biometrics as compared to normalized length measures. The overall average repeatability for geometric biometrics was 0.52, as compared to only 0.40 for normalized length measures. Here again, a notably larger proportion of geometric biometrics exceeded the 0.5 repeatability threshold as compared to normalized length measures.

Thus, geometric biometrics also demonstrate a clear advantage over normalized length measures with respect to overall measure repeatability for topline metrics.

Third is a deeper exploration of the error structures underlying the geometric biometric measures of the topline. From these results, several overarching trends can be observed. With respect to the error structure, it is clear that some traits suffer primarily from lack of robustness to errors in coordinate selection, particularly the inflexion points, which are observed on quite a fine scale relative to other traits (see Table 11 & 12). Other traits, particularly the length traits, however, reveal a fair share of their residual error to be attributed to variations between images. With perhaps the exception of the nares thickness metric, these represent boney traits. Thus, these results suggest that a notable proportion of error observed in topline traits are likely attributable to variations in out-of-plane face angle.

Compared to the performance of eye and muzzle biometrics, topline biometrics show superior performance in terms of repeatability (see Table 14). Traits related to the divergence of the top line from the true line of the face in particular perform well. Amongst the length metrics there is observed quite a bit of volatility in repeatability results. These results indicate that extraction of landmark coordinates using reference lines proved far more reliable than coordinates simply extracted by visual inspection. For landmark point C, the lowest point of the sinus, use of a reference line proved particularly critical to realizing gains in metric repeatability.

Table 11: Proportion of total variability in topline biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps)

	Version 1	Version 2	Version 3	Version 4
Nares Thickness Proportion	0.360 [0.20,0.69]	0.352 [0.30, 0.40]	0.342 [0.29, 0.39]	
Nares Roundness Proportion	0.344 [0.28, 0.40]			
Nares Roundness Point Proportion	0.426 [0.36, 0.48]			
Sinus-Midface Rounding Proportion	0.302 [0.24, 0.35]	0.318 [0.25, 0.37]	0.291 [0.23, 0.34]	0.315 [0.25, 0.36]
Midface-Nose Rounding Proportion	0.277 [0.22, 0.33]	0.270 [0.21, 0.32]	0.232 [0.17, 0.27]	0.234 [0.18, 0.28]
Midface Divergence Proportion	0.159 [0.12, 0.19]	0.138 [0.11, 0.16]	0.134 [0.10, 0.16]	0.120 [0.09, 0.14]
Nose Divergence Proportion	0.191 [0.15, 0.23]	0.192 [0.15, 0.23]		
Nares Divergence Proportion	0.140 [0.11, 0.17]	0.165 [0.13, 0.20]	0.211 [0.16, 0.25]	0.207 [0.16, 0.25]
Midface Inflection Proportion	0.389 [0.32, 0.45]	0.299 [0.24, 0.35]	0.447 [0.37, 0.52]	0.295 [0.23, 0.35]
Midface Inflection Point Proportion	0.496 [0.44, 0.57]	0.615 [0.53, 0.69]	0.518 [0.43, 0.59]	0.635 [0.55, 0.71]
Nose Inflection Proportion	0.471 [0.39, 0.54]	0.449 [0.36, 0.52]		
Nose Inflection Point Proportion	0.621 [0.53, 0.70]	0.645 [0.56, 0.72]		
Nares-Topline Length Proportion	0.333 [0.28, 0.38]			
Nose-Topline Length Proportion	0.561 [0.48, 0.63]	0.530 [0.45, 0.60]		
Midface-Topline Length Proportion	0.238 [0.19, 0.28]	0.575 [0.49, 0.65]	0.329 [0.27, 0.38]	0.612 [0.53, 0.69]
Sinus-Topline Length Proportion	0.204 [0.16, 0.24]	0.562 [0.47, 0.64]		
Nares-Nose Length Proportion	0.390 [0.33, 0.45]	0.435 [0.37, 0.49]		
Upper-Lower Topline Length Proportion	0.454 [0.38, 0.52]	0.519 [0.43, 0.59]		
Sinus-Midface Length Proportion	0.230 [0.18, 0.27]	0.666 [0.57, 0.74]	0.254 [0.21, 0.29]	0.673 [0.58, 0.75]
Midface-Nose Length Proportion	0.384 [0.31, 0.44]	0.818 [0.74, 0.89]	0.521 [0.44, 0.59]	0.587 [0.51, 0.66]
Sinus Projection Proportion	0.239 [0.18, 0.28]			

Table 12: Proportion of total variability in topline biometrics attributed to error in image acquisition (variance between days/photos)

	Version 1	Version 2	Version 3	Version 4
Nares Thickness Proportion	0.411 [0.34,0.50]	0.374 [0.30, 0.45]	0.377 [0.30, 0.46]	
Nares Roundness Proportion	0.229 [0.17, 0.29]			
Nares Roundness Point Proportion	0.292 [0.22, 0.37]			
Sinus-Midface Rounding Proportion	0.143 [0.09, 0.19]	0.151 [0.10, 0.20]	0.140 [0.09, 0.19]	0.155 [0.11, 0.21]
Midface-Nose Rounding Proportion	0.144 [0.09, 0.19]	0.137 [0.09, 0.18]	0.133 [0.09, 0.17]	0.132 [0.09, 0.17]
Midface Divergence Proportion	0.152 [0.11, 0.19]	0.150 [0.10, 0.19]	0.161 [0.12, 0.20]	0.165 [0.12, 0.920]
Nose Divergence Proportion	0.193 [0.14, 0.24]	0.183 [0.13, 0.23]		
Nares Divergence Proportion	0.256 [0.19, 0.31]	0.213 [0.15, 0.26]	0.178 [0.12, 0.23]	0.183 [0.13, 0.23]
Midface Inflection Proportion	0.166 [0.11, 0.22]	0.169 [0.12, 0.22]	0.148 [0.10, 0.22]	0.198 [0.15, 0.25]
Midface Inflection Point Proportion	0.140 [0.08, 0.21]	0.194 [0.13, 0.28]	0.111 [0.06, 0.17]	0.156 [0.09, 0.23]
Nose Inflection Proportion	0.081 [0.03, 0.13]	0.064 [0.02, 0.11]		
Nose Inflection Point Proportion	0.114 [0.05, 0.18]	0.159 [0.09, 0.24]		
Nares-Topline Length Proportion	0.225 [0.16, 0.29]			
Nose-Topline Length Proportion	0.171 [0.11, 0.25]	0.230 [0.16, 0.31]		
Midface-Topline Length Proportion	0.252 [0.19, 0.31]	0.131 [0.07, 0.20]	0.237 [0.17, 0.30]	0.164 [0.10, 0.24]
Sinus-Topline Length Proportion	0.232 [0.17, 0.29]	0.095 [0.04, 0.16]		
Nares-Nose Length Proportion	0.235 [0.17, 0.30]	0.274 [0.20, 0.35]		
Upper-Lower Topline Length Proportion	0.131 [0.07, 0.19]	0.119 [0.06, 0.18]		
Sinus-Midface Length Proportion	0.313 [0.24, 0.38]	0.064 [0.01, 0.13]	0.320 [0.25, 0.39]	0.084 [0.03, 0.16]
Midface-Nose Length Proportion	0.196 [0.14, 0.26]	0.123 [0.05, 0.21]	0.210 [0.15, 0.29]	0.217 [0.15, 0.30]
Sinus Projection Proportion	0.142 [0.09, 0.18]			

Table 13: Repeatability of Geometric Topline Biometrics from a Single Landmark Coordinate Extraction

	Version 1	Version 2	Version 3	Version 4
Nares Thickness Proportion	0.230 [0.14,0.31]	0.273 [0.19, 0.36]	0.281 [0.20, 0.37]	
Nares Roundness Proportion	0.427 [0.35, 0.52]			
Nares Roundness Point Proportion	0.282 [0.20, 0.37]			
Sinus-Midface Rounding Proportion	0.555 [0.49, 0.64]	0.531 [0.46, 0.62]	0.569 [0.45, 0.62]	0.530 [0.45, 0.62]
Midface-Nose Rounding Proportion	0.579 [0.51, 0.66]	0.593 [0.53, 0.67]	0.635 [0.57, 0.72]	0.635 [0.57, 0.71]
Midface Divergence Proportion	0.690 [0.64, 0.76]	0.712 [0.66, 0.78]	0.706 [0.65, 0.78]	0.715 [0.66, 0.79]
Nose Divergence Proportion	0.616 [0.55, 0.70]	0.626 [0.56, 0.71]		
Nares Divergence Proportion	0.604 [0.54, 0.69]	0.532 [0.46, 0.61]	0.405 [0.33, 0.48]	0.601 [0.55, 0.70]
Midface Inflection Proportion	0.445 [0.37, 0.53]	0.191 [0.12, 0.26]	0.371 [0.30, 0.46]	0.507 [0.43, 0.59]
Midface Inflection Point Proportion	0.363 [0.29, 0.45]	0.388 [0.31, 0.46]	0.482 [0.41, 0.54]	0.209 [0.14, 0.28]
Nose Inflection Proportion	0.448 [0.37, 0.54]	0.487 [0.42, 0.57]		
Nose Inflection Point Proportion	0.265 [0.19, 0.35]	0.196 [0.13, 0.27]		
Nares-Topline Length Proportion	0.442 [0.37, 0.53]			
Nose-Topline Length Proportion	0.267 [0.20, 0.35]	0.241 [0.17, 0.32]		
Midface-Topline Length Proportion	0.510 [0.44, 0.61]	0.295 [0.22, 0.38]	0.434 [0.36, 0.53]	0.223 [0.15, 0.30]
Sinus-Topline Length Proportion	0.564 [0.49, 0.65]	0.343 [0.27, 0.42]		
Nares-Nose Length Proportion	0.374 [0.30, 0.46]	0.290 [0.21, 0.38]		
Upper-Lower Topline Length Proportion	0.414 [0.34, 0.50]	0.362 [0.29, 0.45]		
Sinus-Midface Length Proportion	0.457 [0.38, 0.55]	0.270 [0.20, 0.34]	0.426 [0.35, 0.52]	0.242 [0.17, 0.31]
Midface-Nose Length Proportion	0.420 [0.34, 0.51]	0.06 [0.00, 0.10]	0.269 [0.19, 0.35]	0.196 [0.12, 0.26]
Sinus Projection Proportion	0.619 [0.56, 0.70]			

Table 14: Repeatability of Geometric Topline Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction

	Version 1	Version 2	Version 3	Version 4
Nares Thickness Proportion	0.279 [0.17,0.36]	0.233 [0.23, 0.44]	0.340 [0.24, 0.45]	
Nares Roundness Proportion	0.516 [0.43, 0.62]			
Nares Roundness Point Proportion	0.358 [0.26, 0.47]			
Sinus-Midface Rounding Proportion	0.652 [0.58, 0.74]	0.631 [0.56, 0.72]	0.665 [0.60, 0.76]	0.629 [0.55, 0.71]
Midface-Nose Rounding Proportion	0.669 [0.60, 0.75]	0.683 [0.62, 0.77]	0.717 [0.66, 0.79]	0.717 [0.66, 0.79]
Midface Divergence Proportion	0.750 [0.70, 0.82]	0.767 [0.72, 0.83]	0.757 [0.71, 0.82]	0.763 [0.71, 0.82]
Nose Divergence Proportion	0.685 [0.62, 0.77]	0.695 [0.64, 0.77]		
Nares Divergence Proportion	0.648 [0.58, 0.74]	0.677 [0.61, 0.76]	0.682 [0.62, 0.77]	0.680 [0.62, 0.77]
Midface Inflection Proportion	0.553 [0.47, 0.64]	0.626 [0.55, 0.71]	0.521 [0.43, 0.61]	0.595 [0.51, 0.68]
Midface Inflection Point Proportion	0.482 [0.39, 0.58]	0.276 [0.17, 0.36]	0.498 [0.41, 0.60]	0.308 [0.20, 0.41]
Nose Inflection Proportion	0.587 [0.51, 0.69]	0.629 [0.56, 0.72]		
Nose Inflection Point Proportion	0.384 [0.29, 0.49]	0.288 [0.19, 0.39]		
Nares-Topline Length Proportion	0.531 [0.45, 0.62]			
Nose-Topline Length Proportion	0.372 [0.27, 0.48]	0.332 [0.23, 0.43]		
Midface-Topline Length Proportion	0.580 [0.50, 0.68]	0.413 [0.32, 0.52]	0.522 [0.43, 0.63]	0.322 [0.21, 0.41]
Sinus-Topline Length Proportion	0.630 [0.56, 0.72]	0.480 [0.38, 0.58]		
Nares-Nose Length Proportion	0.465 [0.38, 0.56]	0.370 [0.27, 0.48]		
Upper-Lower Topline Length Proportion	0.540 [0.46, 0.64]	0.502 [0.41, 0.60]		
Sinus-Midface Length Proportion	0.517 [0.43, 0.62]	0.407 [0.30, 0.50]	0.489 [0.40, 0.59]	0.365 [0.26, 0.46]
Midface-Nose Length Proportion	0.519 [0.43, 0.61]	0.096 [0.00, 0.16]	0.366 [0.26, 0.46]	0.278 [0.17, 0.36]
Sinus Projection Proportion	0.702 [0.64, 0.79]			

Fourth is to explore the correlation structures between topline biometrics (see Figure 27). Based on the results of the repeatability analysis, only version 1 metrics, which featured landmark points extracted via reference lines, were retained. All other versions were dropped in order to avoid redundancy between biometrics with the same geometric derivations, which would artificially inflate measures of correlation. In comparing the distributions of pairwise correlations between observed biometrics, the two measurement systems are quite similar in shape, though there are perhaps slightly more geometric biometrics occupying the higher ranges of the scale. The overall average correlation for normalized length measures was 0.07, whereas for geometric biometrics this value was 0.08. Thus both measurement systems show minimal pairwise correlation between observed measures, with neither demonstrating a clear advantage.

Table 15: Pairwise Correlation between Geometric Topline Biometrics

	NaresThicknessProportionV1	NaresRoundnessProportion	NaresRoundnessPointProportion	Sinus-MidfaceRoundingProportionV1	Midface-NoseRoundingProportionV1	MidfaceDivergenceProportionV1	NoseDivergenceProportionV1	NaresDivergenceProportionV1	MidfaceInflectionProportionV1	MidfaceInflectionPointProportionV1	NoseInflectionProportionV1	NoseInflectionPointProportionV1	Nares-ToplineLengthProportion	Nose-ToplineLengthProportionV1	Midface-ToplineLengthProportionV1	Sinus-ToplineLengthProportionV1	Nares-NoseLengthProportionV1	Upper-LowerToplineLengthProportionV1	Sinus-MidfaceLengthProportionV1	Midface-NoseLengthProportionV1	SinusProjectionProportion
NaresThicknessProportionV1	1.00	-0.08	-0.24	0.05	-0.12	0.02	0.14	-0.32	0.06	0.07	-0.05	-0.09	-0.25	0.08	0.11	-0.02	-0.22	-0.19	0.08	0.09	0.02
NaresRoundnessProportion	-0.08	1.00	0.04	0.18	-0.07	-0.05	0.03	-0.33	-0.10	0.00	-0.08	0.04	-0.09	0.07	0.04	-0.02	-0.10	-0.04	-0.01	0.02	0.10
NaresRoundnessPointProportion	-0.24	0.04	1.00	0.01	0.00	-0.16	-0.14	0.13	-0.05	0.11	0.00	0.02	0.06	0.02	0.09	-0.12	0.03	0.08	0.08	0.09	-0.04
Sinus-MidfaceRoundingProportionV1	0.05	0.18	0.01	1.00	-0.32	-0.36	0.02	-0.16	-0.15	0.25	-0.10	-0.13	-0.02	-0.08	0.20	-0.15	0.01	-0.10	0.17	0.24	0.03
Midface-NoseRoundingProportionV1	-0.12	-0.07	0.00	-0.32	1.00	0.48	-0.62	-0.16	0.19	-0.13	0.39	0.35	-0.09	0.17	-0.11	0.08	-0.14	0.04	-0.07	-0.18	-0.01
MidfaceDivergenceProportionV1	0.02	-0.05	-0.16	-0.36	0.48	1.00	0.39	-0.13	0.31	-0.04	0.13	0.14	0.10	0.34	0.19	-0.34	-0.06	0.41	0.32	0.06	-0.34
NoseDivergenceProportionV1	0.14	0.03	-0.14	0.02	-0.62	0.39	1.00	0.06	0.07	0.10	-0.30	-0.25	0.18	0.12	0.28	-0.38	0.10	0.32	0.36	0.24	-0.29
NaresDivergenceProportionV1	-0.32	-0.33	0.13	-0.16	-0.16	-0.13	0.06	1.00	-0.13	0.07	-0.22	-0.12	0.22	-0.14	0.11	-0.15	0.22	0.12	0.11	0.16	-0.28
MidfaceInflectionProportionV1	0.06	-0.10	-0.05	-0.15	0.19	0.31	0.07	-0.13	1.00	-0.16	0.02	-0.01	0.00	0.18	-0.09	0.01	-0.07	0.16	0.06	-0.17	-0.05
MidfaceInflectionPointProportionV1	0.07	0.00	0.11	0.25	-0.13	-0.04	0.10	0.07	-0.16	1.00	0.00	0.08	-0.02	-0.07	0.63	-0.54	0.01	-0.09	0.54	0.68	-0.43
NoseInflectionProportionV1	-0.05	-0.08	0.00	-0.10	0.39	0.13	-0.30	-0.22	0.02	0.00	1.00	0.16	-0.09	0.00	-0.05	0.08	-0.06	-0.10	-0.05	-0.05	0.03
NoseInflectionPointProportionV1	-0.09	0.04	0.02	-0.13	0.35	0.14	-0.25	-0.12	-0.01	0.08	0.16	1.00	0.05	0.00	-0.08	0.05	0.04	0.05	-0.11	-0.08	0.02
Nares-ToplineLengthProportion	-0.25	-0.09	0.06	-0.02	-0.09	0.10	0.18	0.22	0.00	-0.02	-0.09	0.05	1.00	-0.50	-0.23	-0.06	0.93	0.66	-0.07	-0.05	-0.15
Nose-ToplineLengthProportionV1	0.08	0.07	0.02	-0.08	0.17	0.34	0.12	-0.14	0.18	-0.07	0.00	0.00	-0.50	1.00	0.27	-0.38	-0.77	0.32	0.33	-0.11	-0.32
Midface-ToplineLengthProportionV1	0.11	0.04	0.09	0.20	-0.11	0.19	0.28	0.11	-0.09	0.63	-0.05	-0.08	-0.23	0.27	1.00	-0.91	-0.28	-0.02	0.91	0.93	-0.74
Sinus-ToplineLengthProportionV1	-0.02	-0.02	-0.12	-0.15	0.08	-0.34	-0.38	-0.15	0.01	-0.54	0.08	0.05	-0.06	-0.38	-0.91	1.00	0.11	-0.40	-0.93	-0.78	0.86
Nares-NoseLengthProportionV1	-0.22	-0.10	0.03	0.01	-0.14	-0.06	0.10	0.22	-0.07	0.01	-0.06	0.04	0.93	-0.77	-0.28	0.11	1.00	0.35	-0.19	0.00	0.03
Upper-LowerToplineLengthProportionV1	-0.19	-0.04	0.08	-0.10	0.04	0.41	0.32	0.12	0.16	-0.09	-0.10	0.05	0.66	0.32	-0.02	-0.40	0.35	1.00	0.22	-0.15	-0.44
Sinus-MidfaceLengthProportionV1	0.08	-0.01	0.08	0.17	-0.07	0.32	0.36	0.11	0.06	0.54	-0.05	-0.11	-0.07	0.33	0.91	-0.93	-0.19	0.22	1.00	0.80	-0.80
Midface-NoseLengthProportionV1	0.09	0.02	0.09	0.24	-0.18	0.06	0.24	0.16	-0.17	0.68	-0.05	-0.08	-0.05	-0.11	0.93	-0.78	0.00	-0.15	0.80	1.00	-0.63
SinusProjectionProportion	0.02	0.10	-0.04	0.03	-0.01	-0.34	-0.29	-0.28	-0.05	-0.43	0.03	0.02	-0.15	-0.32	-0.74	0.86	0.03	-0.44	-0.80	-0.63	1.00

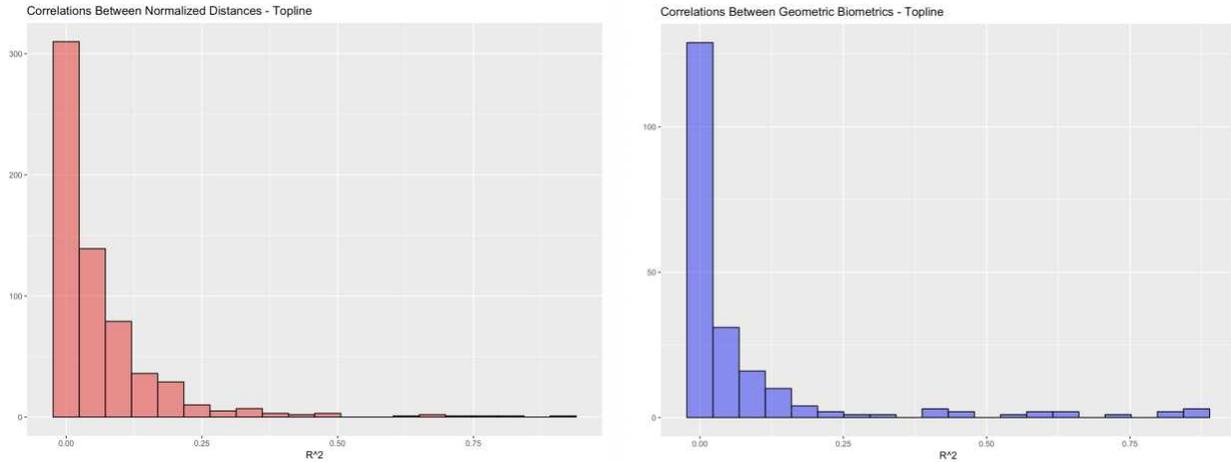


Figure 27: Comparing level of correlation between pairwise combinations of normalized length and geometric topline biometrics

Pairwise correlations amongst error terms for the two measurement systems revealed similar results to correlations between observed biometrics (see Figure 28). The distribution densities for the two measurement systems are visually quite similar, though it should be noted that geometric biometrics occupied a slightly higher scale. The overall average correlation between error terms for geometric biometrics was 0.08, whereas for normalized length measures this value was only 0.04. Thus, normalized length might hold a slight advantage over geometric biometrics in terms of error structure, though for both systems the level of correlation amongst measurement error is effective negligible for most biometrics.

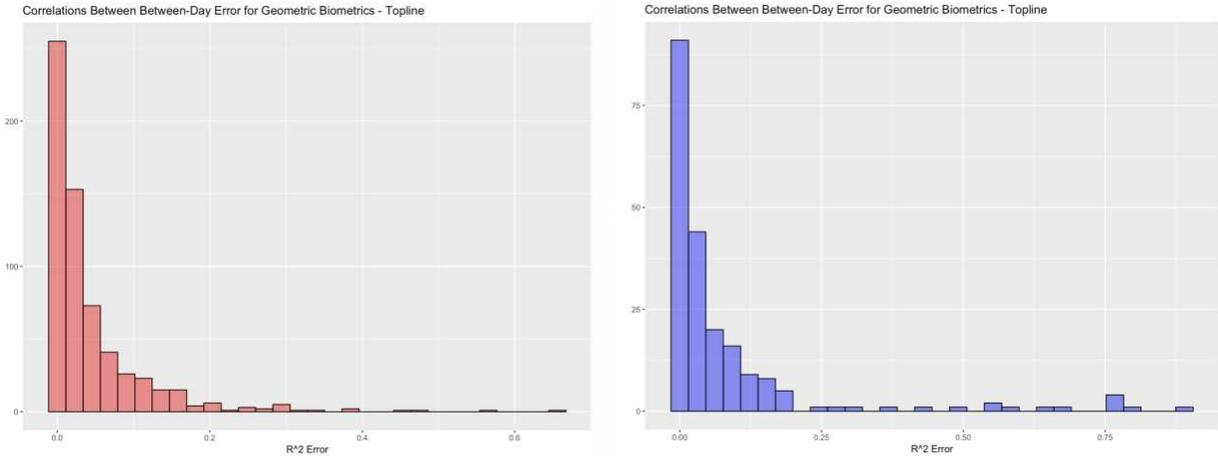


Figure 28: Comparing levels of correlation in error terms between normalized length and geometric topline biometrics

Finally, exploring the proportion of change in biometrics between days that can be linked to changes in image attributes revealed both measurement systems to be quite resistant to variations in image quality for topline biometrics (see Figure 29). With respect to attributes related to image scale and rotation, both measurement systems produced negligible correlations to metric error. Both measurement systems yielded average R^2 values of only 0.01, with all metrics producing correlations to changes in image attributes under 5%. With respect to attributes related to camera position, both measurement systems again produced negligible correlations to metric error. Both measurement systems yielded average R^2 values of only 0.02, with all metrics producing correlations to changes in image attributes under 5%. At these magnitudes, these R^2 likely represent little more than casual correlation, so it cannot be said that either metric system demonstrates a clear advantage over the other with respect to resistance to variability in image quality for topline biometrics.

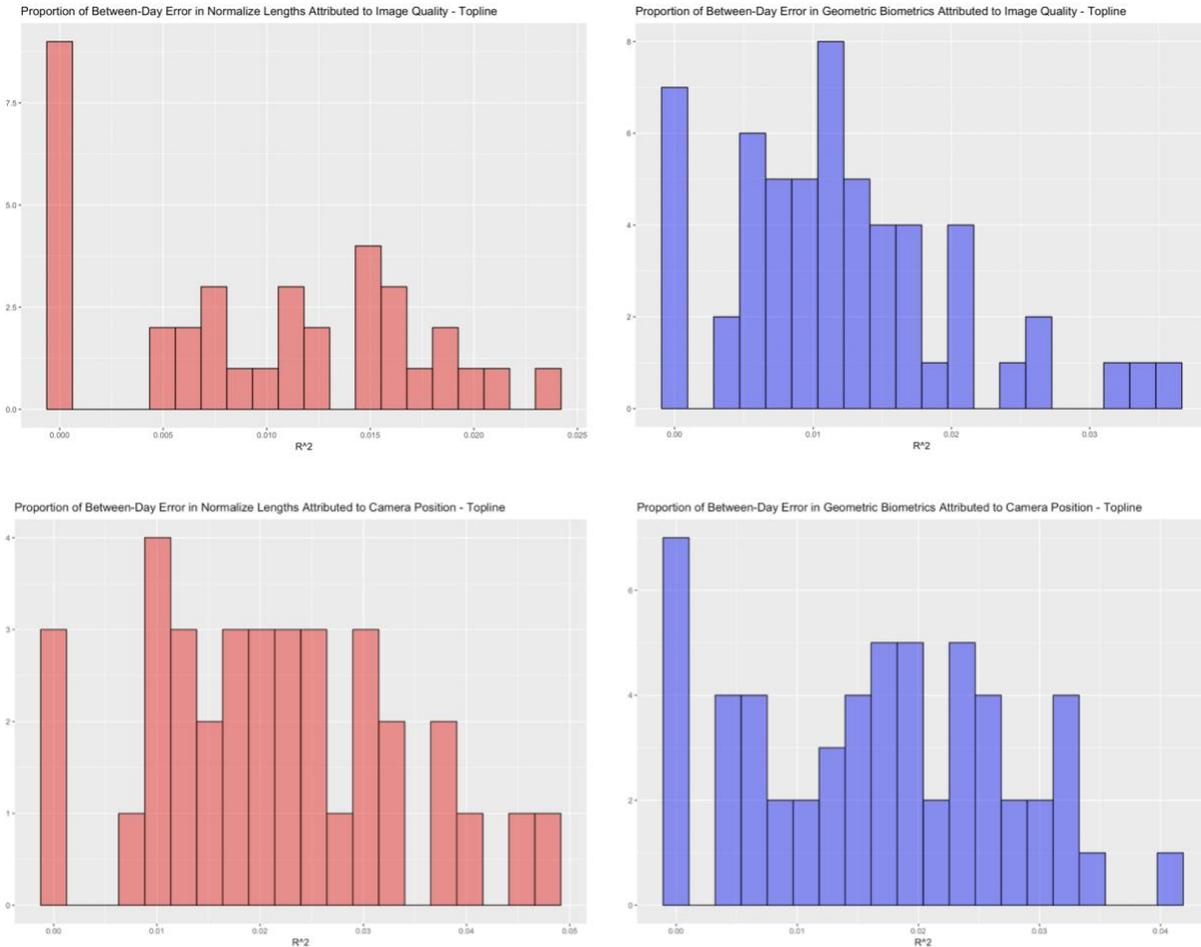


Figure 29: Above – Proportion of error attributed to variations in image scale and rotation;
 Below – Proportion of error attributed to variations in camera position

Results – Forehead & Jaw Biometrics

The shape of the forehead is defined by a total of 14 landmark points: D, S_eye, R, T_slope, T_intercept, T_back, U, M, P, W, X, Y, Z. From these points, a total of 43 candidate geometric biometrics were extracted. Due to the relatively high number of alternative landmark points investigated for comparative efficiency, as many as 42 alternative versions of the same geometric derivation were computed (see Figure 30). Details of these combinations can be found in the derivations of the geometric biometrics in Appendix A. Additionally, downstream quality checks

revealed corruption in an additional photo file – cow 1371, right side, day one – which, like the corrupted files detected after topline coordinate extractions, was also removed from subsequent analyses.

- Canthus Length Proportion
- Canthus Depth Proportion – Depth
- Canthus Depth Proportion – Length
- Chin Length Proportion
- Cheek Nose Size Proportion
- Cranio-Topline Length Ratio
- Canthus Width-to-Height Ratio
- Eye-Cranio Size Ratio
- Eye-Forehead Size Ratio – Linear
- Eye-Forehead Size Ratio – Poly
- Eye Orbital Height-to-Length Ratio
- Eye Orbital-Eye Height Ratio
- Eye Orbital Projection Proportion
- Eye Orbital Roundness Proportion
- Eye Orbital Roundness Point Proportion
- Eye Orbital Thickness Proportion – Poly
- Eye Sinus Size Ratio – Linear
- Eye Sinus Size Ratio – Poly
- Eye-Topline Size Ratio – Linear
- Eye-Topline Size Ratio – Poly
- Forehead-Eye Angle – Slope
- Forehead-Jaw Angle – Slope
- Forehead-Poll Length Ratio
- Forehead-Topline Angle – Slope
- Forehead-Topline Length Ratio
- Forehead Temple Ratio
- Forehead Width-to-Length Ratio
- Forehead-Zygomatic Angle
- Jaw Angle – Slope
- Jowl-Jaw Length Proportion
- Jaw Length Proportion
- Jaw-Midface Size Ratio
- Muzzle Size Proportion
- Midface thickness Proportion
- Midface thickness Proportion
- Nasion Thickness Proportion
- Overall Eye Angle – Angle
- Overall Eye Angle – Slope
- Overall Eye Size
- Poll Depth Proportion – Height
- Poll Depth Proportion – Length
- Poll Height Proportion
- Poll Height Point Proportion

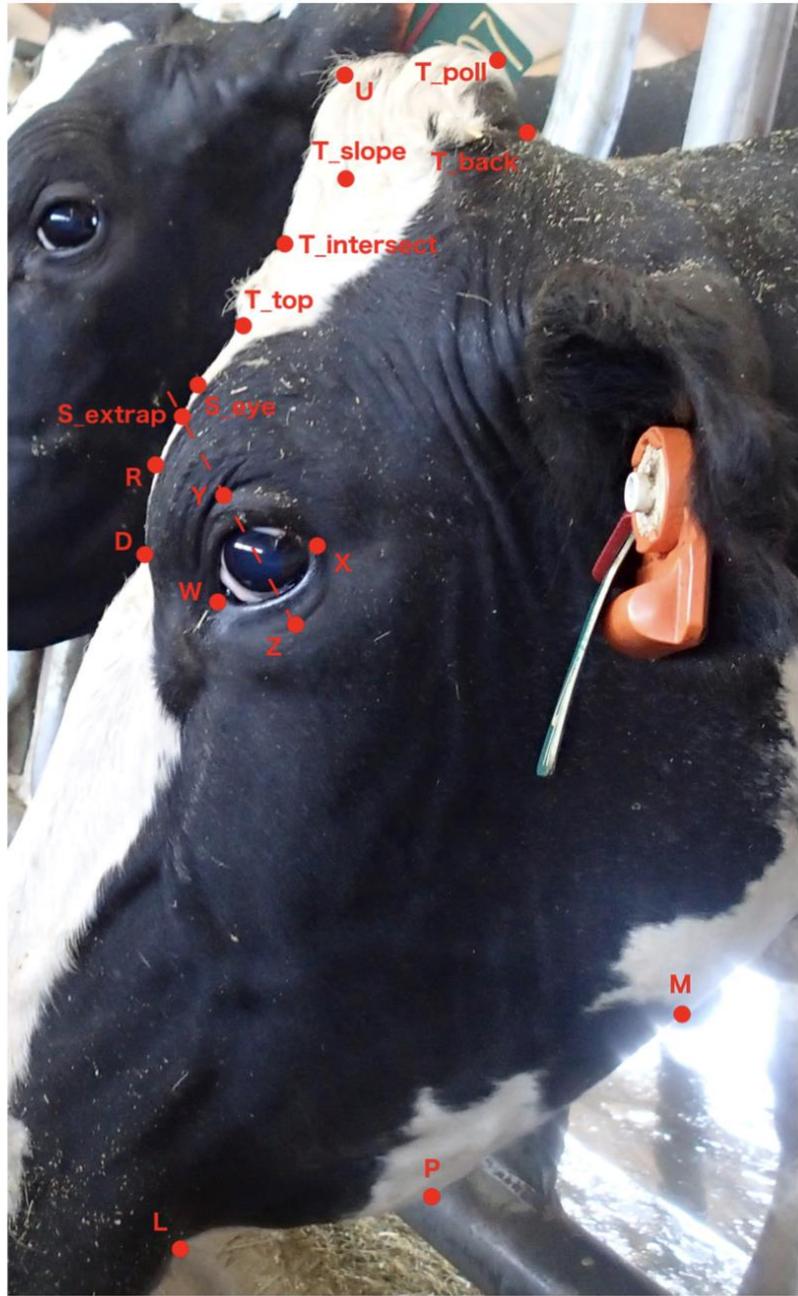


Figure 30: Landmark points and geometric biometrics of the forehead & jaw (further Appendix A)

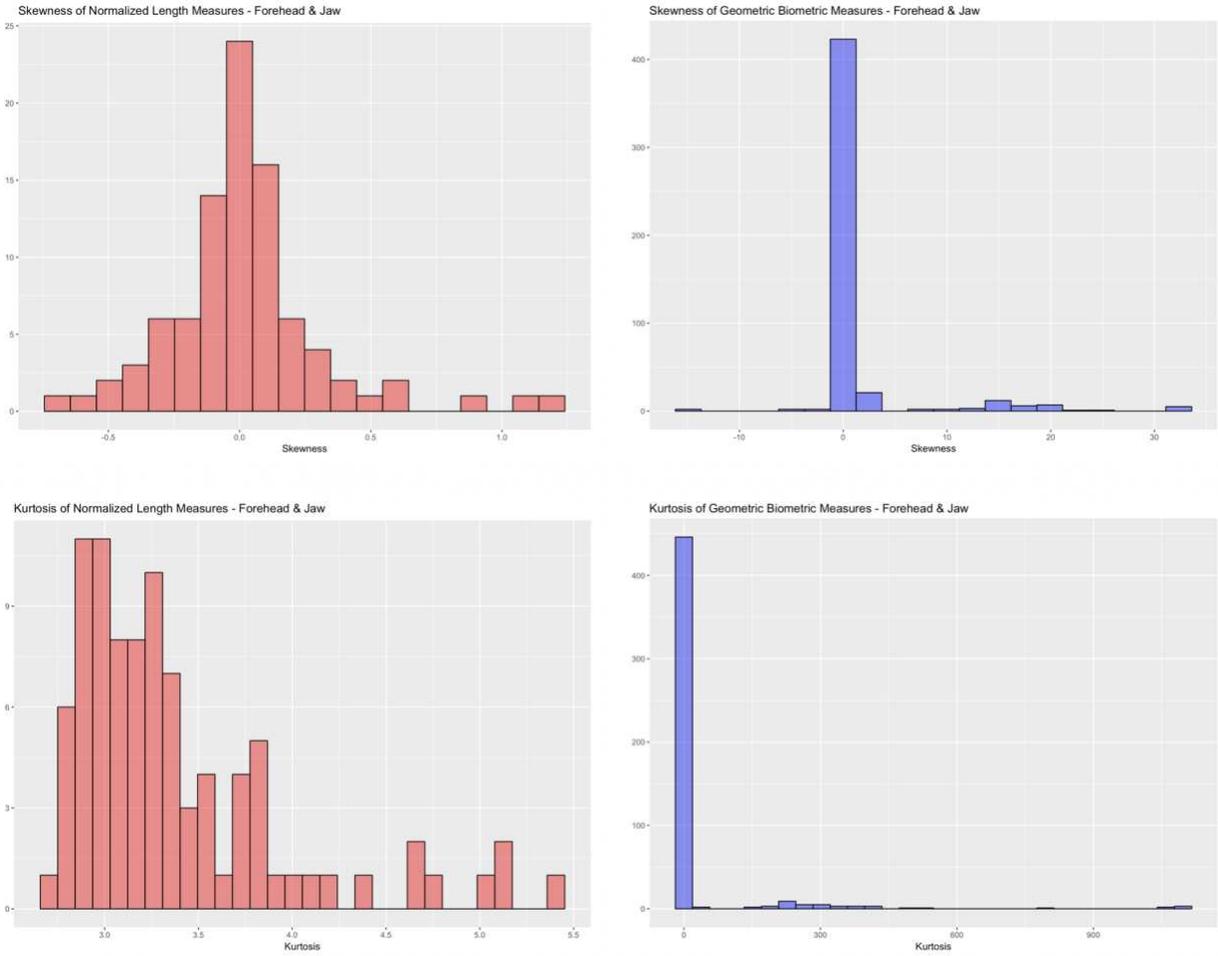


Figure 31: Distribution of 3rd and 4th moments for normalized length and geometric forehead and jaw biometrics

First is an assessment of the global behavior of the higher moments of biometrics in each measurement system (see Figure 31). With respect to measures of skew (3rd moment), normalized length measures show as much tendency to be right skewed as left skewed. Geometric biometrics, on the other hand, had a tendency to produce more extreme values of skew, particularly in the positive direction. With respect to measures of kurtosis (4th moment), normalized length measured showed a slight tendency to produce metric with thick tails, but again geometric biometrics produced comparatively quite extreme values of kurtosis.

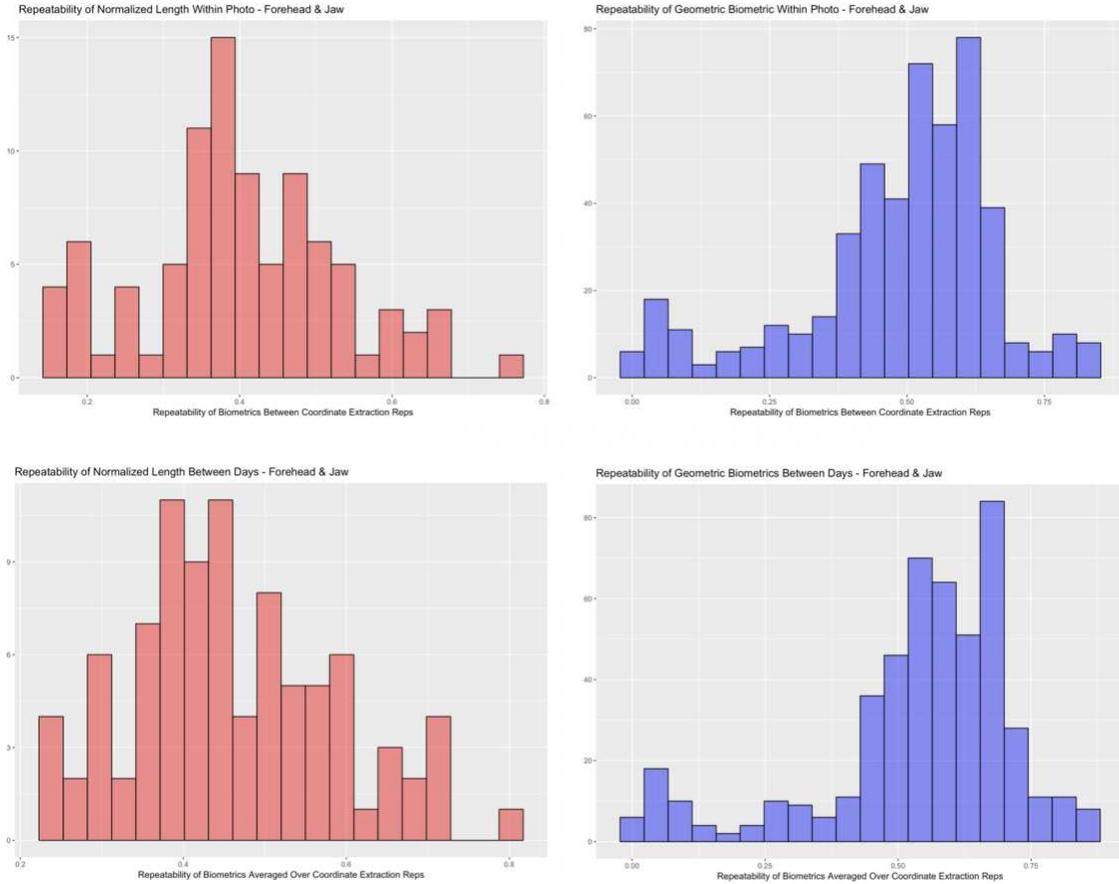


Figure 32: Comparison of Between and Within-Day Repeatability of Normalized and Geometric Forehead and Jaw Biometric Measures

Second is a comparison of the performance of normalized and geometric biometrics in terms of repeatability (see Figure 32). First is the repeatability of these two measurement systems within-photo, which corresponds to their robustness to errors in points selection. Here geometric biometrics clearly out-perform normalized length measures, with far more density shifted towards the upper end of the repeatability scale, and achieving a higher upper range. The overall average repeatability for geometric biometrics was 0.49, whereas for normalized length measures it was only 0.40, and a far greater proportion of geometric biometrics exceeded the 0.5 repeatability threshold as compared to normalized lengths. Thus, in terms of robustness to errors in coordinate extraction, geometric biometrics have a clear advantage for jaw and topline biometrics.

Third is the comparative performance of these measurement systems in terms of between-photo repeatability, or overall measurement repeatability (see Figure 32). Here again the distribution density is clearly shifted right for geometric biometrics as compared to normalized length measures and achieves a higher upper range. The overall average repeatability for geometric biometrics 0.54, as compared to only 0.46 for normalized length measures. Here again, a notably larger proportion of geometric biometrics exceeded the 0.5 repeatability threshold as compared to normalized length measures. Another interesting trend to note is that, for both normalized length and geometric biometrics, the distributions for between and within photo repeatability are quite similar. Thus, geometric biometrics also demonstrate a clear advantage over normalized length measures with respect to overall measure repeatability for forehead and jaw metrics.

Fourth is a deeper exploration of the error structures underlying the geometric biometric measures of the forehead and jaw (see Tables 16 & 17). Based on the relative proportion of errors, it is clear that some traits are predominantly hindered by error in coordinate selection, particularly those traits that occur on a smaller scale relative the overall size of the face. But for many traits, a significant proportion of variability is instead coming from between day error. As with topline, these are boney traits that likely cannot be influenced by variations in facial expression. And again, as with topline traits, the metrics showing the highest proportion of between-day error are those related to relative length measures. This may again suggest that variations in out-of-plane angle of the face may be to blame for this noise.

Table 16: Proportion of total variability in topline biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps) – Part A

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Canthus Length Proportion	0.644 [0.56, 0.72]	0.709 [0.62, 0.79]	0.676 [0.59, 0.76]	0.696 [0.61, 0.77]	0.677 [0.60, 0.76]			
Canthus Depth Proportion – Depth	0.495 [0.42, 0.56]	0.531 [0.45, 0.60]	0.544 [0.47, 0.62]	0.490 [0.42, 0.56]	0.530 [0.45, 0.60]	0.544 [0.47, 0.61]		
Canthus Depth Proportion – Length	0.550 [0.47, 0.62]	0.574 [0.50, 0.65]	0.629 [0.55, 0.71]	0.541 [0.46, 0.61]	0.576 [0.50, 0.65]			
Chin Length Proportion	0.295 [0.24, 0.34]	0.345 [0.29, 0.40]	0.288 [0.23, 0.34]	0.176 [0.14, 0.21]	0.304 [0.25, 0.35]	0.187 [0.14, 0.22]		
Cheek Nose Size Proportion	0.283 [0.23, 0.34]	0.290 [0.23, 0.34]	0.286 [0.23, 0.33]	0.290 [0.23, 0.34]	0.280 [0.22, 0.33]	0.285 [0.23, 0.33]	0.352 [0.29, 0.41]	0.349 [0.28, 0.40]
	0.350 [0.29, 0.41]	0.344 [0.28, 0.44]	0.343 [0.28, 0.40]	0.338 [0.27, 0.39]	0.256 [0.20, 0.30]	0.196 [0.15, 0.23]	0.257 [0.21, 0.30]	0.195 [0.15, 0.23]
	0.254 [0.20, 0.30]	0.193 [0.15, 0.22]						
Cranio-Topline Length Ratio	0.102 [0.08, 0.12]	0.186 [0.15, 0.21]	0.113 [0.09, 0.13]	0.071 [0.06, 0.08]	0.329 [0.28, 0.38]	0.059 [0.05, 0.07]	0.154 [0.13, 0.18]	0.232 [0.19, 0.27]
	0.155 [0.13, 0.18]	0.112 [0.09, 0.13]	0.374 [0.31, 0.43]	0.108 [0.09, 0.13]				
Canthus Width-to-Height Ratio	0.534 [0.45, 0.61]	0.630 [0.55, 0.71]	1.00 [1.00, 1.00]	0.536 [0.46, 0.61]	0.624 [0.54, 0.70]			
Eye-Cranio Size Ratio	0.054 [0.04, 0.06]	0.055 [0.04, 0.07]	0.050 [0.04, 0.06]	0.057 [0.04, 0.07]	0.054 [0.04, 0.06]	0.073 [0.06, 0.09]	0.077 [0.06, 0.09]	0.073 [0.06, 0.09]
	0.077 [0.06, 0.09]	0.074 [0.06, 0.09]	0.098 [0.07, 0.12]	0.106 [0.08, 0.13]	0.106 [0.08, 0.12]	0.106 [0.08, 0.12]	0.100 [0.08, 0.12]	0.056 [0.04, 0.07]
	0.057 [0.04, 0.07]	0.051 [0.04, 0.06]	0.057 [0.04, 0.07]	0.056 [0.04, 0.07]	0.075 [0.06, 0.09]	0.079 [0.06, 0.09]	0.073 [0.06, 0.09]	0.078 [0.06, 0.09]
	0.077 [0.06, 0.09]	0.100 [0.08, 0.12]	0.107 [0.08, 0.13]	0.106 [0.08, 0.12]	0.106 [0.08, 0.12]	0.102 [0.08, 0.12]		
Eye-Forehead Size Ratio – Linear	0.122 [0.09, 0.14]	0.138 [0.11, 0.16]	0.163 [0.13, 0.19]	0.131 [0.10, 0.15]	0.132 [0.10, 0.16]			
Eye-Forehead Size Ratio – Poly	0.099 [0.07, 0.12]	0.101 [0.08, 0.12]	0.098 [0.08, 0.12]	0.101 [0.08, 0.12]	0.099 [0.08, 0.12]	0.109 [0.08, 0.13]	0.112 [0.09, 0.13]	0.109 [0.08, 0.13]
	0.111 [0.08, 0.13]	0.111 [0.08, 0.13]	0.122 [0.09, 0.14]	0.126 [0.10, 0.15]	0.125 [0.10, 0.15]	0.126 [0.10, 0.15]	0.123 [0.09, 0.15]	0.104 [0.08, 0.12]
	0.106 [0.08, 0.13]	0.103 [0.08, 0.12]	0.105 [0.08, 0.12]	0.105 [0.08, 0.12]	0.115 [0.09, 0.14]	0.117 [0.09, 0.14]	0.113 [0.09, 0.14]	0.116 [0.09, 0.14]
	0.117 [0.09, 0.14]	0.128 [0.10, 0.15]	0.131 [0.10, 0.15]	0.129 [0.10, 0.15]	0.130 [0.10, 0.15]	0.129 [0.10, 0.15]		
Eye Orbital Height-to-Length Ratio	0.131 [0.10, 0.15]	0.269 [0.22, 0.31]	0.262 [0.22, 0.30]	0.281 [0.23, 0.32]	0.267 [0.22, 0.31]		0.288 [0.24, 0.33]	0.189 [0.14, 0.22]
	0.336 [0.28, 0.39]	0.367 [0.31, 0.42]	0.331 [0.28, 0.38]	0.321 [0.27, 0.37]	0.366 [0.30, 0.42]	0.312 [0.26, 0.36]		
Eye Orbital-Eye Height Ratio	0.064 [0.05, 0.08]	0.100 [0.08, 0.12]	0.155 [0.12, 0.18]	0.896 [0.85, 0.98]	0.104 [0.08, 0.12]	0.164 [0.13, 0.19]	0.877 [0.83, 0.96]	0.103 [0.08, 0.12]
	0.121 [0.10, 0.14]	0.154 [0.12, 0.18]	0.879 [0.83, 0.96]	0.099 [0.08, 0.12]	0.151 [0.12, 0.18]	0.862 [0.81, 0.95]	0.064 [0.05, 0.08]	0.100 [0.08, 0.12]
	0.155 [0.12, 0.18]	0.909 [0.87, 1.00]	0.103 [0.08, 0.12]	0.164 [0.13, 0.19]	0.893 [0.85, 0.98]	0.111 [0.08, 0.13]	0.120 [0.10, 0.14]	0.155 [0.12, 0.18]
	0.895 [0.86, 0.98]	0.098 [0.08, 0.12]	0.152 [0.12, 0.18]	0.879 [0.83, 0.96]	0.064 [0.05, 0.08]	0.099 [0.08, 0.12]	0.154 [0.12, 0.18]	0.909 [0.87, 1.00]
	0.103 [0.08, 0.12]	0.163 [0.13, 0.19]	0.893 [0.85, 0.98]	0.112 [0.09, 0.13]	0.120 [0.10, 0.14]	0.154 [0.12, 0.18]	0.896 [0.86, 0.98]	0.098 [0.08, 0.12]
	0.151 [0.12, 0.18]	0.880 [0.83, 0.96]						
Eye Orbital Projection Proportion	0.208 [0.16, 0.24]	0.227 [0.18, 0.27]	0.214 [0.17, 0.25]	0.210 [0.16, 0.25]	0.223 [0.17, 0.26]	0.217 [0.17, 0.25]	0.183 [0.14, 0.22]	0.189 [0.15, 0.22]
	0.195 [0.15, 0.23]							
Eye Orbital Roundness Proportion	0.380 [0.32, 0.43]	0.550 [0.48, 0.62]						
Eye Orbital Roundness Point Proportion	0.479 [0.41, 0.54]	0.445 [0.37, 0.51]						
Eye Orbital Thickness Proportion – Poly	0.213 [0.16, 0.25]	0.243 [0.19, 0.28]	0.326 [0.26, 0.38]	0.219 [0.17, 0.26]	0.287 [0.23, 0.33]	0.343 [0.27, 0.40]	0.380 [0.31, 0.44]	0.427 [0.35, 0.49]
	0.360 [0.29, 0.42]	0.375 [0.30, 0.43]						
Eye Sinus Size Ratio – Linear	0.237 [0.18, 0.28]	0.230 [0.18, 0.27]	0.216 [0.17, 0.25]	0.255 [0.20, 0.30]	0.219 [0.17, 0.26]	0.517 [0.43, 0.59]	0.546 [0.46, 0.62]	0.553 [0.46, 0.63]
	0.519 [0.43, 0.59]	0.552 [0.47, 0.63]						
Eye Sinus Size Ratio – Poly	0.096 [0.07, 0.11]	0.101 [0.08, 0.12]	0.103 [0.08, 0.12]	0.104 [0.08, 0.12]	0.100 [0.07, 0.12]	0.094 [0.07, 0.11]	0.100 [0.07, 0.12]	0.102 [0.08, 0.12]
	0.103 [0.08, 0.12]	0.099 [0.07, 0.12]	0.090 [0.07, 0.11]	0.096 [0.07, 0.11]	0.099 [0.07, 0.12]	0.100 [0.07, 0.12]	0.096 [0.07, 0.11]	0.199 [0.11, 0.21]
	0.201 [0.15, 0.24]	0.206 [0.16, 0.24]	0.207 [0.16, 0.24]	0.200 [0.15, 0.24]	0.195 [0.15, 0.23]	0.198 [0.15, 0.23]	0.205 [0.16, 0.24]	0.204 [0.16, 0.24]
	0.199 [0.15, 0.23]	0.192 [0.15, 0.23]	0.196 [0.15, 0.23]	0.205 [0.16, 0.24]	0.203 [0.16, 0.24]	0.198 [0.15, 0.23]		
Eye-Topline Size Ratio – Linear	0.067 [0.05, 0.08]	0.147 [0.11, 0.17]	0.272 [0.21, 0.32]	0.098 [0.07, 0.12]	0.130 [0.09, 0.15]			
Eye-Topline Size Ratio – Poly	0.045 [0.03, 0.05]	0.069 [0.05, 0.08]	0.050 [0.04, 0.06]	0.060 [0.04, 0.07]	0.053 [0.04, 0.06]	0.058 [0.04, 0.07]	0.081 [0.06, 0.10]	0.066 [0.05, 0.08]
	0.071 [0.05, 0.08]	0.067 [0.05, 0.08]	0.074 [0.06, 0.09]	0.097 [0.07, 0.11]	0.092 [0.07, 0.11]	0.088 [0.07, 0.10]	0.082 [0.06, 0.10]	

Table 17: Proportion of total variability in topline biometrics attributed to error in landmark point coordinate extract (variance between coordinate reps) – Part B

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Forehead-Eye Angle – Slope	0.333 [0.27, 0.39]	0.351 [0.29, 0.41]	0.273 [0.22, 0.32]	0.349 [0.28, 0.40]	0.355 [0.29, 0.41]	0.274 [0.22, 0.32]	0.385 [0.31, 0.44]	0.395 [0.33, 0.45]
	0.289 [0.23, 0.34]	0.400 [0.33, 0.46]	0.399 [0.33, 0.46]	0.291 [0.23, 0.34]	0.374 [0.30, 0.43]	0.383 [0.32, 0.44]	0.276 [0.22, 0.32]	0.388 [0.32, 0.45]
Forehead-Jaw Angle – Slope	0.386 [0.32, 0.44]	0.278 [0.22, 0.32]	0.348 [0.28, 0.40]	0.368 [0.30, 0.42]	0.286 [0.23, 0.33]	0.365 [0.30, 0.42]	0.372 [0.31, 0.43]	0.288 [0.23, 0.34]
	0.192 [0.15, 0.23]	0.229 [0.18, 0.27]	0.271 [0.21, 0.32]	0.206 [0.16, 0.24]	0.233 [0.18, 0.27]	0.268 [0.21, 0.31]	0.198 [0.15, 0.23]	0.230 [0.18, 0.27]
Forehead-Poll Length Ratio	0.289 [0.24, 0.34]	0.306 [0.24, 0.36]	0.285 [0.22, 0.33]	0.251 [0.20, 0.29]	0.758 [0.67, 0.84]	0.263 [0.21, 0.31]	0.307 [0.24, 0.36]	0.324 [0.26, 0.38]
	0.305 [0.24, 0.36]	0.280 [0.22, 0.33]	0.944 [0.91, 1.00]	0.289 [0.23, 0.34]				
Forehead-Topline Angle – Slope	0.211 [0.16, 0.25]	0.495 [0.41, 0.57]	0.301 [0.24, 0.35]	0.234 [0.18, 0.28]	0.504 [0.42, 0.57]	0.297 [0.23, 0.35]		
Forehead-Topline Length Ratio	0.189 [0.15, 0.22]	0.227 [0.18, 0.26]	0.194 [0.15, 0.23]	0.249 [0.20, 0.29]	0.287 [0.24, 0.33]	0.252 [0.21, 0.29]		
Forehead Temple Ratio	0.304 [0.24, 0.35]	0.301 [0.24, 0.35]	0.300 [0.24, 0.35]	0.302 [0.24, 0.35]	0.303 [0.24, 0.35]	0.299 [0.24, 0.35]	0.296 [0.24, 0.35]	0.295 [0.23, 0.34]
	0.297 [0.24, 0.35]	0.298 [0.24, 0.35]	0.328 [0.26, 0.38]	0.326 [0.26, 0.38]	0.324 [0.26, 0.38]	0.327 [0.26, 0.38]	0.327 [0.26, 0.38]	0.327 [0.26, 0.39]
Forehead Width-to-Length Ratio	0.335 [0.27, 0.39]	0.334 [0.27, 0.39]	0.336 [0.27, 0.39]	0.336 [0.27, 0.39]	0.333 [0.27, 0.39]	0.331 [0.27, 0.38]	0.330 [0.27, 0.38]	0.332 [0.27, 0.38]
	0.332 [0.27, 0.38]	0.361 [0.29, 0.42]	0.359 [0.29, 0.42]	0.358 [0.29, 0.41]	0.360 [0.29, 0.42]	0.360 [0.29, 0.42]		
Forehead-Zygomatic Angle	0.238 [0.19, 0.28]	0.238 [0.19, 0.28]	0.238 [0.19, 0.28]	0.235 [0.18, 0.28]	0.235 [0.18, 0.28]	0.235 [0.18, 0.28]	0.273 [0.22, 0.32]	0.273 [0.22, 0.32]
	0.273 [0.22, 0.32]	0.272 [0.22, 0.32]	0.272 [0.22, 0.32]	0.272 [0.22, 0.32]	0.348 [0.28, 0.40]	0.348 [0.28, 0.40]	0.348 [0.28, 0.40]	0.350 [0.28, 0.40]
Forehead-Zygomatic Angle	0.350 [0.28, 0.40]	0.350 [0.28, 0.40]						
	0.958 [0.93, 1.00]	0.951 [0.92, 1.00]	0.962 [0.94, 1.00]	0.959 [0.93, 1.00]	0.958 [0.93, 1.00]	0.951 [0.92, 1.00]	0.962 [0.94, 1.00]	0.959 [0.93, 1.00]
Jaw Angle – Slope	0.958 [0.93, 1.00]	0.951 [0.92, 1.00]	0.962 [0.94, 1.00]	0.959 [0.93, 1.00]				
	0.133 [0.11, 0.15]	0.133 [0.11, 0.15]	0.135 [0.11, 0.16]	0.137 [0.11, 0.16]	0.137 [0.11, 0.16]	0.139 [0.11, 0.16]		
Jaw-Jaw Length Proportion	0.281 [0.23, 0.33]	0.253 [0.20, 0.29]	0.294 [0.24, 0.34]	0.298 [0.24, 0.35]	0.269 [0.22, 0.31]	0.254 [0.20, 0.29]	0.275 [0.22, 0.32]	0.288 [0.23, 0.34]
Jaw Length Proportion	0.086 [0.07, 0.10]	0.088 [0.07, 0.10]	0.083 [0.07, 0.10]	0.075 [0.06, 0.09]				
Jaw-Midface Size Ratio	0.071 [0.06, 0.08]	0.083 [0.07, 0.10]	0.079 [0.06, 0.09]	0.092 [0.07, 0.11]	0.089 [0.07, 0.11]	0.101 [0.08, 0.12]		
Muzzle Size Proportion	0.225 [0.18, 0.26]	0.216 [0.17, 0.25]	0.237 [0.19, 0.28]	0.229 [0.18, 0.27]	0.320 [0.26, 0.37]	0.313 [0.25, 0.36]	0.311 [0.25, 0.36]	0.305 [0.25, 0.35]
	0.199 [0.15, 0.23]	0.206 [0.16, 0.24]	0.201 [0.16, 0.24]	0.210 [0.16, 0.25]				
Midface thickness Proportion	0.101 [0.08, 0.12]	0.179 [0.15, 0.21]						
Nasion Thickness Proportion	0.109 [0.08, 0.13]	0.113 [0.09, 0.13]	0.130 [0.10, 0.15]	0.142 [0.11, 0.17]	0.142 [0.11, 0.17]	0.155 [0.12, 0.18]	0.120 [0.09, 0.14]	0.117 [0.09, 0.14]
	0.131 [0.10, 0.15]							
Overall Eye Angle – Angle	0.257 [0.20, 0.30]	0.302 [0.24, 0.35]	0.396 [0.32, 0.46]	0.296 [0.23, 0.35]	0.268 [0.21, 0.31]			
Overall Eye Angle – Slope	0.247 [0.19, 0.29]	0.293 [0.23, 0.34]	0.385 [0.31, 0.45]	0.287 [0.23, 0.34]	0.258 [0.20, 0.30]			
Overall Eye Size	0.061 [0.04, 0.07]	0.067 [0.05, 0.08]	0.060 [0.04, 0.07]	0.065 [0.05, 0.08]	0.064 [0.05, 0.08]	0.078 [0.07, 0.13]	0.086 [0.06, 0.10]	0.080 [0.06, 0.10]
	0.081 [0.06, 0.10]	0.084 [0.06, 0.10]	0.103 [0.07, 0.12]	0.112 [0.08, 0.13]	0.111 [0.08, 0.13]	0.107 [0.08, 0.13]	0.109 [0.08, 0.13]	
Poll Depth Proportion – Height	0.477 [0.39, 0.55]	1.000 [1.00, 1.00]	0.741 [0.66, 0.82]	0.453 [0.37, 0.52]	0.999 [1.00, 1.00]	0.742 [0.66, 0.82]		
Poll Depth Proportion – Length	0.381 [0.31, 0.44]	0.934 [0.90, 1.00]	0.497 [0.43, 0.57]	0.387 [0.32, 0.44]	0.935 [0.90, 1.00]	0.497 [0.43, 0.57]		
Poll Height Proportion	0.366 [0.29, 0.43]	0.996 [0.99, 1.00]	0.471 [0.40, 0.54]	0.308 [0.24, 0.36]	0.945 [0.92, 1.00]	0.358 [0.29, 0.41]	0.373 [0.30, 0.43]	0.999 [1.00, 1.00]
	0.474 [0.40, 0.28]	0.309 [0.24, 0.36]	0.945 [0.92, 1.00]	0.361 [0.30, 0.42]				
Poll Height Point Proportion	0.432 [0.36, 0.50]	0.977 [0.95, 1.00]	0.413 [0.34, 0.47]	0.351 [0.29, 0.40]	0.621 [0.54, 0.70]	0.329 [0.27, 0.28]	0.429 [0.35, 0.49]	0.998 [1.00, 1.00]
	0.411 [0.34, 0.47]	0.350 [0.29, 0.30]	0.626 [0.54, 0.70]	0.328 [0.27, 0.38]				

Table 18: Proportion of total variability in forehead and jaw biometrics attributed to error in image acquisition (variance between days/photos) – Part A

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Canthus Length Proportion	0.153 [0.09, 0.24]	0.124 [0.05, 0.21]	0.095 [0.02, 0.17]	0.123 [0.06, 0.21]	0.134 [0.06, 0.21]			
Canthus Depth Proportion – Depth	0.199 [0.14, 0.27]	0.190 [0.13, 0.27]	0.209 [0.15, 0.29]	0.205 [0.14, 0.28]	0.193 [0.13, 0.27]	0.214 [0.14, 0.29]		
Canthus Depth Proportion – Length	0.200 [0.14, 0.28]	0.194 [0.12, 0.27]	0.176 [0.10, 0.26]	0.206 [0.14, 0.28]	0.189 [0.11, 0.27]			
Chin Length Proportion	0.217 [0.16, 0.27]	0.266 [0.20, 0.33]	0.137 [0.09, 0.18]	0.282 [0.22, 0.34]	0.287 [0.22, 0.35]	0.197 [0.14, 0.25]		
Cheek Nose Size Proportion	0.223 [0.16, 0.28]	0.211 [0.15, 0.27]	0.219 [0.16, 0.27]	0.209 [0.15, 0.26]	0.216 [0.16, 0.27]	0.205 [0.15, 0.26]	0.187 [0.13, 0.25]	0.181 [0.12, 0.24]
	0.182 [0.12, 0.24]	0.177 [0.12, 0.23]	0.181 [0.12, 0.24]	0.175 [0.12, 0.23]	0.225 [0.16, 0.28]	0.262 [0.20, 0.32]	0.222 [0.16, 0.28]	0.259 [0.19, 0.32]
	0.220 [0.16, 0.28]	0.257 [0.19, 0.32]						
Cranio-Topline Length Ratio	0.511 [0.43, 0.59]	0.469 [0.39, 0.55]	0.501 [0.30, 0.48]	0.573 [0.48, 0.66]	0.451 [0.39, 0.53]	0.578 [0.48, 0.66]	0.482 [0.40, 0.56]	0.447 [0.37, 0.52]
	0.481 [0.40, 0.56]	0.554 [0.47, 0.64]	0.421 [0.36, 0.50]	0.552 [0.47, 0.64]				
Canthus Width-to-Height Ratio	0.164 [0.10, 0.24]	0.120 [0.05, 0.14]	0.00 [0.00, 0.00]	0.154 [0.09, 0.23]	0.122 [0.05, 0.20]			
Eye-Cranio Size Ratio	0.315 [0.24, 0.38]	0.341 [0.26, 0.41]	0.342 [0.26, 0.41]	0.317 [0.24, 0.38]	0.345 [0.26, 0.41]	0.310 [0.23, 0.37]	0.338 [0.26, 0.40]	0.336 [0.26, 0.40]
	0.311 [0.24, 0.37]	0.343 [0.26, 0.41]	0.291 [0.22, 0.35]	0.316 [0.24, 0.38]	0.312 [0.24, 0.37]	0.290 [0.22, 0.35]	0.323 [0.25, 0.39]	0.312 [0.24, 0.37]
	0.340 [0.26, 0.40]	0.343 [0.26, 0.41]	0.315 [0.24, 0.38]	0.343 [0.26, 0.41]	0.307 [0.23, 0.37]	0.336 [0.26, 0.40]	0.336 [0.26, 0.40]	0.309 [0.24, 0.37]
	0.340 [0.26, 0.41]	0.288 [0.22, 0.35]	0.313 [0.24, 0.37]	0.313 [0.24, 0.37]	0.287 [0.22, 0.34]	0.320 [0.25, 0.38]		
Eye-Forehead Size Ratio – Linear	0.300 [0.24, 0.36]	0.286 [0.22, 0.35]	0.258 [0.20, 0.31]	0.292 [0.23, 0.35]	0.291 [0.22, 0.35]			
Eye-Forehead Size Ratio – Poly	0.291 [0.22, 0.35]	0.303 [0.23, 0.36]	0.302 [0.23, 0.36]	0.289 [0.22, 0.35]	0.308 [0.24, 0.37]	0.284 [0.22, 0.34]	0.298 [0.23, 0.36]	0.295 [0.22, 0.35]
	0.283 [0.22, 0.34]	0.303 [0.23, 0.36]	0.270 [0.21, 0.33]	0.282 [0.22, 0.34]	0.277 [0.21, 0.33]	0.268 [0.20, 0.32]	0.289 [0.22, 0.35]	0.281 [0.21, 0.34]
	0.293 [0.23, 0.35]	0.296 [0.23, 0.35]	0.280 [0.21, 0.34]	0.298 [0.23, 0.36]	0.274 [0.21, 0.33]	0.288 [0.22, 0.35]	0.288 [0.22, 0.35]	0.273 [0.21, 0.33]
	0.293 [0.22, 0.35]	0.260 [0.20, 0.31]	0.272 [0.21, 0.33]	0.271 [0.53, 0.69]	0.258 [0.19, 0.31]	0.278 [0.21, 0.34]		
Eye Orbital Height-to-Length Ratio	0.212 [0.16, 0.26]	0.333 [0.26, 0.40]	0.347 [0.27, 0.42]	0.367 [0.30, 0.44]	0.328 [0.26, 0.40]	0.335 [0.26, 0.40]	0.357 [0.29, 0.43]	0.155 [0.11, 0.20]
	0.272 [0.21, 0.34]	0.250 [0.19, 0.32]	0.320 [0.25, 0.39]	0.270 [0.20, 0.33]	0.243 [0.18, 0.31]	0.324 [0.26, 0.39]		
Eye Orbital-Eye Height Ratio	0.294 [0.22, 0.35]	0.447 [0.37, 0.53]	0.408 [0.34, 0.48]	0.007 [0.00, 0.01]	0.432 [0.35, 0.51]	0.394 [0.33, 0.47]	0.014 [0.00, 0.03]	0.261 [0.20, 0.72]
	0.451 [0.37, 0.53]	0.428 [0.36, 0.51]	0.012 [0.00, 0.02]	0.451 [0.37, 0.53]	0.420 [0.35, 0.50]	0.019 [0.00, 0.04]	0.296 [0.22, 0.36]	0.448 [0.37, 0.53]
	0.407 [0.34, 0.48]	0.000 [0.00, 0.00]	0.433 [0.35, 0.51]	0.394 [0.33, 0.47]	0.004 [0.00, 0.01]	0.259 [0.20, 0.31]	0.452 [0.38, 0.53]	0.427 [0.36, 0.51]
	0.002 [0.00, 0.00]	0.452 [0.37, 0.53]	0.419 [0.35, 0.50]	0.009 [0.00, 0.02]	0.297 [0.22, 0.36]	0.449 [0.37, 0.53]	0.409 [0.34, 0.48]	0.000 [0.00, 0.00]
	0.434 [0.36, 0.51]	0.396 [0.33, 0.47]	0.004 [0.00, 0.01]	0.259 [0.20, 0.31]	0.454 [0.38, 0.53]	0.430 [0.36, 0.51]	0.002 [0.00, 0.00]	0.453 [0.37, 0.53]
	0.422 [0.35, 0.50]	0.008 [0.00, 0.02]						
Eye Orbital Projection Proportion	0.181 [0.13, 0.23]	0.161 [0.11, 0.21]	0.177 [0.12, 0.22]	0.234 [0.17, 0.29]	0.211 [0.15, 0.26]	0.224 [0.16, 0.28]	0.188 [0.13, 0.23]	0.180 [0.13, 0.23]
	0.189 [0.14, 0.24]							
Eye Orbital Roundness Proportion	0.294 [0.22, 0.37]	0.238 [0.17, 0.32]						
Eye Orbital Roundness Point Proportion	0.343 [0.27, 0.43]	0.211 [0.14, 0.28]						
Eye Orbital Thickness Proportion – Poly	0.166 [0.12, 0.21]	0.191 [0.14, 0.24]	0.136 [0.09, 0.19]	0.194 [0.14, 0.25]	0.151 [0.10, 0.20]	0.129 [0.08, 0.18]	0.115 [0.07, 0.17]	0.087 [0.04, 0.14]
	0.117 [0.07, 0.17]	0.102 [0.05, 0.15]						
Eye Sinus Size Ratio – Linear	0.166 [0.11, 0.21]	0.195 [0.14, 0.25]	0.210 [0.15, 0.26]	0.160 [0.11, 0.21]	0.199 [0.14, 0.25]	0.042 [0.00, 0.07]	0.034 [0.00, 0.07]	0.037 [0.00, 0.07]
	0.044 [0.00, 0.09]	0.026 [0.00, 0.05]						
Eye Sinus Size Ratio – Poly	0.262 [0.20, 0.32]	0.257 [0.19, 0.31]	0.249 [0.19, 0.30]	0.257 [0.19, 0.31]	0.249 [0.19, 0.30]	0.262 [0.20, 0.32]	0.256 [0.19, 0.31]	0.248 [0.18, 0.30]
	0.257 [0.19, 0.31]	0.247 [0.18, 0.30]	0.259 [0.19, 0.31]	0.253 [0.19, 0.31]	0.244 [0.18, 0.30]	0.253 [0.19, 0.31]	0.244 [0.18, 0.30]	0.164 [0.11, 0.21]
	0.161 [0.11, 0.21]	0.155 [0.11, 0.20]	0.163 [0.11, 0.21]	0.153 [0.10, 0.20]	0.168 [0.12, 0.21]	0.165 [0.12, 0.21]	0.158 [0.11, 0.20]	0.167 [0.12, 0.21]
	0.155 [0.11, 0.20]	0.170 [0.12, 0.22]	0.166 [0.12, 0.21]	0.159 [0.11, 0.20]	0.168 [0.12, 0.21]	0.156 [0.11, 0.20]		
Eye-Topline Size Ratio – Linear	0.124 [0.09, 0.15]	0.113 [0.07, 0.15]	0.096 [0.05, 0.14]	0.117 [0.08, 0.15]	0.104 [0.07, 0.13]			
Eye-Topline Size Ratio – Poly	0.212 [0.15, 0.26]	0.207 [0.15, 0.25]	0.211 [0.15, 0.26]	0.204 [0.15, 0.25]	0.218 [0.16, 0.27]	0.222 [0.16, 0.27]	0.222 [0.16, 0.27]	0.225 [0.17, 0.28]
	0.217 [0.16, 0.26]	0.232 [0.17, 0.28]	0.228 [0.17, 0.28]	0.227 [0.17, 0.28]	0.230 [0.17, 0.28]	0.222 [0.16, 0.27]	0.237 [0.18, 0.29]	

Table 19: Proportion of total variability in forehead and jaw biometrics attributed to error in image acquisition (variance between days/photos) – Part B

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Forehead-Eye Angle – Slope	0.211 [0.15, 0.27]	0.250 [0.19, 0.32]	0.210 [0.15, 0.27]	0.204 [0.15, 0.26]	0.244 [0.17, 0.31]	0.209 [0.15, 0.27]	0.183 [0.13, 0.24]	0.219 [0.16, 0.28]
	0.194 [0.14, 0.25]	0.177 [0.12, 0.24]	0.212 [0.16, 0.28]	0.193 [0.14, 0.25]	0.190 [0.13, 0.25]	0.230 [0.18, 0.30]	0.207 [0.15, 0.26]	0.183 [0.13, 0.24]
	0.225 [0.17, 0.29]	0.206 [0.15, 0.26]	0.199 [0.14, 0.26]	0.234 [0.18, 0.30]	0.193 [0.14, 0.25]	0.193 [0.14, 0.25]	0.227 [0.17, 0.29]	0.192 [0.14, 0.25]
Forehead-Jaw Angle – Slope	0.203 [0.15, 0.25]	0.175 [0.12, 0.22]	0.181 [0.13, 0.23]	0.199 [0.14, 0.25]	0.176 [0.12, 0.23]		0.203 [0.14, 0.25]	0.178 [0.13, 0.23]
	0.186 [0.13, 0.24]	0.198 [0.14, 0.25]	0.179 [0.13, 0.23]	0.186 [0.13, 0.24]				
Forehead-Poll Length Ratio	0.143 [0.10, 0.19]	0.135 [0.09, 0.18]	0.147 [0.10, 0.20]	0.188 [0.14, 0.24]	0.076 [0.01, 0.15]	0.176 [0.13, 0.23]	0.144 [0.10, 0.19]	0.135 [0.09, 0.19]
	0.147 [0.10, 0.20]	0.194 [0.14, 0.25]	0.015 [0.00, 0.03]	0.182 [0.13, 0.24]				
Forehead-Topline Angle – Slope	0.117 [0.08, 0.15]	0.106 [0.05, 0.17]	0.087 [0.045, 0.13]	0.117 [0.074, 0.157]	0.104 [0.05, 0.17]	0.090 [0.05, 0.13]		
Forehead-Topline Length Ratio	0.347 [0.28, 0.41]	0.323 [0.26, 0.39]	0.342 [0.28, 0.41]	0.345 [0.28, 0.41]	0.320 [0.26, 0.39]	0.342 [0.28, 0.41]		
Forehead Temple Ratio	0.167 [0.12, 0.22]	0.168 [0.12, 0.22]	0.167 [0.12, 0.22]	0.168 [0.12, 0.22]	0.167 [0.12, 0.22]	0.171 [0.12, 0.22]	0.172 [0.12, 0.22]	0.171 [0.12, 0.22]
	0.172 [0.12, 0.22]	0.171 [0.12, 0.22]	0.157 [0.11, 0.21]	0.158 [0.11, 0.21]	0.158 [0.11, 0.21]	0.158 [0.11, 0.21]	0.157 [0.11, 0.21]	0.196 [0.14, 0.25]
	0.197 [0.14, 0.25]	0.196 [0.14, 0.25]	0.197 [0.14, 0.25]	0.196 [0.14, 0.25]	0.200 [0.14, 0.25]	0.200 [0.14, 0.26]	0.199 [0.14, 0.26]	0.200 [0.14, 0.26]
	0.200 [0.14, 0.26]	0.184 [0.13, 0.24]	0.185 [0.13, 0.24]	0.184 [0.13, 0.24]	0.185 [0.13, 0.24]	0.184 [0.13, 0.25]		
Forehead Width-to-Length Ratio	0.186 [0.13, 0.24]	0.186 [0.13, 0.24]	0.186 [0.13, 0.24]	0.189 [0.13, 0.24]	0.189 [0.13, 0.24]	0.189 [0.13, 0.24]	0.181 [0.13, 0.23]	0.181 [0.13, 0.23]
	0.181 [0.13, 0.23]	0.182 [0.13, 0.23]	0.182 [0.13, 0.23]	0.182 [0.13, 0.23]	0.141 [0.09, 0.19]	0.141 [0.09, 0.19]	0.141 [0.09, 0.19]	0.141 [0.09, 0.19]
	0.141 [0.09, 0.19]	0.141 [0.09, 0.19]						
Forehead-Zygomatic Angle	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]
	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]	0.000 [0.00, 0.00]				
Jaw Angle – Slope	0.349 [0.28, 0.42]	0.349 [0.28, 0.42]	0.345 [0.27, 0.42]	0.355 [0.28, 0.43]	0.355 [0.28, 0.43]	0.350 [0.28, 0.42]		
Jowl-Jaw Length Proportion	0.260 [0.20, 0.32]	0.266 [0.20, 0.33]	0.198 [0.14, 0.25]	0.205 [0.15, 0.26]	0.255 [0.19, 0.32]	0.253 [0.19, 0.31]	0.199 [0.14, 0.25]	0.198 [0.14, 0.25]
Jaw Length Proportion	0.433 [0.35, 0.51]	0.428 [0.34, 0.50]	0.432 [0.35, 0.51]	0.432 [0.35, 0.51]				
Jaw-Midface Size Ratio	0.549 [0.46, 0.63]	0.538 [0.44, 0.62]	0.558 [0.47, 0.64]	0.545 [0.45, 0.63]	0.659 [0.57, 0.75]	0.642 [0.55, 0.73]		
Muzzle Size Proportion	0.214 [0.16, 0.27]	0.218 [0.16, 0.17]	0.218 [0.16, 0.17]	0.221 [0.16, 0.18]	0.164 [0.11, 0.22]	0.167 [0.11, 0.22]	0.179 [0.12, 0.23]	0.182 [0.13, 0.24]
	0.156 [0.11, 0.20]	0.151 [0.11, 0.19]	0.175 [0.12, 0.22]	0.167 [0.12, 0.21]				
Midface thickness Proportion	0.402 [0.32, 0.47]	0.418 [0.34, 0.49]						
Nasion Thickness Proportion	0.362 [0.29, 0.43]	0.362 [0.29, 0.43]	0.343 [0.27, 0.41]	0.391 [0.32, 0.46]	0.386 [0.31, 0.46]	0.356 [0.29, 0.43]	0.378 [0.30, 0.45]	0.362 [0.29, 0.43]
	0.330 [0.26, 0.39]							
Overall Eye Angle – Angle	0.131 [0.08, 0.20]	0.100 [0.06, 0.14]	0.051 [0.01, 0.10]	0.109 [0.06, 0.15]	0.115 [0.07, 0.16]			
Overall Eye Angle – Slope	0.136 [0.09, 0.18]	0.106 [0.06, 0.15]	0.059 [0.01, 0.10]	0.113 [0.07, 0.16]	0.122 [0.08, 0.17]			
Overall Eye Size	0.108 [0.07, 0.17]	0.114 [0.07, 0.15]	0.123 [0.09, 0.15]	0.111 [0.08, 0.14]	0.116 [0.08, 0.14]	0.103 [0.07, 0.13]	0.113 [0.08, 0.14]	0.115 [0.08, 0.14]
	0.107 [0.07, 0.14]	0.114 [0.08, 0.14]	0.098 [0.07, 0.13]	0.104 [0.07, 0.13]	0.104 [0.07, 0.13]	0.099 [0.07, 0.13]	0.108 [0.07, 0.14]	
Poll Depth Proportion – Height	0.069 [0.02, 0.12]	0.00 [0.00, 0.00]	0.151 [0.07, 0.24]	0.075 [0.03, 0.13]	0.000 [0.00, 0.00]	0.151 [0.07, 0.24]		
Poll Depth Proportion – Length	0.185 [0.13, 0.25]	0.000 [0.00, 0.00]	0.216 [0.15, 0.29]	0.188 [0.13, 0.25]	0.000 [0.00, 0.00]	0.217 [0.15, 0.29]		
Poll Height Proportion	0.071 [0.03, 0.12]	0.000 [0.00, 0.00]	0.209 [0.15, 0.28]	0.111 [0.07, 0.16]	0.000 [0.00, 0.00]	0.244 [0.19, 0.31]	0.066 [0.02, 0.11]	0.000 [0.00, 0.00]
	0.212 [0.15, 0.28]	0.106 [0.07, 0.15]	0.000 [0.00, 0.00]	0.248 [0.19, 0.31]				
Poll Height Point Proportion	0.158 [0.10, 0.22]	0.000 [0.00, 0.00]	0.153 [0.10, 0.21]	0.237 [0.17, 0.30]	0.121 [0.06, 0.20]	0.224 [0.16, 0.28]	0.154 [0.10, 0.22]	0.000 [0.00, 0.00]
	0.151 [0.09, 0.22]	0.234 [0.17, 0.30]	0.127 [0.06, 0.20]	0.222 [0.16, 0.28]				

Table 20: Repeatability of Geometric Topline Biometrics from a Single Coordinate Extraction – Part A

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Canthus Length Proportion	0.203 [0.12, 0.26]	0.167 [0.09, 0.23]	0.230 [0.15, 0.30]	0.181 [0.11, 0.24]	0.190 [0.11, 0.25]			
Canthus Depth Proportion – Depth	0.305 [0.22, 0.38]	0.279 [0.20, 0.35]	0.247 [0.16, 0.32]	0.305 [0.22, 0.39]	0.277 [0.20, 0.36]	0.243 [0.16, 0.32]		
Canthus Depth Proportion – Length	0.250 [0.17, 0.32]	0.233 [0.15, 0.31]	0.195 [0.12, 0.26]	0.253 [0.17, 0.32]	0.235 [0.15, 0.31]			
Chin Length Proportion	0.489 [0.41, 0.58]	0.389 [0.31, 0.48]	0.575 [0.51, 0.66]	0.542 [0.47, 0.63]	0.410 [0.33, 0.50]	0.615 [0.55, 0.70]		
Check Nose Size Proportion	0.494 [0.42, 0.58]	0.499 [0.42, 0.59]	0.495 [0.42, 0.58]	0.500 [0.43, 0.59]	0.504 [0.43, 0.59]	0.510 [0.44, 0.60]	0.461 [0.38, 0.55]	0.470 [0.39, 0.56]
	0.468 [0.39, 0.55]	0.478 [0.40, 0.56]	0.476 [0.40, 0.56]	0.486 [0.41, 0.57]	0.520 [0.45, 0.61]	0.543 [0.47, 0.63]	0.522 [0.45, 0.61]	0.546 [0.47, 0.63]
	0.526 [0.45, 0.61]	0.551 [0.48, 0.64]						
Cranio-Topline Length Ratio	0.387 [0.30, 0.48]	0.345 [0.26, 0.44]	0.386 [0.30, 0.48]	0.356 [0.27, 0.46]	0.219 [0.13, 0.30]	0.363 [0.28, 0.46]	0.364 [0.28, 0.46]	0.321 [0.23, 0.41]
	0.363 [0.28, 0.46]	0.334 [0.25, 0.43]	0.204 [0.12, 0.28]	0.339 [0.25, 0.43]				
Canthus Width-to-Height Ratio	0.302 [0.22, 0.38]	0.250 [0.17, 0.32]	0.00 [0.00, 0.00]	0.310 [0.23, 0.39]	0.254 [0.18, 0.33]			
Eye-Cranio Size Ratio	0.631 [0.56, 0.72]	0.603 [0.53, 0.69]	0.607 [0.54, 0.70]	0.626 [0.56, 0.71]	0.601 [0.53, 0.69]	0.618 [0.55, 0.70]	0.585 [0.51, 0.68]	0.591 [0.52, 0.68]
	0.612 [0.54, 0.70]	0.583 [0.51, 0.67]	0.610 [0.54, 0.70]	0.578 [0.51, 0.67]	0.582 [0.51, 0.67]	0.605 [0.54, 0.69]	0.577 [0.51, 0.67]	0.632 [0.56, 0.72]
	0.604 [0.53, 0.69]	0.606 [0.53, 0.70]	0.628 [0.56, 0.71]	0.601 [0.53, 0.69]	0.619 [0.55, 0.70]	0.585 [0.51, 0.68]	0.590 [0.52, 0.68]	0.613 [0.54, 0.70]
	0.583 [0.51, 0.67]	0.612 [0.54, 0.70]	0.579 [0.51, 0.67]	0.581 [0.51, 0.67]	0.606 [0.54, 0.69]	0.578 [0.51, 0.67]		
Eye-Forehead Size Ratio – Linear	0.578 [0.50, 0.66]	0.576 [0.50, 0.66]	0.579 [0.51, 0.67]	0.577 [0.51, 0.66]	0.577 [0.51, 0.67]			
Eye-Forehead Size Ratio – Poly	0.611 [0.54, 0.70]	0.596 [0.52, 0.68]	0.599 [0.53, 0.69]	0.610 [0.54, 0.69]	0.593 [0.52, 0.68]	0.607 [0.54, 0.69]	0.590 [0.52, 0.68]	0.596 [0.53, 0.68]
	0.606 [0.54, 0.69]	0.586 [0.51, 0.68]	0.607 [0.54, 0.69]	0.592 [0.52, 0.68]	0.598 [0.54, 0.69]	0.607 [0.54, 0.69]	0.588 [0.52, 0.68]	0.615 [0.55, 0.70]
	0.601 [0.53, 0.69]	0.601 [0.53, 0.69]	0.615 [0.55, 0.70]	0.596 [0.53, 0.68]	0.611 [0.54, 0.70]	0.595 [0.52, 0.68]	0.598 [0.53, 0.69]	0.611 [0.54, 0.70]
	0.590 [0.52, 0.68]	0.612 [0.54, 0.70]	0.597 [0.53, 0.68]	0.601 [0.53, 0.69]	0.613 [0.54, 0.70]	0.593 [0.52, 0.68]		
Eye Orbital Height-to-Length Ratio	0.657 [0.59, 0.73]	0.398 [0.31, 0.49]	0.391 [0.31, 0.48]	0.353 [0.27, 0.44]	0.405 [0.32, 0.49]	0.393 [0.31, 0.49]	0.355 [0.27, 0.44]	0.656 [0.60, 0.73]
	0.391 [0.31, 0.48]	0.383 [0.30, 0.47]	0.349 [0.26, 0.44]	0.408 [0.33, 0.50]	0.391 [0.31, 0.48]	0.364 [0.28, 0.45]		
Eye Orbital-Eye Height Ratio	0.642 [0.58, 0.73]	0.453 [0.37, 0.54]	0.438 [0.35, 0.53]	0.098 [0.04, 0.14]	0.464 [0.38, 0.56]	0.442 [0.36, 0.53]	0.109 [0.05, 0.15]	0.635 [0.57, 0.72]
	0.429 [0.34, 0.52]	0.418 [0.33, 0.51]	0.109 [0.05, 0.15]	0.451 [0.37, 0.54]	0.429 [0.34, 0.52]	0.119 [0.06, 0.17]	0.640 [0.57, 0.72]	0.453 [0.37, 0.54]
	0.438 [0.35, 0.53]	0.091 [0.04, 0.13]	0.464 [0.38, 0.56]	0.442 [0.36, 0.53]	0.103 [0.05, 0.15]	0.630 [0.56, 0.71]	0.596 [0.52, 0.68]	0.418 [0.33, 0.51]
	0.103 [0.05, 0.15]	0.450 [0.36, 0.54]	0.429 [0.34, 0.52]	0.112 [0.05, 0.16]	0.639 [0.57, 0.72]	0.451 [0.37, 0.54]	0.436 [0.35, 0.53]	0.091 [0.04, 0.13]
	0.463 [0.38, 0.56]	0.440 [0.36, 0.53]	0.102 [0.05, 0.15]	0.629 [0.56, 0.68]	0.427 [0.34, 0.52]	0.417 [0.33, 0.51]	0.102 [0.05, 0.15]	0.449 [0.36, 0.54]
	0.427 [0.34, 0.52]	0.112 [0.05, 0.16]						
Eye Orbital Projection Proportion	0.612 [0.54, 0.69]	0.611 [0.54, 0.70]	0.609 [0.54, 0.69]	0.557 [0.48, 0.64]	0.566 [0.49, 0.65]	0.559 [0.49, 0.64]	0.629 [0.56, 0.71]	0.631 [0.57, 0.71]
	0.616 [0.55, 0.70]							
Eye Orbital Roundness Proportion	0.326 [0.24, 0.41]	0.212 [0.13, 0.29]						
Eye Orbital Roundness Point Proportion	0.178 [0.10, 0.25]	0.345 [0.26, 0.43]						
Eye Orbital Thickness Proportion – Poly	0.621 [0.56, 0.70]	0.566 [0.50, 0.65]	0.538 [0.47, 0.62]	0.586 [0.52, 0.67]	0.562 [0.49, 0.65]	0.529 [0.45, 0.61]	0.505 [0.43, 0.59]	0.487 [0.41, 0.57]
	0.523 [0.45, 0.61]	0.523 [0.45, 0.61]						
Eye Sinus Size Ratio – Linear	0.596 [0.53, 0.68]	0.574 [0.50, 0.66]	0.574 [0.50, 0.66]	0.585 [0.52, 0.67]	0.583 [0.51, 0.67]	0.441 [0.37, 0.52]	0.420 [0.34, 0.50]	0.410 [0.33, 0.49]
	0.436 [0.36, 0.52]	0.422 [0.35, 0.50]						
Eye Sinus Size Ratio – Poly	0.642 [0.58, 0.72]	0.642 [0.58, 0.72]	0.648 [0.58, 0.73]	0.639 [0.57, 0.72]	0.651 [0.59, 0.73]	0.643 [0.58, 0.73]	0.644 [0.58, 0.73]	0.650 [0.59, 0.73]
	0.640 [0.57, 0.72]	0.653 [0.59, 0.73]	0.651 [0.59, 0.73]	0.651 [0.59, 0.73]	0.656 [0.59, 0.74]	0.647 [0.58, 0.73]	0.660 [0.60, 0.74]	0.637 [0.57, 0.72]
	0.639 [0.57, 0.72]	0.639 [0.57, 0.72]	0.630 [0.56, 0.71]	0.647 [0.58, 0.73]	0.637 [0.57, 0.72]	0.638 [0.57, 0.72]	0.637 [0.57, 0.72]	0.629 [0.56, 0.71]
	0.646 [0.58, 0.73]	0.638 [0.57, 0.72]	0.638 [0.57, 0.72]	0.637 [0.57, 0.72]	0.628 [0.56, 0.71]	0.646 [0.58, 0.73]		
Eye-Topline Size Ratio – Linear	0.809 [0.77, 0.86]	0.740 [0.69, 0.81]	0.633 [0.57, 0.71]	0.784 [0.74, 0.84]	0.766 [0.72, 0.83]			
Eye-Topline Size Ratio – Poly	0.743 [0.69, 0.81]	0.725 [0.67, 0.79]	0.739 [0.69, 0.81]	0.736 [0.68, 0.80]	0.729 [0.68, 0.80]	0.720 [0.66, 0.79]	0.697 [0.64, 0.77]	0.709 [0.65, 0.78]
	0.712 [0.66, 0.78]	0.702 [0.64, 0.78]	0.698 [0.64, 0.77]	0.676 [0.62, 0.75]	0.678 [0.62, 0.75]	0.690 [0.63, 0.76]	0.680 [0.62, 0.76]	

Table 21: Repeatability of Geometric Topline Biometrics from a Single Coordinate Extraction – Part B

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Forehead-Eye Angle – Slope	0.456 [0.38, 0.54]	0.399 [0.31, 0.48]	0.517 [0.44, 0.60]	0.447 [0.37, 0.53]	0.401 [0.32, 0.48]	0.517 [0.44, 0.60]	0.432 [0.35, 0.52]	0.386 [0.30, 0.47]
	0.518 [0.44, 0.60]	0.424 [0.34, 0.51]	0.389 [0.31, 0.47]	0.516 [0.44, 0.60]	0.436 [0.36, 0.52]	0.387 [0.30, 0.47]	0.517 [0.44, 0.60]	0.428 [0.35, 0.51]
	0.390 [0.31, 0.47]	0.516 [0.44, 0.60]	0.452 [0.38, 0.54]	0.398 [0.32, 0.48]	0.520 [0.45, 0.61]	0.443 [0.36, 0.53]	0.401 [0.32, 0.48]	0.519 [0.45, 0.61]
Forehead-Jaw Angle – Slope	0.604 [0.54, 0.69]	0.596 [0.53, 0.68]	0.548 [0.48, 0.63]	0.595 [0.53, 0.68]	0.590 [0.52, 0.67]	0.550 [0.48, 0.64]	0.599 [0.53, 0.68]	0.592 [0.52, 0.68]
	0.544 [0.47, 0.63]	0.590 [0.52, 0.67]	0.587 [0.52, 0.67]	0.546 [0.48, 0.63]				
Forehead-Poll Length Ratio	0.567 [0.50, 0.65]	0.559 [0.49, 0.64]	0.567 [0.50, 0.65]	0.561 [0.49, 0.64]	0.166 [0.10, 0.22]	0.561 [0.49, 0.64]	0.549 [0.48, 0.63]	0.541 [0.47, 0.62]
	0.548 [0.47, 0.63]	0.526 [0.45, 0.61]	0.041 [0.00, 0.07]	0.529 [0.45, 0.61]				
Forehead-Topline Angle – Slope	0.672 [0.61, 0.75]	0.400 [0.32, 0.48]	0.612 [0.55, 0.69]	0.649 [0.59, 0.73]	0.392 [0.31, 0.47]	0.613 [0.55, 0.69]		
Forehead-Topline Length Ratio	0.465 [0.38, 0.56]	0.451 [0.37, 0.54]	0.464 [0.38, 0.55]	0.406 [0.32, 0.49]	0.392 [0.31, 0.49]	0.405 [0.32, 0.49]		
Forehead Temple Ratio	0.529 [0.45, 0.61]	0.531 [0.46, 0.62]	0.533 [0.46, 0.62]	0.530 [0.46, 0.61]	0.530 [0.46, 0.61]	0.530 [0.46, 0.62]	0.532 [0.46, 0.62]	0.534 [0.46, 0.62]
	0.531 [0.46, 0.62]	0.531 [0.46, 0.62]	0.515 [0.44, 0.60]	0.517 [0.44, 0.60]	0.518 [0.44, 0.60]	0.516 [0.44, 0.60]	0.516 [0.44, 0.60]	0.467 [0.39, 0.55]
	0.469 [0.39, 0.55]	0.470 [0.39, 0.56]	0.468 [0.39, 0.55]	0.468 [0.39, 0.55]	0.467 [0.39, 0.55]	0.469 [0.39, 0.55]	0.470 [0.39, 0.56]	0.468 [0.39, 0.55]
	0.468 [0.39, 0.55]	0.455 [0.38, 0.54]	0.456 [0.38, 0.54]	0.457 [0.38, 0.54]	0.455 [0.38, 0.54]	0.455 [0.38, 0.54]		
Forehead Width-to-Length Ratio	0.576 [0.51, 0.66]	0.576 [0.51, 0.66]	0.576 [0.51, 0.66]	0.576 [0.51, 0.66]	0.576 [0.51, 0.66]	0.576 [0.51, 0.66]	0.547 [0.48, 0.63]	0.547 [0.48, 0.63]
	0.547 [0.48, 0.63]	0.546 [0.47, 0.63]	0.546 [0.47, 0.63]	0.546 [0.47, 0.63]	0.511 [0.44, 0.60]	0.511 [0.44, 0.60]	0.511 [0.44, 0.60]	0.509 [0.44, 0.59]
	0.509 [0.44, 0.59]	0.509 [0.44, 0.59]						
Forehead-Zygomatic Angle	0.042 [0.00, 0.08]	0.049 [0.00, 0.08]	0.038 [0.00, 0.07]	0.041 [0.00, 0.07]	0.042 [0.00, 0.08]	0.049 [0.00, 0.08]	0.038 [0.00, 0.07]	0.041 [0.00, 0.07]
	0.042 [0.00, 0.08]	0.049 [0.00, 0.08]	0.038 [0.00, 0.07]	0.041 [0.00, 0.07]				
Jaw Angle – Slope	0.518 [0.44, 0.61]	0.518 [0.44, 0.61]	0.520 [0.45, 0.61]	0.508 [0.43, 0.60]	0.508 [0.43, 0.60]	0.511 [0.44, 0.60]		
Jowl-Jaw Length Proportion	0.459 [0.38, 0.55]	0.481 [0.40, 0.57]	0.508 [0.43, 0.59]	0.497 [0.42, 0.58]	0.476 [0.40, 0.56]	0.493 [0.42, 0.58]	0.526 [0.45, 0.61]	0.513 [0.44, 0.60]
Jaw Length Proportion	0.481 [0.40, 0.58]	0.484 [0.40, 0.58]	0.485 [0.41, 0.58]	0.493 [0.41, 0.59]				
Jaw-Midface Size Ratio	0.380 [0.29, 0.48]	0.380 [0.29, 0.48]	0.363 [0.28, 0.46]	0.363 [0.28, 0.46]	0.252 [0.16, 0.35]	0.256 [0.18, 0.35]		
Muzzle Size Proportion	0.561 [0.49, 0.64]	0.566 [0.50, 0.65]	0.545 [0.47, 0.63]	0.550 [0.48, 0.64]	0.516 [0.44, 0.60]	0.519 [0.45, 0.60]	0.510 [0.44, 0.60]	0.514 [0.44, 0.60]
	0.645 [0.58, 0.72]	0.642 [0.58, 0.72]	0.624 [0.56, 0.70]	0.623 [0.56, 0.70]				
Midface thickness Proportion	0.497 [0.42, 0.59]	0.403 [0.32, 0.50]						
Nasion Thickness Proportion	0.529 [0.45, 0.62]	0.525 [0.45, 0.62]	0.528 [0.45, 0.62]	0.466 [0.39, 0.56]	0.471 [0.39, 0.56]	0.489 [0.41, 0.58]	0.502 [0.42, 0.60]	0.521 [0.44, 0.61]
	0.539 [0.46, 0.63]							
Overall Eye Angle – Angle	0.612 [0.55, 0.69]	0.597 [0.53, 0.68]	0.553 [0.48, 0.63]	0.596 [0.53, 0.67]	0.617 [0.55, 0.70]			
Overall Eye Angle – Slope	0.617 [0.55, 0.70]	0.600 [0.53, 0.68]	0.557 [0.49, 0.64]	0.600 [0.53, 0.68]	0.620 [0.56, 0.70]			
Overall Eye Size	0.831 [0.79, 0.88]	0.819 [0.78, 0.87]	0.817 [0.78, 0.87]	0.825 [0.79, 0.88]	0.820 [0.78, 0.87]	0.819 [0.78, 0.87]	0.801 [0.76, 0.86]	0.805 [0.76, 0.85]
	0.812 [0.77, 0.87]	0.802 [0.76, 0.86]	0.799 [0.76, 0.86]	0.784 [0.74, 0.84]	0.785 [0.74, 0.84]	0.794 [0.75, 0.85]	0.783 [0.74, 0.84]	
Poll Depth Proportion – Height	0.454 [0.38, 0.54]	0.00 [0.00, 0.00]	0.108 [0.04, 0.16]	0.472 [0.40, 0.56]	0.001 [0.00, 0.00]	0.107 [0.04, 0.16]		
Poll Depth Proportion – Length	0.435 [0.36, 0.52]	0.066 [0.02, 0.10]	0.287 [0.20, 0.37]	0.424 [0.35, 0.51]	0.065 [0.02, 0.10]	0.287 [0.20, 0.37]		
Poll Height Proportion	0.563 [0.49, 0.64]	0.004 [0.00, 0.01]	0.319 [0.24, 0.40]	0.581 [0.51, 0.66]	0.055 [0.01, 0.09]	0.398 [0.31, 0.48]	0.561 [0.49, 0.64]	0.001 [0.00, 0.00]
	0.314 [0.23, 0.39]	0.585 [0.52, 0.66]	0.055 [0.01, 0.09]	0.391 [0.31, 0.47]				
Poll Height Point Proportion	0.410 [0.33, 0.50]	0.023 [0.00, 0.05]	0.434 [0.36, 0.52]	0.413 [0.33, 0.50]	0.258 [0.18, 0.32]	0.446 [0.37, 0.53]	0.417 [0.34, 0.50]	0.002 [0.00, 0.00]
	0.438 [0.36, 0.52]	0.417 [0.34, 0.50]	0.247 [0.17, 0.31]	0.449 [0.37, 0.54]				

Table 22: Repeatability of Geometric Forehead and Jaw Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction – Part A

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Canthus Length Proportion	0.237 [0.17, 0.31]	0.196 [0.13, 0.26]	0.251 [0.18, 0.30]	0.208 [0.14, 0.27]	0.220 [0.15, 0.29]			
Canthus Depth Proportion – Depth	0.327 [0.22, 0.38]	0.327 [0.20, 0.35]	0.299 [0.23, 0.38]	0.353 [0.28, 0.43]	0.324 [0.25, 0.40]	0.294 [0.22, 0.37]		
Canthus Depth Proportion – Length	0.301 [0.23, 0.38]	0.281 [0.21, 0.36]	0.237 [0.17, 0.31]	0.306 [0.23, 0.38]	0.281 [0.21, 0.36]			
Chin Length Proportion	0.535 [0.46, 0.62]	0.453 [0.38, 0.53]	0.605 [0.54, 0.68]	0.604 [0.54, 0.68]	0.479 [0.407, 0.56]	0.659 [0.60, 0.73]		
Cheek Nose Size Proportion	0.547 [0.47, 0.63]	0.549 [0.48, 0.63]	0.547 [0.48, 0.63]	0.551 [0.49, 0.63]	0.556 [0.49, 0.64]	0.559 [0.49, 0.64]	0.506 [0.43, 0.59]	0.514 [0.44, 0.60]
	0.512 [0.44, 0.59]	0.522 [0.45, 0.60]	0.520 [0.45, 0.60]	0.529 [0.46, 0.61]	0.574 [0.51, 0.65]	0.607 [0.55, 0.63]	0.576 [0.51, 0.65]	0.610 [0.55, 0.68]
	0.579 [0.51, 0.65]	0.614 [0.55, 0.69]						
Cranio-Topline Length Ratio	0.503 [0.43, 0.58]	0.452 [0.38, 0.53]	0.501 [0.43, 0.58]	0.486 [0.42, 0.56]	0.316 [0.24, 0.39]	0.493 [0.42, 0.57]	0.473 [0.40, 0.55]	0.422 [0.35, 0.50]
	0.474 [0.41, 0.55]	0.460 [0.39, 0.54]	0.291 [0.22, 0.37]	0.464 [0.39, 0.54]				
Canthus Width-to-Height Ratio	0.346 [0.27, 0.43]	0.282 [0.21, 0.36]	0.00 [0.00, 0.00]	0.352 [0.28, 0.43]	0.287 [0.21, 0.36]			
Eye-Cranio Size Ratio	0.708 [0.66, 0.77]	0.686 [0.63, 0.76]	0.688 [0.63, 0.76]	0.703 [0.65, 0.77]	0.686 [0.63, 0.75]	0.693 [0.64, 0.76]	0.667 [0.61, 0.74]	0.670 [0.61, 0.74]
	0.687 [0.63, 0.76]	0.667 [0.61, 0.74]	0.680 [0.62, 0.75]	0.653 [0.60, 0.73]	0.654 [0.60, 0.73]	0.674 [0.62, 0.74]	0.655 [0.60, 0.73]	0.708 [0.66, 0.78]
	0.686 [0.63, 0.76]	0.687 [0.63, 0.76]	0.704 [0.65, 0.77]	0.685 [0.63, 0.76]	0.693 [0.64, 0.76]	0.667 [0.61, 0.74]	0.669 [0.61, 0.74]	0.688 [0.63, 0.76]
	0.666 [0.61, 0.74]	0.681 [0.63, 0.75]	0.654 [0.60, 0.73]	0.654 [0.60, 0.73]	0.675 [0.62, 0.75]	0.655 [0.60, 0.75]		
Eye-Forehead Size Ratio – Linear	0.648 [0.59, 0.72]	0.643 [0.58, 0.72]	0.639 [0.58, 0.71]	0.646 [0.59, 0.72]	0.645 [0.58, 0.72]			
Eye-Forehead Size Ratio – Poly	0.681 [0.63, 0.75]	0.669 [0.61, 0.75]	0.671 [0.61, 0.74]	0.679 [0.62, 0.75]	0.668 [0.61, 0.74]	0.675 [0.62, 0.74]	0.662 [0.60, 0.73]	0.666 [0.61, 0.74]
	0.674 [0.62, 0.73]	0.660 [0.60, 0.73]	0.672 [0.61, 0.74]	0.659 [0.60, 0.73]	0.663 [0.61, 0.74]	0.670 [0.61, 0.74]	0.658 [0.60, 0.73]	0.683 [0.63, 0.75]
	0.672 [0.61, 0.74]	0.672 [0.62, 0.74]	0.682 [0.63, 0.75]	0.669 [0.61, 0.74]	0.677 [0.62, 0.75]	0.665 [0.61, 0.74]	0.667 [0.61, 0.74]	0.677 [0.62, 0.75]
	0.662 [0.60, 0.73]	0.675 [0.62, 0.75]	0.663 [0.60, 0.73]	0.665 [0.61, 0.74]	0.674 [0.62, 0.75]	0.660 [0.60, 0.73]		
Eye Orbital Height-to-Length Ratio	0.707 [0.65, 0.77]	0.474 [0.40, 0.56]	0.471 [0.40, 0.55]	0.442 [0.37, 0.52]	0.480 [0.41, 0.56]	0.470 [0.40, 0.55]	0.442 [0.37, 0.52]	0.693 [0.64, 0.76]
	0.453 [0.38, 0.53]	0.439 [0.37, 0.52]	0.424 [0.35, 0.52]	0.469 [0.40, 0.55]	0.445 [0.37, 0.53]	0.441 [0.37, 0.52]		
Eye Orbital-Eye Height Ratio	0.712 [0.66, 0.78]	0.553 [0.49, 0.63]	0.526 [0.45, 0.61]	0.098 [0.04, 0.14]	0.561 [0.49, 0.64]	0.529 [0.46, 0.61]	0.110 [0.05, 0.15]	0.635 [0.57, 0.72]
	0.528 [0.46, 0.61]	0.508 [0.44, 0.59]	0.110 [0.05, 0.15]	0.552 [0.48, 0.63]	0.518 [0.45, 0.60]	0.121 [0.06, 0.16]	0.710 [0.66, 0.78]	0.553 [0.49, 0.63]
	0.526 [0.45, 0.61]	0.091 [0.04, 0.13]	0.561 [0.49, 0.64]	0.528 [0.46, 0.61]	0.103 [0.04, 0.14]	0.691 [0.64, 0.76]	0.529 [0.46, 0.61]	0.507 [0.43, 0.59]
	0.103 [0.04, 0.14]	0.552 [0.48, 0.63]	0.517 [0.44, 0.60]	0.113 [0.05, 0.15]	0.709 [0.66, 0.78]	0.553 [0.48, 0.63]	0.526 [0.45, 0.61]	0.091 [0.03, 0.13]
	0.560 [0.49, 0.64]	0.528 [0.46, 0.61]	0.103 [0.04, 0.14]	0.689 [0.64, 0.76]	0.528 [0.46, 0.61]	0.507 [0.43, 0.59]	0.102 [0.04, 0.14]	0.551 [0.48, 0.63]
	0.517 [0.44, 0.60]	0.112 [0.05, 0.15]						
Eye Orbital Projection Proportion	0.654 [0.59, 0.73]	0.650 [0.59, 0.72]	0.651 [0.59, 0.72]	0.611 [0.55, 0.69]	0.616 [0.55, 0.69]	0.611 [0.55, 0.69]	0.672 [0.62, 0.74]	0.672 [0.62, 0.74]
	0.659 [0.60, 0.73]							
Eye Orbital Roundness Proportion	0.396 [0.32, 0.48]	0.267 [0.20, 0.34]						
Eye Orbital Roundness Point Proportion	0.267 [0.20, 0.34]	0.394 [0.32, 0.47]						
Eye Orbital Thickness Proportion – Poly	0.660 [0.60, 0.73]	0.610 [0.54, 0.69]	0.569 [0.50, 0.65]	0.631 [0.57, 0.71]	0.596 [0.53, 0.68]	0.559 [0.49, 0.64]	0.527 [0.45, 0.61]	0.507 [0.43, 0.59]
	0.548 [0.48, 0.63]	0.547 [0.47, 0.63]						
Eye Sinus Size Ratio – Linear	0.635 [0.57, 0.71]	0.619 [0.55, 0.70]	0.622 [0.56, 0.70]	0.622 [0.56, 0.70]	0.628 [0.57, 0.70]	0.450 [0.37, 0.53]	0.427 [0.35, 0.51]	0.414 [0.33, 0.49]
	0.446 [0.37, 0.53]	0.428 [0.35, 0.51]						
Eye Sinus Size Ratio – Poly	0.705 [0.65, 0.77]	0.704 [0.65, 0.77]	0.708 [0.66, 0.78]	0.701 [0.65, 0.77]	0.711 [0.66, 0.78]	0.707 [0.65, 0.77]	0.705 [0.65, 0.77]	0.709 [0.66, 0.78]
	0.702 [0.65, 0.77]	0.712 [0.66, 0.78]	0.713 [0.66, 0.78]	0.711 [0.66, 0.78]	0.714 [0.66, 0.78]	0.708 [0.65, 0.78]	0.718 [0.67, 0.78]	0.678 [0.62, 0.75]
	0.678 [0.62, 0.75]	0.677 [0.62, 0.75]	0.670 [0.61, 0.74]	0.685 [0.63, 0.76]	0.678 [0.62, 0.75]	0.678 [0.62, 0.75]	0.676 [0.62, 0.75]	0.670 [0.61, 0.74]
	0.684 [0.63, 0.76]	0.680 [0.62, 0.75]	0.679 [0.62, 0.75]	0.676 [0.62, 0.75]	0.670 [0.61, 0.74]	0.684 [0.63, 0.76]		
Eye-Topline Size Ratio – Linear	0.840 [0.81, 0.88]	0.768 [0.72, 0.83]	0.655 [0.59, 0.73]	0.813 [0.76, 0.86]	0.792 [0.75, 0.85]			
Eye-Topline Size Ratio – Poly	0.791 [0.75, 0.85]	0.772 [0.73, 0.83]	0.787 [0.75, 0.84]	0.783 [0.74, 0.84]	0.779 [0.74, 0.83]	0.770 [0.73, 0.83]	0.747 [0.70, 0.81]	0.759 [0.71, 0.82]
	0.761 [0.72, 0.82]	0.754 [0.71, 0.81]	0.749 [0.70, 0.81]	0.727 [0.68, 0.79]	0.729 [0.68, 0.79]	0.740 [0.69, 0.80]	0.733 [0.68, 0.80]	

Table 23: Repeatability of Geometric Forehead and Jaw Biometrics Averaged Over Two Replicates of Landmark Coordinate Extraction – Part B

	Version 1/9/17/25/33	Version 2/10/18/26/34	Version 3/11/19/27/35	Version 4/12/20/28/36	Version 5/13/21/29/37	Version 6/14/22/30/38	Version 7/15/23/31/39	Version 8/16/24/32/40
Forehead-Eye Angle – Slope	0.505 [0.43, 0.59]	0.448 [0.37, 0.53]	0.565 [0.49, 0.65]	0.494 [0.42, 0.58]	0.447 [0.37, 0.53]	0.565 [0.49, 0.65]	0.474 [0.40, 0.55]	0.430 [0.35, 0.51]
	0.562 [0.49, 0.64]	0.464 [0.39, 0.55]	0.429 [0.35, 0.51]	0.560 [0.49, 0.64]	0.480 [0.40, 0.56]	0.431 [0.35, 0.51]	0.564 [0.49, 0.65]	0.470 [0.39, 0.55]
	0.430 [0.35, 0.51]	0.562 [0.49, 0.64]	0.499 [0.42, 0.58]	0.447 [0.37, 0.53]	0.565 [0.49, 0.65]	0.487 [0.41, 0.57]	0.445 [0.37, 0.53]	0.564 [0.49, 0.65]
Forehead-Jaw Angle – Slope	0.652 [0.59, 0.73]	0.638 [0.58, 0.71]	0.591 [0.52, 0.67]	0.641 [0.58, 0.72]	0.633 [0.57, 0.71]	0.594 [0.53, 0.67]	0.646 [0.59, 0.72]	0.635 [0.57, 0.71]
	0.587 [0.52, 0.67]	0.636 [0.57, 0.71]	0.630 [0.57, 0.71]	0.590 [0.52, 0.67]				
Forehead-Poll Length Ratio	0.590 [0.52, 0.67]	0.580 [0.51, 0.66]	0.591 [0.52, 0.67]	0.599 [0.53, 0.68]	0.173 [0.11, 0.23]	0.594 [0.53, 0.67]	0.567 [0.49, 0.64]	0.557 [0.48, 0.63]
	0.567 [0.48, 0.64]	0.569 [0.50, 0.65]	0.042 [0.00, 0.07]	0.566 [0.50, 0.65]				
Forehead-Topline Angle – Slope	0.700 [0.65, 0.77]	0.414 [0.33, 0.49]	0.632 [0.57, 0.71]	0.677 [0.62, 0.75]	0.404 [0.32, 0.48]	0.633 [0.57, 0.71]		
Forehead-Topline Length Ratio	0.547 [0.48, 0.63]	0.527 [0.45, 0.61]	0.545 [0.47, 0.63]	0.490 [0.42, 0.57]	0.470 [0.40, 0.55]	0.489 [0.42, 0.57]		
Forehead Temple Ratio	0.570 [0.50, 0.65]	0.572 [0.50, 0.65]	0.574 [0.50, 0.66]	0.571 [0.50, 0.65]	0.571 [0.50, 0.66]	0.572 [0.50, 0.66]	0.574 [0.51, 0.66]	0.576 [0.51, 0.66]
	0.573 [0.50, 0.66]	0.574 [0.50, 0.66]	0.554 [0.48, 0.64]	0.556 [0.48, 0.64]	0.557 [0.49, 0.64]	0.555 [0.48, 0.64]	0.555 [0.48, 0.64]	0.517 [0.44, 0.60]
	0.519 [0.45, 0.60]	0.520 [0.45, 0.60]	0.518 [0.44, 0.60]	0.518 [0.44, 0.60]	0.518 [0.44, 0.60]	0.520 [0.45, 0.60]	0.521 [0.45, 0.60]	0.519 [0.45, 0.60]
	0.519 [0.45, 0.60]	0.502 [0.43, 0.58]	0.503 [0.43, 0.59]	0.504 [0.43, 0.59]	0.503 [0.43, 0.58]	0.503 [0.43, 0.58]		
Forehead Width-to-Length Ratio	0.619 [0.56, 0.70]	0.588 [0.52, 0.67]	0.588 [0.52, 0.67]					
	0.588 [0.52, 0.67]	0.588 [0.52, 0.67]	0.588 [0.52, 0.67]	0.588 [0.52, 0.67]	0.543 [0.47, 0.62]	0.543 [0.47, 0.62]	0.543 [0.47, 0.62]	0.542 [0.47, 0.62]
	0.542 [0.47, 0.62]	0.542 [0.47, 0.62]						
Forehead-Zygomatic Angle	0.042 [0.00, 0.07]	0.049 [0.00, 0.08]	0.038 [0.00, 0.07]	0.041 [0.00, 0.07]	0.042 [0.00, 0.07]	0.049 [0.00, 0.08]	0.038 [0.00, 0.07]	0.041 [0.00, 0.07]
	0.042 [0.00, 0.07]	0.049 [0.00, 0.08]	0.038 [0.00, 0.07]	0.041 [0.00, 0.07]				
Jaw Angle – Slope	0.607 [0.55, 0.68]	0.607 [0.55, 0.68]	0.607 [0.55, 0.68]	0.598 [0.54, 0.67]	0.598 [0.54, 0.67]	0.599 [0.54, 0.67]		
Jowl-Jaw Length Proportion	0.521 [0.45, 0.60]	0.547 [0.48, 0.63]	0.555 [0.49, 0.64]	0.548 [0.48, 0.63]	0.537 [0.47, 0.62]	0.556 [0.49, 0.64]	0.574 [0.51, 0.65]	0.563 [0.50, 0.60]
Jaw Length Proportion	0.582 [0.52, 0.66]	0.581 [0.52, 0.66]	0.586 [0.52, 0.66]	0.592 [0.53, 0.67]				
Jaw-Midface Size Ratio	0.515 [0.45, 0.58]	0.511 [0.44, 0.58]	0.503 [0.44, 0.57]	0.498 [0.43, 0.57]	0.410 [0.35, 0.48]	0.409 [0.34, 0.48]		
Muzzle Size Proportion	0.612 [0.55, 0.69]	0.618 [0.56, 0.69]	0.597 [0.53, 0.67]	0.603 [0.54, 0.68]	0.556 [0.49, 0.63]	0.560 [0.49, 0.64]	0.554 [0.49, 0.53]	0.558 [0.49, 0.64]
	0.684 [0.63, 0.75]	0.680 [0.62, 0.75]	0.667 [0.61, 0.74]	0.664 [0.61, 0.74]				
Midface thickness Proportion	0.593 [0.53, 0.67]	0.504 [0.44, 0.58]						
Nasion Thickness Proportion	0.614 [0.55, 0.69]	0.610 [0.54, 0.68]	0.606 [0.54, 0.68]	0.559 [0.49, 0.64]	0.561 [0.49, 0.64]	0.570 [0.50, 0.65]	0.591 [0.53, 0.67]	0.606 [0.54, 0.68]
	0.615 [0.55, 0.69]							
Overall Eye Angle – Angle	0.645 [0.59, 0.72]	0.622 [0.56, 0.70]	0.565 [0.49, 0.65]	0.622 [0.56, 0.70]	0.646 [0.59, 0.72]			
Overall Eye Angle – Slope	0.652 [0.59, 0.72]	0.627 [0.57, 0.70]	0.572 [0.50, 0.65]	0.628 [0.57, 0.70]	0.652 [0.59, 0.73]			
Overall Eye Size	0.858 [0.83, 0.90]	0.847 [0.82, 0.89]	0.847 [0.81, 0.89]	0.852 [0.82, 0.89]	0.849 [0.82, 0.89]	0.844 [0.81, 0.89]	0.829 [0.79, 0.88]	0.833 [0.80, 0.88]
	0.838 [0.80, 0.88]	0.830 [0.80, 0.88]	0.823 [0.79, 0.87]	0.809 [0.77, 0.86]	0.809 [0.77, 0.86]	0.818 [0.78, 0.87]	0.809 [0.77, 0.86]	
Poll Depth Proportion – Height	0.470 [0.39, 0.55]	0.00 [0.00, 0.00]	0.139 [0.08, 0.20]	0.489 [0.41, 0.57]	0.001 [0.00, 0.00]	0.138 [0.08, 0.20]		
Poll Depth Proportion – Length	0.479 [0.41, 0.56]	0.066 [0.01, 0.10]	0.337 [0.26, 0.42]	0.469 [0.40, 0.55]	0.065 [0.01, 0.10]	0.579 [0.26, 0.42]		
Poll Height Proportion	0.579 [0.51, 0.66]	0.004 [0.00, 0.01]	0.366 [0.29, 0.45]	0.606 [0.54, 0.68]	0.055 [0.01, 0.09]	0.454 [0.38, 0.54]	0.576 [0.51, 0.66]	0.001 [0.00, 0.00]
	0.360 [0.29, 0.44]	0.608 [0.54, 0.69]	0.055 [0.01, 0.09]	0.447 [0.37, 0.53]				
Poll Height Point Proportion	0.447 [0.37, 0.53]	0.023 [0.00, 0.05]	0.471 [0.40, 0.55]	0.468 [0.40, 0.55]	0.283 [0.21, 0.36]	0.501 [0.43, 0.58]	0.453 [0.38, 0.53]	0.002 [0.00, 0.00]
	0.474 [0.40, 0.56]	0.471 [0.39, 0.55]	0.275 [0.20, 0.35]	0.503 [0.43, 0.58]				

Compared to the performance of eye and muzzle biometrics, largely boney forehead and jaw biometrics as a whole show superior performance in terms of repeatability, on par with that of topline biometrics (see Tables 22 & 23). Unlike with topline biometrics, selection of the coordinate location of the S landmark, the point where the eye orbital ends and forehead begins, proved equally reliable when determined using a reference line as when selected by eye. Forehead traits that used landmark C, however, still performed better when selected using the reference line than by eye. The results also clearly indicate that, among the alternative landmark points for T, the T_top landmark clearly underperformed. This is likely partially attributable to the fact that the T_top landmark was partly influenced by the hair of the poll, which is a seasonally variable non-boney trait. T_back tended to outperform T_poll where these two traits were compared, but T_poll and T_slope showed comparable performance, with some biometric performing better with one over the other. Interestingly, selecting landmark L, the point where the chin meets the jaw, proved more robust when selected from the entire image as compared to the muzzle subframe.

Fifth is to explore the correlation structures between topline biometrics (see Figure 33). Based on the results of the repeatability analysis, metrics retained were versions that utilized landmarks aa, ba, cc, dc, C_extrap, S_extrap. Where a version offered a choice between T_slope and T_poll, the version that either had notably superior repeatability, or else a higher proportion of error attributed to between-day error was selected for use. In comparing the distributions of pairwise correlations between observed biometrics, the density of normalized length metrics were clearly shifted right. This corresponded to an average correlation between pairwise correlations of 0.12, as compared to only 0.07 for geometric biometrics. Though both measurement systems ultimately show modest correlation between metrics, geometric biometrics do demonstrate a slight advantage over normalized length for forehead traits.

Table 24: Pairwise Correlation between Geometric Forehead and Jaw Biometrics

	Canthus.Depth.Proportion_Depth_V3	Canthus.Depth.Proportion_Length_V3	Canthus.Length.Proportion_V3	Canthus.Width.to.Height.Ratio_V3	Cheek.Nose.Size.Proportion_V18	Chin.Length.Proportion_V6	Cranio.Topline.Length.Ratio_V2	Eye.Cranio.Size.Ratio_Poly_V3	Eye.Forehead.Size.Ratio_Linear_V1	Eye.Forehead.Size.Ratio_Poly_V3	Eye.Orbital.Height.to.Length.Ratio_V1	Eye.Orbital.Projection.Proportion_V9	Eye.Orbital.Roundness.Proportion_V1	Eye.Orbital.Roundness.Point.Proportion_V1	Eye.Orbital.Thickness.Proportion_Poly_V5	Eye.Orbital.Eye.Height.Ratio_V36	Eye.Topline.Size.Ratio_Linear_V1	Eye.Topline.Size.Ratio_Poly_V3	Eye.Sinus.Size.Ratio_Linear_V1	Eye.Sinus.Size.Ratio_Poly_V3	Forehead.Eye.Angle_Slope_V3	Forehead.Eye.Angle_Slope_V2	Forehead.Jaw.Angle_Slope_V2	Forehead.Template.Ratio_V3	Forehead.Topline.Angle_Slope_V1	Forehead.Topline.Length.Ratio_V3	Forehead.Width.to.Length.Ratio_V3	Forehead.Zygomatic.Angle_V2	Forehead.Poll.Length.Ratio_V2	Jaw.Angle_Slope_V1	Jaw.Midface.Size.Ratio_V1	Jowel.Jaw.Length.Proportion_V6	Jaw.Length.Proportion_V3	Midface.Thickness.Proportion_V1	Muzzle.Size.Proportion_V9	Nasion.Thickness.Proportion_V7	Overall.Eye.Angle_V1	Overall.Eye.Angle_Slope_V1	Overall.Eye.Size_V3	Poll.Depth.Proportion_Height_V3	Poll.Depth.Proportion_Length_V3	Poll.Height.Point.Proportion_V4	Poll.Height.Proportion_V4
Canthus.Depth.Proportion_Depth_V3	1.00	0.83	0.05	0.15	0.00	-0.04	-0.01	-0.22	-0.14	-0.15	-0.12	0.23	0.13	0.08	0.25	-0.02	-0.16	-0.11	0.25	-0.13	-0.27	0.08	-0.11	0.14	-0.03	-0.24	0.04	-0.03	-0.09	-0.12	-0.23	0.11	-0.22	0.01	0.11	0.42	0.42	-0.15	0.03	0.09	0.03	0.08	
Canthus.Depth.Proportion_Length_V3	0.83	1.00	0.16	0.18	-0.03	0.02	-0.03	-0.16	-0.13	-0.12	-0.11	0.16	0.11	0.08	0.21	-0.02	-0.16	-0.08	0.19	-0.09	-0.15	0.04	-0.10	0.19	-0.04	-0.13	0.06	-0.03	-0.13	-0.10	-0.22	0.13	-0.18	0.02	0.08	0.31	0.32	-0.09	0.02	0.06	0.02	0.07	
Canthus.Length.Proportion_V3	0.05	0.16	1.00	-0.08	-0.05	-0.01	-0.08	-0.12	-0.11	-0.14	-0.13	0.26	-0.07	0.13	0.18	-0.10	-0.17	-0.23	-0.34	0.33	0.33	0.17	-0.05	0.11	0.01	0.26	-0.01	0.05	0.02	0.14	0.09	0.00	0.19	-0.12	0.28	-0.19	-0.19	-0.20	-0.03	0.00	0.01	-0.01	
Canthus.Width.to.Height.Ratio_V3	0.15	0.18	-0.08	1.00	0.05	-0.10	-0.01	-0.06	-0.08	-0.06	-0.04	0.01	-0.05	0.04	0.02	-0.03	-0.08	-0.04	-0.01	0.00	0.02	0.02	-0.05	0.06	-0.01	0.02	-0.01	-0.01	-0.01	0.00	0.01	-0.04	-0.03	-0.03	-0.08	0.03	0.02	-0.08	0.01	0.02	0.01	0.00	
Cheek.Nose.Size.Proportion_V18	0.00	-0.03	-0.05	0.05	1.00	-0.47	0.15	-0.06	-0.33	-0.21	0.25	-0.09	0.15	0.06	0.16	0.34	0.06	0.30	0.32	-0.38	0.02	0.34	-0.20	-0.37	-0.04	-0.06	0.06	-0.16	-0.50	-0.13	-0.21	0.49	0.20	0.34	-0.20	-0.27	-0.26	0.61	-0.10	-0.06	0.01	0.05	
Chin.Length.Proportion_V6	-0.04	0.02	-0.01	-0.10	-0.47	1.00	0.17	0.18	0.60	0.46	0.36	-0.22	0.21	0.00	-0.06	0.35	0.37	0.42	0.01	-0.11	-0.07	-0.65	-0.37	-0.37	-0.04	-0.06	0.06	-0.16	-0.50	-0.13	-0.21	0.49	0.20	0.34	-0.20	-0.27	-0.26	0.61	-0.10	-0.06	0.01	0.05	
Cranio.Topline.Length.Ratio_V2	-0.01	-0.03	-0.08	-0.01	0.15	0.17	1.00	-0.50	-0.28	-0.44	0.36	-0.07	-0.12	-0.04	-0.23	0.21	0.21	0.25	0.04	-0.01	-0.03	-0.41	0.24	0.17	0.59	0.25	0.00	0.09	-0.40	-0.12	0.10	0.69	0.12	0.14	-0.12	-0.09	-0.09	0.02	-0.12	0.05	0.49	0.18	
Eye.Cranio.Size.Ratio_Poly_V3	-0.22	-0.16	-0.12	-0.06	-0.06	0.18	-0.50	1.00	0.65	0.80	0.08	-0.41	0.30	-0.02	-0.10	0.19	0.55	0.47	0.24	-0.32	0.03	-0.03	-0.09	-0.12	-0.47	-0.24	-0.11	-0.25	0.18	-0.01	-0.02	-0.19	-0.02	0.20	-0.37	-0.19	-0.19	0.65	-0.04	-0.29	0.10	-0.21	
Eye.Forehead.Size.Ratio_Linear_V1	-0.14	-0.13	-0.11	-0.08	-0.33	0.60	-0.28	0.65	1.00	0.90	0.32	-0.22	0.30	-0.09	-0.03	0.33	0.65	0.51	0.11	-0.20	-0.12	-0.45	0.38	-0.45	-0.29	-0.36	-0.04	-0.17	-0.14	0.06	0.03	-0.02	0.24	0.32	-0.16	-0.28	-0.28	0.83	-0.17	-0.19	-0.26	0.04	
Eye.Forehead.Size.Ratio_Poly_V3	-0.15	-0.12	-0.14	-0.06	-0.21	0.46	-0.44	0.80	0.90	1.00	0.26	-0.44	0.34	0.03	-0.09	0.33	0.56	0.53	0.18	-0.34	-0.08	0.20	0.08	-0.39	-0.52	-0.31	-0.07	-0.35	0.06	-0.02	0.04	-0.16	0.07	0.32	-0.37	-0.27	-0.26	0.81	-0.14	-0.27	-0.22	-0.12	
Eye.Orbital.Height.to.Length.Ratio_V1	-0.12	-0.11	-0.13	-0.04	0.25	0.36	0.36	0.08	0.32	0.26	1.00	-0.32	0.24	0.04	0.05	0.08	0.88	0.35	0.59	0.15	-0.42	0.01	-0.29	0.32	-0.49	0.17	0.05	0.02	0.10	0.48	-0.42	-0.47	-0.46	0.40	-0.11	-0.10	-0.08	0.04	-0.11	0.02	0.08		
Eye.Orbital.Projection.Proportion_V9	0.23	0.16	0.26	0.01	-0.09	-0.22	-0.07	-0.41	0.30	0.30	0.34	0.24	-0.14	-0.14	-0.01	0.64	-0.24	-0.19	-0.44	-0.08	0.35	-0.13	0.20	-0.11	0.36	0.01	-0.22	-0.01	0.05	-0.05	0.24	-0.13	0.03	0.20	-0.27	0.76	0.41	-0.39	0.08	0.17	-0.01	0.18	
Eye.Orbital.Roundness.Proportion_V1	0.13	0.11	-0.07	-0.05	0.15	0.21	-0.12	0.30	0.30	0.34	0.24	-0.14	0.10	-0.10	0.01	0.41	0.43	0.28	0.43	0.42	-0.56	-0.22	-0.16	-0.03	-0.13	-0.21	-0.29	0.16	-0.18	-0.02	-0.29	-0.11	0.17	-0.29	0.32	-0.17	0.11	0.11	0.37	-0.13	-0.14	-0.01	-0.05
Eye.Orbital.Roundness.Point.Proportion_V1	0.08	0.08	0.13	-0.04	0.06	0.00	-0.04	-0.02	-0.09	-0.03	0.05	-0.01	-0.10	0.10	0.06	0.09	-0.08	0.02	0.03	-0.06	0.05	0.19	-0.11	0.12	-0.07	0.04	-0.08	-0.07	0.10	-0.10	-0.04	0.00	-0.10	-0.01	0.00	0.00	0.00	-0.04	-0.05	-0.01	0.04	0.12	
Eye.Orbital.Thickness.Proportion_Poly_V5	0.25	0.21	0.18	0.02	0.16	-0.06	-0.23	-0.10	-0.03	-0.09	-0.08	0.64	0.41	0.06	1.00	0.02	-0.04	0.00	0.31	-0.30	-0.22	0.13	-0.27	0.20	-0.26	-0.34	0.08	-0.20	-0.05	-0.16	-0.24	0.20	-0.24	0.04	0.42	0.39	0.39	-0.08	0.03	0.02	0.05	0.02	
Eye.Orbital.Eye.Height.Ratio_V36	-0.02	-0.02	-0.10	-0.03	0.34	0.35	0.21	0.19	0.33	0.33	0.88	-0.24	0.43	0.09	0.22	1.00	0.40	0.72	0.40	-0.64	-0.03	-0.28	0.12	-0.40	-0.02	-0.04	0.15	-0.20	-0.05	-0.52	0.08	0.35	-0.32	0.52	-0.45	-0.32	-0.31	0.45	-0.12	-0.15	0.16	-0.05	
Eye.Topline.Size.Ratio_Linear_V1	-0.16	-0.16	-0.17	-0.08	0.36	0.37	0.21	0.55	0.65	0.56	0.35	-0.19	0.28	-0.08	-0.04	0.40	1.00	0.80	0.37	-0.33	-0.10	-0.38	0.20	-0.35	-0.02	-0.27	0.03	-0.16	-0.09	-0.02	0.18	0.36	0.23	0.37	-0.25	-0.18	-0.18	0.83	-0.19	-0.23	0.08	-0.07	
Eye.Topline.Size.Ratio_Poly_V3	-0.11	-0.08	-0.23	-0.04	0.30	0.42	0.25	0.47	0.51	0.53	0.59	-0.44	0.43	0.02	0.00	0.72	0.80	1.00	0.56	-0.66	-0.15	-0.38	0.14	-0.43	-0.02	-0.19	0.09	-0.20	-0.08	-0.48	0.15	0.44	-0.20	0.54	-0.55	-0.21	-0.21	0.78	-0.13	-0.20	0.14	-0.10	
Eye.Sinus.Size.Ratio_Linear_V1	0.25	0.19	-0.34	-0.01	0.32	0.01	0.04	0.24	0.11	0.18	0.15	-0.08	0.42	0.03	0.31	0.40	0.37	0.56	1.00	-0.84	-0.56	-0.02	-0.12	-0.01	-0.08	-0.50	-0.03	-0.13	-0.05	-0.62	-0.34	0.29	-0.70	0.25	-0.27	-0.55	0.55	0.32	0.02	-0.03	0.11	-0.04	
Eye.Sinus.Size.Ratio_Poly_V3	-0.13	-0.09	0.33	0.00	-0.38	-0.11	-0.01	-0.32	-0.20	-0.34	-0.42	0.35	-0.56	-0.06	-0.30	-0.64	-0.33	-0.66	-0.84	1.00	0.43	0.09	0.12	0.16	0.17	0.40	-0.03	0.22	0.01	0.64	0.24	-0.27	0.70	-0.40	0.45	-0.29	-0.29	-0.42	0.04	0.12	-0.12	0.08	
Forehead.Eye.Angle_Slope_V3	-0.27	-0.15	0.33	0.02	0.02	-0.07	-0.03	0.03	-0.12	-0.08	0.01	-0.13	-0.22	0.05	-0.22	-0.03	-0.10	-0.15	-0.56	0.40	1.00	0.13	-0.07	0.20	0.00	0.81	0.08	0.03	-0.06	0.16	0.28	-0.01	0.27	-0.15	-0.15	-0.68	-0.68	-0.20	-0.17	-0.19	0.07	-0.23	
Forehead.Jaw.Angle_Slope_V2	0.08	0.04	0.17	0.02	0.34	-0.65	-0.41	0.03	-0.45	-0.20	-0.29	-0.20	-0.16	0.19	0.13	-0.28	-0.38	-0.38	-0.02	0.09	0.13	1.00	-0.47	0.48	-0.19	0.05	-0.35	0.03	0.56	0.06	0.12	-0.48	-0.19	-0.24	0.17	0.22	0.21	-0.47	0.19	0.04	-0.15	-0.12	
Forehead.Template.Ratio_V3	-0.11	-0.10	-0.05	-0.05	-0.20	0.37	0.24	-0.09	0.38	0.08	0.32	-0.11	-0.03	-0.11	-0.27	0.12	0.20	0.14	-0.12	0.12	0.07	-0.47	1.00	-0.34	0.73	-0.01	-0.03	0.74	-0.24	0.00	0.12	0.10	0.23	0.11	0.03	0.28	-0.28	0.24	0.01	0.27	0.48	-0.42	
Forehead.Topline.Angle_Slope_V1	0.14	0.19	0.11	0.06	-0.07	-0.37	-0.17	-0.12	-0.45	-0.39	-0.49	0.36	-0.13	0.12	0.20	-0.40	-0.35	-0.43	-0.01	0.16	0.20	0.48	-0.34	1.00	-0.07	0.15	-0.04	0.01	-0.19	0.11	-0.32	0.03	-0.08	-0.40	0.24	0.39	0.39	-0.43	0.17	0.08	0.01	0.24	
Forehead.Topline.Length.Ratio_V3	-0.03	-0.04	-0.01	-0.13	-0.04	0.59	0.47	-0.29	-0.52	0.17	0.01	-0.21	-0.07	-0.26	0.02	-0.02	-0.08	0.17	0.00	-0.19	0.73	-0.07	1.00	0.22	0.00	0.84	-0.14	0.09	0.17	0.26	0.07	0.05	0.07	-0.10	-0.20	0.10	0.36	-0.21	0.37				
Forehead.Width.to.Length.Ratio_V3	-0.24	-0.13	0.26	0.02	0.09	-0.06	0.25	-0.24	-0.36	-0.31	0.05	-0.22	-0.29	0.04	-0.34	-0.04	-0.27	-0.19	-0.50	0.40	0.81	0.05	-0.01	0.15	0.22	1.00	0.05	0.10	-0.16	0.03	0.26	0.11	0.14	-0.11	-0.23	-0.61	-0.60	-0.35	0.12	0.16	0.22	-0.12	
Forehead.Zygomatic.Angle_V2	0.04	0.06	-0.01	-0.01	0.05	0.06	0.00	-0.11	-0.04	-0.07	0.02	-0.01	0.16	-0.08	0.0																												

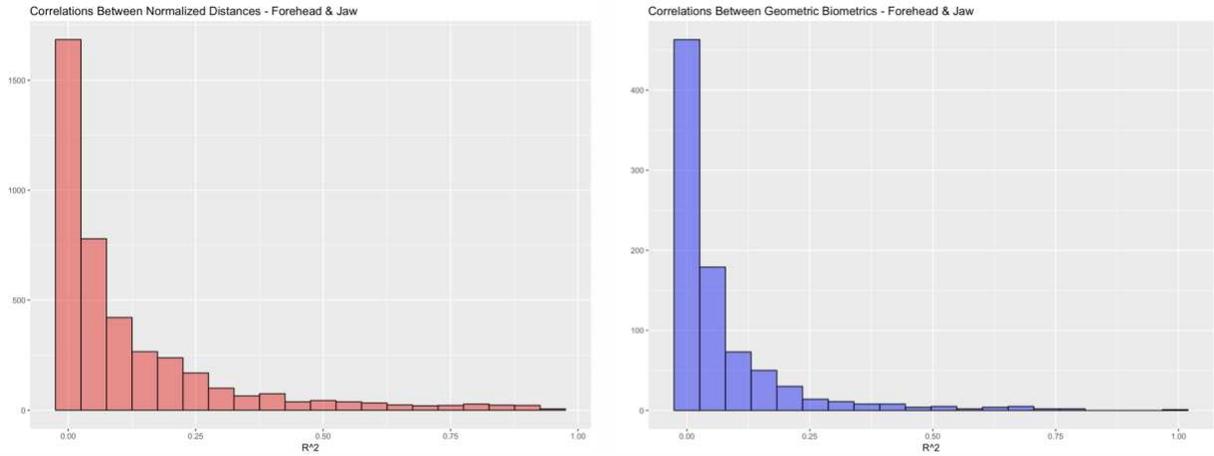


Figure 33: Comparing level of correlation between pairwise combinations of normalized length and geometric forehead and jaw biometrics

Pairwise correlations amongst error terms for the two measurement systems revealed similar results to correlations between observed biometrics (see Figure 34). The density of normalized length metrics were clearly shifted right towards higher R^2 values. The overall average correlation between error terms for normalized length biometrics was 0.13, whereas for geometric biometric this value was only 0.06. Thus, geometric biometrics also hold a slight advantage over normalized length measures in terms of error structure for forehead and jaw metrics.

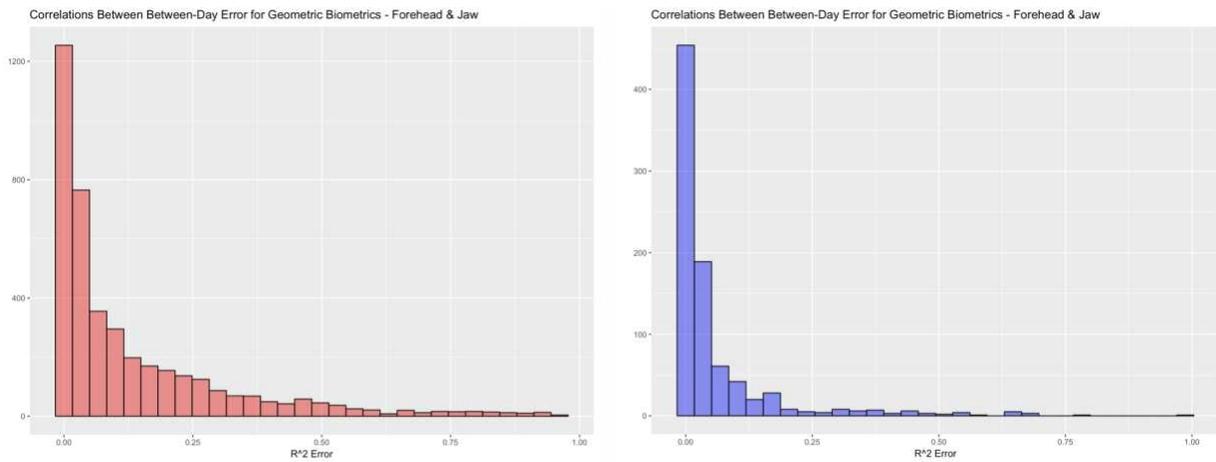


Figure 34: Comparing levels of correlation in error terms between normalized length and geometric forehead and jaw biometrics

Finally, exploring the proportion of change in biometrics between days that can be linked to changes in image attributes revealed some interesting tradeoffs between the two measurement systems for forehead and jaw biometrics (see Figure 35). With respect to attributes related to image scale and rotation, normalized length measures demonstrated a higher range of R^2 values with distribution density being shifted upwards, though geometric biometric did perhaps exhibit a thicker tail. Both measurement systems yielded average R^2 values of only 0.04, with all metrics producing correlations to changes in image attributes under 25%. With respect to attributes related to camera position, both measurement systems again produced largely negligible correlations to metric error. Here, geometric biometric occupied a higher range of R^2 values, but normalized length measures demonstrated a thicker upper tail. The average R^2 value for geometric biometric models was 0.06, and for normalized length measures this value was only 0.04, but again all metrics demonstrated correlations to changes in image attributes under 30%. At these magnitudes, it cannot really be said that either measurement system demonstrates a clear advantage in terms of resistance to variability in image quality for topline biometrics for forehead and jaw biometrics.

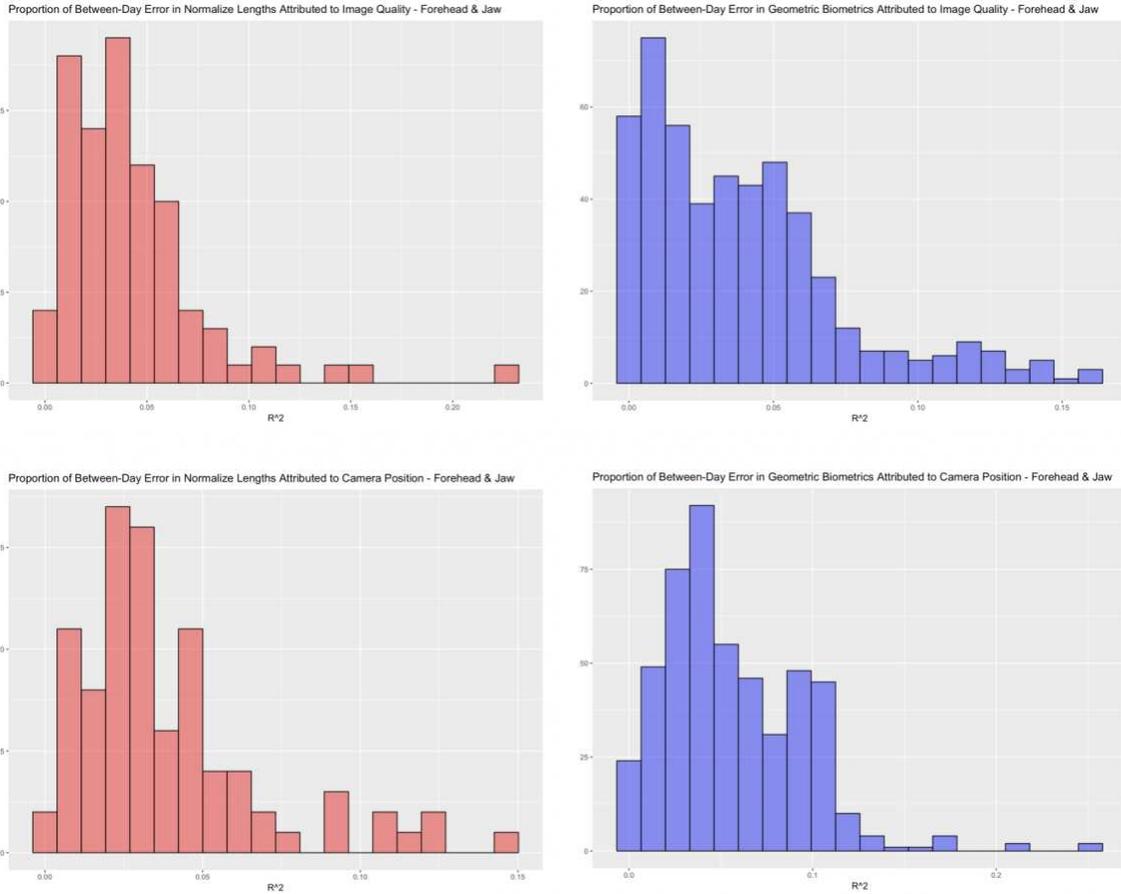


Figure 35: Above – Proportion of error attributed to variations in image scale and rotation;
 Below – Proportion of error attributed to variations in camera position

Discussion

These results show some fairly consistent trends in the comparative merit of geometric biometrics and normalized length measures. With respect to higher moments, geometric biometrics consistently show higher degrees of skew and thicker tails than do normalized length measures. For inclusion in linear models as covariates, this trend may not have much impact, but if geometric biometrics are to be included as a response variable in a linear model, for example, as an indicator trait in genetic evaluations, they will likely need to be transformed to meet assumptions of normality.

With respect to repeatability, it is clear that for both geometric biometrics and normalized length measures, measurement repeatability was quite low. When such metrics are included as covariates in linear models, this noise will bias their corresponding beta estimates towards zero, which may mask the true merit of their contribution to prediction models, particularly if the sample size is small. When included as a response, such as in genetic evaluations, a repeated measures methodology will likely be necessary. For most measures, the majority of measurement error was attributed to errors in point selection. As landmark coordinates selected via use of anatomical reference lines consistently out-performed those selected by eye, it is possible that this source of error could be further reduced by further developments in the methodology dictating point selection. Improvements in image quality – lighting, image resolution, *etc.* – might also improve in the coordinate annotation of smaller traits of the face, particularly those of the eye and muzzle, although this path to improved measurement system performance will be at odds with practical constraints of image acquisition in variable farm environments.

For larger boney features of the face, namely topline and forehead traits, geometric biometric demonstrate a clear advantage over normalized length measures, with roughly 10% gains in both within-photo and between-photo average metric repeatability. For smaller fleshier traits, namely the eye and muzzle metrics, neither measurement system demonstrates a clear or consistent advantage. For these traits, a notable proportion of measurement error is attributed to between-photo error. While out-of-plane variations in face angle could cause this error, changes in facial expression seem a more likely culprit. Future research exploring the use of factor analysis or neural network models to tease apart the influence of structure and expression on the observed value of these traits could yield interesting results for estimation of afferent state.

With respect to correlation structures, the neither metric system demonstrated a clear advantage over the any facial region. For smaller traits in the eye and muzzle regions, observed values of geometric biometrics were slightly less correlated than normalized length traits. While this is not necessarily an advantage if downstream algorithms seek to aggregate information across traits to reduce error, it will produce less variance inflation if these metrics are added directly into linear models as covariates. For the larger boney traits, differences between measurement were largely negligible. With respect to correlation between error terms, geometric biometrics demonstrated a consistent advantage, though again this distinction was often only marginal. If covariance structures are to be leveraged in downstream analysis, this trend is essential to avoid concentration of noise.

Finally, both measurement systems appear reasonably robust to variations in image quality. For topline traits, correlations to image attributes were virtually negligible, and for forehead traits relative performance of geometric biometrics and normalized length traits were roughly equivalent. For eye biometrics, geometric biometrics significantly out-performance normalized length biometrics with respect to attributes related to image scale, though neither were notably effected by variations in camera position. On the other hand, geometric biometric showed a slight advantage with respect to variation in camera position for muzzle traits, with correlations to image scale attributed being effectively equivalent.

Conclusions

As a whole, geometric biometrics demonstrated measurement characteristics either on par or preferable to normalized length measures. While improvements were not always substantial, the efficiency with which geometric biometrics can be computed still makes this approach an

attractive alternative to normalized length biometrics. Beyond quantitative traits, however, geometric biometrics proved far more tractable conceptually, allowing interesting anomalies in the results to be more directly investigated than could be done with highly inter-connected normalized length traits. For the purposes of linear modeling, the importance of this advantage in intuition should not be undervalued in the development of insightful and robust predictive models. Thus, geometric biometrics are on par or superior to established techniques for livestock phenotype extraction.

REFERENCES

- Aldridge, K., George, I.D., Cole, K.K., Austin, J.R., Takahashi, N., Duan, Y., Miles, J.H. "Facial phenotypes in subgroups of prepubertal boys with autism spectrum disorders are correlated with clinical phenotypes." 2011. *Molecular Autism*. 2:15.
- Bates, D., Mächler, M., Bolker, B.M., Walker, S.C. "Fitting Linear Mixed-Effects Models using lme4" *Journal of Statistical Software*. 2015. 67:1.
- Beveridge, J.R. Draper, B.A., Chang, J.M., Kirby, M., Kley, H., Peterson, C. "Principal Angles Separate Subject Illumination Spaces in YDB and CMU-PIE." 2009. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 31:2. 351-356.
- Caiafa, C.F., Proto, A.N., Vergani, D., Stanganelli, Z. "Development of individual recognition of female southern elephant seals, *Mirounga leonina*, from Punta Norte Peninsula Valdes, applying principal components analysis" 2005. *Journal of Biogeography*. 32: 1257-1266.
- Cole JB, Manyama M, Kimwaga E, Mathayo J, Larson JR, Liberton DK, et al. (2016) Genomewide Association Study of African Children Identifies Association of SCHIP1 and PDE8A with Facial Size and Shape. *PLoS Genet* 12(8): e1006174. doi:10.1371/journal.pgen.1006174
- Hurlbert, Stuart. "A gentle depilation of the niche: Dicean resource sets in resource hyperspace." 1981. *Evolutionary Theory* 5:177-184.
- Johnson, R.A., Wichern, D.W. "Applied Multivariate Statistical Analysis" 6th Edition. 2007. Pearson Education Inc. USA.
- Kirby, Michael. "Geometric Data Analysis: An Empirical Approach to Dimensionality Reduction and the Study of Patterns." 2001. John Wiley & Sons Inc. USA.
- Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W. "Applied Linear Statistical Models." 5th Edition. 2005. The McGraw-Hill Companies, Inc. New York, USA.
- Lukasz Komsta and Frederick Novomestky. "moments: Moments, cumulants, skewness, kurtosis and related tests". 2015 R package version 0.14. <https://CRAN.R-project.org/package=moments>
- Measurement Systems Analysis Work Group. "Measurement System Analysis Reference Manual" 4th Edition. 2010. Chrysler Group LLC, Ford Motor Company, and General Motors Corporation. USA.

- Meyer, Karin. "Estimates of genetic parameters for weaning weight of beef cattle accounting for direct-maternal environmental covariances" 1997. *Livestock Production Science*. 52. 187-199.
- Mrode, R., Thompson, R. "Linear Models for the Prediction of Animal Breeding Values." 3rd Edition. 2014. CAB International. UK
- Nielson, H.B. Jarmer, H.O. "Genome Wide Association Study Identifying Determinants of Facial Characteristics For Facial Image Generation." 2011. World Intellectual Property Organization. PCT/DK2010/050325.
- Obafemi-Ajayi, T., Miles, J.H., Takahashi, N., Qi, W., Aldridge, K. Zhang, M. Xin, S.Q., He, Y., Duan, Y. "Facial Structure Analysis Separates Autism Spectrum Disorders into Meaningful Clinical Subgroups" 2015. *Journal of Autism and Developmental Disorders*. 45:5. 1302-1317.
- Ojiem, J.O., de Ridder, N., Vanlauwe, B., Giller, K.E. "Socio-ecological niche: a conceptual framework for integration of legumes in smallholder farming systems." 2006. *International Journal of Agricultural Sustainability*. 4:1. 79-93.
- Rashid, M., Gu, X., Lee, Y.J. "Interspecies Knowledge Transfer for Facial Keypoint Detection." 2017. Preprint. <https://arxiv.org/abs/1704.04023>
- Revelle, W. (2017) *psych: Procedures for Personality and Psychological Research*, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych/>
- RSPI Vision. "Deep Learning and Convolutional Neural Networks." 2017. Online. <https://www.rsipvision.com/exploring-deep-learning/>
- R Core Team. "R: A language and environment for statistical computing". 2016 R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Singh, Bhupendra Pratap. "Imaging Applications of Charge Coupled Devices (CCDs) for Cherenkov Telescope" 2012. Project Report Bhabha Atomic Research Centre. Online. file:///Users/catie/Downloads/Bhupendra_report.pdf
- Van Heel, Marin. "Multivariate Statistical Classification of Noisy Images (Randomly Oriented Biological Macromolecules)". 1984. *Ultramicroscopy*. 13: 165-184.
- Venables, W. N. & Ripley, B. D. "Modern Applied Statistics with S". 4th Edition. 2002. Springer, New York.

CHAPTER 2 – FACIAL BIOMETRICS AS PREDICTORS OF GENOMIC MERIT

Introduction

As highlighted in the review of existing literature, what studies have found significant relationships between behavior, health, and facial morphology have failed to elucidate a consistent biological mechanism underlying such statistical trends. Genetic, epigenetic, and environmental influences on both prenatal and neonatal development have all been proposed, and future research may yet reveal facial variation to be responsive to a combination of biological drivers. However, without a strong theoretical foundation on which to develop a targeted experimental design, a more exploratory statistical approach is necessary. The goal of this analysis was two-fold. The first was to determine what types of response variables show robust relationships to facial biometrics in intensively managed livestock populations. The second was to discern the most appropriate statistical model to employ in order to adequately mirror the underlying relationships between biometrics and any auxiliary traits relevant to the predictive model. The simplest place to start was with direct genetic effects. The overall goal of this chapter was to determine if any facial biometrics demonstrate strong and consistent associations with traits commonly reported on standard dairy genetic evaluations. This question was explored using multiple classes of models in order to systematically explore what level of statistical complexity is necessary to capture the underlying biology driving these associations.

Materials and Methods

Image Acquisition and Computation

Data for this analysis was generated from an intensively managed dairy herd consisting of 10,000 registered Holstein cattle located in Central Florida. A complete set of facial photographs consisting of both front and side profile images were generated from a total of 640 unique animals over the course of a three-day data collection period. Images were acquired using two standard quality 2D cameras: an Olympus TG-2 iHS 12MP Waterproof Camera and a Canon PowerShot SX530 HS. Roughly half of these images were acquired while the animals were standing in stocks for routine hoof care. The remaining images were acquired while animals were restrained for milking in an indoor 40 double-parallel milking parlor. Animals included in this sample represent a convenience sample of mature lactating cows based on either the accessibility of their stall position in the parlor or their eligibility for normally scheduled hoof maintenance based on their days in milk. Any animals being treated for pink eye, or who presented with apparent injuries to the face, as determined by skin lesions, scarring, or significant discoloration of the hair not consistent with the coat pattern, were excluded from this study.

Side profile images of each cow were annotated with the full set of facial landmark points (see Appendix A for details) and the coordinate locations extracted using MatLab's image analysis tools. Coordinate point extraction was repeated in two separate runs within each facial region in order to produce two independent replicates of image annotations for each individual cow. From these two sets of landmark coordinates, two independent calculations of each geometric biometric measure were computed and then averaged in order to reduce measurement error due to inconsistencies in photo annotation. Over the course of the image annotation process, several additional cows were excluded from the sample due to issues with image quality – shadows,

blurriness, physical obstruction of anatomical points – not initially identified in the process of cataloguing images, resulting in a final sample size of 594 fully annotated images. Based on the results from analyses performed in Chapter 1, the following subset of 60 geometric biometrics with superior measurement error were selected as candidate variables:

Z1_front = Eye Depth Proportion - Front Eye Length
Z2 = Eye Depth Point Proportion
Z3_front = Eye Height Proportion - Front Eye Length
Z4 = Eye Height Point Proportion
Z7_length = Eye Length Proportion - Length
Z9_poly = Eye Roundness Proportion Upper Back - Area
Z10_poly = Eye Roundness Proportion Lower Back - Area
Z11_poly = Eye Roundness Proportion Lower Front - Area
Z11_linear = Eye Roundness Proportion Lower Front - Linear
NFP_LF = Nostril Flare Proportion - Lower Front
NFPP_LF = Nostril Flare Point Proportion - Lower Front
NPA = Nostril Position Angle
ULRP_V1 = Upper Lip Roundness Proportion_V1
MTP_V1 = Mouth Thickness Proportion_V1
CTP_V1 = Chin Thickness Proportion_V1
CLTR_V1 = Chin-to-Lip Thickness Ratio_V1
NRP = Nares Roundness Proportion
SMRP_V1 = Sinus-Midface Roundness Proportion_V1
MNRP_V1 = Midface-Nose Roundness Proportion_V1
MDP_V1 = Midface Divergence Proportion_V1
NDP_V1 = Nose Divergence Proportion_V1
NaDP_V1 = Nares Divergence Proportion_V1
MIP_V1 = Midface Inflection Proportion_V1
MIPP_V1 = Midface Inflection Point Proportion_V1
NIP_V1 = Nose Inflection Proportion_V1
NaTLP = Nares-Topline Length Proportion
MTLP_V1 = Midface-Topline Length Proportion_V1
STLP_V1 = Sinus-Topline Length Proportion_V1
ULTLP_V1 = Upper-Lower Topline Length Proportion_V1
SMLP_V1 = Sinus-Midface Length Proportion_V1
MNLP_V1 = Midface-Nose Length Proportion_V1
EFSRP_V1 = Eye-Forehead Size Ratio_Poly
EFSRL_V = Eye-Forehead Size Ratio_Linear
ECSR_P_V1 = Eye-Cranio Size Ratio_Poly
ETSRP_V1 = Eye-Topline Size Ratio_Poly
ETSRL_V1 = Eye-Topline Size Ratio_Linear
ESSRP_V1 = Eye-Sinus Size Ratio_Poly
ESSRL_V1 = Eye-Sinus Size Ratio_Linear_V1

MTP_V1 = Midface Thickness Proportion_V1
OES_V1 = Overall Eye Size_V1
OEAA_V1 = Overall Eye Angle_Angle_V1
MSP_V9 = Muzzle Size Proportion_V9
CNSP_V14 = Cheek-Nose Size Proportion_V14
CLP_V6 = Chin Length Proportion_V6
JJLP_V7 = Jowl-Jaw Length Proportion_V1
JAS_V1 = Jaw Angle_Slope_V1
JMSR_V1 = Jaw Midface Size Ratio_V1
EOPP_V1 = Eye Orbital Projection Proportion_V1
NsTP_V1 = Nasion Thickness Proportion_V1
EOEHR_V1 = Eye Orbital-Eye Height Ratio_V1
EOTPP_V1 = Eye Orbital Thickness Proportion_Poly_V1
FTR_V1 = Forehead Temple Ratio_V1
FEAS_V3 = Forehead-Eye Angle_Slope_V3
FTAS_V3 = Forehead-Topline Angle_Slope_V3
FJAS_V3 = Forehead-Jaw Angle_Slope_V3
FTLR_V3 = Forehead Topline Length Ratio_V3
CTRL_V3 = Cranio-Topline Length Ratio_V3
FPLP_V3 = Forehead-Poll Length Ratio_V3
PHP_V4 = Poll Height Proportion_V4
FWLP_V3 = Forehead Width-to-Length Ratio_V3

Data Compilation

Preliminary analysis revealed the majority of these biometric traits to be reasonably normally distributed. Five biometric traits - Z3_front, Z9_poly, Z10_poly, ESSRP_V1, EOEHR_V1 - proved to be quite skewed, and were recoded as a log transform. The hat matrix for all biometric values was then computed and a leverage value for each cow plotted in a histogram. Seven animals displayed leverage values above 0.3, but they were neither notably removed from population density for leverage values, nor did closer analysis did not reveal any apparent error in the image files themselves. It was decided to retain these individuals in the analysis as unusual but not erroneous biological outliers. With this final list of candidate animals, genomically-enhanced PTA evaluation records were requested from Holstein Association USA, with permission of the farm owner. A total of 573 animals had complete genetic evaluations on record for standard PTA traits, including genomically-enhanced estimates for the following type traits.

PTA.Type = Predicted Transmitting Ability Type
UDC = Udder Composite
FLC = Feet Leg Composite
GPTA.STA = Stature
GPTA.STR = Strength
GPTA.BDE = Body Depth
GPTA.DRM = Dairy Form
GPTA.RPA = Rump Angle
GPTA.TRW = Thurl Width
GPTA.RLS = Rear Legs - Side
GPTA.RLR = Rear Legs - Rear
GPTA.FTA = Foot Angle
GPTA.FLS = Feet Leg Score
GPTA.FUA = Fore Udder Attachment
GPTA.RUH = Rear Udder Height
GPTA.RUW = Rear Udder Width
GPTA.UCL = Udder Cleft
GPTA.UDP = Udder Depth
GPTA.FTP = Front Teat Placement
GPTA.RTP = Rear Teat Placement
GPTA.TLG = Teat Length
GPTA.BSC = Body Score

Of these animals with complete PTA evaluations, 344 cows also had genomic evaluations on record for the Zoetis Clarified Plus Dairy Wellness panel, which features a range of cow and calf health traits. From this additional data, a total of 22 genetic traits were selected as candidate response variables on which to build predictive models:

CTPI = Cow Total Performance Index
NM = Net Merit Index
PTAM = Predicted Transmitting Ability Milk
PTAF = Predicted Transmitting Ability Fat
PTAP = Predicted Transmitting Ability Protein
FeedEff = Feed Efficiency
PL = Productive Life
HCR = Heifer Conception Rate
CCR = Cow Conception Rate
SCS = Somatic Cell Count
Fert.Index = Fertility Index
CE = Calving Ease
SB = Still Births
Z_Mast = Clarified Mastitis
Z_Lame = Clarified Lameness

Z_Met = Clarified Metritits
Z_RP = Clarified Retained Placenta
Z_Ket = Clarified Ketosis
Z_DA = Clarified Displace Abomasum
Z_Calf_Liv = Clarified Calf Livability
Z_Calf_Scours = Clarified Calf Scours
Z_Calf_Resp = Clarified Calf Respiratory

Data Analysis

Model building began by exploring the simplest possible relationship between facial biometrics and genetic performance traits: a simple additive linear effect. This was done by fitting LASSO models using the *glmnet* package in the R programming environment (Friedman *et al* 2010). Amongst the animals with complete PTA information, 103 individuals were randomly held out as a validation set, leaving 470 animals for model training. Amongst the cows with additional genomic evaluations for health traits, 54 animals were randomly held out as a validation set, leaving 290 individuals in the training set. First, base models for each response variable were developed using only the genomic conformation trait values. Next, a model for each response featuring only biometric candidate variables was fit. Finally, full models were fit using the full complement of both candidate biometric values and genomic PTA type trait values. For standard PTA response variables, models were weighted by the overall reliability of the performance evaluations (PREL). For genomic health traits, reliability of the health evaluation were not reported, so the overall reliability of the PTA type evaluation (TREL) was used to weight the model. The lambda level was selected using the k-fold cross validation utility built into the *glmnet* package. For the larger PTA training set, 10-fold cross-validation was employed, whereas for the smaller health training self 5-fold cross-validation was used. For all response values a grid of 10,000 lambda levels, selected internally by the cross validation utility on a log scale, was used to explore the tuning parameter space. To assess the efficacy of facial biometric parameters retained

in the model, coefficient of determination (R^2) was computed for both the full and baseline model, and the gain in variance explained due to addition of facial biometrics calculated as $R^2_{\text{Full}} - R^2_{\text{Base}}$ for each response. Fitted models were then applied to the holdout data and the R^2 values and gain in R^2 recalculated on the validation set in order to assess model stability. Finally, coefficient values for all candidate values are reported for each response model, where parameters shrunk to zero by the LASSO optimization are considered insignificant (see Appendix C). The frequency with which each individual candidate variable was retained across response model was also tracked in order to identify predictor variables that might be more broadly predictive of health and performance traits.

As a simple additive linear effect is ultimately quite a restrictive modeling assumption, the next logical step to take in assessing model adequacy was to search for nonlinear trends between facial biometrics and the genetic response variables considered here. As here the principal question is that of model structure and not predictive performance, a holdout validation set was not utilized, and all complete observations were utilized to train the model. First, the baseline model featuring only structure traits was again fit using the same LASSO modeling procedure described above, tuned via a cross validation approach, to each of the genetic response variables, with only the non-zero covariates retained as significant predictors. In order to identify significant non-linear trends, each biometric was added individually to these baseline models in the form of a penalized smoothing spline using the *mgcv* package in the R programming environment (Wood 2003; Wood 2004; Wood 2016). Each smoothing spline model was tuned using the generalized cross validation routine built into the *gam* function in the *mgcv* package, which utilizes a thin-plate spline basis. The effective degrees freedom and p-value associated with the spline component of each biometric coefficient was recorded, as was the gain in R^2 above the baseline model, and reported for each

genetic response models. Additionally, a running count of the number of significant spline components was kept across all genetic response models to again assess if any facial biometric demonstrated a more global impact on health and performance.

Finally, regression trees were used to explore the importance of interaction terms and general nonlinearity in the relationship between facial biometrics and genetic traits available in this data set using the *gbm* package in the R programming environment (Ridgeway *et al* 2017). As individual regression trees tend to overfit a model, even when pruned by cross validation, bagging was utilized to both avoid overfitting and more effectively explore the full parameter space (James *et al* 2013). For PTA traits, the same holdout data set as randomly selected in the LASSO analysis was utilized as a validation set. For the health traits, which were fewer in number, no holdout dataset was utilized and all complete observations were used to train the model. For each response variable, boosted regression trees were built for depths 1-5 out to a chain length of 7,000 using a slow learning rate ($\lambda = 0.001$). The cross-validation error was recorded for each step in these chains, with PTA traits utilizing 10-fold cross validation and the health traits utilizing 5-fold cross-validation. The optimal chain length for each tree depth was selected as the shortest chain length within 0.5% of the smallest cv-error on the entire chain, rounded to the nearest hundredths, with a minimum chain length value of 100. The optimal chain depth was then selected as the shallowest depth within 0.5% of the absolute minimum cross-validation error for all depths for the optimal chain lengths selected for each corresponding tree depth. For the final depth and chain length parameters selected, cross validation error was plotted and visually inspected for adequate convergence. Both baseline structure models and full models featuring both structure and biometric traits were fitted in this manner, and their R^2 values recorded, as well as the gain in R^2 realized from addition of biometric values over structure traits. For the optimized full model, the

top 15 covariates in terms of variable importance measures were recorded, and a running count kept across the range of genetic response models. For PTA traits, the fitted full and baseline models for each genetic trait were then applied to the holdout dataset, and the both the test R^2 value of the full model and gain in R^2 over the baseline model for the validation set were calculated and recorded.

For more details on this methodology, see code provided in Appendix C.

Results

Results of the LASSO model indicated that facial biometric values provided very little additional predictive value over standard conformation traits for this class of models (see Table 25). Gains in R^2 value above the baseline conformation trait models for training data seldom exceeded 5% of total model variance, and gains in R^2 values proved quite volatile for the validation data. The R^2 values observed for the models fit exclusively with biometric values were themselves low; however, given that they are higher than the R^2 gains seen between the full and base models, this suggests there is some overlap in the information provided by biometric and type traits. This observation is further supported by results of the coefficient values returned by the optimized LASSO results for full covariate models (see Table 26). While type traits are, as a whole, more frequently retained across LASSO models, a number of biometric values also consistently provided significant contributions to response variables. In fact, five biometrics were retained in roughly half of the response models explored in this analysis: Eye Roundness Proportion Lower Back (Area), Nostril Flare Point Proportion - Lower Front, Chin Thickness Proportion, and Midface Inflection Proportion. Additionally, closer appraisal of coefficient results also reveal models fit to the Clarified health traits to consistently retain fewer coefficient variables than the

PTA traits from the standard genomic panel. This may simply be a reflection of the smaller training set size, or perhaps indicate the presence of a greater amount of noise present in these newer and less robust genomic evaluations. It could also be an indication of lack of model sufficiency.

Table 25: LASSO Model Performance on Training and Validation Data

	R ² Full Training Model	R ² Structure Training Model	R ² Biometric Training Model	R ² Biometric Gain Training	R ² Full Test Model	R ² Biometric Gain Test
CTPI	0.629	0.589	0.215	0.039	0.423	-0.025
Net Merit	0.491	0.449	0.174	0.043	0.269	-0.035
PTA Milk	0.548	0.521	0.146	0.027	0.45	-0.044
PTA Fat	0.394	0.369	0.119	0.026	0.25	-0.034
PTA Protein	0.636	0.601	0.167	0.035	0.483	0.01
Feed Efficiency	0.555	0.512	0.148	0.044	0.396	-0.01
Productive Life	0.425	0.42	0.101	0.005	0.186	-0.048
Heifer Conception Rate	0.226	0.16	0	0.065	0.061	-0.014
Cow Conception Rate	0.246	0.201	0.06	0.045	0.047	-0.051
Somatic Cell Score	0.155	0.179	0.053	-0.024	0.044	-0.05
Fertility Index	0.287	0.256	0.076	0.032	0.063	-0.089
Calving Ease	0.201	0.194	0.069	0.007	0.074	-0.011
Still Births	0.185	0.146	0.1	0.039	0.028	-0.042
Mastitis	0.161	0.282	0	-0.121	0.092	-0.171
Lameness	0.113	0.158	0.089	-0.046	0.241	0.109
Metritis	0.031	0.117	0.16	-0.086	0.015	-0.13
Retained Placenta	0.171	0.145	0.132	0.026	0.033	-0.021
Ketosis	0.161	0.193	0.035	-0.032	0	-0.029
Displaced Abomasum	0.203	0.265	0.083	-0.062	0.113	-0.057
Calf Livability	0.05	0.046	0.024	0.003	0.01	-0.019
Calf Scours	0.157	0.144	0	0.013	0.103	0.04
Calf Respiratory	0.073	0.045	0.051	0.028	0	-0.003

Table 26: Count of Influential Variables Across Response Variables for LASSO Models

Type Coefficient	Model Inclusion Count	Biometric Coefficient	Model Inclusion Count
PTA.Type	9	Z1_front	4
UDC	0	Z2	7
FLC	1	Z3_front	7
GPTA.STA	11	Z4	0
GPTA.STR	9	Z7_length	6
GPTA.BDE	12	Z9_poly	8
GPTA.DRM	16	Z10_poly	10
GPTA.RPA	11	Z11_poly	7
GPTA.TRW	10	Z11_linear	6
GPTA.RLS	14	NFP_LF	6
GPTA.RLR	4	NFPP_LF	10
GPTA.FTA	10	NPA	9
GPTA.FLS	9	ULRP_V1	5
GPTA.FUA	7	MTP_V1	4
GPTA.RUH	11	CTP_V1	12
PTA.RUW	8	CLTR_V1	5
GPTA.UCL	7	NRP	12
GPTA.UDP	11	SMRP_V1	4
GPTA.FTP	8	MNRP_V1	6
GPTA.RTP	6	MDP_V1	7
GPTA.TLG	13	NDP_V1	5
GPTA.BSC	2	NaDP_V1	8
		MIP_V1	12
		MIPP_V1	6
		NIP_V1	8
		NaTLP	7
		MTLP_V1	6
		STLP_V1	4
		ULTLP_V1	1
		SMLP_V1	5
		MNLP_V1	4
		EFSRP_V1	3
		EFSRL_V1	2
		ECSRP_V1	7
		ETSRP_V1	5
		ETSRL_V1	4
		ESSRP_V1	1
		ESSRL_V1	7
		MTP_V1,1	4
		OES_V1	2
		OEAA_V1	5
		MSP_V9	1
		CNSP_V14	2
		CLP_V6	7
		JJLP_V7	6
		JAS_V1	5
		JMSR_V1	2
		EOPP_V1	5
		NsTP_V1	6
		EOEHR_V1	7
		EOTPP_V1	8
		FTR_V1	4
		FEAS_V3	2
		FTAS_V3	8
		FJAS_V3	1
		FTLR_V3	2
		CTLR_V3	8
		FPLP_V3	5
		PHP_V4	8
		FWLP_V3	7

Results of the spline analysis revealed that, while some biometrics do demonstrate non-linear tendencies in their associations with the various genetic traits considered here, relaxation of linear modeling assumptions did not result in major gains in R^2 values for any one individual biometric for any response model (see Table 27). Individual biometrics seldom accounted for more than a 3% gain in R^2 above the baseline structure model, which seems to agree with the findings of earlier research, that the predictive potential of individual facial traits tend to be significant but small in magnitude (Haselhuhn *et al* 2015). Among significant spline terms, the majority of these models demonstrate an effective degree freedom of one, indicating that for many of these traits linear terms are sufficient to adequately capture their relationship with the corresponding genetic trait. A number of these significant spline terms, however, demonstrate higher effective degrees freedom. Among these, many are still fairly low order splines, suggesting that there may simply be optimal levels of these biometrics for their corresponding genetic trait; however, several of these terms demonstrate higher effective degrees of freedom (>4). Additionally, there are a number of spline components in these models optimized to higher effective degrees freedom that approached but did not necessarily reach a significant p-value. This could suggest that there are multiple optimal values for these biometric values for the corresponding genetic trait, in which case conversion to a categorical response may be a more appropriate encoding method. Alternatively, this could reflect that multiple levels of a given biometric might interact synergistically with levels of other traits, in which case a simple additive model may not adequately reflect the true relationship between biometrics and these genetic values.

Looking beyond the results of the individual spline components, additional trends emerge on a more global level of this data. At the nominal alpha level of 0.05, we should expect with independent tests for 60 biometrics roughly 3 false positive spline results per response model,

though the cross-validation approach used to optimize the smoothing spline should further reduce this risk below the nominal rate. Several response models, however, returned a higher rate of significant spline terms, most notably: Heifer Conception Rate, Calving Ease, Still Birth, Feed Efficiency, Displaced Abomasum, and Calf Scours. Looking closer at these individual models, there is some agreement between significant biometric terms. Eye Roundness Proportion Lower Back (Area) was significant for both Calving Ease and Still Births, as well as Retained Placenta. Nostril Position Angle was significant for Feed Efficiency, Heifer Conception Rate, and Still Births. Mouth Thickness Proportion was significant for Feed Efficiency, Still Births, and Displaced Abomasum. Nares Roundness Proportion was significant for both Feed Efficiency and Calf Scours. Sinus-Topline Length Proportion was significant for Heifer Conception Rate and Still Births. Sinus-Midface Length Proportion was significant for Feed Efficiency, Heifer Conception Rate, and Still Births. Eye-Topline Size Ratio (Poly) was significant for Heifer Conception Rate, Calving Ease, and Displaced Abomasum. Jaw Midface Size Ratio was significant for Calving Ease and Displaced Abomasum. Forehead Temple Ratio was significant for Displaced Abomasum and Calf Scours. Forehead-Eye Angle (Slope) was predictive for Calving Ease and Still Births. Forehead-Topline Angle (Slope) was predictive for both Displaced Abomasum and Calving Ease. Among the lower-performing response models, it is also notable that Nostril Roundness Proportion was significant for all three calf traits.

Table 27: Significance of Spline Components Across Genomic Response Models

	CTPI	NM	PTAM	PTAF	PTAP	FeedEff	PL	HCR	CCR	SCS	Fertility Index	CE	SB	Mastitis	Lame	Metritis	RP	Ketosis	DA	Calf Liv.	Calf Scours	Calf Resp.	Total Significant Terms
Z1_front	0.83	0.70	0.09	0.81	0.14	0.68	0.43	0.22	0.23	0.69	0.34	0.99	0.17	0.28	0.08	0.31	0.54	0.47	0.11	0.35	0.63	0.13	0
Z2	0.69	0.84	0.08	0.58	0.43	0.79	0.62	0.46	0.17	0.68	0.29	0.21	0.03	0.54	0.16	0.30	0.73	0.72	0.21	0.40	0.74	0.48	1
Z3_front	0.08	0.07	0.94	0.15	0.02	0.01	0.19	0.99	0.33	0.29	0.60	0.12	0.28	0.61	0.87	0.60	0.07	0.94	0.10	0.30	0.28	0.65	2
Z4	0.60	0.48	0.76	0.64	0.13	0.14	0.53	0.47	0.29	0.03	0.38	0.46	0.89	0.63	0.57	0.17	0.26	0.85	0.46	0.54	0.59	0.80	1
Z7_length	0.79	0.94	0.52	0.94	0.54	0.72	0.74	0.14	0.22	0.08	0.47	0.35	0.04	0.35	0.70	0.19	0.33	0.97	0.85	0.23	0.74	0.38	1
Z9_poly	0.45	0.77	0.77	0.60	0.42	0.74	0.31	0.12	0.19	0.00	0.11	0.98	0.68	0.08	0.41	0.02	0.49	0.24	0.45	0.11	0.45	0.70	2
Z10_poly	0.25	0.38	0.26	0.45	0.04	0.34	0.26	0.36	0.17	0.72	0.16	0.00	0.02	0.51	0.04	0.31	0.04	0.86	0.85	0.62	0.40	0.98	5
Z11_poly	0.96	0.84	0.66	0.40	0.29	0.93	0.17	0.27	0.64	0.39	0.58	0.68	0.34	0.58	0.41	0.93	0.35	0.43	0.14	0.66	0.49	0.02	1
Z11_linear	0.76	0.70	0.23	0.71	0.73	0.89	0.35	0.93	0.65	0.47	0.72	0.22	0.04	0.79	0.38	0.65	0.12	0.46	0.29	0.62	0.18	0.55	1
NFP_LF	0.08	0.15	0.46	0.30	0.03	0.22	0.58	0.99	0.64	0.14	0.72	0.25	0.93	0.28	0.37	0.09	0.21	0.83	0.00	0.10	0.46	0.61	2
NFPP_LF	0.06	0.06	0.67	0.48	0.51	0.63	0.05	0.28	0.45	0.20	0.44	0.25	0.47	0.13	0.15	0.42	0.14	0.47	0.41	0.51	0.70	0.32	1
NPA	0.02	0.02	0.97	0.05	0.22	0.02	0.11	0.05	0.09	0.76	0.09	0.15	0.01	0.18	0.59	0.07	0.12	0.45	0.43	0.93	0.18	0.28	6
ULRP_V1	0.39	0.64	0.41	0.62	0.38	0.46	0.99	0.25	0.55	0.26	0.32	0.54	0.41	0.02	0.79	0.04	0.17	0.01	0.01	0.58	0.28	0.53	4
MTP_V1	0.09	0.20	0.84	0.08	0.18	0.04	0.32	0.28	0.39	0.26	0.42	0.18	0.05	0.03	0.63	0.11	0.10	0.36	0.00	0.29	0.34	0.37	4
CTP_V1	0.11	0.25	0.12	0.51	0.19	0.58	0.40	0.03	0.02	0.37	0.02	0.34	0.06	0.30	0.30	0.22	0.71	0.34	0.88	0.91	0.65	0.51	3
CLTR_V1	0.60	0.65	0.62	0.43	0.75	0.55	0.26	0.47	0.15	0.42	0.18	0.32	0.52	0.03	0.72	0.64	0.30	0.57	0.06	0.29	0.66	0.52	1
NRP	0.07	0.13	0.10	0.00	0.14	0.02	0.46	0.82	0.15	0.58	0.16	0.67	0.56	0.09	0.54	0.62	0.15	0.08	0.28	0.05	0.02	0.04	5
SMRP_V1	0.60	0.46	0.26	0.20	0.76	0.16	0.91	0.14	0.69	0.61	0.68	0.83	0.23	0.64	0.02	0.54	0.15	0.80	0.54	0.22	0.47	0.04	2
MNRP_V1	0.23	0.26	0.44	0.46	0.09	0.36	0.62	0.23	0.59	0.74	0.43	0.60	0.69	0.27	0.71	0.32	0.04	0.93	0.72	0.38	0.25	0.26	1
MDP_V1	0.72	0.96	0.92	0.84	0.55	0.77	0.96	0.23	0.41	0.52	0.31	0.66	0.01	0.91	0.17	0.68	0.15	0.66	0.12	0.71	0.87	0.06	1
NDP_V1	0.73	0.86	0.47	0.77	0.11	0.49	0.60	0.28	0.45	0.86	0.34	0.85	0.01	0.61	0.52	0.39	0.00	0.69	0.18	0.51	0.29	0.79	2
NaDP_V1	0.21	0.26	0.41	0.82	0.18	0.27	0.37	0.15	0.10	0.72	0.23	0.00	0.38	0.93	0.10	0.45	0.29	0.66	0.29	0.46	0.77	0.09	1
MIP_V1	0.11	0.21	0.04	0.53	0.36	0.43	0.53	0.15	0.11	0.56	0.10	0.46	0.29	0.87	0.76	0.12	0.59	0.39	0.01	0.77	0.33	0.69	2
MIPP_V1	0.77	0.24	0.48	0.18	0.87	0.31	1.00	0.11	0.90	0.03	1.00	0.80	0.45	0.61	0.25	0.18	0.95	0.02	0.54	1.00	0.18	0.87	2
NIP_V1	0.03	0.04	0.75	0.24	0.54	0.30	0.02	0.11	0.14	0.07	0.13	0.62	0.31	0.31	0.98	0.07	0.26	0.15	0.06	0.60	0.52	0.13	3
NaTLP	0.24	0.52	0.61	0.30	0.30	0.16	0.31	0.30	0.53	0.24	0.25	0.71	0.82	0.15	0.29	0.44	0.92	0.14	0.48	0.72	0.18	0.16	0
MTLP_V1	0.70	0.84	0.98	0.45	0.36	0.60	0.36	0.08	0.26	0.68	0.27	0.59	0.04	0.40	0.85	0.46	0.98	0.24	0.33	0.81	0.96	0.68	1
STLP_V1	0.50	0.33	0.98	0.06	0.52	0.15	0.09	0.00	0.09	0.46	0.11	0.41	0.04	0.43	0.76	0.30	0.94	0.04	0.06	0.56	0.41	0.10	3
ULTLP_V1	0.50	0.71	0.94	0.56	0.89	0.83	0.74	0.08	0.27	0.64	0.11	0.43	0.16	0.67	0.95	0.17	0.86	0.02	0.02	0.15	0.18	0.35	2
SMLP_V1	0.13	0.09	0.75	0.01	0.51	0.01	0.84	0.04	0.36	0.64	0.28	0.52	0.04	0.44	0.54	0.24	0.88	0.08	0.13	0.53	0.67	0.40	4
MNLP_V1	0.23	0.27	0.79	0.99	0.85	0.68	0.25	0.31	0.41	0.98	0.59	0.43	0.79	0.20	0.31	0.23	0.85	0.25	0.48	0.32	0.96	0.44	0
EFSP_V1	0.35	0.27	0.36	0.98	0.09	0.36	0.30	0.29	0.70	0.13	0.58	0.41	0.26	0.94	0.59	0.67	0.32	0.37	0.57	0.06	0.31	0.39	0
EFSL_V1	0.35	0.31	0.24	0.41	0.25	0.44	0.34	0.58	0.99	0.74	0.85	0.56	0.14	0.91	0.35	0.42	0.21	0.36	0.16	0.06	0.89	0.35	0
ECSR_V1	0.29	0.60	0.87	0.29	0.20	0.37	0.45	0.04	0.15	0.67	0.10	0.34	0.58	0.95	0.31	0.69	0.26	0.36	0.31	0.20	0.44	0.52	1
ETSR_V1	0.36	0.64	0.37	0.29	0.51	0.46	0.84	0.01	0.19	0.48	0.06	0.05	0.15	0.18	0.31	0.25	0.31	0.42	0.04	0.12	0.27	0.23	3
ETSL_V1	0.55	0.63	0.17	0.79	0.47	0.41	0.47	0.41	0.86	0.14	0.58	0.20	0.21	0.47	0.06	0.30	0.40	0.23	0.71	0.35	0.65	0.48	0
ESSR_V1	0.64	0.62	0.76	0.50	0.59	0.70	0.86	0.10	0.25	0.70	0.33	0.17	0.28	0.13	0.34	0.55	0.89	0.73	0.23	0.40	0.77	0.08	0
ESSL_V1	0.17	0.33	0.11	0.30	0.32	0.48	0.25	0.09	0.06	0.22	0.04	0.30	0.29	0.27	0.37	0.33	0.85	0.23	0.05	0.12	0.69	0.14	2
MTP_V1.1	0.41	0.32	0.84	0.13	0.37	0.07	0.40	0.31	0.16	0.66	0.21	0.38	0.43	0.18	0.27	0.20	0.29	0.34	0.19	0.76	0.42	0.48	0
OES_V1	0.90	0.91	0.87	0.60	0.72	0.97	0.49	0.85	0.65	0.74	0.94	0.35	0.30	0.68	0.34	0.66	0.71	0.13	0.99	0.82	0.62	0.74	0
OEAA_V1	0.41	0.31	0.53	0.06	0.44	0.10	0.96	0.36	0.50	0.72	0.29	0.08	0.00	0.11	0.69	0.05	0.35	0.81	0.11	0.27	0.12	0.50	2
MSP_V9	0.24	0.48	0.17	0.22	0.49	0.22	0.54	0.21	0.37	0.36	0.16	0.52	0.81	0.40	0.15	0.18	0.49	0.68	0.50	0.53	0.05	0.11	0
CNSP_V14	0.40	0.77	0.34	0.74	0.29	0.49	0.82	0.06	0.08	0.75	0.08	0.02	0.24	0.86	0.95	0.16	0.75	0.33	0.44	0.08	0.27	0.11	1
CLP_V6	0.41	0.22	0.06	0.17	0.22	0.19	0.57	0.24	0.20	0.86	0.22	0.28	0.02	0.32	0.50	0.15	0.46	0.66	0.56	0.17	0.51	0.86	1
JULP_V7	0.30	0.35	0.11	0.90	0.76	0.75	0.15	0.18	0.25	0.28	0.53	0.23	0.05	0.24	0.13	0.07	0.88	0.46	0.05	0.19	0.89	0.49	1
JAS_V1	0.48	0.67	0.65	0.91	0.76	0.99	0.55	0.57	0.07	0.29	0.10	0.69	0.13	0.46	0.61	0.21	0.21	0.10	0.30	0.46	0.53	0.54	0
JMSR_V1	0.07	0.10	0.70	0.32	0.09	0.17	0.37	0.31	0.21	0.42	0.27	0.05	0.84	0.07	0.52	0.25	0.34	0.35	0.02	0.58	0.11	0.49	2
EOPP_V1	0.46	0.29	0.66	0.60	0.35	0.31	0.45	0.47	0.96	0.59	0.78	0.17	0.95	0.04	0.07	0.78	0.17	0.27	0.43	0.51	0.15	0.45	1
NaTP_V1	0.61	0.72	0.82	0.13	0.31	0.56	0.74	0.93	0.87	0.82	0.66	0.08	0.61	0.29	0.08	0.15	0.29	0.07	0.11	0.50	0.02	0.13	1
EOEHR_V1	0.84	0.87	0.49	0.63	0.72	0.97	0.90	0.76	0.07	0.88	0.12	0.03	0.56	0.26	0.43	0.14	0.54	0.11	0.87	0.77	0.41	0.20	1
EOTPP_V1	0.59	0.57	0.42	0.57	0.17	0.74	0.35	0.96	0.93	0.12	0.69	0.40	0.11	0.23	0.16	0.97	0.29	0.09	0.81	0.41	0.07	0.20	0
FTR_V1	0.89	0.46	0.36	0.35	0.80	0.93	0.47	0.27	0.42	0.13	0.38	0.42	0.40	0.28	0.88	0.18	0.52	0.27	0.05	0.90	0.04	0.92	2
FEAS_V3	0.38	0.57	0.27	0.17	0.45	0.20	0.21	0.38	0.93	0.91	0.69	0.00	0.02	0.45	0.21	0.33	0.72	0.80	0.44	0.76	0.29	0.86	2
FTAS_V3	0.81	0.87	0.48	0.83	0.96	0.95	0.29	0.20	0.34	0.45	0.18	0.01	0.54	0.30	0.10	0.16	0.39	0.37	0.05	0.34	0.56	0.27	2
FJAS_V3	0.87	0.65	0.37	0.94	0.91	0.94	0.47	0.75	0.26	0.60	0.31	0.05	0.62	0.16	0.17	0.99	0.81	0.20	0.12	0.76	0.41	0.07	1
FTLR_V3	0.45	0.70	0.52	0.78	0.18	0.51	0.65	0.30	0.53	0.11	0.47	0.61	0.34	0.51</									

Finally, results of the regression tree analysis largely echo the results of prior models. While there is some evidence of significant interaction depth in a handful of the biometric models, the majority of these genetic response variables were sufficiently modeled using stumps, suggesting that as a whole assumptions of a simple additive response were largely sufficient for the relationship between facial biometrics and genetic merit (see Table 28). Interestingly, the baseline type models showed a greater tendency for deeper interaction depths than the full models. This may simply be a reflection of the significant increase in scale of candidate coefficients considered relative to training data present with the addition of biometrics into the model, or it might indicate that biometric values provided information to the model that simplified the relationship between structure traits. It should also be noted, however, that the largest gains in R^2 observed from addition of biometric values came from models utilizing deeper regression trees. Thus, an additional explanation for this trend may simply be that noise in the biometric data itself may have obscured finer details in the interactions between the underlying signals.

Lending further merit to the suspicion that noise in the covariates themselves may have limited the information gleaned from these models is their poor performance relative to the more constrained LASSO models. The boost in model performance often realized with any boosted modeling approach (James *et al* 2013) is not observed in this data set, and in fact several of the optimized full models under-performed relative to the LASSO performance results. Among the PTA response models, any gains in R^2 realized in the training set were not upheld with the training data. While this may simply be a reflection of over-fitting, the volatility observed across models suggests a more fundamental lack of suitability of this data within this modeling approach.

In appraising the performance of the individual covariates themselves, results of the regression trees mirror those of the previous simpler models (see Table 29). Type traits clearly out-performed biometric values with respect to consistency of importance to these tree models. Several biometrics, however, still proved influential across a number of these genetic response variables. Eye Roundness Proportion Lower Back (Area), Nostril Position Angle, Mouth Thickness Proportion, Nostril Roundness Proportion, Sinus-Midface Length Proportion, Cranio-Topline Length Ratio, and Forehead-Topline Angle (Slope) all showed evidence of consistent persistency in the simpler models, and did so again with the regression tree models. Here also Forehead Width-to-Length ratio also revealed some consistent behavior seen slightly in the LASSO models but not in the spline results.

Table 28: Performance of Boosted Regression Models

	Depth Full Model	N Trees Full Model	Depth Base Model	N Trees Base Model	R ² Training Full Model	R ² Gain Training	R ² Test Full Model	R ² Gain Test
CTPI	1	3900	1	5200	0.54	-0.01	0.39	-0.00
Net Merit	2	4200	1	5700	0.53	0.13	0.23	-0.01
PTA Milk	1	4800	2	4400	0.47	-0.06	0.38	-0.03
PTA Fat	1	4800	2	3000	0.40	0.01	0.20	-0.02
PTA Protein	1	5700	3	4500	0.53	-0.10	0.38	-0.02
Feed Efficiency	1	5500	2	4700	0.48	-0.05	0.34	-0.02
Productive Life	2	3600	1	5400	0.51	0.13	0.11	-0.00
Heifer Conception Rate	1	1800	1	2100	0.23	0.07	0.02	0.01
Cow Conception Rate	1	3700	1	3200	0.32	0.07	0.06	-0.00
Somatic Cell Score	1	2100	1	1500	0.24	0.07	0.03	-0.01
Fertility Index	1	3200	1	3000	0.32	0.05	0.10	0.00
Calving Ease	1	3100	1	2200	0.28	0.08	0.07	-0.00
Still Births	1	2700	2	900	0.30	0.10	0.03	-0.03
Mastitis	1	3300	1	2400	0.35	0.08	NA	NA
Lameness	1	900	1	1100	0.20	0.04	NA	NA
Metritis	2	1100	1	100	0.51	0.36	NA	NA
Retained Placenta	2	1400	1	300	0.44	0.35	NA	NA
Ketosis	1	100	1	300	0.18	0.06	NA	NA
Displaced Abomasum	1	2900	2	1700	0.34	0.01	NA	NA
Calf Livability	1	100	1	100	0.16	0.06	NA	NA
Calf Scours	1	1500	1	1100	0.24	0.06	NA	NA
Calf Respiratory	1	100	1	100	0.20	0.11	NA	NA

Table 29: Count of Influential Variables Across Response Variables for Regression Tree Models

Type Coefficient	Model Inclusion Count	Biometric Coefficient	Model Inclusion Count
PTA.Type	9	Z1_front	1
UDC	6	Z2	5
FLC	8	Z3_front	6
GPTA.STA	8	Z4	1
GPTA.STR	7	Z7_length	3
GPTA.BDE	13	Z9_poly	4
GPTA.DRM	17	Z10_poly	6
GPTA.RPA	8	Z11_poly	3
GPTA.TRW	11	Z11_linear	2
GPTA.RLS	13	NFP_LF	3
GPTA.RLR	14	NFPP_LF	5
GPTA.FTA	11	NPA	9
GPTA.FLS	9	ULRP_V1	4
GPTA.FUA	6	MTP_V1	7
GPTA.RUH	12	CTP_V1	7
PTA.RUW	5	CLTR_V1	5
GPTA.UCL	2	NRP	9
GPTA.UDP	15	SMRP_V1	4
GPTA.FTP	9	MNRP_V1	5
GPTA.RTP	5	MDP_V1	3
GPTA.TLG	11	NDP_V1	3
GPTA.BSC	7	NaDP_V1	2
		MIP_V1	4
		MIPP_V1	1
		NIP_V1	3
		NaTLP	3
		MTLP_V1	4
		STLP_V1	4
		ULTLP_V1	0
		SMLP_V1	6
		MNLP_V1	5
		EFSRP_V1	2
		EFSRL_V1	3
		ECSR_V1	3
		ETSRP_V1	5
		ETSRL_V1	4
		ESSRP_V1	2
		ESSRL_V1	3
		MTP_V1,1	1
		OES_V1	1
		OEAA_V1	5
		MSP_V9	5
		CNSP_V14	2
		CLP_V6	2
		JJLP_V7	3
		JAS_V1	4
		JMSR_V1	2
		EOPP_V1	3
		NsTP_V1	5
		EOEHR_V1	4
		EOTPP_V1	2
		FTR_V1	4
		FEAS_V3	3
		FTAS_V3	8
		FJAS_V3	5
		FTLR_V3	4
		CTLR_V3	6
		FPLP_V3	2
		PHP_V4	3
		FWLP_V3	6

Discussion

While several facial biometrics were identified as significant covariates across a range of response variables and model types, facial biometrics values collectively did not yield practically meaningful improvements in the prediction of dairy genetic evaluation parameters above and beyond the standard type traits, nor did biometric-augmented models prove reproducible when applied to validation data. A number of biological and methodological explanations potentially lend themselves to these results. The first and perhaps the most straight forward would be that variations in facial biometrics are simply responsive to the same developmental drivers of variations in body structure. In this case, facial biometrics could be used interchangeably with correlated physical measurements, and thus would not provide any additional information to the prediction of genetic merit beyond the standard type traits. Correlations between type traits and facial biometrics, however, proved quite low. The average magnitude of correlation between type and biometric traits was only 0.05, with the largest observed magnitude of correlation failing to exceed 0.20. Thus, it seems unlikely that facial biometrics simply represent redundant information with respect to type traits.

Another potential explanation for these results would be confounding effects with genetic trend. As the herd used in this study makes heavy use of genomic testing, embryo transfer, and even a customized selection index, the rate of genetic improvement is undoubtedly quite high, and thus a genetic trend is likely detectable even within the limited range of ages represented by cows in this study. Continuous improvements to herd management protocols and the farm environment, particularly to the calf rearing protocols, have also been made during this time period. If variations in facial biometrics are in fact controlled by developmental factors driven by environmental or epigenetic factors, or if they are driven by genes that are responsive to environmental conditions

(ie – a genetic by environmental effect), then systematic changes in herd environment and management might be reflected as trends in individual or combinations of facial biometric values over time, which may in turn spuriously correlate to genetic trends if PTA traits driven by aggressive selection schemes. Were simple confounding effects to be the underlying mechanism producing significant associations between biometric values and measure of genetic merit, however, we should expect to see consistent patterns in significant results for individual coefficients across models. In other words, if say a given biometric value correlates to the genetic trends in say PTAM, HCR, and RP, then any other biometric that significantly correlates to one of these response values should consistently correlate to this same set of response values. Given that a consistent pattern in groupings of significant response terms are not observed across the range of biometric candidate variables in either the LASSO or spline models, systematic correlations to underlying genetic trends likely cannot fully explain the results returned by these analyses.

Another explanation for these results may simply be that the information contained in facial biometrics is not well represented by genomic data. As underscored in the review of literature, both historic and modern work on facial traits suggest that any relationship between facial biometrics and dairy performance is likely driven by correlations to behavior. But behavioral traits have not traditionally been considered good targets for genomic estimation due to their complex and often polygenetic nature (Rittschof and Hughes 2012). Thus, inclusion of quantitative trait loci related to behavior as indirect influencers of health and performance may be limited in current commercial genomic tests. Additionally, while polygenetic traits can be estimated using traditional genetic evaluations, genetic correlations between temperament and production traits are typically only moderate in magnitude (Haskill *et al* 2014). Thus, even among PTA estimates reinforced by phenotypic records, the amount of information present in these values representing the influence

of behavior on the corresponding performance trait is still likely quite limited. While such biological and technological limitations inherent to this dataset almost certainly limit the predictive value of biometric values, they do not necessarily explain why several biometric values returned significant results across a range of response models, especially among the purely genomic health traits, nor would they fully account for the model volatility observed between training and validation results.

Perhaps the most comprehensive explanation for results obtained in this analysis would be noise within the biometric values themselves. In any form of least squares regression, covariate values are assumed to be measured without error. Practical divergence from this assumption manifests by biasing the coefficient corresponding to a noise covariate towards zero (Kutner *et al* 2005). Noise in explanatory variables can also lead to model instability, particularly when utilized in multiplicative interaction terms as with the regression tree models. Results of Chapter 1 revealed that, while geometric biometric demonstrated clear advantages over simple Euclidean distances for use in linear models, they are still quite noisy. Additionally, even if facial biometrics are in fact heritable, they would still be partially influenced by environmental factors, just as with any biological trait. When regressed against genetic response traits, this component of variance in observed biometric values due to environment would thus represent an additional source of noise. Thus, results of this analysis might be interpreted as evidence of significant associations between facial biometrics and genetic traits, but the practical value of these relationships may be obscured by the lack of signal strength.

Conclusions

While results of this analysis did not reveal facial biometrics to be practically useful in the prediction of genetic merit for health and performance traits, neither did they reveal facial biometrics to be completely inept components of such models. This suggests that further analysis may be warranted in order to reconcile the contradictory findings of both intra-model predictive volatility and inter-model consistency in covariate structure. Future work is warranted to explore improvement of the signal-to-noise ratio within facial biometric traits to improve their predictive performance in models trained against measures of genetic merit.

Implications

While measurement error, or perhaps other yet unidentified factors, limit the predictive potential of facial biometrics, the significant associations between facial biometrics may still prove useful in ongoing efforts to breed sounder and more profitable dairy animals. Future research might explore the incorporation of facial biometrics into genomic or traditional genetic evaluations as indicator traits to improve the estimated accuracies of target health and performance traits. Additionally, as several facial biometric demonstrated significant correlations across a range of response traits, these features may prove with further study indicative of broadly adaptive traits, potentially making uniquely suited to serve as indicator traits for index estimates. Finally, while a number of facial biometric traits did not show significant associations to genomic traits, this does not preclude the possibility that these trait may serve as indicators for environmental influence on prenatal and neonatal development. Thus, future study may also be warranted into the use of facial biometrics as fixed effects in genomic and genetic estimation models to further help control for non-genetic effects at the individual level.

REFERENCES

- Friedman, J., Hastie, T., Tibshirani, R. "Regularization Paths for Generalized Linear Models via Coordinate Descent." 2010. *Journal of Statistical Software*. 33:1. 1-22. URL <http://www.jstatsoft.org/v33/i01/>.
- Haselhuhn, M.P., Ormiston, M.E., Wong, E.M. "Men's Facial Width-to-Height Ratio Predicts Aggression: A Meta-Analysis." *PLOS One*. 2015.
- Haskill, M.J., Simm, G., Turner, S. "Genetic Selection for temperament traits in dairy and beef cattle." 2014. *Frontiers in Genetics*. 5:1. 1-18.
- James, G., Witten, D., Hastie, T., Tibshirani, R. "An Introduction to Statistical Learning: with Applications in R." 2013. Springer. New York, USA.
- Greg Ridgeway with contributions from others. "gbm: Generalized Boosted Regression Models." 2017. R package version 2.1.3. <https://CRAN.R-project.org/package=gbm>
- Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W. "Applied Linear Statistical Models: 5th Edition." 2005. McGraw-Hill. New York, USA.
- Rittschof, C.C., Hughes, K.A. "Advancing behavioral genomics by considering timescale." *Nature Communications*. 2018. 9:489. 1-11.
- Wood, S.N. "Thin-plate regression splines." 2003. *Journal of the Royal Statistical Society B*. 65:1. 95-114.
- Wood, S.N. "Stable and efficient multiple smoothing parameter estimation for generalized additive models." 2004. *Journal of the American Statistical Association*. 99. 673-686.
- Wood, S.N. "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models." 2011. *Journal of the Royal Statistical Society B*. 73:1. 3-36.
- Wood, S.N. "Smoothing parameter and model selection for general smooth models." 2016. *Journal of the American Statistics Association*. 111. 1548-1575.
- Wood, S.N. "Generalized Additive Models: An Introduction with R." 2017. Chapman and Hall/CRC

CHAPTER 3 – FACIAL BIOMETRICS AS PREDICTORS OF SOCIAL RANK

Introduction

In recent years, dairy geneticists have identified negative correlations between multiple traits related to productivity and measures of dairy longevity, a robust association that appears to have played no small role in the downward trends readily observed in health and fertility traits after decades of selection programs that emphasize milk yield and components (Oltenucu and Oltenucu 2010). In order to reverse, or at the very least stall, loss of dairy robustness, selection indices for many breeds have been adjusted to place more weight on longevity traits, and research resources have been redirected towards developing genetic tools for more targeted selection of dairy health traits (Oltenucu and Oltenucu 2010; Egger-Danner et al 2014). While more holistic approaches to selective breeding will no doubt benefit the modern dairy cow, history has also repeatedly shown that, with the introduction of any new target for selective breeding, comes new risks for indirect influence on unmeasured traits, namely those related to animal behavior (Grandin and Deesing 2014; Turner et al 2006; Turner and Lawrence 2007). This risk has been underscored in several recent preliminary studies. In 2009, Gibbons reported that amongst primiparous British Holstein cattle milked in standard commercial dairy environments, animals whose sires ranked highly on the robustness component of the standard \pounds PLI index responded more aggressively to confrontation events at the feed bunk as compared to animals whose sires were low-ranking for robustness traits. Similarly, Llonch et al (2018) reported that frequent feed bunk visits and high total time spent at the feed bunk were associated not only with phenotypes for high feed efficiency and low methane emissions, but were also behaviors observed more frequently among dominant cattle, as established through observations of feed bunk displacement behaviors.

These results suggest that, if selection schemes for dairy cattle are adjusted to incorporate phenotypic information related to health outcomes and measures of feed efficiency, farmers may indirectly select for more aggressive and dominant animals. While dominance may provide an individual animal with a competitive advantage with regards to resource allocation within a normal herd composition, broadly increasing the dominance levels of animals via genetic selection could disrupt normal herd structure and result in social instability that would impart a negative impact on health, production and welfare (Gibbons 2007). Thus, the question explored in this chapter is whether facial biometrics, if they are in fact correlated with dimensions of personality, might be correlated with measures of social dominance, and may in turn be utilized in genetic selection schemes to distinguish between highly robust animals that have gained this advantage through dominance or those who have more broadly adaptive genes.

Materials & Methods

Description of Data

To explore potential relationships between facial biometrics and social dominance, milking order data was acquired for the herd from which facial images were collected for the analyses described in Chapter 1 (Soffié *et al* 1975). Animals were both primiparous and multiparous, and represented predominantly Holstein genetics. As this group of animals were simultaneously participating in a nutrition trial for a supplemental fat additive, this data represents a closed herd, with no new cows introduced after the initial enrollment period. The trial began on January 16, 2017 with 203 animals and ended on July 16, 2017. Individuals were added to the herd as eligible animals calved, with the majority of cows having been introduced within 50 days from the start of the trial. During the initial period of the trial, animals remained in their home pen. At a stocking

ratio of <1 (space:cow), this environment provided animals ample access to bunk space and sand bedded, as well as free access to a large outdoor dry lot area. On April 24th cows were granted access overnight to a nearby pasture for the spring and summer grazing periods. Any animals that developed health conditions requiring extended treatment were removed to a hospital pen for closer observation and not returned to the herd.

Cows were milked three times a day in a rotary milking parlor, upon entry of which their RFID numbers were logged and the recorded milking order dumped to csv files once a day. Due to the setup of the program, order data was only recorded for cows registered to the trial pen, and only in the relative order to pen mates. This means that if cows somehow strayed into another milking group, their relative order would be distorted. Similarly, if a cow from another pen found its way into the trial group, its relative position within the herd was not recorded. While this data extraction method did pose risks of distorting the recorded milk order from that which actually occurred, this risk was deemed minimal due to the highly hands-on nature of the concurrent feed trial. Cows were checked daily by the graduate student administering the protocol, and it was anecdotally reported that in a given week seldom more than a handful of cows had to be removed from the trial herd or else tracked down within the regular milking herd.

Exploratory Data Analysis

Milk order records were obtained from a total of 150 milkings within the trial period. Preliminary exploratory data analyses were then conducted to get a sense of underlying patterns within this data set. The daily quantile rank of each individual cow over the course of the trial period was visualized using a loess line feature in the *ggplot2* package in the *R* programming environment in order to explore stability of herd rank over time (Wickham 2016). Next, in order to get a sense of underlying herd structure, principal component analysis was performed such that

each data point represented an individual cow whose quantile rankings in each of the milk order observations were embedded into a lower dimensional Euclidean space. For this visualization, milk order observations were dropped for the first 55 days and final 10 days of the trial to ensure data quality, as were any cows who either appeared in fewer than 85% of the observed milkings or else were absent from two or more consecutive milkings. This resulted in a complete data set consisting of 160 cows observed over 100 milkings. This data was subsequently analyzed utilizing the *princomp* function in *base R* (R Core Team 2018) and visualized using the *ggplot2* package (Wickham 2016). Additionally, in order to visually assess differences in herd structure before and after access to pasture was granted, PCA analysis was repeated using the same data set where here each data point represented an individual milking observation within which quantile ranks were assigned to a set domain of cows.

Rank Order Analysis

Based on the results of exploratory data analysis, it was determined that milking observations recorded during the early part of the trial where cows remained in their home pen and milking observations recorded after access to pasture was granted were sufficiently different to warrant separate analyses. A number of techniques have been employed to estimate the rank order of cattle from milk order observations. The simplest method is to simply average the nominal milk order values for each cow over the range of days observed, and then order animals using these values (Soffie *et al* 1976). As the preliminary visualizations revealed milk order to be fairly dynamic over a wider observation window, it was determined that this approach was not appropriate for this data set. The simplest alternative to working with nominal milk order values is to evaluate the data using pairwise relationships between animals, where entrance of a given cow into the parlor directly ahead of another individual was recorded as a didactic instance of a

dominance-submission behavior. Analyses of this type traditionally take one of two forms. The most common method used in estimation of dairy social structures is Angular Dominance Value (Hasegawa *et al* 1997; Phillips and Rind 2002; Soffie *et al* 1976). However, the creators of this technique expressly warn against its application to data with ambiguous pairwise interactions (Beilharz and Zeeb 1982). As there was certainly noise present within the observed milk order data due to pushing at the entrance of the milking parlor and other environmental factors, this was not deemed a suitable approach for this data set. Alternatively there is the Bradley & Terry model, which seeks to estimate a strictly linear dominance hierarchy by maximizing a likelihood function based on pairwise interactions between all animals (Bradley and Terry 1952). This model, however, tends to perform poorly when data contains sparse and highly localized pairwise interaction observations (Fushing *et al* 2011). Provided that 80% of possible pairwise dominance interactions were unobserved within the pasture dataset and 91% of possible pairwise dominance interactions were unobserved within the pen dataset, this data again proved poorly suited to this analytical method.

In order to better accommodate the noisy and sparse nature of interaction data generated from milking data, rank orders were estimated using a network-based technique originally developed for analysis of primate social hierarchies implements in the *Perc* package in the R programming environment (Fujii *et al* 2016). Among the benefits of applying this methodology to estimate rank order include: the use of a beta random field to estimate dominance probabilities, which allows for ambiguity to exist between cows with little or no interaction; use of a percolation algorithm to estimate the consistency of information flow through the weighted interaction network, allowing for augmentation of the adjacency matrix with information from indirect dominance information; and finally, estimation of an optimal ranking is generated by applying an

annealing algorithm to a large ensemble of random draws from the dominance probability matrix, which accounts for ambiguity between animals in the process of extracting a linear hierarchy from poorly defined or non-linear sections of the dominance network.

First adjacency matrices were calculated separately for subsets of the milking observations coming from the pre and post pasture access intervals. For the pen subset, milking observations recorded prior to 50 days into the start of the trial were thrown out in order to ensure that the majority of cows had been introduced to the herd and that all animals had an opportunity to establish themselves within the hierarchy. Additionally, milking observations from the final 10 days of the trial, the 5 days preceding access to pasture, and 10 days following access to pasture were all dropped in order to ensure that any changes in management schedule did not unduly disrupt the movement of the cows into the milking parlor. As the overwhelming majority of milk order records came from morning milkings, data from later milking were thrown out to further ensure consistency within the data. Finally, five consecutive milk order observations in late June identified as outliers via results of the principal component analysis were excluded from the pasture data subset, as a finite cause for this change could not be identified. This left 31 milk order observations from the pen period and 50 observations from the pasture period from which to build respective adjacency matrices.

To avoid recording erroneous dominance dyads resulting from trial cows accidentally finding their ways into other milking groups, information on the first and last two cows in any milking observation were excluded in adding dyads to the weighted adjacency matrices. Finally, if any cows were recorded in less than 50% of the milking observations in each of the respective observation windows, that animal was dropped from the adjacency matrix. Of the 191 animals that remained in the herd past the freshening period, 186 were retained in the adjacency matrix for the

pen period, and of these animals 182 were retained with sufficient records for the pasture period. The adjacency matrices created from the pen and pasture data subset were then separately augmented with information from indirect dominance relationships using the conductance functionality in the *Perc* package using a max chain length of 3. The *simRankOrder* function was then separately applied to the resulting probability matrices using the default optimization parameters to estimate overall rank order from both datasets. Output of rank order simulations were analyzed in two ways. First, heatmaps were constructed of the probability matrices resulting from either data set, re-ordered to reflect estimated rank orders. From these, qualitative inferences could be gleaned with respect to overall herd structure, as well as impressions of the overall adequacy of the rank order simulation. Additionally, linear rank estimations from each data set were reduced to the largest common subset of animals, and then ordinal rank values compared using both the standard Pearson and Kendall Tau Rank correlation.

Predictive Modeling of Rank Order

Facial biometric values for each cow were calculated in several ways using the output of nested mixed models for repeated measures described in Chapter 1. First, an overall biometric value was calculated using information from all photos collected on a given cow by applying the *coef* function to the main cow effect (Bates *et al* 2015). Additionally, literature on traditional face reading techniques suggest that one side of the face may be more reflective of social preferences than the other (Bridges 2012; Haner 2008). This seems to echo recent research results providing some preliminary evidence for laterality in social interaction amongst several livestock species (Camerlink *et al* 2018; Phillips *et al* 2015; Robins & Phillips 2010). To account for the potential impact of laterality on prospective relationships between facial morphology and social dominance, facial biometrics were also separately computed for the left and right sides of the face for each

cow by also applying the *coef* function to the cow x side effect. Any animals that did not have at least one suitable photo for each side of the face were subsequently excluded from this analysis, leaving 107 animals with biometric values, of which 92 animals had more than one photo for either side of the face. Additionally, age in days at the start of the trial was calculated from herd records as an explanatory variable in order to account for potential influence of seniority on herd hierarchy.

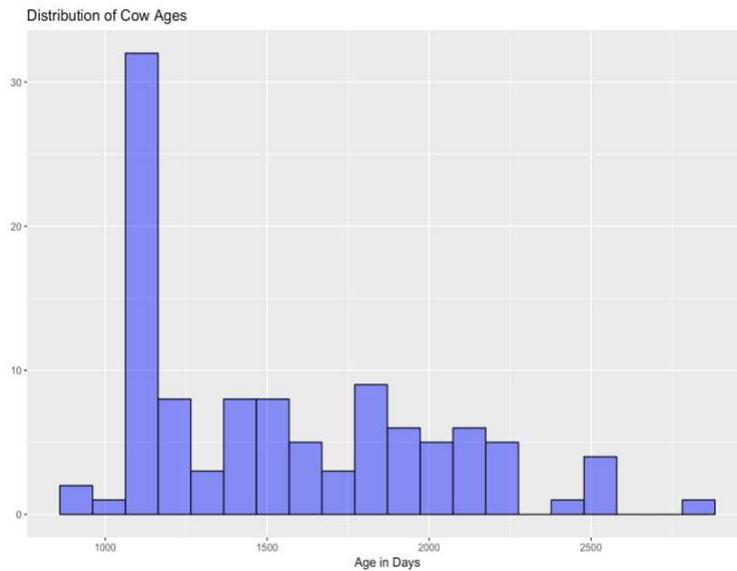


Figure 36: Distribution of cows ages at start of trial

Two statistical techniques were employed to identify potential relationships between facial biometrics and social dominance. First, Kendal Tau correlation estimates were calculated between biometric variables and rank estimates generated from both the pen and pasture data sets. Two-sided tests for significant correlations were subsequently performed, and the resulting p-values used not only to identify potential predictor variables for dominance, but also to assess the comparative performance the right vs left biometric subsets. Next, boosted regression trees were used to explore the collective predictive value of facial biometrics against rank order estimates. Separate models were fit for left, right, and overall facial biometric data sets against both the pen

and pasture rank order estimates. Optimization methodology mirrored that described in Chapter 2, but with several accommodations for the smaller number of observations. Five-fold cross validation was utilized over the standard ten-fold cross validation default. Additionally, the cross validation error threshold was increased to 1%, and the minimum model size reduced to only 5 trees. Finally, variable importance values for the best 20 candidate biometrics were reported for all regression tree models that accounted for a substantial proportion of variance in rank order, as determined by estimated R^2 values.

Social Network Analysis

Finally, while estimations of rank order may be a fairly straight-forward representation of social dominance, they are unlikely to capture more complex dimensions of herd structure that are non-linear in nature. Thus, in order to further explore social information contained in this data set beyond the simplifying assumption of linear rank order, adjacency matrices generated from milk order data were also assessed using the *igraph* package in the *R* programming environment (Csardi & Nepusz 2006) in order to extract several common social network metrics. First, weighted directed graphs were separately generated for the pen and pasture datasets. Measures of betweenness were calculated for each vertex in the graph, and merged with the biometric datasets. Boosted regression trees were then used to assess the predictive potential of biometrics for a given cows measure of betweenness within the network using the methodology described above. Finally, the assortativity of each biometric within the graph was calculated, and a p-value for the significance of these results estimated by comparing the magnitude of the observed graph to assortativity values calculated from graphs where biometric values assigned to nodes had been randomly permuted over 5000 iterations.

Results

Visualization of the quantile ranks of individual cows across all milk order observations revealed rank order to be inconsistent over time, with some animals even showing a sinusoidal trend in milk order across the duration of the trial (see Figure 37). While entry into the milk parlor itself was not consistent, discernable patterns in milk order were observed, some of which seemed distinctive to groups of cows with similar ear tag numbers. As similar ear tag numbers are indicative that cows were born and likely raised together, this finding suggested that subhierarchies might exist within this data set, and that these subgroups might not themselves follow a strictly linear pattern relative to the whole herd. To view quantile plots for all animals, see Appendix D.

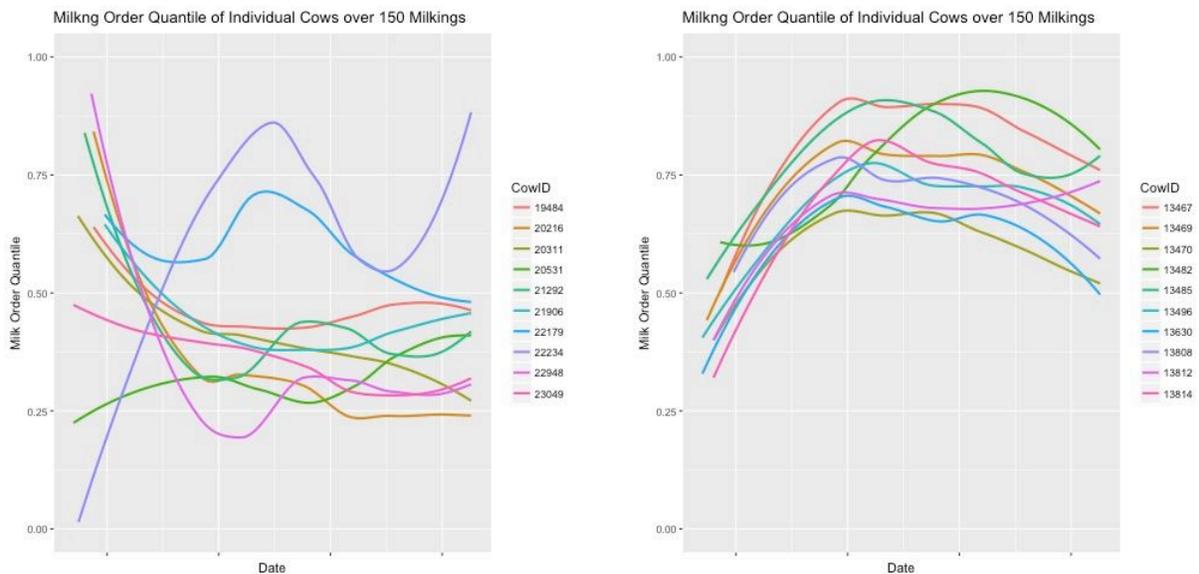


Figure 37: Examples of quantile plots of milk order over time by individual cow demonstrating both highly variable (left) but potentially cohesive (right) patterns in milk order

Results of the PCA analysis by cow indicated that the milk order data existed in a relatively low dimensional space, with the scree plot of eigen values dropping off quickly after only the first two components (see Figure 38). And yet, to capture the majority of variance present in the data,

required a substantial number of dimensions. This could simply reflect the inherent noisiness of milk order data, as the underlying dominance structure would certainly be distorted by day-to-day variations in crowding at the entrance to the milking parlor and handling by stockpersons. Alternatively, however, it could suggest that dominance structures are poorly modeled using linear techniques, and that nonlinear tendencies are found in these milk order observations. In visualizing the first two principal components themselves, herd order appears to be fairly linear, but with greater variability observed in the middle ranks of the herd (see Figure 39).

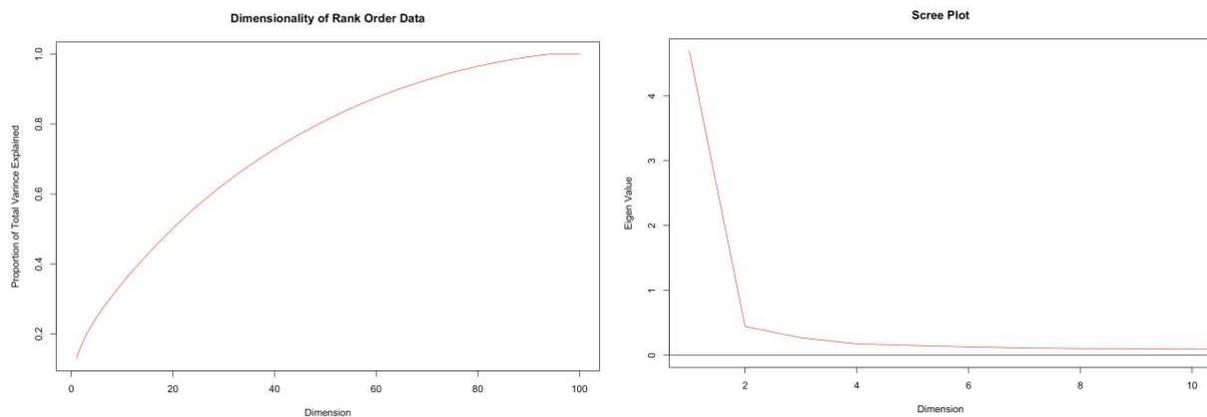


Figure 38: Visualizations of dimensionality results from PC analysis by cow

Results of the PCA analysis seeking to embed the individual milking observations into a lower dimensional space revealed some surprising details hiding within this large data set (see Figure 40). First, pen and pasture milking order observations appeared well separated along the first three principal axes. This suggests that, provided a significant change in herd environment, the milk order of the cows clearly shifted to match the new conditions. Whether this represents an adaptation of the underlying social structure, or just a change in how the social structure manifested itself within the milking parlor is a question that cannot be directly answered from this data.

However, it is also interesting to note the presence of five significant outliers along the first principal axis, which were highly distinct from all remaining milking observations but were themselves quite tightly clustered on the first and second principal axis, and otherwise indistinguishable from the remaining milk observations along the third principal axis. These five outliers represent milking observations from five consecutive days in late June. More surprisingly, following this small handful of outliers the milking order returned suddenly to that seen in the preceding pasture data. While the exact cause of this change in herd structure could not be identified, it provides evidence of a surprising degree of dynamic behavior within milk order data.

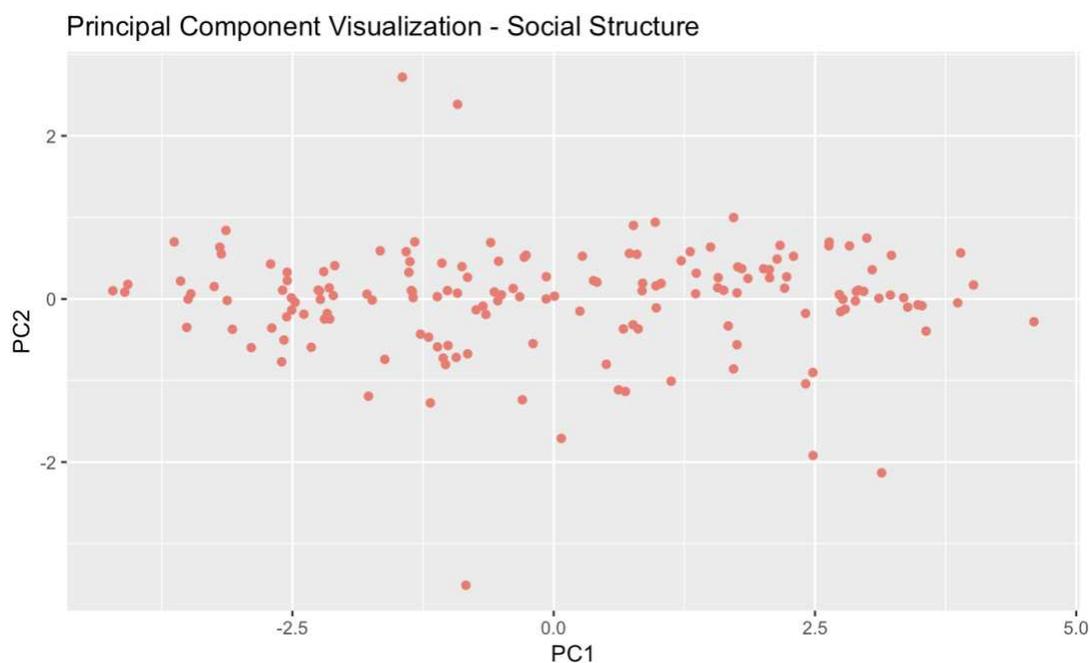


Figure 39: Results of PCA to approximate and visualize overall herd structure

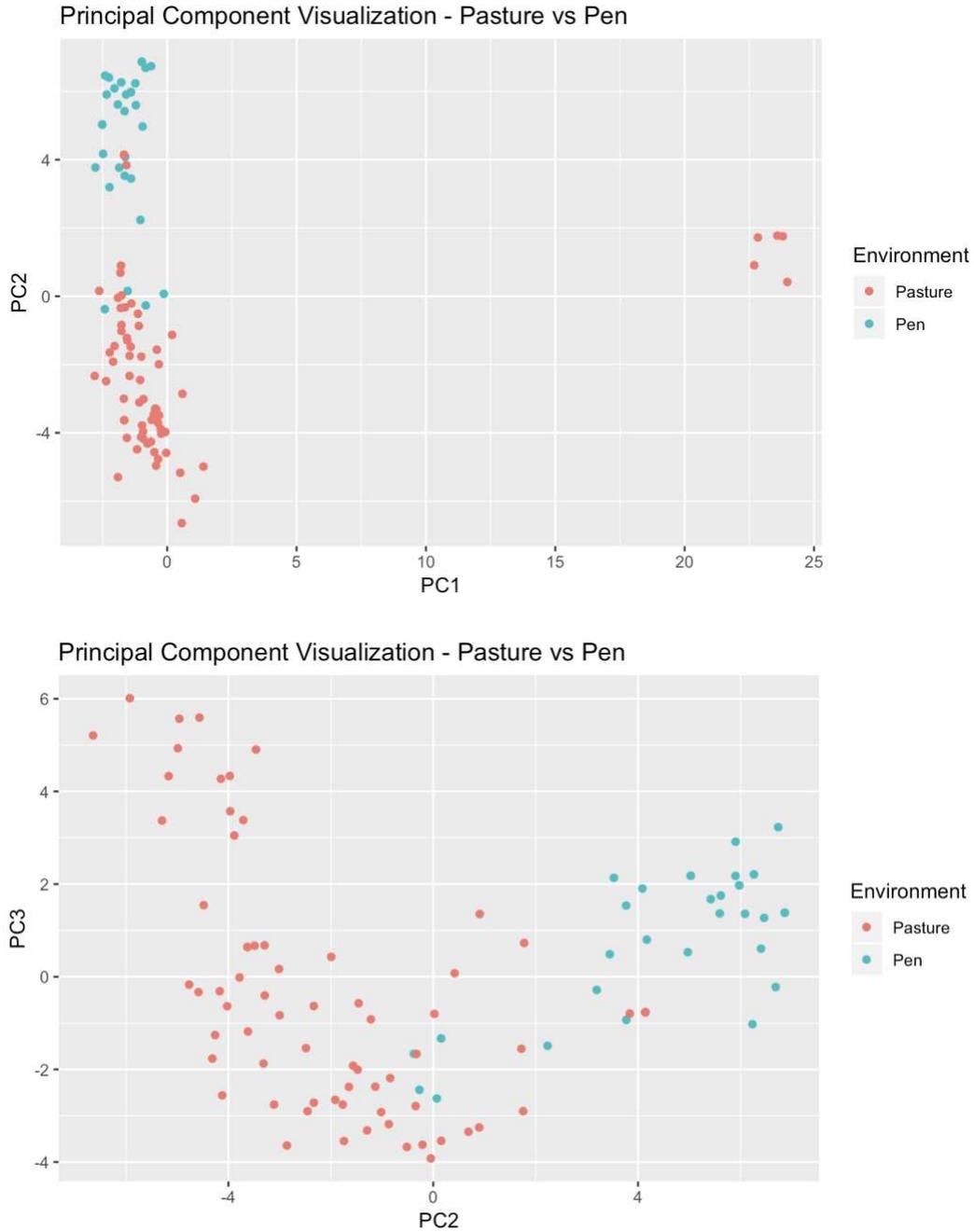


Figure 40: Visualization of the first three principal components embedding milking observations

Visualization of the dominance probability results (see Figure 41) did not reveal any clear blocks of ambiguous dominance relationships, which would indicate the presence of a multi-tiered dominance hierarchy (Vandeleest *et al* 2016). Dominance relationships were reasonably confident

along the diagonal but fairly ambiguous elsewhere, which could indicate that dominance hierarchies seem quite localized within the herd, or simply be a reflection of the inherently linear nature of milk order observations. Comparison of rank order estimates generated separately from the pen and pasture data sets yielded a Pearson correlation value of 0.004 ($p = 0.48$, one sided) and a Kendal Tau correlation of only 0.003 ($p = 0.47$, one sided). Given the separation between pasture and pen milk order observations in the PCA embedding, this result is not necessarily surprising; however, if these rank order estimations are reasonable approximations of the true rank order, this result suggests that reorganization of milk order in response to pasture access did not follow any type of simple systematic pattern, such as a simple inversion that might be seen if dominant animals in the pen inversed their motivation to enter the milking parlor between pen and pasture environments.

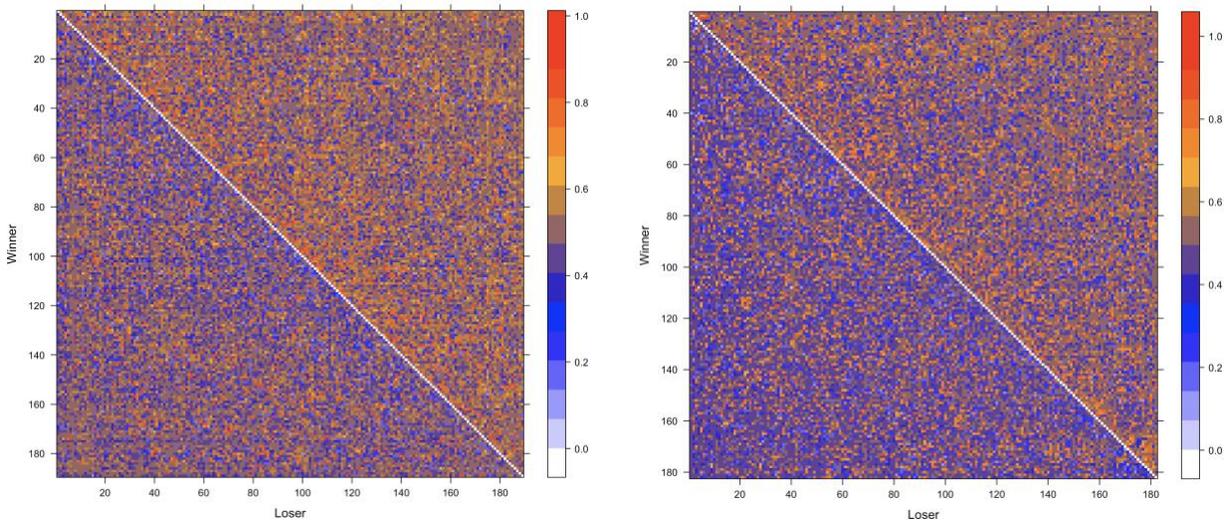


Figure 41: Heatmap visualization of pair-wise dominance probabilities ordered by descending rank order for pen data (left) and pasture data (right)

Results of the Kendal Tau analyses identified several facial biometrics with significant correlations to rank order estimates (see Table 30). By far the most highly significant correlation was between Nostril Position Angle (NPA) and rank order in the pen environment, with the remaining significant results occurring nearer the typical $\alpha = 0.05$ threshold. Overall, unilateral biometric values produced more significant correlations than those computed using photos from both sides of the face. Pasture rank was significantly correlated with unilateral biometrics from both the left and right sides of the face, but no bilateral biometrics. By contrast, pen rank was predominantly correlated with unilateral biometrics from the right side of the face. Both unilateral estimates for NPA demonstrated modest p-values for significant correlations in the same direction, which may indicate why the bilateral version of this biometric proved so significantly correlated. In contrast, midface-to-nose length proportion (MNLP) yielded significant correlations against pasture rank for both the left and right unilateral biometric estimates, but in opposite directions, effects that seemed to negate each other when combined in the bilateral biometric estimate. While pen and pasture rank estimates demonstrated roughly the same number of significant correlates, no single facial biometric demonstrated a significant correlation for both pen and pasture rank estimates.

Table 30: Kendal Tau Estimates for correlation between rank order and facial biometrics

Biom	Overall Facial Biometrics		Left Facial Biometrics		Right Facial Biometric	
	Pasture	Pen	Pasture	Pen	Pasture	Pen
Age	-0.03 (p = 0.60)	-0.05 (p = 0.49)	-0.03 (p = 0.60)	-0.05 (p = 0.49)	-0.03 (p = 0.60)	-0.05 (p = 0.49)
Z1_front_6	0.07 (p = 0.29)	0.07 (p = 0.29)	0.05 (p = 0.43)	0 (p = 0.94)	0.04 (p = 0.52)	0.12 (p = 0.06)
Z2_6	-0.06 (p = 0.34)	-0.02 (p = 0.81)	-0.05 (p = 0.48)	0.08 (p = 0.25)	-0.03 (p = 0.64)	-0.15 (p = 0.03)
Z3_front_6	0.03 (p = 0.60)	0.08 (p = 0.20)	-0.07 (p = 0.27)	0.04 (p = 0.52)	0.08 (p = 0.20)	0.07 (p = 0.33)
Z4_6	-0.04 (p = 0.58)	-0.07 (p = 0.30)	0.04 (p = 0.58)	0.03 (p = 0.68)	-0.06 (p = 0.37)	-0.08 (p = 0.22)
Z7_length_6	0.07 (p = 0.32)	0.07 (p = 0.33)	-0.04 (p = 0.53)	-0.03 (p = 0.70)	0.09 (p = 0.19)	0.12 (p = 0.07)
Z9_poly_6	-0.01 (p = 0.84)	-0.06 (p = 0.38)	0.02 (p = 0.72)	-0.06 (p = 0.34)	-0.06 (p = 0.38)	-0.01 (p = 0.89)
Z10_poly_6	-0.01 (p = 0.84)	0.03 (p = 0.65)	0.01 (p = 0.90)	0.08 (p = 0.26)	-0.1 (p = 0.11)	-0.01 (p = 0.83)
Z11_poly_6	0.04 (p = 0.57)	0.05 (p = 0.50)	0.03 (p = 0.69)	-0.06 (p = 0.37)	0 (p = 0.99)	0.16 (p = 0.02)
Z11_linear_6	0.05 (p = 0.48)	-0.01 (p = 0.90)	0.02 (p = 0.80)	-0.11 (p = 0.09)	0.01 (p = 0.84)	0.15 (p = 0.02)
NFP_LF	0.10 (p = 0.11)	0.01 (p = 0.92)	0.14 (p = 0.03)	0.05 (p = 0.47)	-0.01 (p = 0.86)	-0.07 (p = 0.29)
NFP_LF	0.04 (p = 0.54)	0.01 (p = 0.86)	-0.04 (p = 0.51)	0.05 (p = 0.49)	0.06 (p = 0.35)	-0.03 (p = 0.71)
NPA	-0.09 (p = 0.15)	0.25 (p = 0.00)	-0.07 (p = 0.27)	0.09 (p = 0.16)	-0.01 (p = 0.87)	0.16 (p = 0.02)
ULRP_V1	0.02 (p = 0.79)	0.02 (p = 0.71)	0.01 (p = 0.86)	0.10 (p = 0.13)	0.01 (p = 0.84)	-0.09 (p = 0.16)
MTP_V1	0.07 (p = 0.30)	-0.11 (p = 0.10)	0.08 (p = 0.23)	-0.01 (p = 0.91)	0.01 (p = 0.85)	-0.09 (p = 0.20)
CTP_V1	0.07 (p = 0.32)	0 (p = 0.95)	-0.05 (p = 0.48)	0.01 (p = 0.84)	0.10 (p = 0.14)	-0.03 (p = 0.64)
CLTR_V1	0 (p = 0.95)	0.09 (p = 0.18)	-0.06 (p = 0.34)	0.05 (p = 0.41)	0.09 (p = 0.17)	0.03 (p = 0.66)
NRP	-0.03 (p = 0.61)	-0.06 (p = 0.33)	0.02 (p = 0.81)	0.05 (p = 0.48)	-0.01 (p = 0.90)	-0.11 (p = 0.09)
SMRP_V1	-0.08 (p = 0.23)	-0.10 (p = 0.12)	-0.08 (p = 0.22)	-0.04 (p = 0.50)	0.06 (p = 0.40)	0 (p = 0.96)
MNRP_V1	-0.07 (p = 0.28)	0.09 (p = 0.20)	0.10 (p = 0.12)	-0.08 (p = 0.21)	-0.11 (p = 0.09)	0.11 (p = 0.11)
MDP_V1	0.04 (p = 0.58)	0.10 (p = 0.11)	0.07 (p = 0.27)	-0.09 (p = 0.20)	-0.05 (p = 0.45)	0.18 (p = 0.01)
NDP_V1	0.06 (p = 0.34)	0.02 (p = 0.74)	0.01 (p = 0.91)	0 (p = 0.98)	0.04 (p = 0.59)	0.08 (p = 0.21)
NaDP_V1	-0.06 (p = 0.38)	0.06 (p = 0.34)	-0.03 (p = 0.60)	-0.05 (p = 0.48)	0 (p = 0.98)	0.12 (p = 0.08)
MIP_V1	0.07 (p = 0.30)	-0.01 (p = 0.91)	0.09 (p = 0.15)	0.04 (p = 0.51)	-0.03 (p = 0.60)	-0.03 (p = 0.63)
MIPP_V1	-0.10 (p = 0.11)	0.01 (p = 0.92)
NIP_V1	-0.03 (p = 0.70)	0.04 (p = 0.58)	-0.03 (p = 0.68)	-0.01 (p = 0.93)	-0.07 (p = 0.29)	0.02 (p = 0.80)
NaTLP	-0.01 (p = 0.91)	-0.02 (p = 0.79)	0.05 (p = 0.45)	-0.05 (p = 0.49)	-0.02 (p = 0.77)	0.02 (p = 0.76)
MTLP_V1	0.03 (p = 0.67)	0.04 (p = 0.51)	-0.09 (p = 0.16)	0.04 (p = 0.57)	0.13 (p = 0.05)	0.05 (p = 0.45)
STLP_V1	-0.04 (p = 0.54)	-0.07 (p = 0.33)	0.02 (p = 0.76)	-0.01 (p = 0.85)	-0.07 (p = 0.30)	-0.05 (p = 0.45)
ULTLP_V1	0.12 (p = 0.06)	0 (p = 0.96)	0.15 (p = 0.02)	-0.01 (p = 0.89)	-0.07 (p = 0.27)	0.04 (p = 0.57)
SMLP_V1	0.03 (p = 0.66)	0.06 (p = 0.38)	-0.03 (p = 0.64)	0.05 (p = 0.46)	0.06 (p = 0.33)	0.06 (p = 0.39)
MNLP_V1	-0.02 (p = 0.80)	0.05 (p = 0.49)	-0.13 (p = 0.05)	0.03 (p = 0.69)	0.15 (p = 0.02)	0.05 (p = 0.43)
EFSRP_V1	-0.04 (p = 0.57)	0.01 (p = 0.89)	-0.01 (p = 0.84)	-0.01 (p = 0.92)	0.01 (p = 0.84)	-0.02 (p = 0.81)
EFSRL_V1	0 (p = 0.95)	0.01 (p = 0.92)	-0.03 (p = 0.68)	-0.02 (p = 0.80)	0.04 (p = 0.53)	-0.01 (p = 0.94)
ECSR_V1	-0.04 (p = 0.51)	0.01 (p = 0.84)	-0.03 (p = 0.62)	-0.02 (p = 0.82)	0.01 (p = 0.89)	0.03 (p = 0.65)
ETSRL_V1	0 (p = 0.96)	0.04 (p = 0.57)	0.02 (p = 0.77)	0.07 (p = 0.28)	-0.04 (p = 0.59)	-0.04 (p = 0.52)
ETSRL_V1	-0.02 (p = 0.75)	-0.04 (p = 0.59)	0.01 (p = 0.85)	0.04 (p = 0.56)	-0.01 (p = 0.88)	-0.06 (p = 0.40)
ESSRP_V1	-0.03 (p = 0.65)	-0.08 (p = 0.23)	0 (p = 0.98)	0.01 (p = 0.82)	-0.02 (p = 0.71)	-0.06 (p = 0.34)
ESSRL_V1	0.01 (p = 0.88)	0.10 (p = 0.12)	0.02 (p = 0.81)	0.08 (p = 0.26)	0.03 (p = 0.67)	0.05 (p = 0.46)
MTP_V1.1	-0.09 (p = 0.19)	-0.10 (p = 0.12)	-0.02 (p = 0.75)	-0.02 (p = 0.72)	-0.05 (p = 0.43)	-0.05 (p = 0.46)
OES_V1	-0.04 (p = 0.55)	0.06 (p = 0.35)	-0.01 (p = 0.90)	0.07 (p = 0.30)	-0.01 (p = 0.93)	-0.04 (p = 0.53)
OEAA_V1	-0.01 (p = 0.93)	0.06 (p = 0.34)	-0.06 (p = 0.35)	0.02 (p = 0.79)	0.04 (p = 0.53)	0.01 (p = 0.90)
MSP_V9	-0.01 (p = 0.82)	-0.06 (p = 0.39)	0.05 (p = 0.41)	-0.08 (p = 0.20)	-0.02 (p = 0.81)	0.03 (p = 0.62)
CNSP_V14	0.02 (p = 0.77)	-0.07 (p = 0.27)	0.02 (p = 0.74)	-0.07 (p = 0.28)	-0.02 (p = 0.81)	-0.03 (p = 0.63)
CLP_V6	-0.02 (p = 0.81)	0.01 (p = 0.91)	0.05 (p = 0.41)	0.12 (p = 0.06)	-0.06 (p = 0.38)	-0.08 (p = 0.21)
JJLP_V7	-0.02 (p = 0.72)	-0.09 (p = 0.18)	-0.02 (p = 0.79)	-0.05 (p = 0.44)	-0.03 (p = 0.63)	-0.05 (p = 0.43)
JAS_V1	-0.01 (p = 0.91)	-0.06 (p = 0.37)	-0.06 (p = 0.34)	0.04 (p = 0.53)	0.07 (p = 0.27)	-0.12 (p = 0.07)
JMSR_V1	-0.09 (p = 0.17)	-0.06 (p = 0.40)	-0.06 (p = 0.33)	-0.05 (p = 0.46)	-0.07 (p = 0.31)	-0.01 (p = 0.92)
EOPP_V1	-0.03 (p = 0.66)	-0.09 (p = 0.16)	-0.03 (p = 0.61)	0.02 (p = 0.80)	-0.03 (p = 0.68)	-0.08 (p = 0.21)
NsTP_V1	-0.01 (p = 0.83)	-0.11 (p = 0.11)	-0.04 (p = 0.51)	-0.02 (p = 0.82)	0.03 (p = 0.64)	-0.05 (p = 0.45)
EOHR_V1	0.09 (p = 0.17)	0.01 (p = 0.85)	0.05 (p = 0.44)	0.01 (p = 0.84)	0.04 (p = 0.59)	-0.01 (p = 0.87)
EOTPP_V1	0 (p = 0.98)	-0.07 (p = 0.30)	-0.08 (p = 0.24)	-0.03 (p = 0.63)	0.05 (p = 0.49)	-0.02 (p = 0.72)
FTR_V1	0.05 (p = 0.41)	0.10 (p = 0.14)	0.12 (p = 0.08)	-0.01 (p = 0.82)	-0.11 (p = 0.11)	0.05 (p = 0.47)
FEAS_V3	-0.01 (p = 0.89)	-0.04 (p = 0.54)	0.01 (p = 0.87)	-0.02 (p = 0.77)	-0.04 (p = 0.50)	-0.04 (p = 0.53)
FTAS_V3	-0.06 (p = 0.38)	0.03 (p = 0.61)	-0.04 (p = 0.58)	0.03 (p = 0.69)	-0.02 (p = 0.81)	0 (p = 0.94)
FJAS_V3	-0.05 (p = 0.44)	-0.05 (p = 0.50)	-0.09 (p = 0.19)	0.07 (p = 0.29)	0.08 (p = 0.24)	-0.10 (p = 0.14)
FTLR_V3	0.02 (p = 0.72)	0.10 (p = 0.15)	0.11 (p = 0.11)	0.03 (p = 0.70)	-0.08 (p = 0.21)	0.04 (p = 0.57)
CTLR_V3	0.02 (p = 0.79)	0.05 (p = 0.49)	0.09 (p = 0.17)	-0.02 (p = 0.76)	-0.05 (p = 0.41)	0.04 (p = 0.57)
FPLP_V3	0.03 (p = 0.70)	0.11 (p = 0.10)
PHP_V4	-0.02 (p = 0.71)	0.09 (p = 0.16)	0.03 (p = 0.60)	-0.03 (p = 0.65)	-0.04 (p = 0.58)	0.05 (p = 0.44)
FWLP_V3	0.02 (p = 0.73)	-0.04 (p = 0.53)	0.08 (p = 0.20)	-0.10 (p = 0.13)	-0.10 (p = 0.12)	0.05 (p = 0.48)

Boosted regression tree models produced only modest predictive performance for estimates of rank order (see Table 31). Despite their superior performance with the Kendal Tau tests for significant correlations, boosted regression tree models utilizing unilateral facial biometrics performed poorly, with very little improvement in cross validation error with addition biometric information above a simple intercept model. As expected from the Kendal Tau results, very little variance in pasture rank was explained by bilateral biometric traits for pasture rank estimates, but nearly half the variance in pen rank estimates was accounted for in the optimized model using bilateral facial biometrics, namely nostril position angle (NPA). Provided that the repeatability of pen rank estimates is likely less than 1, this model likely accounts for quite a large proportion of variance in rank dominance in a pen environment (see Figure 42).

Table 31: Performance of boosted regression tree models against pen and pasture rank

	Overall Facial Biometrics		Left Facial Biometrics		Right Facial Biometric	
	Pasture	Pen	Pasture	Pen	Pasture	Pen
Tree Depth	1	1	1	1	1	1
Number of Trees	5	219	5	5	5	5
R-Squared	0.089	0.492	0.168	0.112	0.154	0.165

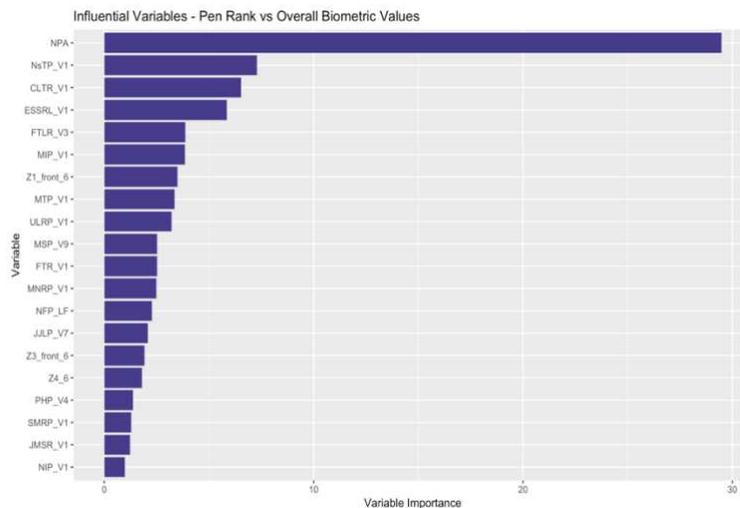


Figure 42: Comparison of variable importance values from boosted regression tree results (see Chapter 2 Methods for Biometric Abbreviations)

Networks generated from the pasture and pen datasets had a diameter of 4 and 6 nodes respectively, and reciprocities of 0.366 and 0.27. Betweenness measures proved quite normally distributed, and when plotted against estimated rank order, appeared randomly scattered for the pen dataset, but demonstrated some quadratic and heteroskedastic tendency for the pasture dataset, where the very highest and lowest ranking animals proved less central to the graph (see Figure 43). Attempts to estimate the betweenness of a cow within the graph using boosted regression trees proved largely ineffective with little if any improvement in cross-validation error, even with rank order included as a covariate (see Table 32). Similarly, assortativity analysis returned no highly significant results (see Table 33). With only four biometrics returning p-values below the standard $\alpha = 0.05$ significance level, these significant results could simply reflect compounding error from the large number of explanatory variables tested. However, it is important to note that Eye Length Proportion (*Z7_length_6*) was significant for both pen and pasture networks. Given the separation observed in PCA embeddings of pen and pasture milk order observations, and the subsequent lack of correlation between pen and pasture rank orders, this could be viewed as roughly independent results, which would decrease the chance of false positive result. The consistent negative assortativity result could suggest that, while changes in environment resulted in significant changes in milk order, cows consistently strove to space out animals with similar eye length proportions. Additionally, it is interesting to note that for the pen network, eye biometrics consistently yielded significant or near-significant p-values for assortativity results.

Table 32: Performance of boosted regression tree model against estimates of vertex betweenness

	Overall Facial Biometrics		Left Facial Biometrics		Right Facial Biometric	
	Pasture	Pen	Pasture	Pen	Pasture	Pen
Tree Depth	1	1	1	1	1	1
Number of Trees	5	5	5	5	5	5
R-Squared	0,083	0,216	0,205	0,228	0,061	0,178

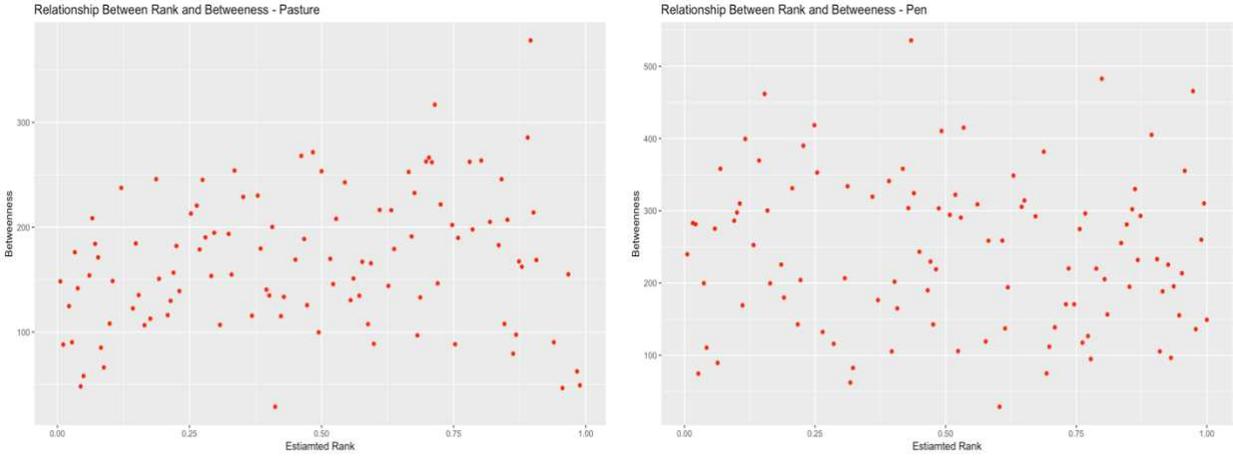


Figure 43: Relationship between estimated rank and vertex betweenness

Table 33: Assortativity of facial biometrics within the pen and pasture networks

Biometric	Pasture Network		Pen Network	
	Assortativity	P-Value	Assortativity	P-Value
Age	0.021	0.33	-0.006	0.82
CLP_V6	-0.012	0.59	-0.001	0.98
CLTR_V1	-0.010	0.64	0.018	0.52
CNSP_V14	-0.044	0.02	-0.012	0.66
CTLR_V3	0.021	0.31	-0.005	0.85
CTP_V1	-0.011	0.63	0.006	0.84
ECSR_V1	-0.002	0.92	-0.036	0.19
EFSRL_V1	-0.015	0.48	-0.005	0.84
EFSRP_V1	-0.018	0.39	-0.018	0.50
EOEHR_V1	-0.003	0.89	-0.036	0.19
EOPP_V1	0.002	0.91	-0.006	0.84
EOTPP_V1	-0.003	0.90	-0.008	0.78
ESSRL_V1	0.015	0.47	0.017	0.52
ESSRP_V1	0.024	0.26	0.028	0.30
ETSRL_V1	0.006	0.77	0.025	0.37
ETSRP_V1	0.005	0.82	-0.029	0.30
FEAS_V3	0.022	0.29	-0.018	0.51
FJAS_V3	0.001	0.97	-0.014	0.62
FPLP_V3	0.018	0.40	0.014	0.61
FTAS_V3	-0.005	0.80	0.006	0.83
FTRL_V3	0.009	0.69	-0.005	0.85
FTR_V1	0.009	0.68	0.005	0.87
FWLP_V3	0.015	0.50	-0.015	0.58
JAS_V1	-0.001	0.96	-0.021	0.45
JJLP_V7	-0.004	0.84	-0.018	0.50
JMSR_V1	-0.022	0.30	-0.033	0.24
MDP_V1	0.012	0.58	-0.022	0.45
MIP_V1	-0.037	0.07	0.002	0.95
MIPP_V1	-0.011	0.62	0.007	0.79
MNLP_V1	-0.013	0.54	-0.003	0.91
MNRP_V1	-0.011	0.62	-0.041	0.14
MSP_V9	-0.022	0.30	-0.003	0.91
MTLP_V1	-0.010	0.64	-0.001	0.97
MTLP_V1.1	-0.010	0.65	-0.001	0.97
MTP_V1	-0.025	0.23	0.006	0.82
MTP_V1.1	-0.012	0.57	-0.002	0.93
NaDP_V1	-0.025	0.22	0.000	0.99
NaTLP	-0.018	0.41	-0.011	0.68
NDP_V1	0.012	0.57	-0.018	0.50
NFP_LF	-0.002	0.94	-0.016	0.58
NFPP_LF	0.004	0.86	0.033	0.23
NIP_V1	0.010	0.64	-0.029	0.29
NPA	-0.028	0.18	-0.033	0.23
NRP	-0.034	0.10	-0.018	0.52
NsTP_V1	-0.016	0.46	-0.029	0.30
OEAA_V1	0.004	0.84	-0.037	0.18
OES_V1	0.004	0.85	-0.004	0.89
PHP_V4	-0.009	0.67	-0.024	0.38
SMLP_V1	-0.003	0.89	-0.001	0.97
SMRP_V1	0.009	0.69	-0.014	0.63
STLP_V1	0.009	0.68	0.006	0.84
ULRP_V1	-0.029	0.16	-0.033	0.23
ULTLP_V1	0.013	0.56	0.000	1.00
Z1_front_6	-0.022	0.29	-0.011	0.68
Z10_poly_6	-0.024	0.20	-0.049	0.03
Z11_linear_6	-0.005	0.81	-0.016	0.55
Z11_poly_6	0.002	0.92	-0.047	0.07
Z2_6	-0.016	0.44	0.000	1.00
Z3_front_6	-0.015	0.49	-0.040	0.14
Z4_6	-0.031	0.14	-0.048	0.08
Z7_length_6	-0.045	0.02	-0.057	0.03
Z9_poly_6	0.017	0.44	-0.030	0.27

Discussions

Before further exploring these results, it seems prudent to first examine the limitations of this data set and the impact they could have on the analytic tools utilized. While milk order data represents a promising avenue by which to explore herd structures, by virtue of both the volume at which it is produced and the ease with which it may be obtained, it is undoubtedly a nosier source of dyadic interaction data than would likely be obtained either by noninvasively observing herd interactions in their home environment or else by performing controlled behavioral tests on pairs of animals. Further, as these two sources of social interaction data have never been directly compared with data collected from milk order data, it cannot be conclusively stated that milk order data reliably mirrors underlying herd structures. And while it seems reasonable to assume that dominant animals would want to enter the milking machine first in order to have first access to post-milking TMR, there is no formal research indicating how herd hierarchy is expressed entering the milking parlor, nor how that motivation may change in different environments and under different management conditions. The results of the percolation and conductance analysis to estimate rank order clearly reflected the artificially linearized nature of social interactions extracted from milk order data, calling into question just how much of the complex social structure present in the herd can be extracted from this source of data. Given these limitations, these analyses should be regarded largely as a pilot study and the insights gleaned more the start of a conversations than the last word.

Bearing in mind these limitations, these network-based analyses of milk order data did yield several interesting results. Perhaps the most interesting, and one likely warranting further study, are multiple results underscoring the dynamic nature of milk order data. PCA visualizations revealed a clear shift in herd dynamics upon entering the milk parlor with changing environments.

This change in milk order could simply be a reflection of non-social variables, such differences in the willingness of cows to leave the pasture relative the pen, or else reflect the substantial difference in walking distance to the parlor between pen and pasture environments. To have five adjacent days visually identifiable as outliers within the pasture period, and then to have the milk order immediately shift back within the range of previous days, however, gives the impression that changes in milk order may be quite adaptive. Further, it is interesting to note the complete lack of correlation between rank order estimates generated from pen and pasture data. If changes in milk order were attributable to changes in motivations, such as a decreased willingness to enter the milk parlor coming from pasture compared to pen, it seems reasonable to assume that this sentiment would be shared among a significant proportion of animals, which should have been detected as a modest negative correlation. Assuming rank order estimates generated from this data are not completely nonsensical, this result suggests shifts in milk order in response to changing environment are more complex in nature.

With respect to results relating to prediction of rank order, there do appear to be some non-trivial relationships between facial biometrics and rank determined by milking order. The most significant association found in both the Kendal Tau and boosted regression tree variable importance analyses was between pen rank order and nostril position angle (NPA). In Chapter2, this biometric was also identified as a potential predictor of Feed Efficiency, Heifer Conception Rate, and Still Births. Provided that in antiquated face reading techniques practiced by gypsy horsemen, sloping noses are associated with “a dominant character” and “a strong tendency to test each new rider”, which aligns correctly with the direction of correlation observed here, this result potentially raises red flags as to potential unintended consequences of selection for these robustness traits (Tellington-Jones 1995). Another interesting result observed in the Kendal Tau

results was the increased significance of several biometric correlations when assessed using photos from only one side of the face, despite the increase in measurement error realized in using fewer observations to generate the BLUP estimate as compared to bilateral biometric traits. While this could simply be a figment of spurious correlations, it could also indicate that unilateral traits provided an improvement in estimation of traits related to social dominance as to outweigh the subsequent loss of measurement repeatability. This later explanation may be lent further support by the observation that pen rank was particularly well predicted by unilateral biometrics on the right side of the face. The left hemisphere of the brain, which is connected to the right eye is associated with established patterns of behavior typically found in non-stressful environments (Rogers 2010). Analogously, in traditional Chinese face reading, the right side of the face is typically associated with public expression of emotion and social traits. Thus, improved predictive performance of unilateral biometrics from the right side of the face for a response measure predominantly relating to social interaction would seem to agree with both traditional face reading techniques and more modern scientific literature, though further research would be needed to fully confirm this trend.

While unilateral traits may have out-performed bilaterally estimated facial biometrics on univariate tests of significance, the only boosted regression tree model to demonstrate notable predictive ability was for bilateral facial biometrics against estimates of pen rank. This may prove a contradictory result to the conclusion drawn from the Kendal Tau tests that unilateral biometrics provide more targeted information on which to draw social inferences, or it may simply be a reflection of the volatility undoubtedly created in this cross-validation optimization from the higher proportion of measurement error present in unilateral biometric estimates. Regardless, bilateral biometric traits were capable of inferring nearly 50% of the total variability in rank order

estimates. In assessing the performance of this model, however, it should be noted that a key assumption was violated: independence. Given the lack of correlation in rank order estimates for a given cow in this study between home pen environments, it does not seem reasonable to assume that a cow's rank would not also be influenced by different social environments. In other words, a cow that ranked moderately high in a pen environment in this herd might rank near the bottom of the herd in the same environment in a group containing more dominant animals. While feed order has been shown to be reasonably consistent for swine in a dynamic group environment (Horback & Parsons 2016), it does not appear that the repeatability of milk order position in cows has been assessed. Thus, it is not possible to determine what proportion of variance in the current rank order estimates may be inherent to the cow, and thus what proportion of variance observed in rank order estimates should in turn be potentially predicted by facial biometrics. Additional work following individual cows across multiple herd compositions would be needed to fully assess the robustness of this modeling result.

Finally, while analysis of traits demonstrated by networks generated from pen and pasture datasets proved largely insignificant when considered against facial biometric traits, there are several interesting trends observed in the assortativity analyses warranting further examination. First, while very few biometrics demonstrated statistically significant trends, those that were demonstrated only negative assortativity values. Further, even among insignificant assortativity results, the notable majority were negative. This would suggest that birds of a feather generally do not flock together within the herd, and that cows demonstrating similar facial biometric traits tend to space themselves out within the network. While this trend is only subtle for this analysis, it may warrant further analysis with perhaps higher quality didactic information than that which can be extracted from milk order data. Additionally, Eye Length Proportion demonstrated

significant negative assortativity values for both pen and pasture networks, a surprising result given the stark differences in rank orders generated by these two datasets. Provided again that in antiquated face reading techniques practiced by gypsy horsemen, a tight short eye is associated with “high degree of anxiety and tension”, and that a long almond-shaped eye may correspond to an animal that is “introverted and slightly standoffish”, a negative assortativity value for this trait seems to agree with the antiquated beliefs (Tellington-Jones 1995). Unlike with Nostril Position Angle, however, Eye Length Proportion was not identified as a consistent predictor of health traits, perhaps because this trait is not traditionally associated specifically with a personality trait that would necessarily result in a competitive advantage. This result may suggest that herd structure is not simply predicated only on dominance, and that milk order records may be able to pick up on some of these subtler social structures.

Conclusions

These analyses revealed milk order records, and subsequent estimates of rank order generated from them, to be surprisingly dynamic in response to changing environments, a result that likely merits further study. While results did reveal some potential to predict various facets of herd social structure, these relationships seem somewhat obscured by noise within both explanatory variables and the response. At least one feature, Nostril Position Angle, appeared robustly correlated with rank order in a pen environment. As a significant for several dairy genomic traits, this biometric may be representative of undesirable dimensions of dairy robustness. Results also indicate that unilateral biometrics may offer an advantage over bilateral estimates of facial structure for social traits, a result that should be explored further to further refine future models utilizing facial biometric information.

Implications

With further validation, results of these analyses could have two important implications for dairy management. The first major result, that milk order appears to change in response to changes in environment, could potentially be used to identify otherwise undetected changes in the herd environment. While here environment pertained to a major change in the physical location of the cows, it is possible that this result might generalize to other potentially subtler changes, such as sudden changes in feed or water quality, or perhaps outbreaks of disease. Additionally, while this study could not identify any obvious patterns in the shift between rank orders estimated from pen and pasture data, closer analysis across a wider range of environmental changes might reveal the presence of more complex strategies underlying such reorganizations, which may in turn yield insights into the coping strategies of dairy cattle to management stress and how farmers may be able to better support such responses.

The next major result, that there appears to be at least one facial biometric that correlates with both estimates of genomic merit and dominance as indicated by milk order, lends further credence to concerns over the unintended consequences of increased selection for dairy robustness traits (Gibbons *et al* 2009; Llonch *et al* 2018). While further research is needed to confirm the robustness of this result across a broader range of herd environments and compositions, this correlation could potentially be leveraged in selection indices to penalize cows whose genomic advantages for fitness traits may be attributable to dominance. This would allow for differentiation in fitness phenotypes attributable to dominance and more broadly adaptable genes without requiring data inputs from time consuming behavioral tests. This application seems particularly promising if biometric values and information from milk order data could potentially be combined to yield an even more robust estimate of dominance in genetic analyses.

REFERENCES

- Bates, D., Maechler, M., Bolker, B., Walker, S. "Fitting Linear Mixed-Effects Models Using lme4." 2015 Journal of Statistical Software. 67:1. 1-48.
- Beilharz, R.G., Zeeb, K. "Social Dominance in Dairy Cattle." 1982. Applied Ethology. 8. 79-97.
- Belyaev, D.K. Destabilizing selection as a factor in domestication. 1979. J. Hered. 70, 301-308.
- Bradley, R. A. & Terry, M. E. "Rank analysis of incomplete block designs". I. The method of paired comparison. 1952. Biometrika. 39. 324-345.
- Camerlink, I., Menneson, S., Turner, S.P., Farish, M., Arnott, G. "Lateralization influences contest behaviour in domestic pigs." 2018. Nature Scientific Reports. 8:12116. 1-9.
- Csardi, G., Nepusz, T. "The igraph software package for complex network research." 2006. Inter Journal Complex Systems. 1695.
- Egger-Danner, C., Col, J.B., Pryce, J.E., Gengler, N., Heringstad, B., Brandley, A., Stock, K.F. "Invited review: overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits." 2014. Animal. 1-17.
- Fujii, K., Jin, J., Shev, A., Beisner, B., McCowan, B., Fushing, H. "Perc: Using Percolation and Conductance to Find Information Flow Certainty in a Direct Network." 2016. R package version 0.1.2.
- Fushing, H., McAssey, M. P., Beisner, B. & McCowan, B. "Ranking network of a captive rhesus macaque society: a sophisticated corporative kingdom." 2011. PLoS ONE 6, e17817.
- Fushing, H., McAssey, M. P., Beisner, B. & McCowan, B. "Computing a ranking network with confidence bounds from a graph-based Beta random field." 2011. Proceeding Royal Society A.
- Gibbons, J.M. "The Effect of Selecting for "Robustness" on Temperament in Dairy Cows" 2009. Thesis: University of Edinburgh.
- Grandin, T. Deesing, M.J. "Genetics and Animal Welfare" In Genetics and the Behavior of Domestic Animals. 2014. Elsevier Inc.
- Hasegawa, N., Nishiwaki, A., Sugawara, K., Ito, I. "The effects of social exchange between two groups of lactating primiparous heifers on milk production, dominance order, behavior and adrenocortical response." 1997. Applied Animal Behavior. 51. 15-27.

- Llonch, P., Somarriba, M., Duthie, C.A., Troy, S., Roehe, R., Rooke, J., Haskell, M.J., Turner, S.P. "Temperament and dominance relate to feeding behaviour and activity in beef cattle: implications for performance and methane emissions." 2019. *Animal*. 1-10.
- Oltenacu, P.A., Broom, D.M. "The impact of genetic selection for increased milk yield on the welfare of dairy cows." 2010. *Animal Welfare*. 19:S. 39-49.
- Phillips, C.J.C., Oevermans, H., Syrett, K.L., Jespersen, A.Y., Pearce, G.P. "Lateralization of behavior in dairy cows in response to conspecifics and novel persons." 2015. *Dairy Sci*. 98:4. 2389 – 2400
- Phillips, C.J.C., Rind, M.I. "The Effects of Social Dominance on the Production and Behavior of Grazing Dairy Cows Offered Forage Supplements." 2002. *Dairy Science*. 85. 51-59.
- R Core Team. "R: A language and environment for statistical computing." 2018. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Robins, A., Phillips, C. "Lateralised visual processing in domestic cattle herds responding to novel and familiar stimuli" 2010. *LATERALITY*. 15:5. 514 - 534
- Rogers, L.J. "Relevance of the brain and behavioral lateralization to animal welfare." 2010. *Applied Animal Behavior*. 127. 1-11.
- Soffie, M., Thines, G., De Marneffe, G. "Relationship between milking order and dominance in group of dairy cows." 1976. *Applied Ethology*. 2. 271-276
- Tellington-Jones, Linda. "Getting in TTouch: Understanding and Influence Your Horse's Personality." 1995. Trafalgar Square Publishing: North Pomfret, Vermont.
- Turner, S.P., Lawrence, A.B. "Relationship between maternal defensive aggression, fear of handling and other maternal care traits in beef cows." 2007. *Livest. Sci*. 106, 182-188.
- Turner, S.P., White, I.M.S., Brotherstone, S., Farnworth, M.J., Knap, P.W., Penny, P., Mendle, M., Lawrence, A.B. "Heritability of post-mixing aggressiveness in grower-stage pigs and its relationship with production traits." 2006. *Animal Science*. 82. 615-620.
- Vandeleest, J.J., Beisner, B., Hannibal, D.L., Nathman, A.C., Capitano, J.P., Hsieh, H., Atwill, E.R., McCowan, B. "Decoupling social status and status certainty effects on health in macaques: a network approach." 2016. *Peer J*.
- Wickham, H. "ggplot2: Elegant Graphics for Data Analysis." 2016 Springer-Verlag New York.

EXECUTIVE SUMMARY

In the agricultural sciences, history can often be as rich a source of inspiration for innovation as scientific canon. One technique that has persisted since antiquity in small pockets of the modern livestock industry is the visual assessment of facial phenotypes as indicators of an animal's temperament, health, and productive potential. The purpose of this project was to more rigorously test the veracity of this proposed biological relationship. The ultimate goal was to determine if measures of facial morphology demonstrated robust correlations to estimates of cattle quality already heavily utilized within the industry, such as genomic merit, and also to traits for which effective means of measurement have not been established, such as temperament and social behaviors. To do this, novel image analysis techniques were combined with powerful statistical learning tools to provide a more holistic assessment of both the quality and quantity of predicative information contained in facial structures.

The first step was to develop and validate image analysis tools capable of comprehensively quantifying subtle variations in bovine facial structures. Measurement system validation was performed on images acquired in a working farm environment, with minimal animal restraint or disruption of normal chore schedules, using images whose quality would be achievable with any standard digital camera or smart phone device. Images were first annotated with a standardized set of anatomical reference points. The pixel coordinate locations of these reference points were then analyzed using novel geometric biometrics to produce quantitative measures of facial shape. Measurement system validation revealed geometric biometrics to be more resistant to variations in image resolution and quality than simple linear measures. Geometric biometrics also showed improved signal-to-noise ratios, and less correlation in error structures. These results indicate that

a geometric approach to biometric extraction yields metrics more suitable to the assumptions of standard predictive modeling tools than simple linear measures.

The next goal was to identify prospective relationships facial biometrics and measures of genomic merit for standard dairy production, health, and fertility traits. While predictive models featuring biometrics proved to be somewhat volatile, a result likely attributable to the noisiness of facial biometrics calculated from a single image, several models retained a notable number of biometric values, particularly calving ease, still births, feed efficiency, and several of the Zoetis Clarified health traits. Further, several of the biometrics appeared in models for multiple genomic response variables, suggesting they might be indicative of more broadly adaptive traits. Modeling also revealed that simple linear models may not fully capture the relationship between biometrics and measures of genomic merit, with some evidence provided for presence of non-linearity and the importance of interaction terms. These results indicate that while single-shot biometrics may not provide reliable predictions of genomic merit on their own, they may still have value as indicator traits to improve the reliability of genetic merit estimates.

Finally, exploratory analyses were conducted using network-based techniques to identify potential correlations between facial biometrics and social behaviors in lactating dairy cattle. Information on herd structure was generating using milking order data in a closed herd of cattle observed in both a pen and pasture environment. Milking order proved surprising dynamic over time across environments. Rank order estimates from pen and pasture data proved completely uncorrelated. One biometric result, which had proved predictive of genomic measures for productivity and health, also demonstrated a highly significant correlation to rank in the pen environment. There was also some evidence for laterality in the predictive potential of facial biometrics for rank order, with facial biometrics proving particularly well suited to estimating rank

in a pen environment. Additional analysis of network characteristics also revealed a fairly consistent tendency for cows with similar biometric values to space themselves out within the herd network, suggesting there might be a general preference highly heterogenous social structures. One biometric value also proved significantly predictive for their heterogenous tendency in both pen and pasture environments. These result give some evidence that facial biometrics may also be used to predict behavioral traits that are otherwise difficult to measure in a production environment, and may prove uniquely suited to correcting for the influence of social aggression and dominance in outcome-based genetic selection parameters.