

Generalized additive mixed model on the analysis of total transport losses of market-weight pigs¹

Tiago L. Passafaro,[†] Denise Van de Stroet,[‡] Nora M. Bello,^{||,◊} Noel H. Williams,[§] and Guilherme J. M. Rosa^{†,¶,2}

[†]Department of Animal Sciences, University of Wisconsin, Madison, WI 53706; and [‡]PIC, Pig Improvement Company, Des Moines, IA 50105; ^{||}Department of Statistics, Kansas State University, Manhattan, KS 66506; [§]Iowa Select Farms, Iowa Falls, IA 50126; [¶]Department of Biostatistics & Medical Informatics, University of Wisconsin, Madison, WI 53706

ABSTRACT: Transportation losses of market-weight pigs are an animal welfare concern, and result in direct economic impact for producers and abattoirs. Such losses are related to multiple factors including pig genetics, human handling, management, and weather conditions. Understanding the factors associated with total transport losses (TTL) is important to the swine industry because it can aid decision-making, and help in the development of transportation strategies to minimize the risk of losses. Hence, the objective of this study was to investigate factors associated with TTL on market-weight pigs in typically field conditions for Midwestern United States using a generalized additive mixed model (GAMM). The final quasi-binomial GAMM included the fixed (main and interactions) effects of abattoir of destination, type of driver, average market weight, distance traveled, wind speed, precipitation, and

temperature-humidity index (THI), as well as the random effects of truck companies and the combination of site of origin and period of the year. Results indicate significant associations between TTL and the main effect of all explanatory variables ($P < 0.05$), except for wind speed and precipitation. Interactions of average market weight \times abattoir, and wind speed \times precipitation were also significant. A complex nonlinear relationship between TTL and model covariates were observed for distance traveled, THI, and interaction terms. This study showed that TTL of market-weight pigs are caused by a complex system involving multiple interacting factors, which can be potentially managed to mitigate the risk of losses. In addition, the GAMM showed to be a simple and flexible approach to model TTL because it can capture nonlinear relationships, handle non-normal data, and can potentially accommodate data structure.

Key words: generalized additive mixed model, market-weight pigs, total transport losses

© The Author(s) 2019. Published by Oxford University Press on behalf of the American Society of Animal Science. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

J. Anim. Sci. 2019.97:2025–2034
doi: 10.1093/jas/skz087

INTRODUCTION

The transport of market-weight pigs is a critical and integral part of the modern pig production

in the United States. In 2016, more than 118 million pigs were transported to the market in the United States (USDA-NASS, 2018). However, transportation is a stressful event in the pigs' life, which might result in undesirable losses. Such losses are caused either by deaths during the transportation, termed as dead on arrival (DOA), or by non-ambulatory pigs (NAP), which are unable to continue with their own group during the transportation process (Anderson et al., 2002) due to various reasons including lameness, traumas, and

¹The authors would like to acknowledge Iowa Select Farms for providing the data, and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the financial support of the first author.

²Corresponding author: grosa@wisc.edu

Received January 16, 2019.

Accepted March 8, 2019.

injuries. The sum of DOA and NAP is commonly referred to as total transport losses (TTL).

TTL are a big problem for commercial pig production in the United States, with a direct economic impact on producers and abattoirs. Such economic impact is magnified as transport losses occur at the last step of the production cycle, thus encompassing all aggregated costs to produce a pig. Annual transportation losses in the United States have been estimated at 50 to 100 million dollars (Ellis et al., 2003; Ritter et al., 2009). TTL are also a concern from an animal welfare and well-being perspective. DOA and NAP are useful indicators of animal welfare and often reflect adverse conditions during transportation from farms to abattoirs (Vitali et al., 2014). Moreover, higher prevalence of TTL may result in negative publicity for pig companies and abattoirs. Hence, regulations and procedures to improve animal welfare during transportation are continuously debated by the swine industry (NPB, 2017).

TTL are related to multiple factors including pig genetics, human handling, management decisions, truck type, facilities design, and weather conditions (Rademacher and Davies, 2005; Ritter et al., 2006; Haley et al., 2008a; Fitzgerald et al., 2009; Sutherland et al., 2009; Vitali et al., 2014; NPB, 2017). These studies helped the swine industry in developing techniques and standard procedures to mitigate the risk of losses, but many other contributing factors to TTL remain unexplained (Fitzgerald et al., 2009). In this context, observational studies are extremely important as they provide data from real field conditions, which can be very helpful to investigate additional factors associated with TTL and thus, help inform management decisions.

Transport losses for each truckload, expressed as the number of pigs lost over the total number of pigs in the shipment, follow a binomial process. Binomial response variables can be analyzed using generalized linear models, such as logistic regression (Nelder and Wedderburn, 1972). This method uses a linear predictor to describe the relationship between the response variable and the covariates, in the logistic scale (Hastie and Tibshirani, 1986). To accommodate potential nonlinear relationships, an alternative is to use generalized additive models (GAM; Hastie and Tibshirani, 1986; Wood, 2017). Whenever a GAM incorporates both fixed and random effects, it is termed as a generalized additive mixed model (GAMM). Such GAMM model might be appropriate to study TTL of market-weight pigs because of the binomial nature of

this trait and its potentially nonlinear relationships with explanatory variables. Therefore, the objective of this study was to investigate factors associated with TTL on market-weight pigs in typical field conditions in Midwestern United States using a GAMM approach.

MATERIALS AND METHODS

Animal care and use approval were not obtained for this study because the data provided for statistical analysis came from an existent database.

Data Description

Shipments reports of market-weight pigs collected from July 2014 to December 2015 were provided by Iowa Select Farms, Iowa Falls, IA. For each shipment of pigs, available data included information about farm of origin, abattoir of destination, date of shipment, management group (i.e., pigs were raised in wean-to-finishing or finishing groups), transportation company, type of driver (i.e. driver is the owner of the truck or employee of transportation company), the number of pigs loaded, average market weight per shipment, and the latitude and longitude of all sites and abattoirs (Table 1). Latitude and longitude were used to estimate the distance traveled between farms and abattoirs, by using the R package *gmapsdistance* (Melo et al., 2018). Moreover, the daily average measurements for temperature, relative humidity, wind speed, and precipitation were retrieved from the National Oceanic and Atmospheric Administration (NOAA) government organization. In addition, the temperature humidity Index (THI) was calculated using the following formula (NOAA, 1976):

$$THI = T - [0.55 - (0.0055 \times RH)] \times [T - 14.5],$$

where T is the temperature in degree Celsius, and RH is the relative humidity in percentage. The exact time of the transportation was not available in the dataset. Hence, the date of the shipment was used to estimate the weather conditions for each load by applying the weighted K-nearest neighborhood algorithm available in the R package *kknn* (Hechenbichler and Schliep, 2016), using the information of 22 weather stations. In this approach, the weather condition for each shipment was estimated similar to a geometric mean weighted by the distance between farms and weather stations.

Table 1. Description of explanatory variables observed on 26,919 shipments, transporting 4,567,514 market-weight pigs to two abattoirs from July 2014 to December 2015

Explanatory variables	Description	Units	Number of levels
Abattoir	The packing plant of destination	—	2
Average market weight	The average body weight of the pigs in the shipment	kg	—
Date ¹	The date of the shipment	—	—
Latitude ²	The north–south orientation of the farm and abattoir location	Degrees	—
Longitude ²	The east–west orientation of the farm and abattoir location	Degrees	—
Management group	Pigs were raised in finishing or wean-to-finishing groups	—	2
Pigs	The number of pigs in each shipment	Pigs	—
Precipitation	The estimated average daily precipitation of each shipment	mm	—
Relative humidity ³	The estimated average daily relative humidity of each shipment	%	—
Site	The farm of origin	—	420
Temperature ³	The estimated average daily temperature of each shipment	°C	—
THI	The estimated average daily temperature-humidity index of each shipment	—	—
Traveled distance	The distance between farms and abattoirs	km	—
Truck company	The company responsible for the transportation	—	78
Type of driver	The driver is the truck owner or employee	—	2
Wind speed	The estimated average daily wind speed of each shipment	m/s	—

¹Provided information on the day, month, and year of each shipment.

²Latitude and longitude were used to estimate the travel distance between farms and abattoirs.

³Temperature and relative humidity was used to create a temperature-humidity index (THI).

Data editing consisted of removing observations for which: farm of origin, abattoir of destination, or the number of animals in the shipment was not reported; truck companies with missing information or <20 shipments; loadings with <100 or >210 pigs; and any combination of site and quarter of the year with less than five observations. After this procedure, the final dataset included a total of 4,567,514 pigs from 420 farms loaded in 26,819 shipments delivered to two abattoirs.

Generalized Additive Models

A GAM is a generalized linear model in which part of the linear predictor is expressed in terms of the sum of smooth functions (Hastie and Tibshirani, 1986; Wood, 2017). For any response variable y_i ($i = 1, \dots, n$) with a probability distribution from the exponential family with mean $\mu_i = E(y_i)$ and scale parameter φ , i.e., $y_i \sim \text{EF}(\mu_i, \varphi)$, the linear predictor of a GAM is often represented by the following structure:

$$g(\mu_i) = \mathbf{a}_i \boldsymbol{\theta} + \sum_j f_j(x_{ij}),$$

where g is a known link function, \mathbf{a}_i is the i th row of an incidence matrix, $\boldsymbol{\theta}$ is the vector of corresponding parameters, and f_j is the j th ($j = 1, \dots, k$) unknown smooth function of covariates x_{ij} . The smooth functions, f_j , are commonly depicted by

reduced rank smoothing splines (Wood et al., 2017), including different kind of polynomials such as the P-spline, adaptive variants, tensor products, thin plate, and cubic splines (Eilers and Marx, 1996; Wood, 2006; Wood et al., 2017). Any reduced rank smoothing spline can be represented as $f_j = \mathbf{X}_j \boldsymbol{\beta}_j$, in which \mathbf{X}_j is an $n \times p_j$ incidence matrix containing the smooth spline basis functions evaluated at vector \mathbf{x}_j , and $\boldsymbol{\beta}_j$ is the corresponding regression coefficient vector.

A common challenge of spline regression is to define the type and size of the basis functions in a way to prevent model overfitting. This problem can be addressed by adding a penalization term in the model likelihood. Hence, the estimation of the unknown parameters of the model seeks to maximize $\hat{\boldsymbol{\beta}} = \arg \max_{\boldsymbol{\beta}} l(\boldsymbol{\beta}) - \boldsymbol{\beta}^T \mathbf{S}_\lambda \boldsymbol{\beta} / 2$, where $l(\boldsymbol{\beta})$ is the model $\log^{\boldsymbol{\beta}}$ likelihood, $\boldsymbol{\beta}^T$ is $\boldsymbol{\beta}$ transposed, and $\mathbf{S}_\lambda = \sum_j \lambda_j \mathbf{S}_j$, in which \mathbf{S}_j is a matrix with known coefficients and λ_j are the smoothing parameters, which control the tradeoff between goodness of fit and smoothness of the model. When $\lambda_j \rightarrow 0$, the penalty term has no effect on the j th smoothing spline function, which will closely interpolate the data observations, leading to a rougher adjustment. On the other hand, when $\lambda_j \rightarrow \infty$ the penalty term has a substantial effect on the j th smoothing spline function, which become almost perfectly smooth, approximating to a straight line. The estimation of

the unknown parameters of a GAM can be performed by penalized iteratively re-weighted least squares, given λ_j . In practice, the smoothing parameters are unknown but they can be estimated by generalized cross-validation (Craven and Wahba, 1979) or by restricted maximum likelihood estimation (Wood, 2011).

Generalized additive models can be extended to accommodate random effects by defining the model in an empirical Bayesian approach (Wood, 2017). In this framework, smooth functions are believed to be more smooth than wiggly, and this can be formalized by specifying an improper prior distribution for β proportional to $\exp[(\lambda\beta^T\mathbf{S}_\lambda\beta)/2]$. Re-parametrizations of the model are critical so that simple priors can be used, such as a uniform distribution for the fixed effects, and a Gaussian distribution for the random components, i.e., $N(0, \mathbf{I}\lambda^{-1})$. The large sample approximation can be used to obtain credible intervals for β using a Gaussian posterior distribution $N(\hat{\beta}, [\mathbf{X}^T\mathbf{W}\mathbf{X}/\varphi + \mathbf{S}_\lambda]^{-1})$, where $\mathbf{W} = \text{diag}(w_i)$ and $w_i = \text{variance}(\hat{\mu}_i)\hat{g}'(\hat{\mu}_i)^2$. Such intervals are calibrated with the confidence intervals obtained by frequentist inference (Wood, 2013).

Statistical Analysis

TTL (y_i) were fit with a GAMM with the R package *mgcv* (Wood, 2011, 2017). Individual shipments were the observation units of this study, and TTL were modeled as $y_i \sim \text{Binomial}(n_i, p_i)$, where n_i is the number of hogs in shipment i and p_i is the probability of loss for that specific shipment. The variance of the Binomial distribution is a function of its own mean, making overdispersion a common challenge for logistic regression models in which $n_i > 1$. Hence, a quasi-likelihood approach was considered to accommodate potential overdispersion in the data. Linear covariates were scaled to display mean zero and variance one and to make regression coefficients more interpretable. The model was determined gradually by taking the following steps: (1) introduction of random effects to specify data architecture; (2) fit of a base GLMM including the fixed effects of abattoir, type of driver, management group, distance traveled, average market weight, wind speed, precipitation, and THI, as well as the random effects of truck companies and combination of site and quarter of year; (3) test two-way interactions between all fixed effects with a forward stepwise procedure considering the model deviance (the lowest value) and the statistical significance of

an interaction as criterion to introduce interactions into the base model; (4) replacement of all significant linear covariates and interactions by reduced ranking smoothing cubic splines; and (5) removal of variables for which the main effect and interactions displayed no significance. In this study, a 5% level of significance was adopted and the Pearson's correlations and the variation inflation factor were used to investigate issues of collinearity among covariates. Regression coefficients and confidence intervals for all explanatory variables were exponentiated to be interpreted as odds ratios. After the forward selection procedure, the final model for TTL was specified as follows:

$$(y_i | \alpha_{k[i]}, \gamma_{j[i]}) \sim \text{Overdispersed Binomial}(n_i, p_i, \varphi)$$

$$\text{logit}(p_i) = \alpha_{k[i]} + \gamma_{j[i]} + a_{b[i]} + d_{o[i]} + \sum_{e=1}^9 b_{t_e}(t_i)\beta_{t_e}$$

$$+ \sum_{g=1}^9 b_{s_g}(s_i)\beta_{s_g} + \sum_{h=1}^9 b_{r_h}(r_i)\beta_{r_h}$$

$$+ \sum_{p=1}^9 b_{u_p}(u_i)\beta_{u_p} + \sum_{c=1}^9 b_{w_c}(w_i)_{[a_{b[i]}}]\beta_{w_c.[a_{b[i]}}$$

$$+ \sum_{l=1}^5 \sum_{m=1}^5 \beta_{s_l, r_m} b_{s_l}(s_i) b_{r_m}(r_i)$$

where $a_{b[i]}$ is the fixed effect of abattoir, with two levels; $d_{o[i]}$ is the fixed effect of the type of driver (employee or owner); $b_{t_e}(t_i)$, $b_{s_g}(s_i)$, $b_{r_h}(r_i)$, and $b_{u_p}(u_i)$ are the reduced ranking cubic splines basis with knots e , g , h , and p placed at their corresponding 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th quantiles for distance traveled (i.e., t_i), wind speed (i.e., s_i), precipitation (i.e., r_i), and THI (i.e., u_i), respectively; β_{t_e} , β_{s_g} , β_{r_h} , and β_{u_p} are the regression coefficients for distance traveled, wind speed, precipitation, and THI, respectively; $b_{w_c}(w_i)_{[a_{b[i]}}$ is the reduced ranking cubic spline basis for average market weight with knots c placed at 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th quantiles nested within the abattoir b ; $\beta_{w_c.[a_{b[i]}}$ is the regression coefficients for average market weight nested within the abattoir b ; β_{s_l, r_m} is the regression coefficient for the interaction between wind speed and precipitation which is described by a tensor product of the cubic splines $b_{s_l}(s_i)$ and $b_{r_m}(r_i)$ with knots l and m placed at their corresponding 10th, 30th, 50th, 70th, and 90th quantiles; $\alpha_{k[i]}$ is the random effect of the k th level of the combination of sites and quarters of the year; and $\gamma_{j[i]}$ is the random effect of the j th truck company. Random

effects were assumed to be independent from each other with distributions $\alpha_{k[j]} \sim N(0, \lambda_1^{-1})$ and $\gamma_{j[i]} \sim N(0, \lambda_2^{-1})$, where λ_1^{-1} and λ_2^{-1} are the smoothing parameters for the combination of sites and quarters of year and truck companies, respectively. The $\text{logit}(p_i) = \log(p_i/(1-p_i))$ is the link function, and φ is the scale parameter which is estimated based on the Pearson χ^2 statistics and quantifies data dispersion. When $\varphi > 1$ the data are underdispersed, for $\varphi < 1$ the data are overdispersed, and for $\varphi = 1$ the statistical model reduces to a binomial logistic regression. The cubic splines basis functions were obtained as follows: $b_1(x) = 1$, $b_2(x) = x$ and $b_{z+2}(x) = R(x, x_z^*)$ for $z = 1, 2, 3, \dots, q$, where x_z^* is the location of each knot z ; x is the observed value; and $R(x, x_z^*) = \left[(x_z^* - \frac{1}{2})^2 - \frac{1}{12} \right] \left[(x - \frac{1}{2})^2 - \frac{1}{12} \right] / 4 - \left[(|x - x_z^*| - \frac{1}{2})^4 - \frac{1}{2} (|x - x_z^*| - \frac{1}{2})^2 + \frac{7}{240} \right] / 24$.

RESULTS AND DISCUSSION

Descriptive Statistics of TTL

The overall TTL of market-weight pigs from July 2014 to December 2015 was 0.76%, in which 0.19% and 0.57% were attributed to DOA and NAP, respectively (Table 2). Fitzgerald et al. (2009) found incidences of 0.85%, 0.25%, and 0.60% for TTL, DOA, and NAP, respectively. In a review paper, Ritter et al. (2009) reported an average of TTL of 0.62%, with 0.25% and 0.37% of DOA and NAP. Overall, the incidence of TTL in the present study was within the industry range of 0.14% to 2.39% (Ritter et al., 2009), with observed percentage of DOA numerically smaller than those found in previous studies (Fitzgerald et al., 2009; Ritter et al., 2009), while the occurrence of NAP was numerically higher than those reported by Ritter et al. (2009). These differences in the incidence of TTL, DOA, and NAP may be explained by various reasons including differences across pig production systems, personnel, management procedures, companies' logistics, period when the data was collected, and pig genetics.

Weather Variables

The THI was significantly associated with TTL ($P < 0.0001$, Table 3). The risk of loss per shipment was fairly constant for THI values between -10 and

10 , with major increases below and above this interval (Figure 1A). This result indicates two critical values for TTL at extreme conditions of the THI, suggesting that additional care must be considered in these cases. For instance, for THI greater than 10 , pigs should be transported at the coolest time of the day and cooling devices may be installed in trucks (i.e., fans), while for THI lower than -10 , trucks must be provided with some kind of bedding (Sutherland et al., 2009; Vitali et al., 2014). Different from our study, Fitzgerald et al. (2009) modeled THI with a quadratic polynomial function and reported major increases of TTL above 1 unit of THI, indicating that the probability of TTL drastically increased after this point. Differences in the shape of the curve and in the number of critical values are explained by the polynomial function used across studies. Penalized cubic splines are more flexible than quadratic polynomials, justifying part of such difference.

The interaction between wind speed and precipitation was associated with TTL ($P = 0.0209$, Table 3). For example, for higher values of precipitation, the risk of losses decreased with wind speed above 4.5 ms^{-1} , while it increased for lower values of wind speed (Figure 1B). On the other hand, in days with no precipitation, TTL seems larger with lower values of wind speed. This result indicates a complex relationship between transportation losses, wind speed, and precipitation. Outside wind speed can affect the air circulation inside trucks, especially for vehicles relying on passive ventilation system, which is typically used in commercial pig transportation in the United States. McGlone et al. (2014) studied different boarding levels to quantify transportation losses relative to the amount of ventilation inside trucks in the Midwestern United States. These authors found increased losses for low boarding levels (0% to 30% of truck side-vents are closed) and temperatures below $-5 \text{ }^\circ\text{C}$, highlighting the importance of adjust ventilation inside trucks. In addition, precipitation may provide some cooling effect, decreasing the temperature inside trailers. Therefore, wind speed and precipitation have an influence on trucks micro-environmental conditions, which can potentially affect pigs' thermo-regulation mechanisms. Furthermore, wind speed and precipitation can affect TTL by changing road conditions. For instance, in harsh transportation conditions, trucks may become more instable resulting in additional stress for pigs, which can potentially increase the risk of transport losses. Hence, the effect of wind speed and precipitation

Table 2. Descriptive statistics of sample proportions of transportation losses, and numerical variables observed on 26,919 shipments, transporting 4,567,514 market-weight pigs to 2 abattoirs from July 2014 to December 2015

Variable	Mean	SD	Minimum	Maximum
DOA, %	0.19	0.45	0.00	7.69
NAP, %	0.57	0.84	0.00	8.82
Total losses, %	0.76	1.04	0.00	9.88
Number of pigs per shipment	170.31	8.44	100.00	201.00
Average market weight, kg	125.50	5.77	104.60	145.09
Distance traveled, km	136.64	63.44	35.62	396.50
Average daily wind speed, ms ⁻¹	4.21	1.78	0.52	10.98
Average daily precipitation, mm	2.28	5.91	0.00	58.10
Average daily THI	9.71	9.63	-16.53	26.27

DOA = dead on arrival; NAP = non-ambulatory pigs; THI = temperature-humidity index.

Table 3. Parameter estimates and approximate significance level of smoothing functions on total transport losses of market-weight pigs using a generalized additive mixed model

Parameters	Estimates	Odds ratios	Confidence interval (odds ratios)		P-value
			Lower limit	Upper limit	
Intercept	-4.945	0.007	0.006	0.008	<0.0001
Abattoir: B	-0.323	0.720	0.719	0.722	<0.0001
Driver: Owner	-0.142	0.868	0.866	0.870	<0.0001
Smoothing functions			EDF ¹	Ref. DF ²	P-value
Distance travelled			3.2862	9	0.0034
THI			5.0547	9	<0.0001
Wind speed			0.0004	9	0.9998
Precipitation			1.0021	9	0.8837
Average market weight × abattoir A			4.9742	9	<0.0001
Average market weight × abattoir B			5.0547	9	<0.0001
Wind speed × precipitation			1.5295	16	0.0209

EDF = effective degree of freedom; Ref. DF = reference number of degrees of freedom.

on TTL need to be investigate in further studies for a better understand of such interaction.

Abattoir, Type of Driver, Distance Traveled, and Average Market Weight

TTL were negatively associated with the type of driver ($P < 0.0001$, Table 3). The estimated risk of losses was 0.867 (95% CI: 0.865 to 0.870) times lower when trucks were driven by owners instead of by employees. This result might be interpreted as truck owners having a vested interest, thus being more careful at handling and moving pigs during the whole process of the transportation, compared with contracted haulers. Fitzgerald et al. (2009) found a significant association of drivers with TTL, while Haley et al. (2008b) reported that ~8% of transport losses variation was attributed to differences between drivers. To the best of the authors'

knowledge, no previous studies have reported the influence of the type of driver on the risk of TTL. These results suggest that understanding in greater depth the differences between drivers may further explain transportation losses, and thus be useful to the swine industry to contract, hire, or train drivers.

The distance traveled was associated with TTL ($P = 0.0034$, Table 3). The risk of losses increased with distances up to 125 km and decreased from this point to 397 km (Figure 1C). The stress incurred during the loading period may result in the increase in TTL observed for shipments traveling <125 km, as pigs do not have enough time to recover prior to arriving at the abattoir. Ritter et al. (2006) found that the vast majority of stressed and fatigued pigs can reestablish normal conditions if they are allowed to rest 2 to 3 h. In addition, Sutherland et al. (2009) hypothesized that longer trips may allow extra time for pigs to recover and rest from the

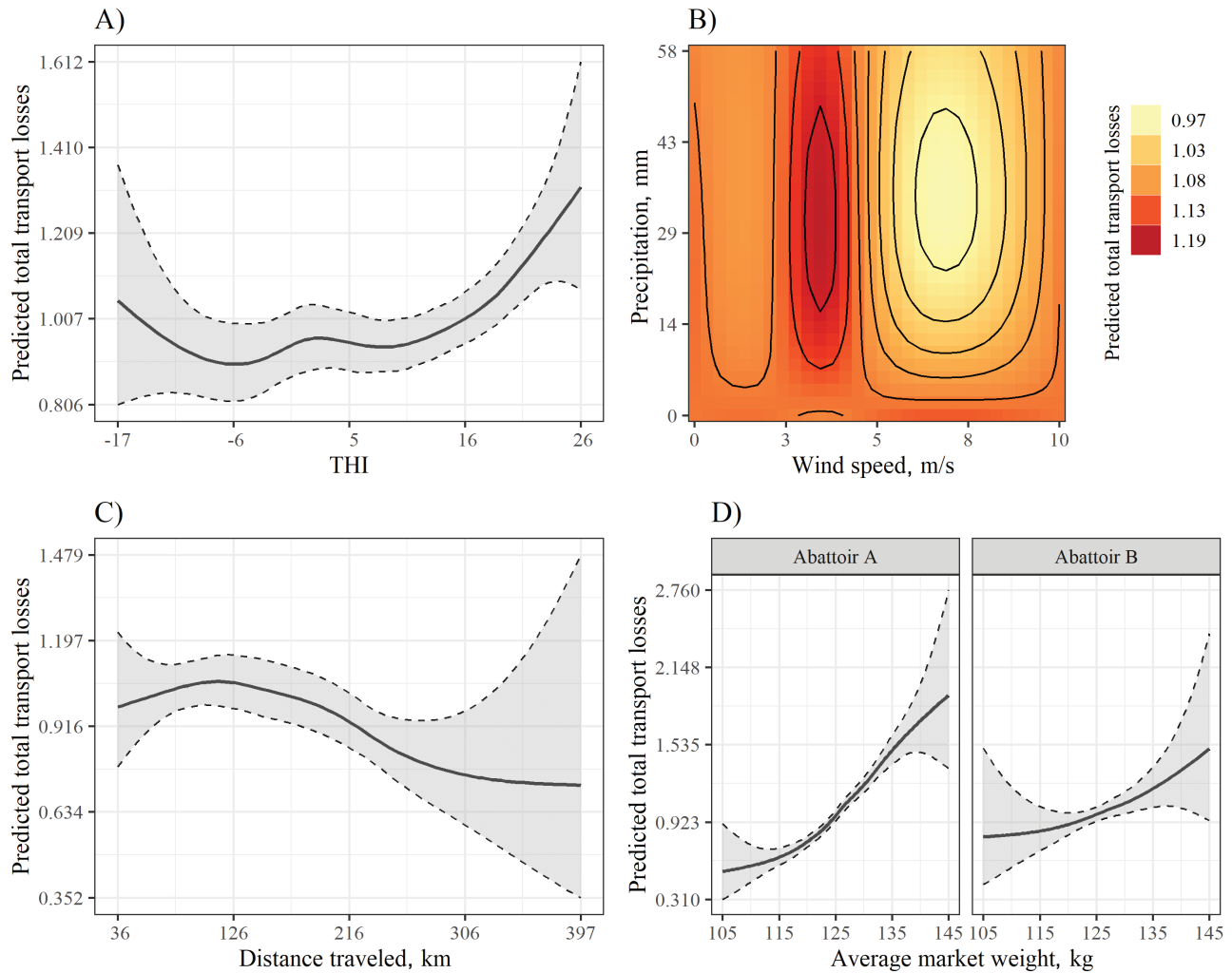


Figure 1. Predicted total transport losses of market-weight pigs on the odds ratios scale for each significant reduced ranking cubic splines: (A) THI, (B) interaction between precipitation \times wind speed, (C) distance traveled, and (D) interaction between average market weight \times abattoir.

early stress caused by the loading process, decreasing the overall transportation losses. These findings are consistent with Canadian studies that found a greater risk of TTL in shorter than longer distances (Haley et al., 2008a,b). Loading approaches including restricted use of electric prods, loading groups of four to six pigs, and avoid mixing of pigs from different pens (Goumon and Faucitano, 2017) can be used to mitigate the stress incurred at loading. Therefore, such strategies can help to reduce TTL, especially in short distance trips.

The interaction between average market weight and abattoir was associated with TTL ($P < 0.0001$, Table 3). This result leads to different estimates of losses through the values of live weights within each abattoir (Figure 1D). TTL increased along the values of average market weight in both abattoirs, with a faster increment in the risk of losses in one facility relative to the other. A linear positive association between average market weight and TTL

has been reported in other studies (Ellis and Ritter, 2005; Rademacher and Davies, 2005). According to Ellis and Ritter (2005), heavier pigs may show a greater metabolic response compared to lighter pigs during transportation, as a consequence of handling procedures. However, Hamilton et al. (2004) reported minimal influence of live weight (104 vs. 128 kg) on blood acid–base status under different handling intensity in finishing pigs. The stock density inside trucks can be used to adjust the average market weight within shipments, reducing the risk of TTL. Floor space of 0.48 m² is recommended for each pig with about 129 kg to reduce transportation losses (Ritter et al., 2006). In addition, another strategy to monitor the average market weight is to use split marketing. This approach consists in removing the heaviest 25% to 50% pigs from a pen to market 1 to 2 weeks earlier than their corresponding pen-mates (Scroggs et al., 2002; Conte et al., 2012), with the intent to

deliver pigs with optimal weights to the abattoir. The benefits of split marketing include reduction in the production cost, and shipments with more uniform weights (Scroggs et al., 2002; Conte et al., 2012). Despite these advantages, economic and logistic aspects of the pig production system should be taken into consideration before a farm adopts the split marketing strategy.

The abattoir of destination is also an important source of variation for TTL, indicating differences in the risk of losses across plants. This result may reflect differences in policies (i.e., discard of non-ambulatory pigs, and grid payment), facilities (i.e., unloading deck and ramp), management procedures (i.e., waiting and unloading time), and personnel of each abattoir. Haley et al. (2008b) studied the random effect of 33 abattoirs on transport losses of market-weight pigs delivered in Canada and the United States. These authors found that ~16% of transport losses variation was attributed to differences among abattoirs, highlighting the importance of this factor in the risk of TTL.

Ultimately, economic or logistics analysis to quantify the benefits of employing management strategies related to factors associated with TTL were not investigated in our study. For instance, suppose that the swine industry is willing to contract only drivers that are the owners of their trucks, as an attempt to decrease TTL. Such decision, however, may incur in an extra shipping cost which could potentially lowered the net economic benefit of adopting such strategy. In addition, this strategy could reduce the overall availability of such type of driver (i.e., owner) compared with the industry demand. Hence, the adoption of such practice may not be justified by economic or logistics aspects alone. However, TTL is an animal welfare issue, and thus, adopting management practices to reduce the risk of loss should still be considered. Improving animal welfare in pig production is crucial to ensure better quality of life for pigs, as well as to improve the overall consumer perception about the good practices employed in the swine industry to enhance animal welfare.

Generalized Additive Mixed Models

Mixed models have the benefit of accommodating random effects into the statistical analysis, thereby accounting for the hierarchical organization of the data, which is naturally inherited from the data collection process. Taking into consideration, the structure of the data allows the partition of the total random variation across multiple

levels of the data architecture, such as farms and transport companies, providing a more calibrated inference for hypothesis testing (Tempelman, 2009; Bello and Renter, 2018). In this study, we treated the combination of sites and quarters of the year, as well as truck companies as random effects, and they represented 21.2% and 0.59% of the total random variation of TTL, respectively. These results suggest that the differences among truck companies are small and may reflect a uniformity in the industry on standard management procedures and truck designs between pig carriers. Conversely, the combination of sites and quarters of the year accounted for a considerable amount of variance, indicating large differences across sites and seasons of the year. Farms facilities (e.g., barn design), handling techniques, and personnel may partially explain the variability of sites and period of the year combination. This finding is consistent with Haley et al. (2008b), who reported that 25% of the total variation in in-transit losses was attributed to differences among farms.

GAMMs are a simple and flexible statistical regression technique with the capability to capture nonlinear relationships between explanatory and response variables. This feature of GAMM is often investigated by graphical visualization of the smoothing function since the interpretation of a simple linear coefficient is not meaningful in GAMM. The visual inspection of the model covariates clearly displayed a nonlinear pattern for distance traveled, THI, and both interactions of average market weight and abattoir, and wind speed and precipitation (Figure 1C, A, D, and B). Moreover, the graphical inspection of such curves has the advantage to potentially identify regions in the curve where the risk of transport losses substantially increases. For instance, the smoothing function of THI showed major increases on the risk of TTL between -10 and 10. Another benefit of GAMMs is to mitigate issues of overfitting, very common for example with linear models involving higher order polynomials. To avoid this problem, GAMMs introduce penalties in the maximum likelihood estimation so that the resulting curves are smoother and can potentially increase the model ability to predict new observations (Gareth et al., 2007; Wood, 2017).

Statistical inference using GAMMs is useful to provide guidance to support decision-making. The estimates of regression coefficients and smoothing functions can be used to define strategies to mitigate the risk of TTL. For instance, imagine a producer transporting pigs in the summer to an

abattoir specialized in processed meats (i.e., ham). To make these products, heavier pigs are demanded, resulting in an expected increase in the risk of TTL with higher values of average market weight. In this scenario, the producer could contract drivers who are the owner of their trucks and try to transport pigs at the coolest times of the day in vehicles containing cooling devices (e.g., fans). Therefore, according to specific situations, decisions can be tuned to balance the risk of TTL.

It is important to notice though that, without some additional assumptions, the information conveyed by regression models reflects only associations between the response and explanatory variables, with not necessarily a direct causal interpretation (Rosa and Valente, 2013; Bello et al., 2018). The existence of confounders not accounted for in the model can generate spurious associations limiting the claim of causal mechanisms. Rosa and Valente (2013) and Bello et al. (2018) reviewed the potential use and implications of inferring causal effects from observational data in livestock production. These authors discussed specific statistical methods applied to causal inference, the challenges, and assumptions considered to claim a causal mechanism from observation data, as well as how causal interventions could improve management decisions on livestock production. Although GAMMs are not specifically designed to infer causal effects from observational data, such modeling approach is still important to support management decisions, in the sense that it provides predictive information of TTL for new observations.

CONCLUSIONS

TTL of market-weight pigs are caused by a complex system involving multiple interacting factors, and nonlinear relationships. In the current study, the risk of TTL was associated with several factors including type of driver, distance traveled, THI, and two interactions terms (e.g., average market weight \times abattoir and wind speed \times precipitation). GAMMs showed to be an efficient approach to investigate factors associated to TTL, as it can handle non-normal distribution of the response variable (e.g., Binomial, Poisson, and Gamma), accommodate nonlinearity between predictors and response variable, and take in consideration the hierarchical structure of the data by introducing random effects. Therefore, GAMMs provide a powerful prediction tool for the swine industry to forecast potential losses during transportation. In addition, GAMMs

can contribute with a better understanding of the factors associated with TTL, which can aid decision-making and development of management strategies to try to minimize such losses. However, as mentioned before, the GAMM analysis performed in this study was implemented on a predictive context, and thus, the causal interpretation of parameters is mostly speculative. Therefore, further analyses investigating causal inference on TTL of market-weight pigs are a subject of our future work.

Conflict of interest statement. None declared.

LITERATURE CITED

- Anderson, D. B., D. J. Ivers, M. E. Benjamin, H. W. Gonyou, D. J. Jones, K. D. Miller, R. K. McGuffey, T. A. Armstrong, D. H. Mowrey, L. F. Richardson, et al. 2002. Physiological responses of market hogs to different handling practices. *Proceedings of American Association of Swine Veterinarians, Kansas City, MO*; p. 399–400.
- Bello, N. M., V. C. Ferreira, D. Gianola, and G. J. M. Rosa. 2018. Conceptual framework for investigating causal effects from observational data in livestock. *J. Anim. Sci.* 96:4045–4062. doi:10.1093/jas/sky277
- Bello, N. M., and D. G. Renter. 2018. Invited review: reproducible research from noisy data: revisiting key statistical principles for the animal sciences. *J. Dairy Sci.* 101:5679–5701. doi:10.3168/jds.2017-13978
- Conte, S., P. G. Lawlor, N. O’Connell, and L. A. Boyle. 2012. Effect of split marketing on the welfare, performance, and carcass traits of finishing pigs. *J. Anim. Sci.* 90:373–380. doi:10.2527/jas.2010-3453
- Craven, P., and G. Wahba. 1979. Smoothing noisy data with spline functions. *Numer. Math.* 31:377–403. doi:10.1007/BF01404567
- Eilers, P. H. C., and B. D. Marx. 1996. Flexible smoothing with B-splines and penalties. *Stat. Sci.* 11:89–121. doi:10.1214/ss/1038425655
- Ellis, M., F. McKeith, D. Hamilton, T. Bertol, and M. Ritter. 2003. Analysis of the current situation: what do downers cost the industry and what can we do about it? In: *Proceedings of 4th American Meat Science Association Pork Quality Symposium, Columbia, MO*; p. 1–3.
- Ellis, M., and M. Ritter. 2005. Transport losses : causes and solutions. In: *Allen D. Leman Swine Conference. Saint Paul, MN*; p. 176–178.
- Fitzgerald, R. F., K. J. Stalder, J. O. Matthews, C. M. Schultz Kaster, and A. K. Johnson. 2009. Factors associated with fatigued, injured, and dead pig frequency during transport and lairage at a commercial abattoir. *J. Anim. Sci.* 87:1156–1166. doi:10.2527/jas.2008-1270
- Gareth, J., W. Daniela, H. Trevor, and T. Robert. 2007. *An introduction to statistical learning with applications in R.* 1th ed. New York, NY: Springer.
- Goumon, S., and L. Faucitano. 2017. Influence of loading handling and facilities on the subsequent response to pre-slaughter stress in pigs. *Livest. Sci.* 200:6–13. doi:10.1016/j.livsci.2017.03.021
- Haley, C., C. E. Dewey, T. Widowski, and R. Friendship. 2008a. Association between in-transit loss, internal trailer

- temperature, and distance traveled by Ontario market hogs. *Can. J. Vet. Res.* 72:385–389.
- Haley, C., C. E. Dewey, T. Widowski, Z. Poljak, and R. Friendship. 2008b. Factors associated with in-transit losses of market hogs in Ontario in 2001. *Can. J. Vet. Res.* 72:377–384.
- Hamilton, D. N., M. Ellis, T. M. Bertol, and K. D. Miller. 2004. Effects of handling intensity and live weight on blood acid-base status in finishing pigs. *J. Anim. Sci.* 82:2405–2409. doi:10.2527/2004.8282405x
- Hastie, T., and R. Tibshirani. 1986. Generalized additive models. *Stat. Sci.* 1:297–318. doi:10.1214/ss/1177013604
- Hechenbichler, K., and K. Schliep. 2016. kkn: weighted k-nearest neighbors. R package version 1.3.1. <https://CRAN.R-project.org/package=kkn> – [accessed 1 February 2018].
- McGlone, J., A. Sapkota, A. Johnson, and R. Kephart. 2014. Establishing trailer ventilation (boarding) requirements for finishing pigs during transport. *Animals (Basel)*. 4:515–523. doi:10.3390/ani4030515
- Melo, R. A., D. T. Rodriguez, and D. Zarruk. 2018. gmapsdistance: distance and travel time between two points from Google Maps. R package version 3.4. <https://CRAN.R-project.org/package=gmapsdistance> – [accessed 1 March 2018].
- Nelder, J. A., and R. W. M. Wedderburn. 1972. Generalized linear models. *J. R. Stat. Soc. A.* 135:370–384. doi:10.2307/2344614
- NOAA. 1976. Livestock hot weather stress. Regional operations manual letter C-31-76. Kansas City, MO: NOAA.
- NPB. 2017. Transport quality assurance handbook. 6th ed. Des Moines, IA: National Pork Board.
- Rademacher, C., and P. Davies. 2005. Factors associated with the incidence of mortality during transport of market hogs. In: Allen D. Leman Swine Conference. Saint Paul, MN; p. 186–191.
- Ritter, M. J., M. Ellis, J. Brinkmann, J. M. DeDecker, K. K. Keffaber, M. E. Kocher, B. A. Peterson, J. M. Schlipf, and B. F. Wolter. 2006. Effect of floor space during transport of market-weight pigs on the incidence of transport losses at the packing plant and the relationships between transport conditions and losses. *J. Anim. Sci.* 84:2856–2864. doi:10.2527/jas.2005-577
- Ritter, M. J., M. Ellis, N. L. Berry, S. E. Curtis, L. Anil, E. Berg, M. Benjamin, D. Butler, C. Dewey, B. Driessen, et al. 2009. Review: transport losses in market weight pigs: I. A review of definitions, incidence, and economic impact. *Prof. Anim. Sci.* 25:404–414. doi:10.15232/S1080-7446(15)30735-X
- Rosa, G. J. M., and B. D. Valente. 2013. Breeding and genetics symposium: inferring causal effects from observational data in livestock. *J. Anim. Sci.* 91:553–564. doi:10.2527/jas.2012-5840
- Scroggs, L. V., H. G. Kattesh, J. L. Morrow, K. J. Stalder, J. W. Dailey, M. P. Roberts, J. F. Schneider, and A. M. Saxton. 2002. The effects of split marketing on the physiology, behavior, and performance of finishing swine. *J. Anim. Sci.* 80:338–345. doi:10.2527/2002.802338x
- Sutherland, M. A., A. McDonald, and J. J. McGlone. 2009. Effects of variations in the environment, length of journey and type of trailer on the mortality and morbidity of pigs being transported to slaughter. *Vet. Rec.* 165:13–18. doi:10.1136/vetrec.165.1.13
- Tempelman, R. J. 2009. Invited review: assessing experimental designs for research conducted on commercial dairies. *J. Dairy Sci.* 92:1–15. doi:10.3168/jds.2008-1404
- USDA-NASS. 2018. Agricultural statistics 2017. 1th ed. Washington, DC: National Agriculture Statistics Services.
- Vitali, A., E. Lana, M. Amadori, U. Bernabucci, A. Nardone, and N. Lacetera. 2014. Analysis of factors associated with mortality of heavy slaughter pigs during transport and lairage. *J. Anim. Sci.* 92:5134–5141. doi:10.2527/jas.2014-7670
- Wood, S. N. 2006. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics* 62:1025–1036. doi:10.1111/j.1541-0420.2006.00574.x
- Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. B.* 73:3–36. doi:10.1111/j.1467-9868.2010.00749.x
- Wood, S. N. 2013. On p-values for smooth components of an extended generalized additive model. *Biometrika*. 100:221–228. doi:10.1093/biomet/ass048
- Wood, S. N. 2017. Generalized additive models: an introduction with R. 2th ed. Boca Raton, FL: CRC Press.
- Wood, S. N., Z. Li, G. Shaddick, and N. H. Augustin. 2017. Generalized additive models for gigadata: modeling the U.K. black smoke network daily data. *J. Am. Stat. Assoc.* 112:1199–1210. doi:10.1080/01621459.2016.1195744