Computers and Electronics in Agriculture 151 (2018) 286–294

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Machine learning algorithms to predict core, skin, and hair-coat temperatures of piglets

Michael T. Gorczyca[a,*], Hugo Fernando Maia Milan[a], Alex Sandro Campos Maia[b], Kifle G. Gebremedhin[a]

[a] Department of Biological and Environmental Engineering, Cornell University, Ithaca, NY 14853, United States
[b] Laboratory of Animal Biometeorology, Department of Animal Science, Faculty of Agricultural and Veterinary Sciences, State University of São Paulo, Jaboticabal, SP 14884-900, Brazil

ARTICLE INFO

Keywords:
Bioenergetics
Machine learning
Piglets
Precision livestock farming
Temperature

ABSTRACT

Internal-body (core) and surface temperatures of livestock are important information that indicate heat stress status and comfort of animals. Previous studies focused on developing mechanistic and empirical models to predict these temperatures. Mechanistic models based on bioenergetics of animals often require parameters that may be difficult to obtain (e.g., thickness of internal tissues). Empirical models, on the other hand, are data-based and often assume linear relationships between predictor (e.g., air temperature) and response (e.g., internal-body temperature) variables although, from the theory of bioenergetics, the relationship between the predictor and the response variables is non-linear. One alternative to consider non-linearity is to use machine learning algorithms to predict physiological temperatures. Unlike mechanistic models, machine learning algorithms do not depend on biophysical parameters, and, unlike linear empirical models, machine learning algorithms automatically select the predictor variables and find non-linear functions between predictor and response variables. In this paper, we tested four different machine learning algorithms to predict rectal (T_r), skin-surface (T_s), and hair-coat surface (T_h) temperatures of piglets based on environmental data. From the four algorithms considered, *deep neural networks* provided the best prediction for T_r with an error of 0.36%, *gradient boosted machines* provided the best prediction for T_s with an error of 0.62%, and *random forests* provided the best predictions for T_h with an error of 1.35%. These three algorithms were robust for a wide range of inputs. The fourth algorithm, *generalized linear regression*, predicted at higher errors and was not robust for a wide range of inputs. This study supports the use of machine learning algorithms (specifically deep neural networks, gradient boosted machines, and random forests) to predict physiological temperature responses of piglets.

1. Introduction

One of the current challenges in agriculture is to increase food production to feed the world's growing population while considering environmental responsibilities and the comfort of the biological object (livestock; Hunter et al., 2017). In animal production, the challenge is in developing precision livestock farming techniques (Van Hertem et al., 2017; Wathes et al., 2008) to increase animal comfort and production. These techniques (Guarino et al., 2017) are focused on continuous monitoring of animal health, comfort, and production indicators, such as internal-body and skin-surface temperature. These temperatures indicate the health status and production levels of animals (Da Silva and Maia, 2013; Soerensen and Pedersen, 2015), as well as their heat stress level, estimated to cost the swine industry $300 million each year (St-Pierre et al., 2003).

Heat stress is a major issue that decreases animal welfare (Silanikove, 2000), production (Nienaber et al., 1999), reproduction (Wolfenson et al., 2000), and growth potential (Collin et al., 2001). To cope with heat stress, pigs rely on behavioral (Vasdal et al., 2009) and physiological (Brown-Brandl et al., 2001, 2014; Robertshaw, 2006) responses. Because of the importance of monitoring heat stress of pigs (Shao and Xin, 2008), and the difficulty of measuring the necessary parameters that indicate heat stress (McCafferty et al., 2015), two classical approaches are used to estimate heat stress of animals: (1) mechanistic modelling, and (2) empirical modelling.

Mechanistic models are based on the biophysical understanding of conservation of energy, momentum, and mass in live animals (Collier and Gebremedhin, 2015; DeShazer, 2009). Using conservation equations, a governing equation for the problem is formulated and solved analytically or numerically. The limitations of analytical and numerical

* Corresponding author.
E-mail address: mtg62@cornell.edu (M.T. Gorczyca).

https://doi.org/10.1016/j.compag.2018.06.028
Received 7 March 2018; Received in revised form 4 June 2018; Accepted 10 June 2018
Available online 15 June 2018

models are the assumption that internal and/or superficial temperatures are known, or a simple mathematical relationship exists between them, and/or some of the parameters are also difficult to obtain (e.g., thickness of internal tissues, etc.). Furthermore, mechanistic models reveal that the relationship between environmental and physiological responses are non-linear (Hensley et al., 2013; Milan and Gebremedhin, 2016a,b; McArthur, 1981).

Empirical models are data-based and usually assume a linear relationship between predictor variables (e.g., air temperature) and the response variable (e.g., internal-body temperature). These relationships are chosen by the researcher and has a considerable impact on the accuracy of the model (Mostaço et al., 2015; Pathak et al., 2009; Ramirez, 2017; Soerensen and Pedersen, 2015).

A third approach that is receiving increased attention from swine researchers are machine learning and computer vision algorithms (Kamilaris and Prenafeta-Boldú, 2018). Recent applications include monitoring animal behavior (Cross et al., in press; Lao et al., 2016; Nasirahmadi et al., 2017; Shao and Xin, 2008), and weight (Kashiha et al., 2014; Shi et al., 2016; Wongsriworaphon et al., 2015). In this paper, we propose the use of machine learning algorithms to predict internal-body temperature, skin-surface temperature, and hair-coat surface temperature of piglets from environmental variables. The advantage of this approach compared to mechanistic models is that it does not rely on biophysical parameters. The advantage of this approach compared to empirical models is that it automatically finds a non-linear function from the data, removing the subjectivity from the researcher choosing the relationship between predictor and response variables. To the best of our knowledge, this is the first study that applies machine learning algorithms to predict physiological temperatures of swine.

2. Materials and methods

2.1. Experimental measurements

Animal use and research protocol were approved by the Animal Care and Use Committee from São Paulo State University (FAPESP Proc. 17.519/14). The experiment was conducted in Jaboticabal, São Paulo, Brazil (21°15′40″ South Latitude and 595 m elevation) for five consecutive days. Ten 5-days-old piglets (weight = 3.76 ± 0.41 kg, mean ± SEM) from the commercial lineage "Large White" were randomly selected from the same farrowing. The farrowing was not provided with supplemental heat. The selected piglets were randomly separated into 5 groups (2 piglets in each group) and managed inside a brooder (1.0 × 1.0 × 1.0 m^3) from 3 a.m. to 8 a.m. Physiological measurements were performed hourly and started one hour after the piglets were inside the brooder (i.e., from 4 a.m. to 8 a.m.) to allow for adaptation to the environment. Four of the five groups were provided with supplemental heat (lamps) with intensities of 60 W, 100 W, 160 W, or 200 W. The fifth group (control) was not provided with supplemental heat.

Skin-surface temperature (T_s, °C) at the upper leg of the animal was measured with a skin- temperature probe (MLT422/AL, ADInstruments, accuracy ± 0.2 °C) and rectal temperature (T_r, °C) was measured with a rectal temperature probe (MLT1403, ADInstruments, accuracy ± 0.2 °C). These probes were connected to thermistor pods (ML309, ADInstruments), and the pods were connected to a data acquisition system (PL3516/P, PowerLab 16/35 and LabChart Pro, ADInstruments) that recorded data every second for approximately 5 min. Hair-coat-surface temperature (T_h, °C) at the upper leg was measured with an infrared thermometer (Model 568, Fluke, accuracy ± 1 °C). Air temperature (T_a, °C) and relative humidity (RH, %) inside the brooder were measured every minute (HOBO U12 Temp/RH, Onset, accuracy ± 0.35 °C and ± 2.5%). Black globe temperature (T_g, °C) inside the brooder was measured using a 15-cm dia. black globe installed 10 cm above the ground (thermocouple TMC20-HD, datalogger U12-013, accuracy ± 0.35 °C, Onset).

2.2. Model development

2.2.1. Data processing

The experiment was designed to provide 200 data points. Each individual data point contained the time of measurement (in hours), intensity of the supplemental heat, T_a, RH, T_g, T_r, T_s, and T_h. Time of measurement, intensity of supplemental heat, T_a, and T_g were used as predictors of T_r, T_s, and T_h. RH was not used as a predictor variable because 22% of the data was lost due to sensor failure. Further technical problems led to a reduction in the number of collected datapoints from 200 to 173. Correlations of the variables, mean and standard error of the mean were calculated. The univariate number of the outliers in the dataset was calculated using the z-score method at 2.5 standard deviations above or below the mean (Cousineau and Chartier, 2010).

The dataset was divided into training and testing datasets (Hastie et al., 2003). The training dataset was used to develop the machine learning models and the testing dataset was used to evaluate the predictive performance of the models. The training dataset consisted of 130 data points (75% of the dataset) and the testing dataset consisted of 43 points (25% of the dataset). The testing dataset was first obtained using stratified random sampling for each combination of time of measurement/intensity of supplemental heat (strata). This approach ensured that the testing dataset contained at least two data points from each stratum. Mean values were calculated for each strata of the dataset (yielding 20 data points) to determine the mean percentage error of each model for every stratum.

2.2.2. Overview of machine learning models

The machine learning algorithms used in this study were generalized linear regression model with elastic net regularization (GLM; Zou and Hastie, 2005), random forests (RF; Breiman, 2001), gradient boosted machines (GBM; Natekin and Knoll, 2013), and deep neural networks (feedforward neural networks) with the ReLU activation function (DNN; Goodfellow et al., 2016). Each algorithm has hyperparameters that influence the model learned from the data.

GLM is ordinary linear regression with penalty terms in the L_1 (sum of magnitudes) and L_2 (sum of squares) norms of the linear regression coefficients. The penalties shrink irrelevant regression coefficients and limit the impact of collinearity between the predictor variables (Zou and Hastie, 2005). The objective function of the GLM model is described as

$$\min_{\beta,\beta_0} \frac{1}{2N} \sum_{i=1}^{N} (x_i^T \beta + \beta_0 - y_i)^2 + \lambda \left[\alpha \, \|\beta\|_1 + \frac{1}{2}(1-\alpha)\|\beta\|_2^2 \right]$$

where β, β_0 are regression coefficients, the summation represents the squared residual errors, x_i is the predictor variable from the i^{th} row of data, y_i is the predicted variable from the i^{th} row of data, λ is the severity of penalty applied, and α distributes the penalty between $L_1(\|\beta\|_1)$ and squared $L_2(\|\beta\|_2^2)$ norms of the regression coefficients. The hyperparameters are λ and α.

The RF and GBM models rely on decision trees, which are simple predictive models that stratify the input data space into output areas. The output-area prediction of decision trees is the mean of the response variables from the training dataset that fall in that output area (Fig. 1). For RF, several decision trees are developed independently from different subsets of the training dataset as well as from the different predictor variables. The prediction of the RF is the average of the predictions from all decision trees. The hyperparameters for RF are number of decision trees, minimum number of observations in a leaf, number of variables used to develop each split in a decision tree, and the maximum depth of the decision trees. For the GBM model, decision trees are developed sequentially, where each new decision tree is designed to improve on the predictive performance of the previous decision trees. The hyperparameters of the GBM are nearly the same as the hyperparameters for the RF, except GBM uses all predictor variables in a

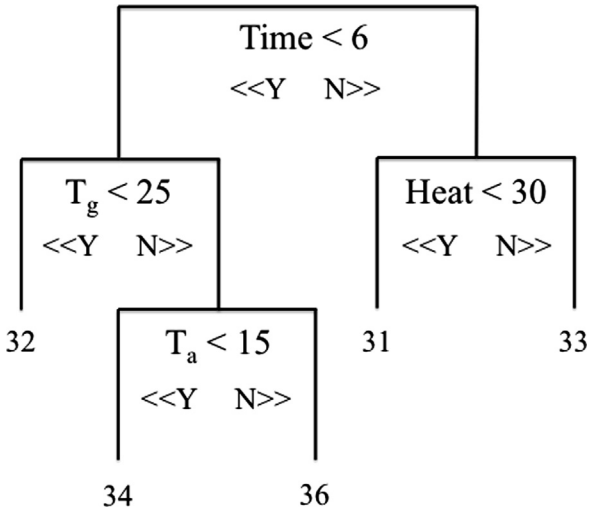

Fig. 1. Example of a decision tree for predicting hair-coat surface temperature. A decision tree is developed by segmenting the input space into structured outputs. Each decision (e.g., Time < 6) represents a split of the tree. A leaf is the end node of the tree (e.g., the node with the value of 31 for Time ≥ 6 and Heat < 30). Random forests are based on creating several decision trees and averaging their output. Gradient boosted machines are based on creating several sequential decision trees, where new trees focus on improving the prediction accuracy of previous trees, and linearly combining the predictions of these trees. Time: time of measurement (hours); Heat: intensity of supplemental heat (W); T_a: air temperature (°C); T_g: black globe temperature (°C).

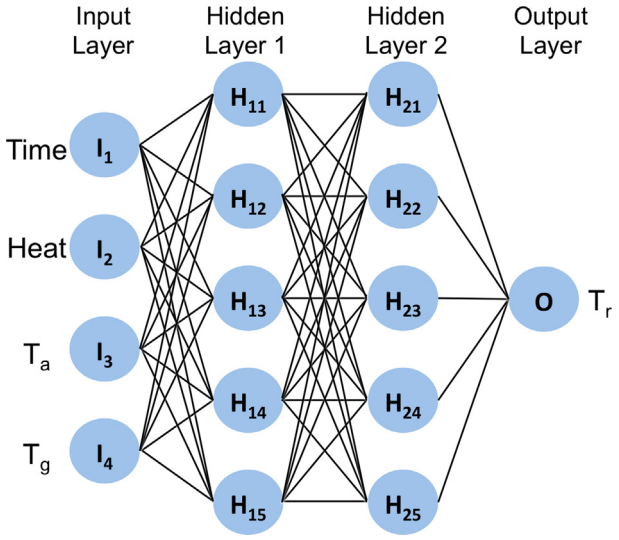

Fig. 2. Feedforward neural network. Each input variable represents one neuron (I_n) that connects to every hidden neuron in the first hidden layer (H_{1m}). Each hidden neuron is a non-linear function (activation function), where the outputs of the hidden neurons in the previous hidden layer are inputs to the hidden neurons in the next hidden layer. The outputs of the last hidden layer are inputs to the output neuron (O), which provides the prediction of the neural network. Time: time of measurement (hours); Heat: intensity of supplemental heat (W); T_a: air temperature (°C); T_g: black globe temperature (°C); I_n: input neuron n; H_{nm}: hidden neuron m of hidden layer n; O: output neuron; T_r: rectal temperature (°C).

dataset for each split. GBM also has the learning rate of the sequential trees and an annealing rate (the influence of sequential trees on the final prediction output) as hyperparameters.

DNN algorithms provide flexible and robust approaches to develop nonlinear machine learning models. Feedforward neural networks, the type of DNN used in this study, consist of an input layer, hidden layers of unobserved variables, and an output layer (Fig. 2). Given an input vector x, the output of hidden layer h is computed as follows:

$$h = f(\theta + Wx)$$

where θ is a vector of offsets, W is a matrix of weights, and f is a user selected activation (non-linear) function (ReLU was used in this study). The output from f (i.e., h) is an input for the next layer. This process is repeated until the output layer is reached. The variable calculated for the output layer, the prediction of the feedforward neural network, is calculated similarly as h but with a different activation function. In this study, the activation function for the output layer was the identity function, which is equivalent to linear regression with the variables of the last hidden layer. The hyperparameters for the DNN model are the number of hidden layers, number of neurons in each hidden layer, mini-batch size (the number of observations used in each iteration in the model optimization process), epochs (the number of times the whole training dataset is used in training), dropout percentage (the percentage of weights not updated during a mini-batch iteration to avoid overfitting; Srivastava et al., 2014), and ρ and ε (hyperparameters of the ADADELTA optimization framework; Zeiler, 2012).

2.2.3. Training and testing machine learning models

The objective of this paper was to develop machine learning models to predict T_r, T_s, and T_h using T_a, T_g, time of measurement, and intensity of supplemental heat as predictors. The machine learning models were trained in R (R Core Team, 2017) using the H2O package (The H2O.ai Team, 2017) with modular 5-fold cross-validation (Hastie et al., 2003). To develop the machine learning models, a random search for hyperparameter optimization (Bergstra and Bengio, 2012) was performed on the hyperparameter space described in Table 1. For GLM, RF, and GBM, 1000 random searches were performed (resulting in 1000 trained models for each of these algorithms). For DNN, because of its inherently larger hyperparameter space, 2000 random searches were performed (resulting in 2000 trained deep neural network models). Computations were performed on an Oryx Pro from System76, with Pop-OS 17.10, 512 GB PCIe M.2 SSD, 64 GB DDR4 RAM memory (2133 MHz), i7-6820HK (3.6 GHz), 8 GB GeForce GTX 980 M.

The mean squared error (MSE) was used as the evaluation metric

Table 1
Hyperparameter space used to sample hyperparameters for training the machine learning algorithms.

Hyperparameter	Distribution[a]	Hyperparameter	Distribution[a]
Random Forests		Generalized Linear Model	
#Trees	$\mathcal{U}_d(10, 250)$	λ	$10^{\mathcal{U}(-10,0)}$
MNOL[b]	$\mathcal{U}_d(1, 30)$	α	$\mathcal{U}(0, 1)$
NVS[c]	$\mathcal{U}_d(1, 4)$	Deep Neural Network	
Max. TreeDepth	$\mathcal{U}_d(1, 100)$	#Hidden layers	$\mathcal{U}_d(1, 4)$
Gradient Boosted Machines		#Neurons	$\mathcal{U}_d(1, 250)$
#Trees	$\mathcal{U}_d(1, 100)$	Dropout percentage	$\mathcal{U}(0, 0.33)$
MNOL[b]	$\mathcal{U}_d(1, 20)$	Epochs	$\mathcal{U}_d(1, 10^4)$
Max. Tree Depth	$\mathcal{U}_d(1, 100)$	Mini-batch size[d]	$\mathcal{U}_d(1, 130)$
Learning rate	$\mathcal{U}(0.001, 1)$	ρ	$\mathcal{U}(0.75, 0.999)$
Annealing	$\mathcal{U}(0.8, 1)$	ε	$10^{\mathcal{U}(-10,-6)}$

[a] $\mathcal{U}_d$ stands for uniform discrete random distribution from a to b; $\mathcal{U}$ stands for uniform random distribution from a to b.
[b] MNOL: minimum number of observations in a leaf.
[c] NVS: number of variables used in each split.
[d] The maximum number corresponds to the number of observations in the training dataset.

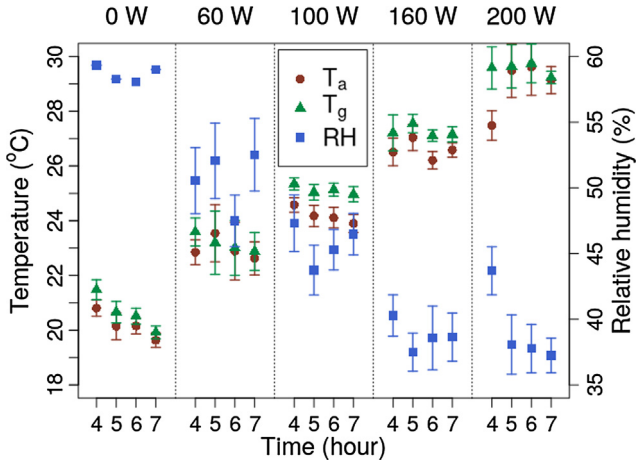

Fig. 3. Experimental data (mean ± standard error of the mean) for air temperature (T_a), black-globe temperature (T_g), and relative humidity (RH) separated by time of measurement and intensity of supplemental heat.

(Hastie et al., 2003). We used cross-validation MSE to select the best performing model from each algorithm. Of these best performing models, the overall best model was the one that minimized MSE on the testing dataset.

Robustness and generalization of the best models were tested using partial dependence plots (Friedman, 2001) for 5 artificial datasets. Each dataset was designed to test how the models would perform under different conditions. Four artificial datasets had T_a = [−20, 100] °C, T_g = [−20, 100] °C, time of measurement = [0, 24] h, or intensity of supplemental heat = [0, 1000] W, while keeping the remaining predictor variables at their mean values. The fifth artificial dataset consisted of 10,000 random combinations of these artificial values to further test how change in the predictor variables would affect the prediction from the machine learning models.

3. Results and discussion

3.1. Environmental data

Fig. 3 shows the measured environmental data from the dataset stratified for the different time of measurement and intensity of supplemental heat while Table 2 shows the coefficients of correlation, mean, standard error of the mean, and number of outliers. As expected (Monteith and Unsworth, 2013), T_a and T_g increased when the intensity of the supplemental heat increased while RH decreased.

3.2. Performance of machine learning models

Fig. 4 shows the MSE of the best performing machine learning models (that minimized cross-validation MSE) using cross-validation, training dataset, and testing dataset. Table 3 shows the hyperparameters for these models. Fig. 5 shows the training and testing MSE of these models for the training iterations. The GLM model converged at one iteration, but the other models required more than ten iterations to converge. The best overall model (based on minimum testing MSE), was DNN for T_r, GBM for T_s, and RF for T_h.

Fig. 6a shows the prediction output from the best machine learning algorithms using the mean dataset and Fig. 6b shows the absolute percentage error. The model predictions are very close to the measured values. The observed mean absolute errors of T_r, T_s, and T_h, were 0.36%, 0.62% and 1.35%, respectively (Fig. 6b). These errors are lower than those previously reported from either statistical or mechanistic models. Mostaço et al. (2015) predicted rectal temperatures of pigs with 2.5% error using multiple linear regression for air enthalpy and tympanic temperature (known to be correlated with internal body temperature; Korthals et al., 1995). Costa et al. (2010) predicted surface temperature of piglets with 5.5% error using a linear regression model. Loughmiller et al. (2001) predicted mean body-surface temperature of pigs with 3.5% error using a linear regression model. Turnpenny et al. (2000a,b) developed a mechanistic model and the resulting error was 7% for predicting skin-surface temperature of pigs.

3.3. Test of robustness and generalization of the best machine learning models

Figs. 7–9 show the partial dependence plots (Friedman, 2001) from the effect of changing one predictor variable (while keeping the remaining predictor variables at their mean values) on T_r, T_s, and T_h, respectively. These figures show, with the exception of GLM, that the machine learning models were robust with respect to the input variables because they did not produce unexpected predictions. GLM, which fits linear functions for the predictor variables, however, produced relationships that are counter to expectations, such as decreasing T_s and T_r while increasing T_g.

Fig. 10 shows the effect of randomly changing all predictor variables on T_r, T_s, and T_h, which are predicted by the best performing machine learning models. This figure shows that temperature predictions using GLM resulted in higher variance, which means that GLM is not robust to changes in the predictor variables. The predictions from RF, GBM, and DNN were, however, closer to the mean measured values and the variance of their predictions was lower, which means that these algorithms are robust to changes in the predictor variables.

Table 2
Correlation coefficients, mean and standard error of the mean, and number of univariate outliers of the measured data. The number of outliers is displayed on the rightmost column, the mean and standard error of each data variable is displayed on the main diagonal of the table, and the correlation coefficients are displayed on the remaining entries of the table. No outliers were removed from training and testing datasets.

Var.[a]	Hour	Heat	T_a	T_g	RH[b]	T_r	T_s	T_h	#Outliers
Hour	5.490 ± 1.149	−0.009	−0.020	−0.075	−0.110	−0.497	−0.699	−0.214	0
Heat	−0.009	101.850 ± 72.441	0.891	0.912	−0.706	0.330	0.225	0.642	0
T_a	−0.020	0.891	24.455 ± 3.424	0.969	−0.590	0.287	0.258	0.743	3
T_g	−0.075	0.912	0.969	25.068 ± 3.464	−0.596	0.306	0.282	0.740	0
RH[b]	−0.110	−0.706	−0.590	−0.596	43.888 ± 7.833	−0.063	0.071	−0.275	0
T_r	−0.497	0.330	0.287	0.306	−0.063	37.917 ± 0.637	0.493	0.418	5
T_s	−0.699	0.225	0.258	0.282	0.071	0.493	32.803 ± 1.506	0.428	2
T_h	−0.214	0.642	0.743	0.740	−0.275	0.418	0.428	33.836 ± 1.864	3

[a] Variables: Hour: time of measurement (hour); Heat: intensity of supplemental heat (W); T_a: air temperature (°C); T_g: black globe temperature (°C); RH: relative humidity (%); T_r: rectal temperature (°C); T_s: skin-surface temperature (°C); T_h: hair-coat surface temperature (°C).

[b] Number of samples for RH was 132.

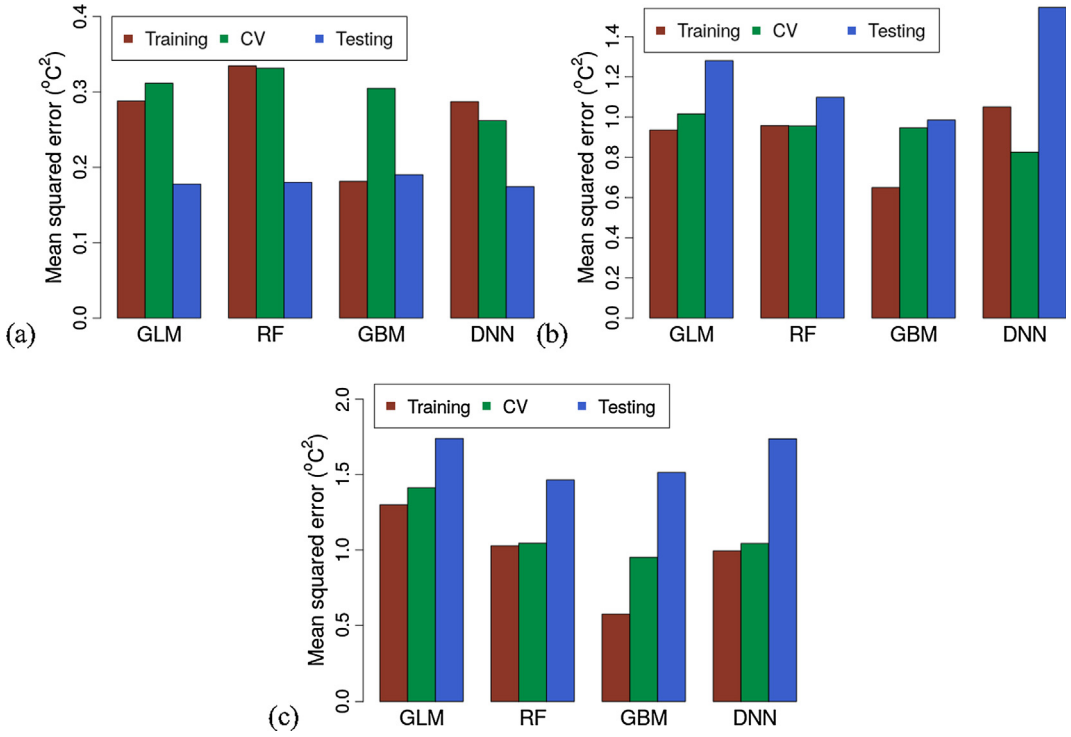

Fig. 4. Performance of the best machine learning models for predicting rectal (T_r; a), skin-surface (T_s; b), and hair-coat surface (T_h; c) temperatures. GLM: generalized linear regression model with elastic net regularization; RF: random forests; GBM: gradient boosted machines; DNN: deep neural network with ReLU activation function.

Table 3
Hyperparameters of the best machine learning models.

Hyperparameter[a]	T_r	T_s	T_h
Random Forests			
#Trees	61	46	236
MNOL[b]	4	6	2
NVS[c]	1	3	2
Max. Tree Depth	72	26	82
Gradient Boosted Machines			
#Trees	80	60	25
MNOL[b]	20	12	10
Max. Tree Depth	29	2	41
Learning Rate	0.351	0.504	0.398
Annealing	0.976	0.882	0.808
Generalized Linear Model			
λ	1.632×10^{-10}	0.240	8.749×10^{-7}
α	0.244	0.453	0.409
Deep Neural Network			
#Hidden Layers	2	4	2
#Neurons[d]	(242, 190)	(11, 53, 241, 230)	(65, 20)
Dropout Percentage[d]	(0.13, 0.19)	(0.06, 0.19, 0.17, 0.06)	(0.03, 0.04)
Epochs	14	14.567	12.129
Mini-Batch Size	53	128	82
ρ	0.876	0.946	0.914
ε	2.855×10^{-7}	5.604×10^{-7}	5.646×10^{-8}

[a] Hyperparameters of the best machine learning models to predict rectal temperature (T_r, °C), skin-surface temperature (T_s, °C), and hair-coat surface temperature (T_h, °C).
[b] MNOL: minimum number of observations in a leaf.
[c] NVS: number of variables used in each split.
[d] The numbers in parenthesis represent the value used for each hidden layer.

3.4. Limitations and potential applications of machine learning models

The main limitations of machine learning models are that they are data-based as well as time-consuming and computationally expensive to train. In addition, if the training dataset is noisy or the model is trained inappropriately, then, the model may "learn" noise instead of the non-linear relationships that may exist between the predictor variables and the response variable (Natekin and Knoll, 2013). We showed in Section 3.3 that all algorithms considered in this study, except GLM, were robust to changes in the predictor variables. It should be noted, however, that the models were trained and tested from the same data population. This means that the models proposed in this study should not be applied to different data sets obtained from other livestock species. If a model is, however, trained with a larger dataset obtained from several livestock species, it would provide accurate predictions within the population represented by the dataset. It is also important to note that in this study, T_a and T_g were the only environmental predictor variables. Future studies may include other environmental predictor variables (e.g., relative humidity and heat stress indices) and spatio-temporal parameters (e.g., time of the year), which could improve model performance.

Training and validation of the four machine learning models considered in this study took ~ 9.5 h to complete. Most of this time was spent on training the models (~8 h in total; GLM = 50 min; RF = 35 min; GBM = 30 min; DNN = 6 h). The time to compute one prediction was ~0.3 ms, which is faster than the computing time required for analytical or numerical models (Milan and Gebremedhin, 2016b).

Our results suggest that machine learning algorithms, particularly RF, GBM, and DNN were found to be accurate in predicting rectal

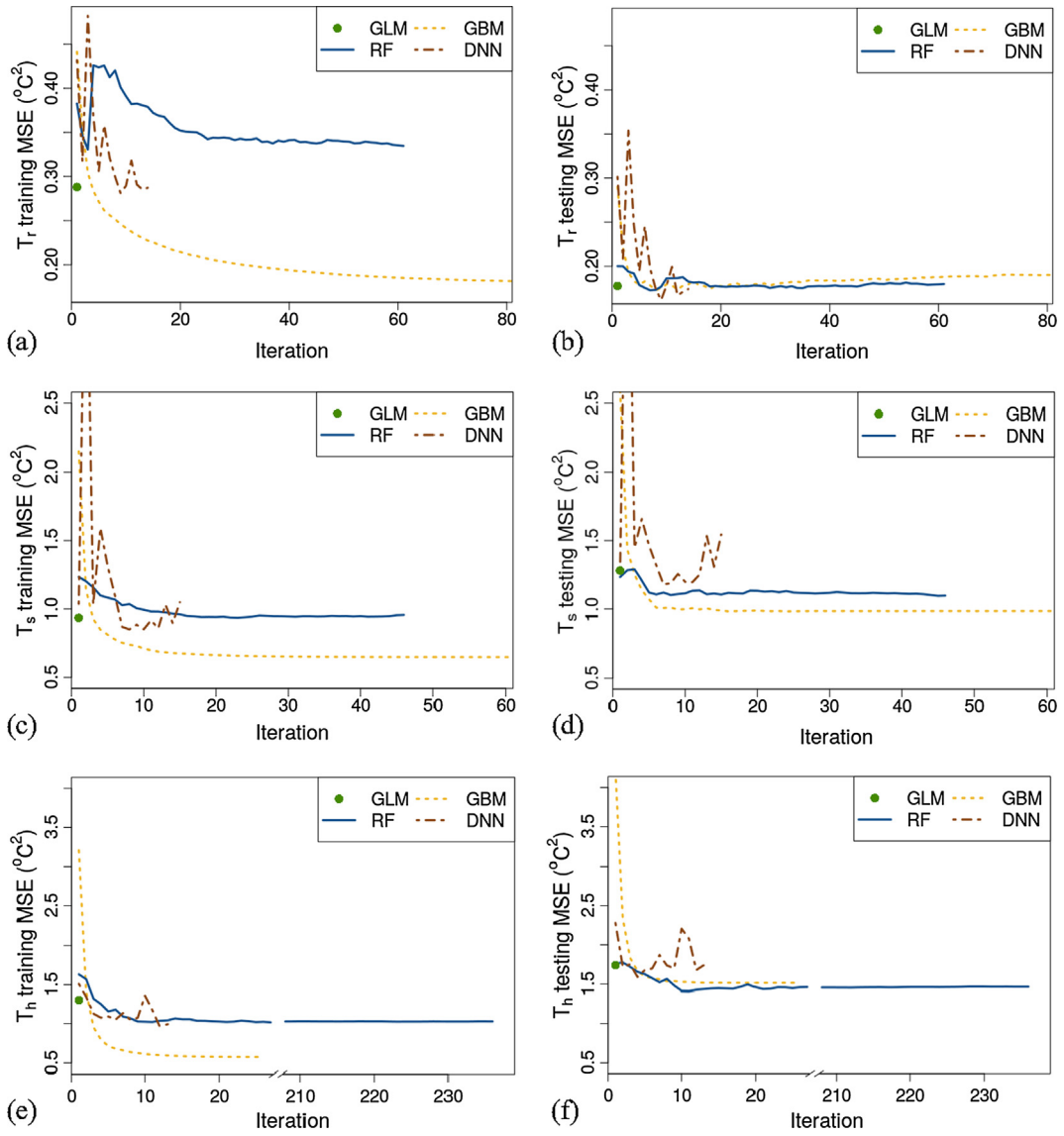

Fig. 5. Mean squared error (MSE) on the training (a, c, e) and testing (b, d, f) datasets for predicting rectal (T_r; a, b), skin-surface (T_s; c, d), and hair-coat surface (T_h; e, f) temperatures using the best performing machine learning models.

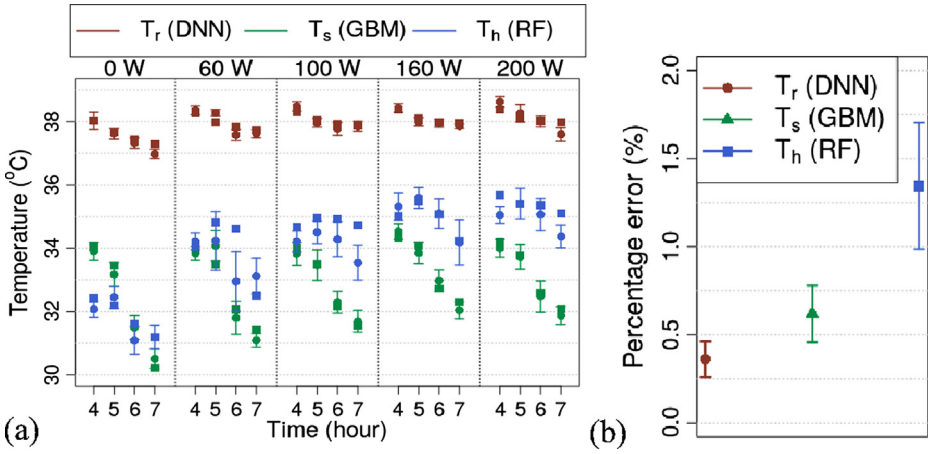

Fig. 6. Measured (●) and predicted (■) rectal (T_r), skin-surface (T_s), and hair-coat surface (T_h) temperatures for the mean dataset stratified by (a) time of measurement and intensity of supplemental heat, and (b) absolute percentage errors of the predicted temperatures. Measured values and absolute percentage errors are presented as mean ± standard error of the mean. Temperatures were predicted from the best performing machine learning models. RF: random forests; GBM: gradient boosted machines; DNN: deep neural network with ReLU activation function.

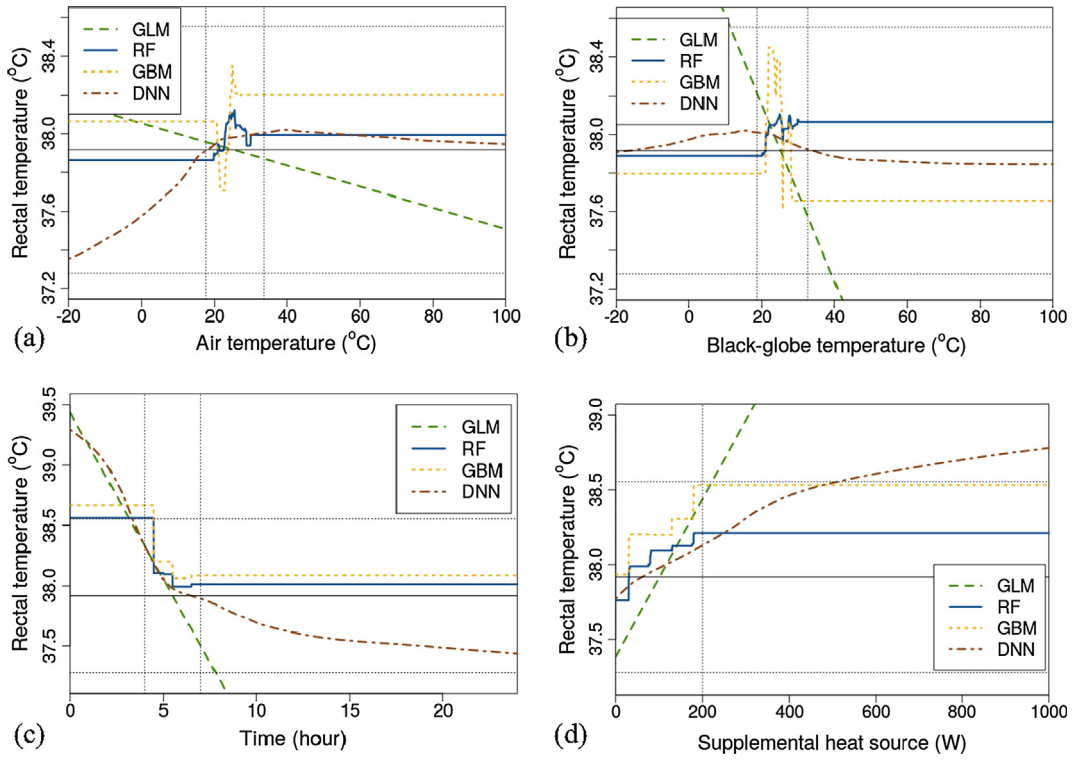

Fig. 7. Test of robustness and generalization of the best machine learning models in predicting rectal temperature when changing (a) air temperature, (b) black-globe temperature, (c) time of measurement, or (d) intensity of supplemental heat, while keeping the remaining predictor variables at their mean values. The vertical dashed lines represent the range of the measured predictor variable. The horizontal solid line represents the mean rectal temperature, and the horizontal dashed lines represent the mean rectal temperature ± one standard deviation from the mean.

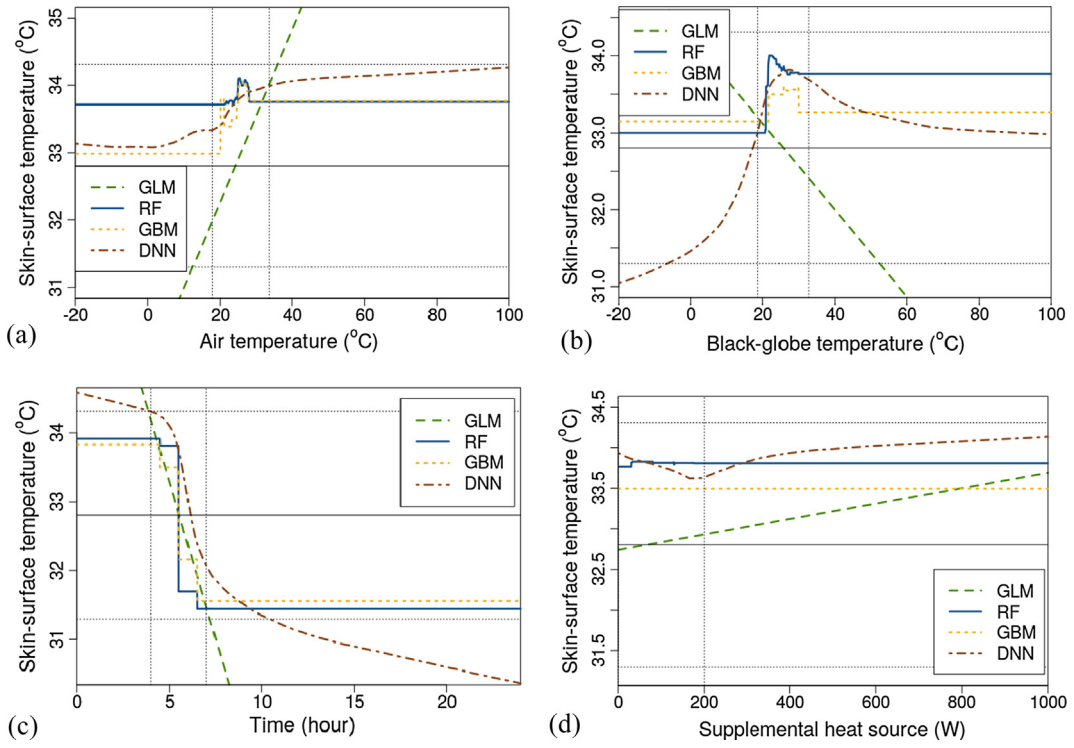

Fig. 8. Test of robustness and generalization of the best machine learning models in predicting skin-surface temperature when changing (a) air temperature, (b) black-globe temperature, (c) time of measurement, or (d) intensity of supplemental heat, while keeping the remaining predictor variables at their mean values. The vertical dashed lines represent the range of the measured predictor variable. The horizontal solid line represents the mean skin-surface temperature, and the horizontal dashed lines represent the mean skin-surface temperature ± one standard deviation from the mean.

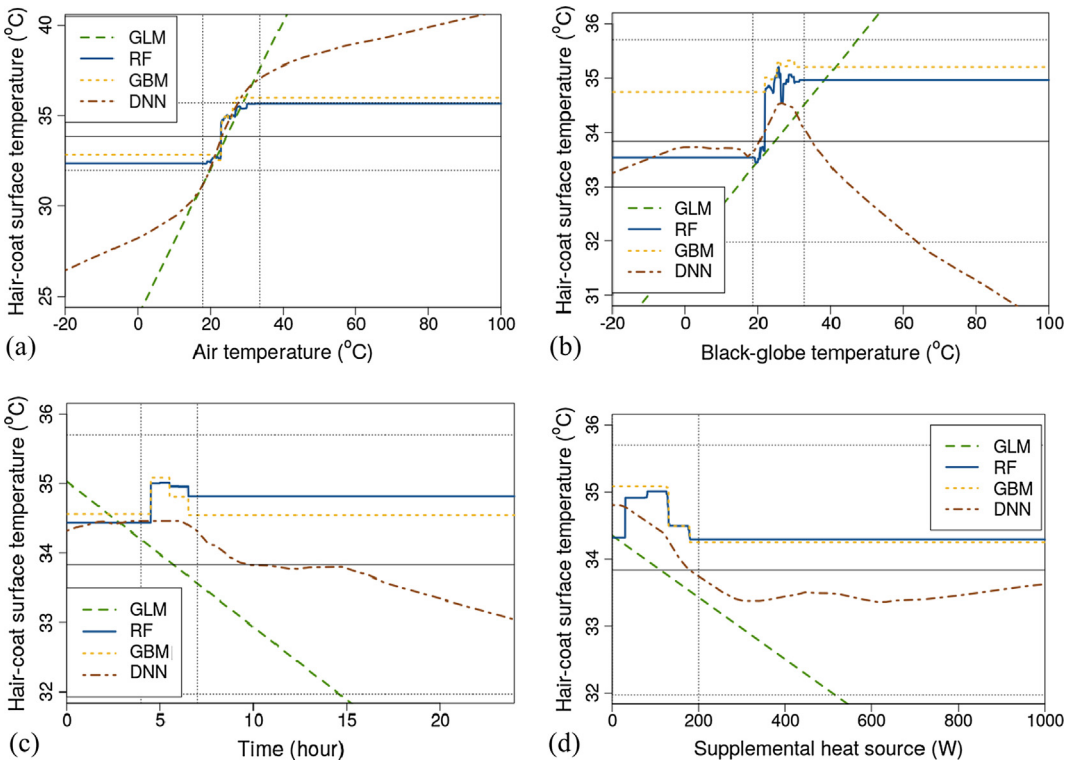

Fig. 9. Test of robustness and generalization of the best machine learning models in predicting hair-coat surface temperature when changing (a) air temperature, (b) black-globe temperature, (c) time of measurement, or (d) intensity of supplemental heat, while keeping the remaining predictor variables at their mean values. The vertical dashed lines represent the range of the measured predictor variable. The horizontal solid line represents the mean hair-coat surface temperature, and the horizontal dashed lines represent the mean hair-coat surface temperature ± one standard deviation from the mean.

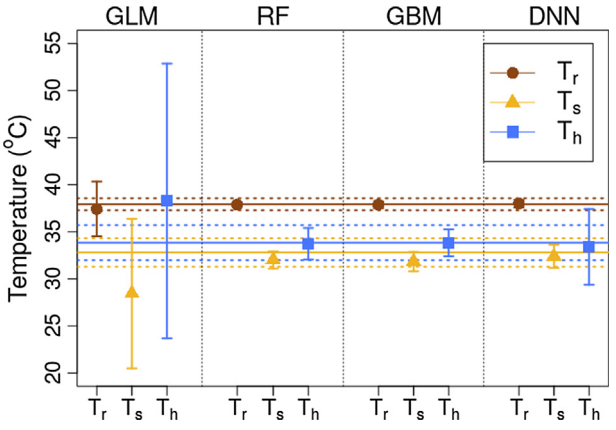

Fig. 10. Test of robustness and generalization of the best machine learning models in predicting rectal (T_r), skin-surface (T_s), and hair-coat surface (T_h) temperatures when randomly changing air temperature, black-globe temperature, time of measurement, and intensity of supplemental heat. Points represent mean ± one standard deviation of the mean (10,000 samples). Horizontal solid lines represent mean temperatures, and horizontal dashed lines represent mean ± one standard deviation of the mean.

temperature (T_r), skin-surface temperature (T_s), and hair-coat surface (T_h) temperature, but not GLM. The main advantage of machine learning models is that only data is needed to train the non-linearity of the data. For mechanistic models, the non-linearity comes from the assumptions made in solving the conservation equations. Since machine learning algorithms predict temperatures that are necessary to solve mechanistic models, one possible application of machine learning algorithms would be to provide inputs to mechanistic models.

4. Conclusions

Four machine learning algorithms were trained to predict rectal temperature, skin-surface temperature, and hair-coat surface temperature of piglets based on environmental data. Deep neural networks, gradient boosted machines, and random forests were the best algorithms, based on the lowest mean squared error on the testing dataset, to predict rectal, skin-surface, and hair coat-surface temperatures, respectively. The mean absolute percentage errors calculated using the mean dataset were 0.36% for rectal temperature, 0.62% for skin-surface temperature and 1.35% for hair coat-surface temperature. These three algorithms, different from generalized linear regression models, were robust to a wide range of inputs. The data supports the use of machine learning algorithms to predict physiological temperatures of livestock, and these temperature predictions can be used as inputs to mechanistic models. The combination of mechanistic and machine learning algorithms has the potential to provide more information about thermal-comfort status of livestock.

Acknowledgements

Funding: Brazilian National Council of Technological and Scientific Development (CNPq, Proc. 203312/2014-7), São Paulo Research Foundation (FAPESP, Proc. 17.519/14), and the USDA/Hatch (Washington, DC) funds as part of the W-3173 Regional Project through Cornell University.

References

Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, 281–305.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Brown-Brandl, T.M., Eigenberg, R.A., Nienaber, J.A., Kachman, S.D., 2001.

Thermoregulatory profile of a newer genetic line of pigs. Livest. Prod. Sci. 71, 253–260.
Brown-Brandl, T.M., Hayes, M.D., Xin, H., Nienaber, J.A., Li, H., Eigenberg, R.A., Stinn, J.P., Shepherd, T., 2014. Heat and moisture production of modern swine. ASHRAE Trans. 120 (1), 469–489.
Collier, R.J., Gebremedhin, K.G., 2015. Thermal biology of domestic animals. Ann. Rev. Anim. Biosci. 3, 513–532.
Collin, A., van Milgen, J., Dubois, S., Noblet, J., 2001. Effect of high temperature on feeding behaviour and heat production in group-housed young pigs. Br. J. Nutr. 86, 63–70.
Costa, L.N., Redaelli, V., Magnani, D., Cafazzo, S., Amadori, M., Razzauoli, E., Verga, M., Luzi, F., 2010. Preliminary study of the relationship between skin temperature of piglets measured by infrared thermography and environmental temperature in a vehicle in transit. In: LXIV Ann. Meet. Ital. Soc. Vet. Sci., pp. 193–197.
Cousineau, D., Chartier, S., 2010. Outliers detection and treatment: a review. Int. J. Psychol. Res. 3 (1), 58–67.
Cross, A.J., Rohrer, G.A., Brown-Brandl, T.M., Cassady, J.P., Keel, B.N., 2018;al., in press. Feed-forward and generalised regression neural networks in modelling feeding behaviour of pigs in the grow-finish phase. Biosyst. Eng in press.
Da Silva, R.G., Maia, A.S.C., 2013. Principles of Animal Biometeorology. Springer, New York.
DeShazer, J.A., 2009. Livestock Energetics and Thermal Environmental Management. ASABE.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, Massachusetts.
Guarino, M., Norton, T., Berckmans, D., Vranken, E., Berckmans, D., 2017. A blueprint for developing and applying precision livestock farming tools: a key output of the EU-PLF project. Anim. Front. 7 (1), 12–17.
Hastie, T., Tibshirani, R., Friedman, J., 2003. The Elements of Statistical Learning. Springer, New York.
Hensley, D.W., Mark, A.E., Abella, J.R., Netscher, G.M., Wissler, E.H., Diller, K.R., 2013. 50 years of computer simulation of the human thermoregulatory system. J. Biomech. Eng. 135, 021006.
Hunter, M.C., Smith, R.G., Schipanski, M.E., Atwood, L.W., Mortense, D.A., 2017. Agriculture in 2050: recalibrating targets for sustainable intensification. BioScience 67 (4), 386–391.
Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90.
Kashiha, M., Bahr, C., Ott, S., Moons, C.P.H., Niewold, T.A., Ödberg, F.O., Berckmans, D., 2014. Automatic weight estimation of individual pigs using image analysis. Comput. Electron. Agric. 107, 38–44.
Korthals, R.L., Eigenberg, R.A., Hahn, G.L., Nienaber, J.A., 1995. Measurements and spectral analysis of tympanic temperature regulation in swine. Trans. ASAE 33 (3), 905–909.
Lao, F., Brown-Brandl, T., Stinn, J.P., Liu, K., Teng, G., Xin, H., 2016. Automatic recognition of lactating sow behaviors through depth image processing. Comput. Electron. Agric. 125, 56–62.
Loughmiller, J.A., Spire, M.F., Dritz, S.S., Fenwick, B.W., Hosni, M.H., Hogge, S.B., 2001. Relationship between mean body surface temperature measured by use of infrared thermography and ambient temperature in clinically normal pigs and pigs inoculated with *Actinobacillus pleuropneumoniae*. Am. J. Vet. Res. 62 (5), 676–681.
McArthur, A.J., 1981. Thermal resistance and sensible heat loss from animals. J. Therm. Biol. 6, 43–47.
McCafferty, D.J., Gallon, S., Nord, A., 2015. Challenges of measuring body temperatures of free-ranging birds and mammals. Anim. Biotele. 3, 33.
Milan, H.F.M., Gebremedhin, K.G., 2016a. Triangular node for Transmission-Line Modeling (TLM) applied to bio-heat transfer. J. Therm. Biol. 62, 116–122.
Milan, H.F.M., Gebremedhin, K.G., 2016b. Tetrahedral node for Transmission-Line Modeling (TLM) applied to bio-heat transfer. Comp. Biol. Med. 79, 243–249.
Monteith, J., Unsworth, M., 2013. Principles of Environmental Physics: Plants, Animals, and the Atmosphere. Academic Press, Massachusetts.
Mostaço, G.M., Miranda, K.O.S., Condotta, I.C.F., Salgado, D.D.A., 2015. Determination of piglets' rectal temperature and respiratory rate through skin surface temperature under climatic chamber conditions. J. Braz. Assoc. Agric. Eng. 35 (6), 979–989.
Nasirahmadi, A., Edwards, S.A., Sturm, B., 2017. Implementation of machine vision for detecting behaviour of cattle and pigs. Livest. Sci. 202, 25–38.
Natekin, A., Knoll, A., 2013. Gradient boosting machines, a tutorial. Front. Neurorobot 7.
Nienaber, J.A., Hahn, G.L., Eigenberg, R.A., 1999. Quantifying livestock responses for heat stress management: a review. Int. J. Biometeorol. 42, 183–188.
Pathak, M., Parkhurst, A.M., Arias, R.A., Mader, T.L., 2009. Comparative study of time series and multiple regression for modeling dependence of cattle body temperature on environmental variables during heat stress. In: Ann. 21st Conf. Appl. Statist. Agric.
R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Ramirez, B.C., 2017. A novel approach to measure, understand, and assess the thermal environment in grow-finish swine facilities. Iowa State University. Graduate Theses and Dissertations, 16201.
Robertshaw, D., 2006. Mechanisms for the control of respiratory evaporative heat loss in panting animals. J. Appl. Physiol. 101, 664–668.
Shao, B., Xin, H., 2008. A real-time computer vision assessment and control of thermal comfort for group-housed pigs. Comput. Electron. Agric. 62, 15–21.
Shi, C., Teng, G., Li, Z., 2016. An approach of pig weight estimation using binocular stereo system based on LabVIEW. Comput. Electron. Agric. 129, 37–43.
Soerensen, D.D., Pedersen, J., 2015. Infrared skin temperature measurements for monitoring health in pigs: a review. Acta Vet. Scand. 57, 5.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.
St-Pierre, N.R., Cobanov, B., Schnitkey, G., 2003. Economic losses from heat stress by US livestock industries. J. Dairy Sci. 86, E52–E77.
Silanikove, N., 2000. Effects of heat stress on the welfare of extensively managed domestic ruminants. Livest. Prod. Sci. 67, 1–18.
The H2O.ai team, 2017. H2O: R Interface for H2O, version 3.16.0.2.
Turnpenny, J.R., McArthur, A.J., Clark, J.A., Wathes, C.M., 2000a. Thermal balance of livestock 1. A parsimonious model. Agr. Forest. Meteorol. 101, 15–27.
Turnpenny, J.R., Wathes, C.M., Clarck, J.A., McArthur, A.J., 2000b. Thermal balance of livestock 2. Application of a parsimonious model. Agr. Forest. Meteorol. 101, 29–52.
Van Hertem, T., Rooijakkers, L., Berckmans, D., Peña Fernández, A., Norton, T., Berckmans, D., Vranken, E., 2017. Appropriate data visualization is key to precision livestock farming acceptance. Comput. Electron. Agric. 138, 1–10.
Vasdal, G., Wheeler, E.F., Boe, K.E., 2009. Effect of infrared temperature on thermoregulatory behaviour in usckling piglets. Animal 3 (10), 1449–1454.
Wathes, C.M., Kristensen, H.H., Aerts, J.-M., Berckmans, D., 2008. Is precision livestock farming an engineer's daydream or nightmare, an animal's friend or foe, and a farmer's panacea or pitfall? Comput. Electron. Agric. 64, 2–10.
Wolfenson, D., Roth, Z., Meidan, R., 2000. Impaired reproduction in heat-stressed cattle: basic and applied aspects. Anim. Reprod. Sci. 60, 535–547.
Wongsriworaphon, A., Arnonkijpanich, B., Pathumnakul, S., 2015. An approach based on digital image analysis to estimate the live weight of pigs in farm environments. Comput. Electron. Agric. 115, 26–33.
Zeiler, M.D., 2012 Adadelta: An Adaptive Learning Rate Method. Also Available at: arXiv:1212.5701.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. Roy. Stat. Soc. Ser. B 67 (2), 301–320.