

Kansas State University Libraries

New Prairie Press

---

Conference on Applied Statistics in Agriculture

2011 - 23rd Annual Conference Proceedings

---

## ESTIMATING AREA AND LAG ASSOCIATED WITH THERMAL HYSTERESIS IN CATTLE

F. Yang

A. M. Parkhurst

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

---

### Recommended Citation

Yang, F. and Parkhurst, A. M. (2011). "ESTIMATING AREA AND LAG ASSOCIATED WITH THERMAL HYSTERESIS IN CATTLE," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1055>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact [cads@k-state.edu](mailto:cads@k-state.edu).

# Estimating Area and Lag Associated with Thermal Hysteresis in Cattle

F. Yang and A. M. Parkhurst

University of Nebraska – Lincoln

## Abstract

Thermal hysteresis in cattle becomes visible when the phase diagram of body temperature ( $T_b$ ) vs ambient temperature ( $T_a$ ) exhibits a loop. The hysteresis loop shows a rotated elliptical pattern which depends on the lag between  $T_b$  and  $T_a$ . The area of the loop can be used to quantify the amount of heat stress during thermal challenge. Three methods to estimate the area and lag of the elliptical hysteresis loop are: linear least squares method, ellipse-specific nonlinear least squares method, and Lapshin's analytical method. Linear least squares method uses residual least squares to estimate the coefficients of the ellipse for which the sum of the squares of the distances to the observations is minimal. The estimated coefficients can be used to calculate both the rotated angle and area of the ellipse. The ellipse-specific method is based on quadratic constrained least mean squares fitting to simultaneously determine the best elliptical fit for a set of scattered data. It provides estimates of the rotated angle and semi-major and semi-minor axes to calculate the area of the ellipse. Lapshin's analytical method is a two-stage procedure that fits a sinusoidal function to the input and then the output. It provides parameters in addition to lag and area which further characterize the hysteresis loop. The area and lag along with their standard errors are compared for the three methods using the delta method and bootstrapping. The delta method is used to calculate the standard errors of the derived parameter estimates and bootstrapping is used to assess the appropriateness of the delta method.

**Key words:** Delay-Relay Model,  $T_a$ - $T_b$  phase diagram, Thermo-regulatory response, Eigenvalue-eigenvector, Heat load,  $T_b$ - $T_a$  correlation, Energy dissipation, Farm animals.

## 1. Introduction

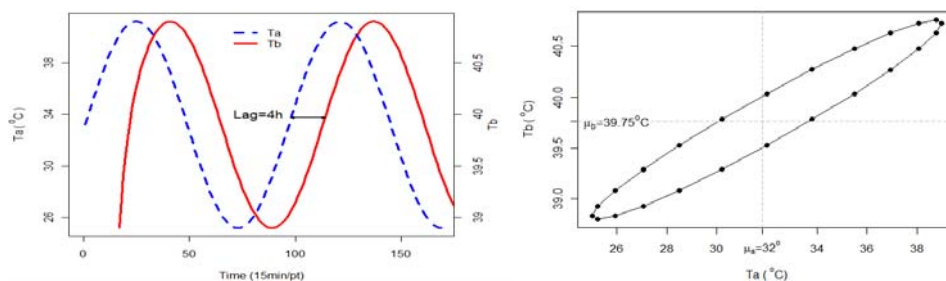
The phenomenon of hysteresis can be defined in terms of the relationship between the output of a system and the inputs. The three fundamental properties of the output from a hysteretic process are: memory, rate independence, and initial state of the process. Memory is the most distinguishing feature of hysteresis. It is exemplified by the existence of a trajectory (or path profile) that displays the dependence of the current value on past history. The behavior of a hysteretic process is irreversible. Rate independence indicates the output value does not depend on the time scale used to

measure the input. In many situations, the initial state of the process is needed to describe the start of the hysteretic process.

Thermal hysteresis is one of cyclic hysteresis. In animal science, when an animal is thermally challenged, physiological processes in the body become activated in order to dissipate the increasing heat load. Body temperature ( $T_b$ ) is one process that shows a dramatic response to heat stress, (HS). Studies show  $T_b$  is essentially unresponsive to air temperature ( $T_a$ ) until a threshold is reached and then there is a dramatic response to increasing air temperature (Hahn, Parkhurst, and Gaughan 1997; Mader et al. 2001; Parkhurst 2010).

During heat stress, the dynamics can be characterized by thermal hysteresis which follows an elliptical loop. It measures and depicts the amount of heat stress an animal experiences in a hot environment. A phase diagram for  $T_b$  versus  $T_a$  indicates the time-dependence and memory associated with the hysteretic behavior; thus, plotting a phase diagram between  $T_b$  and  $T_a$  provides a way to show the existence of thermal hysteresis. Thermal hysteresis becomes prominent once a delay (or lag) between  $T_b$  and  $T_a$  is noticeable in the  $T_b$ - $T_a$  diagram and the supposition that  $T_a$  is the driving force when an animal is thermally challenged appears plausible. (Fig 1)

**Fig 1.** Left: A time series plot shows a 4 hour lag between  $T_b$  and  $T_a$  for an animal in the temperature controlled chamber. Right: The diagram of  $T_b \sim T_a$  for this animal shows a closed loop of the hysteresis.



Two features of the hysteretic ellipse, angle of rotation and area, help characterize the dynamics of heat stress. In this study, the angle of rotation of an ellipse is the angle between ellipse's major axis and the  $T_a$  (horizontal)-axis. It indicates the length of the time delay; i.e. how long  $T_b$  lags  $T_a$ . The lag can be estimated from the angle of rotation of an elliptical loop. In thermodynamics, the area inside the hysteresis loop equals the work done in one period or cycle (Brokate and Sprekels 1996). Thus, the area of the ellipse formed during thermal hysteresis is an indication of the animal's heat load. In other words, heat load can be assessed by estimating the area within the hysteresis loop. As a result, it is possible to fit an ellipse for not only predicting  $T_b$  by identifying the lag of  $T_b$  on the current  $T_a$ , but also for estimating the amount of heat load.

The objectives of this project are 1) to identify three methods to characterize thermal hysteresis, 2) to fit parameters of thermal hysteresis ellipse and estimate lag and area, 3) to compare parameters from three methods in simulated data, and 4) to compare estimated area and lag along with their standard errors.

## 2. Methods and Models

An ellipse can be represented using a pair of first-order differential equations:

$$\frac{dy}{dt} = rx \quad \frac{dx}{dt} = -\frac{1}{r}y$$

By solving the differential equations and reparameterizing the results, an ellipse centered at the origin can be obtained parametrically as

$$\alpha(t) = (r_1 \cos(t) \quad r_2 \sin(t)) \quad (2.1)$$

where  $r_1$  is the length of axis along the x-coordinate (major axis),  $r_2$  is the length of axis along the y-coordinate (minor axis),  $t$  is the time. (Horn and Weldon 1986)

After rotating the ellipse count-clockwise over an angle  $\theta$  and translating to the centroid  $(c_x, c_y)$ , the ellipse can be written parametrically as

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} r_1 \cos(\theta) \cos(t) - r_2 \sin(\theta) \sin(t) + c_x \\ r_1 \sin(\theta) \cos(t) + r_2 \cos(\theta) \sin(t) + c_y \end{pmatrix} \quad (2.2)$$

### 2.1 Linear Least Squares Method

Any simple closed curve  $(x, y)$  can be written algebraically as

$$f(x, y) = 0$$

For a rotated ellipse,

$$f(x, y) = a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 = 0 \quad (2.3)$$

where  $a_1 \neq 0, a_3 \neq 0$ .

Hence, a linear statistical model based on (2.3) can be constructed to achieve  $a$  coefficients. For  $i=1, \dots, n$ ,

$$\xi_i = a_1x_i^2 + a_2x_iy_i + a_3y_i^2 + a_4x_i + a_5y_i + a_6 + \varepsilon_i \quad (2.4)$$

The expected surface of the model is  $E(\xi) = a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 = 0$ .

Then, the 5 parameters of a rotated ellipse can be obtained using the coefficients from (2.4) as

$$\begin{aligned} c_x &= \frac{2a_3a_4 - a_2a_5}{a_2^2 - 4a_1a_3} & r_1 &= \sqrt{\frac{a_1c_x^2 + a_2c_xc_y + a_3c_y^2 - a_6}{a_1 \cos^2 \theta + a_2 \sin \theta \cos \theta + a_3 \sin^2 \theta}} \\ c_y &= \frac{2a_1a_5 - a_2a_4}{a_2^2 - 4a_1a_3} & r_2 &= \sqrt{\frac{a_1c_x^2 + a_2c_xc_y + a_3c_y^2 - a_6}{a_1 \sin^2 \theta - a_2 \sin \theta \cos \theta + a_3 \cos^2 \theta}} & \theta &= \frac{1}{2} \left( \arctan \frac{a_2}{a_1 - a_3} \right) \end{aligned} \quad (2.5)$$

All the  $a$ 's coefficients in (2.5) can be found using least squares method. That is,

$$\frac{\partial \xi}{\partial a_i} = 0, \quad i = 1, \dots, 6 \text{ with constraints } a_1 \neq 0, a_3 \neq 0.$$

After the 5 parameters were obtained from the linear least squares method, the other 4 parameters: area, lag, retention, and coercion can be derived as

$$\begin{aligned} A_{HL} &= \text{Area of ellipse} = \pi * r_1 * r_2 \\ \text{lag} &= \arctan\left(\frac{r_1 \sin(\theta)}{r_2 \cos(\theta)}\right) - \arctan\left(\frac{-r_1 \cos(\theta)}{r_2 \sin(\theta)}\right) \\ R &= \sqrt{r_1^2 \cos^2(\theta) + r_2^2 \sin^2(\theta)} \\ C &= \sqrt{r_1^2 \sin^2(\theta) + r_2^2 \cos^2(\theta)} \end{aligned} \quad (2.6)$$

## 2.2 Ellipse-Specific Nonlinear Least Squares Method

In (2.2), a rotated ellipse with rotation angle  $\theta$  and centered at  $(c_x, c_y)$  was written parametrically by both input (x) and output (y) functions. Therefore, a nonlinear statistical model of 5 parameters can be used to describe the relationship between the input and output series.

$$\xi = \left(\frac{\cos(\theta)(x - c_x) + \sin(\theta)(y - c_y)}{r_1}\right)^2 + \left(\frac{-\sin(\theta)(x - c_x) + \cos(\theta)(y - c_y)}{r_2}\right)^2 - 1 + \varepsilon \quad (2.7)$$

Equation 2.7 is called the parameterized ellipse model. The expected surface of the model is

$$E(\xi) = \left(\frac{\cos(\theta)(x - c_x) + \sin(\theta)(y - c_y)}{r_1}\right)^2 + \left(\frac{-\sin(\theta)(x - c_x) + \cos(\theta)(y - c_y)}{r_2}\right)^2 - 1 = 0$$

Therefore, the 5 parameters can be found directly from the nonlinear model. The other parameters we are interested in can be derived as in (2.6).

### 2.2.1 Starting Values and Self-starter function

When fitting a nonlinear regression model, one of the most important things is to find appropriate starting values for the model to converge. Finding starting values for the parameters by guessing requires some skill, or experience from previous analyses, to come up with a successful guess. A grid search provides a more systematic approach. However, both methods are time-consuming. The use of self-starter functions provide an automated way of obtaining starting values and is indispensable, for repeated use of the same nonlinear regression model. (Ritz and Streibig 2008)

In this study, a self-starter function was built based for the ellipse-specific direct least squares method. The method was introduced by Pilu, Fitzgibbon and Fisher (1996); and it was developed by Fitzgibbon, Pilu and Fisher (1999), Halir and Flusser (1998)

and O’Leary and Zsombor-Murray (2004). In 2010, Yang and Parkhurst applied this method to describe the trajectories of an elliptical hysteresis loop of heat stressed animals.

In the ellipse-specific method, the algebraic distance is minimized over the set of  $n$  points in the least squares sense, that is

$$\hat{d}^2 = \min_a \left\{ \sum_{i=1}^N \left( x_i \ y_i \ 1 \right) \begin{pmatrix} a_1 & a_2/2 & a_4/2 \\ & a_3 & a_5/2 \\ \text{Symmetric} & & a_6 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \right\} \quad (2.8)$$

With the constraint  $a_2^2 - 4a_1a_3 < 0$ .

Hence, it can be represented as a generalized eigen-system:

$$a^t D^t D a - \lambda a^t C a = 0 \quad (2.9)$$

where  $D = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n & 1 \end{pmatrix}$ ,  $C = \begin{pmatrix} 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ ,  $a = \begin{pmatrix} a_1 \\ \vdots \\ a_6 \end{pmatrix}$ .

Then,  $\hat{a}$  is the eigenvector corresponding to the negative eigenvalue in the eigen-system  $D^t D a = \lambda C a$ . That is,

$$\hat{a} = \text{Kernel} \{ \lambda < 0 : \det | D^t D - \lambda C | \} \quad (2.10)$$

Therefore, the starting values for 5 parameters can be obtained from  $\hat{a}$ .

$$\begin{aligned} c_{x\_start} &= \frac{2a_3a_4 - a_2a_5}{a_2^2 - 4a_1a_3} & r_{1\_start} &= \sqrt{\frac{a_1c_x^2 + a_2c_xc_y + a_3c_y^2 - a_6}{a_1 \cos^2 \theta + a_2 \sin \theta \cos \theta + a_3 \sin^2 \theta}} \\ c_{y\_start} &= \frac{2a_1a_5 - a_2a_4}{a_2^2 - 4a_1a_3} & r_{2\_start} &= \sqrt{\frac{a_1c_x^2 + a_2c_xc_y + a_3c_y^2 - a_6}{a_1 \sin^2 \theta - a_2 \sin \theta \cos \theta + a_3 \cos^2 \theta}} \end{aligned} \quad \theta_{start} = \frac{1}{2} \left( \arctan \frac{a_2}{a_1 - a_3} \right) \quad (2.11)$$

### 2.3 Lapshin Method

The third method used to describe the elliptical hysteresis loops is due to Lapshin (1995). It is a two-stage procedure which describes input and output separately.

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} b_a \cos(2\pi\omega t + \phi_a) + c_x + \varepsilon_a \\ b_b \cos(2\pi\omega t + \phi_a) - a \sin(2\pi\omega t + \phi_a) + c_y + \varepsilon_b \end{pmatrix} \quad (2.12)$$

where  $\phi_a$  is the phase angle of input ( $x$ );  $a$  is the split point coordinate;  $b_a, b_b$  are the

saturation point coordinates;  $c_x$  and  $c_y$  are center coordinates;  $\varepsilon_a$  and  $\varepsilon_b$  are errors.

After sequentially fitting two linear regression models for both input (x) and output (y), Lapshin parameters can be used to describe elliptical parameters. Represented the intercepts of two linear regression models as  $c_x$  and  $c_y$ , the lengths of two axes and angle of rotation are

$$\begin{aligned} r_1 &= \sqrt{(b_a \cos(\phi_a))^2 + (b_b \cos(\phi_a) - a \sin(\phi_a))^2} \\ r_2 &= \sqrt{(b_a \sin(\phi_a))^2 + (b_b \sin(\phi_a) + a \cos(\phi_a))^2} \end{aligned}, \quad \theta = \arctan\left(\frac{b_b \cos(\phi_a) - a \sin(\phi_a)}{b_a \cos(\phi_a)}\right)$$

(2.13)

Then, the lag which is the difference between the phase angles of  $T_a$  and  $T_b$  is:

$$lag = \arctan(a / b_b) \tag{2.14}$$

The area of the hysteresis loop in Lapshin model can be derived using Green's

formula:  $Area = \frac{1}{2} \int_0^{period} \left( y \frac{dx}{dt} - x \frac{dy}{dt} \right) dt.$

Hence,

$$area = \pi b_y \sqrt{b_x^2 + a^2} \sin(lag) \tag{2.15}$$

Retention, R, and coercion, C, are

$$\begin{aligned} R &= a \\ C &= b_a \sin(lag) \end{aligned} \tag{2.16}$$

## 2.4 Model Comparisons

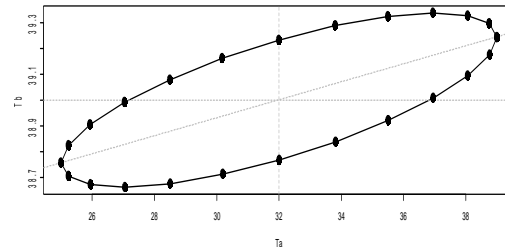
### 2.4.1. Simulations

To find the best procedure estimates, a simulation study was constructed. Realistic numbers for the model parameters based on values reported in previous literature were used, Table 1. An input heat challenge of  $32 \pm 7$  °C is applied to an output signal of  $39 \pm 0.233$  °C with a 2.9 hr lag. The phase diagram for these values is presented in Figure 2. The 1000 datasets were simulated by adding errors to both  $T_a$  ( $\varepsilon_a$ ) and  $T_b$  ( $\varepsilon_b$ ) such that  $\varepsilon_a \sim N(0, 0.2)$  and  $\varepsilon_b \sim N(0, 0.04)$ . Several statistics derived from the simulated datasets, standard errors, MSE, percentage of bias, and 95% coverage levels were compared, and Shapiro-Wilk tests were applied to check the normality for all parameters in a simulation study. The method that produces the smallest standard errors, MSE, percentage bias, and close to 95% coverage, with normally distributed estimates is the preferred procedure.

**Table 1.** Parameter values for simulated dataset for a steer in a sinusoidal Ta environment ( $32 \pm 7$  °C) with Tb ( $39 \pm 0.233$  °C) lagging Ta by 2.9 h.

<i>Parameter</i>	<i>Simulated</i>
$r_1$	7
$r_2$	0.233
$c_x$	32
$c_y$	39
$\theta$	2
Lag	2.904
Area	5.124
R	6.996
C	0.337

**Fig 2.** TaTb phase diagram for a steer in a sinusoidal Ta environment ( $32 \pm 7$  °C) with Tb ( $39 \pm 0.233$  °C) lagging Ta by 2.9 h.



#### 2.4.2. Bias

To compare the difference between the expected value and the true value of each parameter, percentage of bias (%Bias) is calculated.

$$\%Bias = (E(\mathbf{p}) - \mathbf{p}) / \mathbf{p}100\% \quad (2.17)$$

where  $\mathbf{p}$  stands for the true parameter value. The lower the %Bias, the closer the expected value is to the truth. A bias higher than 1% is considered to be significant.

#### 2.4.3. Standard Errors and Coverage

When fitting the linear least square model, the standard errors (SEs) for parameters ( $r_1$ ,  $r_2$ ,  $c_x$ ,  $c_y$ ,  $\theta$ ) were obtained. For the derived parameters, lag, area, coercion, and retention, the delta method was used to calculate the SEs. Similarly when fitting the ellipse-specific nonlinear least square model, the SEs for the parameters ( $r_1$ ,  $r_2$ ,  $c_x$ ,  $c_y$ ,  $\theta$ ) were obtained directly from the nonlinear model, but the other standard errors (lag, area, coercion, and retention) were obtained by the delta method. In Lapshin model, SEs of centroids, retention and lag can be obtained from the model directly; SEs of lag can be obtained from delta method; but since two stages are used to fit the Lapshin model, SEs for other parameters cannot be obtained from the model or delta method.

To compare standard errors of parameters with different units, the coefficient of variation (CV) was used as a measure of dispersion.

$$CV = |Standard\ error / Estimate| \quad (2.18)$$

The Wald confidence interval (CI) can be obtained using the expected value and the SE for each parameter. Then for each parameter, the Wald coverage indicates the proportion of simulations that the true value falls in the Wald CI, based on the 1000 simulations.



For the ellipse-specific nonlinear model, two other algorithms: Sandwich method and T-profile method can be used to obtain SEs. The sandwich estimation procedure is a general method for estimating the covariance matrix of parameter estimates. The method asymptotically adjusts covariance matrix in order to obtain consistent estimates of the standard error even if some of the distributional assumptions are misspecified (Carroll and Ruppert 1988). Suitable implementations (*R package: Sandwich*) are available in the R system for statistical computing for Sandwich estimation (Zeileis 2006). T-profile is a method based on the idea of marginal likelihood intervals for nonlinear model parameters (Bates and Watts 1988). It begins by relating a linear model interval to the sum of squares functions and a profile t function is then developed. (*R function: confint.default*) With these algorithms, two new sets of SEs and the corresponding CIs and 95% coverages for parameters ( $r_1$ ,  $r_2$ ,  $c_x$ ,  $c_y$ ,  $\theta$ ) can be obtained.

#### 2.4.4. Bootstrapping

Neither Wald nor Sandwich nor T-profile can give SEs for all parameters for all three models. Bootstrapping gives SEs for all parameters in all three models. Hence, to make comparisons the parameter estimates from each model, the bootstrapping method was applied. Bootstrapping developed by Efron (1979; 1993) provides an alternative way of calculating SEs. It helps to learn about the sample characteristics by resampling and using the information to infer the population distribution. (*R package: Boot*) By resampling, subset of each simulated realization was drawn randomly with replacement. For each of the 1000 simulated realizations, 1000 bootstrapping iterations were performed on linear least squares method, ellipse-specific nonlinear least squares method, and Lapshin method. Bootstrapping estimates, SEs, and coverages, for each parameter were obtained by averaging over the total number of iterations. At the end of bootstrapping, models were compared using the bootstrapping estimates, bootstrapping SEs, and bootstrapping coverages.

### 3. Results

#### 3.1 Linear least squares method

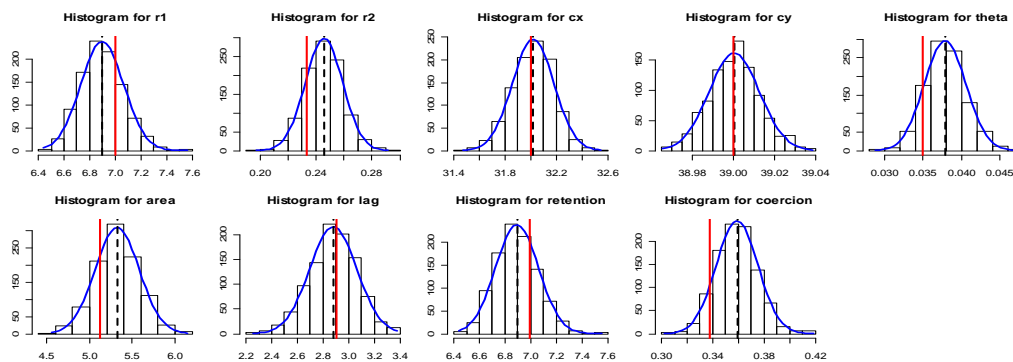
All the estimated parameters have small percentages of bias (<.09%). The standard errors of the length of major axis ( $r_1$ ), centroid on the input axis ( $c_x$ ), area, and coercion (C) are larger compared with other estimates, but the SE's of these parameters, with the exception of coercion are still comparable, because the coefficients of variation (CV's) are <7%. The standard error of coercion is serious because of the magnitude of the CV (75%). The Wald coverages for the estimates are close to 95% for only 3 of the 9 parameters. Coverage is worrisome for the length of minor axis ( $r_2$ ), angle of rotation (theta), area, lag and retention (R). The coverage for coercion is extremely high, which may be caused by the large standard error. (Table 2)

**Table 2.** Linear Least Squares Method: True Values, Estimates, %Bias, Standard Errors, and 95% Wald Coverages

Parameter	TRUE	Estimate	%Bias	SE	Coverage
$r_1$	7	6.898	-1.45E-02	.270	.994
$r_2$	.233	.246	5.48E-02	.010	.738
$c_x$	32	32.019	5.96E-04	.225	.991
$c_y$	39	39.000	1.22E-05	.013	.948
$\theta$	.035	.038	8.50E-02	.003	.829
area	5.124	5.323	3.88E-02	.256	.891
lag	2.904	2.880	-8.14E-03	.047	.426
R	6.996	6.893	-1.47E-02	.016	.126
C	.337	.359	6.36E-02	.270	1

The histograms are given in Fig 3. The expected values for centroids ( $c_x$ ,  $c_y$ ) and lag overlap their true values in the histograms. The expected values for other parameters are either lower or higher than their true values. Hence, in linear least square method, centroids and lag are symmetric about the true value; length of minor axis ( $r_2$ ), angle of rotation (theta), area, and coercion tend to under-estimate the parameter; length of major axis ( $r_1$ ), and retention tend to over-estimated the parameter.

**Fig 3.** Linear least squares method: Distribution plots of estimates for all parameters based on 1000 simulations, solid line-true values; dashed line-expected values.



### 3.2 Ellipse-specific nonlinear least squares method

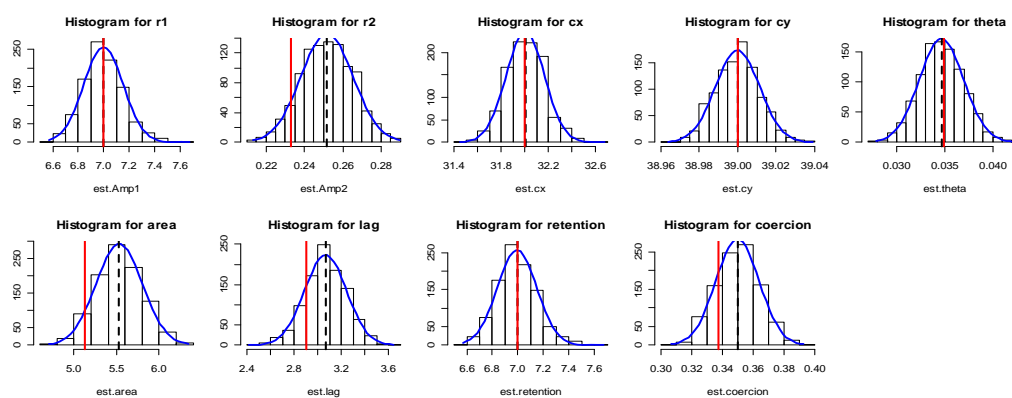
For the ellipse-specific nonlinear method, all the estimates appear unbiased. (<.09%) (Table 3) The standard errors of the length of major axis ( $r_1$ ), centroid on the input axis ( $c_x$ ), area, and retention (R) are large compared with other estimates, but the CV's are <7%. The Wald coverages for all estimates are close to 95% for 6 of the 9 parameters, but length of minor axis ( $r_2$ ), area, lag, and coercion (R) are worrisome. The coverage for the sandwich method for only 5 of the 9 parameters shows no improvement. The length of minor axis ( $r_2$ ) is very low compared with others. Nor does the T-profile method improve the results. The coverage of length of minor axis ( $r_2$ ) is still worrisome.

**Table 3.** Ellipse-specific nonlinear least squares method: true values, estimates, %bias, standard errors, 95% Wald coverages, 95% Sandwich coverages, and 95% T-profile coverages

Parameter	True	Estimate	%Bias	SE	Coverage		
					Wald	Sandwich	T-profile
$r_1$	7	6.997	-4E-4	.263	.999	.952	.998
$r_2$	.233	.252	.081	.011	.587	.595	.549
$c_x$	32	32.004	1E-4	.218	.991	.941	.989
$c_y$	39	39.000	1 E-6	.011	.960	.916	.949
$\theta$	.035	.035	-.007	.002	.946	.902	.934
area	5.124	5.532	.080	.254	.692		
lag	2.904	3.067	.056	.047	.881		
R	6.996	6.993	-4E-4	.263	.999		
C	.337	.350	.037	.014	.908		

The histograms for ellipse-specific nonlinear least squares method are given in Fig 4. The expected values for length of major axis ( $r_1$ ), centroids ( $c_x$ ,  $c_y$ ), angle of rotation (theta), and retention overlap their true values in the histograms. Other expected values are either lower or higher than their true values. Hence, in ellipse-specific nonlinear method, length of major axis, centroids, angle of rotation, and retention are symmetric about the true values; length of minor axis, area, and coercion greatly under-estimate the true value; the lag tends to be highly over-estimated.

**Fig 4.** Ellipse-specific nonlinear least squares method: Distribution plots of estimates for all parameters based on 1000 simulations- solid line-true values; dashed line-expected values.



### 3.3 Lapshin model:

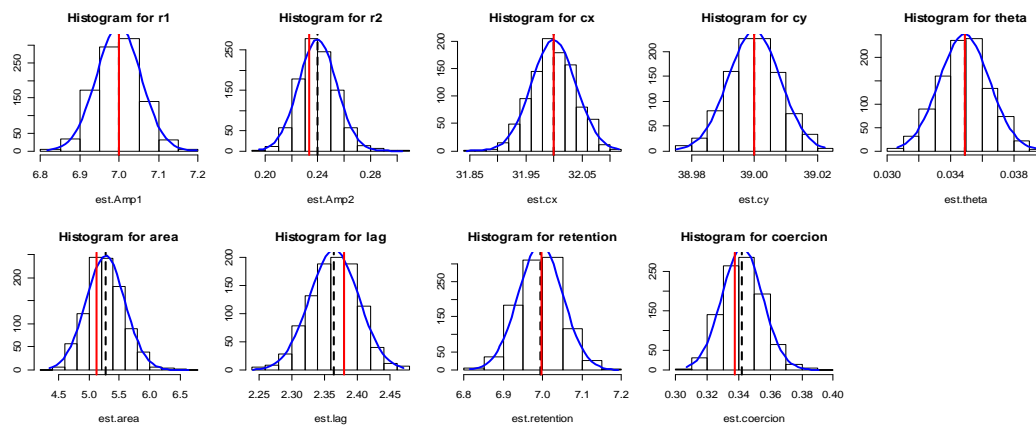
For the Lapshin method, all the estimated parameters have small percentages of bias(<.09%). (Table 4) Only standard errors of centroids ( $c_x$ ,  $c_y$ ) can be obtained directly from the two linear models, where  $SE(c_x)=0.039$  and  $SE(c_y)=0.008$ . Since Lapshin method is a two-stage estimation procedure, delta method, sandwich and T-profile are not applicable for obtaining standard errors.

The histograms for Lapshin method are given in Fig 5. The expected values for length of major axis ( $r_1$ ), centroids ( $c_x$ ,  $c_y$ ), angle of rotation (theta), and retention overlap their true values in the histograms. Other expected values are either lower or higher than their true values. Hence, in Lapshin method, length of major axis, centroids, angle of rotation, and retention are symmetric about the true values; length of minor axis ( $r_2$ ), area, and coercion under-estimate the parameter; lag over-estimates the parameter.

**Table 4.** Lapshin method: true values, estimates, and %bias

Parameter	TRUE	Estimate	%Bias
$r_1$	7	6.999	-2E-04
$r_2$	.233	.240	.029
$c_x$	32	31.999	-2E-05
$c_y$	39	39.000	-2E-06
$\theta$	.035	.035	-6E-04
area	5.124	5.271	.029
lag	2.904	2.959	.019
R	6.996	6.994	-2E-04
C	.337	.342	.014

**Fig 5.** Lapshin method: Distribution plots of estimates for all parameters based on 1000 simulated datasets . Solid-true values; dashed line-expected values.



### 3.4 Normality

Linear least squares method produced only three normally distributed parameters based on 1000 simulations: length of minor axis ( $r_2$ ), centroid ( $c_y$ ), and area. On the other hand, the ellipse-specific nonlinear method produced only two non-normal distributed parameters: length of major axis ( $r_1$ ) and retention (R). The Lapshin method produced three non-normal distributed parameters: length of minor axis ( $r_2$ ), area and coercion (C). (Table 5)

**Table 5.** Shapiro-Wilk test for normality of each parameter in three methods

	Shapiro p-value		
	LS	NLS	Lap
$r_1$	< .0001	< .0001	.9017
$r_2$	.6047	.4069	< .0001
$c_x$	.0500	.1510	.4169
$c_y$	.1448	.7414	.5425
$\theta$	.0367	.9524	.3130
area	.3369	.6606	< .0001
lag	.0313	.2022	.1574
$R$	< .0001	< .0001	.8901
$C$	< .0001	.3612	< .0001

### 3.5 Comparison of Bootstrapping Results

Since the assumption of normality for the area and other estimates was questionable, bootstrapping was performed. All the estimated bootstrapping parameters have small percentages of bias (<.09%) in all three methods. Although standard errors of some parameters are large compared with others in three method, CV's are all reasonable (<7%). Comparisons of the 95% bootstrapping coverages show that linear least squares method produces poor coverages for two parameters, coercion and angle of rotation; the ellipse-specific nonlinear least squares method produced poor coverages for length of minor axis and lag; coverages of all parameters by Lapshin method were acceptable, close to 95%. (Table 6)

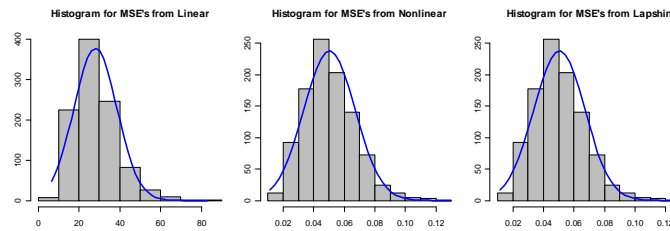
**Table 6.** Comparisons of Bootstrapping estimates, standard errors, coverages

Param eter	TRUE	Boot Estimate			Boot %Bias			Boot SE			Boot 95% Coverage		
		LS	NLS	Lap	LS	NLS	Lap	LS	NLS	Lap	LS	NLS	Lap
$r_1$	7	6.898	6.872	6.999	-.015	-.018	-2E-4	.161	.147	.054	.928	.892	.961
$r_2$	.233	.246	.246	.240	.055	.055	.029	.013	.012	.014	.886	.851	.941
$c_x$	32	32.019	32.014	31.999	6E-4	4E-4	-2E-5	.155	.152	.039	.954	.956	.956
$c_y$	39	39.000	39.000	39.000	1E-5	8E-6	-2E-6	.012	.011	.008	.945	.960	.944
$\theta$	.035	.038	.035	.035	.082	-.009	-.003	.002	.002	.002	.806	.956	.968
area	5.124	5.323	5.303	5.271	.039	.035	.029	.252	.250	.317	.906	.909	.943
lag	2.904	2.880	3.055	2.365	-.008	.052	-.186	.175	.170	.037	.948	.883	.935
$R$	6.996	6.893	6.868	6.994	-.015	-.018	3E-4	.161	.147	.054	.929	.892	.961
$C$	.337	.359	.342	.342	.065	.016	.016	.015	.013	.013	.724	.955	.942

### 3.6. Comparisons of MSEs

The comparisons of the MSEs among three methods show that ellipse-specific nonlinear method has the smallest averaged MSE (0.0503), but Linear least squares method has the largest (28.1037); MSEs from 1000 simulations are normally distributed only in Lapshin method (p-value =0.1396). (Fig 6, Table 7)

**Fig 6.** Distribution plots of MSEs for three methods based on 1000 simulations



**Table 7.** Comparisons of the averaged and the normality of MSEs based on 1000 simulations among three methods

	LS	NLS	Lapshin
MSE	28.1037	0.0503	0.1423
Shapiro P-value	<.0001	<.0001	0.1396

#### 4. Conclusion

Thermal hysteresis elliptical loops can be estimated using three methods, linear least squares, ellipse-specific nonlinear least squares, and the two-stage Lapshin regression. In the linear least squares method, a linear model with  $T_a$  and  $T_b$  is directly fitted. In the ellipse-specific nonlinear least squares method, the relationship between  $T_a$  and  $T_b$  can be described using a nonlinear model; a self-started function based on the ellipse-specific least squares method can be used to produce the starting values to facilitate model convergence. In Lapshin method,  $T_a$  and  $T_b$  can be analyzed using a two-stage procedure which models  $T_a$  and  $T_b$  sequentially.

Five elliptical parameters, lengths of major and minor axes, centroids, and angle of rotation, and four derived parameters lag, area, retention, and coercion can be estimated by each method. Lag and area of an elliptical hysteresis loop are two features used to characterize the dynamics of heat stress. The longer the lag between  $T_b$  and  $T_a$ , the larger area of the elliptical hysteretic loop, which indicated more heat needs to be dissipated during one cycle.

The simulation study showed there are differences among the methods in the quality of the parameter estimates. Linear least squares method generally produced low Wald coverages and non-normal estimates. The CV for coercion is a troubling 75%. The ellipse-specific nonlinear least squares method produced poor Wald coverages on three parameters: length of minor axis, area, lag, and coercion. Neither sandwich nor T-profile methods substantially improve coverages on the 5 elliptical parameters. Five estimates: length of major axis, centroids, angle of rotation, and retention, are symmetric about the true value. In the Lapshin method, all %bias are small but it is difficult to assess coverage. Once again five estimates: length of major axis, centroids, angle of rotation, and retention are symmetric about the true value.

The estimation of the derived parameters, lag and area, in the linear least squares

method was poor. The ellipse-specific nonlinear least squares method improved the estimation of lag and area, but the Wald coverages for the two parameters did not improve. The Lapshin method produced the smallest of bias on lag and area compared to the other two methods. However, the area estimates by Lapshin method were not normally distributed. The lack of normality leads to a bootstrapping study.

Comparisons of bootstrapping statistics for lag and area show that Lapshin method had the best coverage. In addition, the Lapshin bootstrapping estimate had the smallest bias for area and the smallest SE.

Overall, the Lapshin method produced small bootstrapping bias, small bootstrapping standard errors, and good 95% bootstrapping coverage for all parameters. Also, the Lapshin method produced the smallest averaged MSE and it was normally distributed. In conclusion, the Lapshin method is the preferred method for fitting the elliptical hysteresis loop and estimating the hysteretic parameters.

## References

- Bates, D.M. and D.G. Watts. 1988. *Nonlinear regression analysis and its applications*. Wiley Online Library.
- Brokate, M. and J. Sprekels. 1996. *Hysteresis and Phase Transitions*. Springer Verlag.
- Carroll, R.J. and D. Ruppert. 1988. *Transformation and weighting in regression*. Chapman Hall CRC.
- Carroll, RJ, S. Wang, DG Simpson, AJ Stromberg, and D. Ruppert. 1998. The sandwich (robust covariance matrix) estimator. *Unpublished manuscript*.
- Efron, B. 1979. Bootstrap methods: another look at the jackknife. *The annals of Statistics* 7(1):1-26.
- Efron, B. and R. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman Hall CRC.
- Fitzgibbon, A.W., M. Pilu, and R.B. Fisher. 1999. Direct Least Squares Fitting of Ellipses. *Pattern Recognition, 1996., Proceedings of the 13th International Conference on* 1:253-257.
- Hahn, G.L., A.M. Parkhurst, and J.B. Gaughan. 1997. Cattle Respiration Rate as A Function of Ambient Temperature. *Transactions of American Society of Agricultural Engineering* 40:97-121.
- Halir, R. and J. Flusser. 1998. Numerically Stable Direct Least Squares Fitting of Ellipses. *The Sixth International Conference in Central Europe on Computer Graphics and Visualization* 59-108.
- Horn, B.K.P and E.J. Weldon. 1986. Filtering closed curves. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* (5):665-668.

- Lapshin, R.V. 1995. Analytical Model for the Approximation of Hysteresis Loop and Its Application to the Scanning Tunneling Microscope. *Review of Scientific Instruments* 66(9):4718-4730.
- Mader, T.L., S.M. Holt, J.B. Gaughan, G.L. Hahn, M.S. Davis, A.M. Parkhurst, and D.E. Spiers. 2001. Heat Load Management for Feedlot Cattle. *Proceedings, Sixth International Livestock Environment Symposium. Louisville, Kentucky* 147-153.
- O'Leary, P. and P. Zsombor-Murray. 2004. Direct and Specific Least-Square Fitting of Hyperbole and Ellipses. *Journal of Electronic Imaging* 13:492.
- Parkhurst, A.M. 2010. Model for Understanding Thermal Hysteresis During Heat Stress: A Matter of Direction. *International Journal of Biometeorology* 54:637.
- Pilu, M., A.W. Fitzgibbon, and R.B. Fisher. 1996. Ellipse-Specific Direct Least-Square Fitting. *Image Processing, 1996. Proceedings., International Conference on* 3:599-602.
- Ritz, C. and J.C. Streibig. 2008. *Nonlinear regression with R*. Springer Verlag.
- Yang, F. and A.M. Parkhurst. 2010. Characterizing Thermal Hysteresis In Body Temperature Of Heat Stressed Steers. *Proc. 22nd Annual Kansas State University Conference on Applied Statistics in Agriculture. Kansas State Univ. Manhattan, KS.*
- Zeileis, A. 2006. Object-oriented computation of sandwich estimators. *Journal of Statistical Software* 16(9):1-16.