COMPARATIVE STUDY OF TIME SERIES AND MULTIPLE REGRESSION FOR MODELING DEPENDENCE OF CATTLE BODY TEMPERATURE ON ENVIRONMENTAL VARIABLES DURING HEAT STRESS

M. Pathak

A. M. Parkhurst

R. A. Arias

T. L. Mader

See next page for additional authors

Author Information

M. Pathak, A. M. Parkhurst, R. A. Arias, and T. L. Mader

COMPARATIVE STUDY OF TIME SERIES AND MULTIPLE REGRESSION FOR MODELING DEPENDENCE of CATTLE BODY TEMPERATURE ON ENVIRONMENTAL VARIABLES DURING HEAT STRESS

M. Pathak[1], A. M. Parkhurst[1], R. A. Arias[2] and T.L. Mader[2]
[1]University of Nebraska – Lincoln, Department of Statistics
[2]University of Nebraska – Lincoln, Department of Animal Science

Abstract

During the summer, a challenging thermal environment is known to cause a significant reduction in food intake, growth, milk production, reproduction and even death in cattle. In this study, we attempt to characterize the relationship of cattle body temperature with several environmental variables, such as air temperature, soil surface temperature, relative humidity, solar radiation, wind speed, incoming and outgoing short and long wave radiation. For these variables, the measurements taken over time are correlated. This places severe restrictions on the applicability of many conventional statistical methods that depend on the assumption of independent and identically distributed errors. In addition to these assumptions, there is serious collinearity among several weather variables and the variables are not stationary. Commonly used multiple regression models can be misleading when predictor variables are stochastic and issues of collinearity and non-stationary are ignored. In this paper, time series analysis is used as a tool to investigate the adequacy of classical regression models. Various aspects of dynamics of cattle body temperature and its relationship to environmental variables are discussed using the frequency and time domain analysis. Finally, we present a detailed approach for fitting cattle body temperature using a transfer function model with multiple environmental variables as inputs.

1. Introduction

Hot weather has negative effects on cattle welfare and performance. Economic losses in the feedlot industry alone averaged between \$10 million to \$20 million/year as a result of adverse climatic conditions (Mader, 2003). During hot weather, cattle not adapted to hot climates suffer from excessive heat load leading to heat stress (Gaughan et al., 2008).Reduction in food intake, growth and fertility and increase of respiratory and mortality rate are some of the animal responses to hot weather (Hahn & Mader, 1997). During this period, key environmental factors such as air temperature, soil surface temperature, relative humidity, and solar radiation are relatively high, causing heat waves that place cattle at risk and pose serious threats to performance, productivity and health of cattle during summer (Hahn, 1999, Hahn & Mader, 1997).

The general health status, animal comfort and thermal balance are often assessed by core body temperature (Finch, 1986; Mader et al., 2002; Mader et al., 2005). Body temperature of healthy adult cattle ranges between 37.7 and 40.2ºC (Lindley & Whitaker, 1996; Mader et al., 2002).

New Prairie Press
https://newprairiepress.org/agstatconference/2009/proceedings/7

Study has shown that body temperature presents small circadian fluctuations during the day, which follow the same pattern of changes observed in some environmental variables (da Silva and Minomo, 1995, Feng et al., 2000). In order to prevent susceptibility to hyperthermia and improve overall summertime feedlot performance, management strategies designed to alter the peak and/or pattern of body temperature must be implemented (Davis et al. 2003). Knowledge of how cattle body temperature responses to hot weather is extremely important for management strategies. Modeling the dynamics of the dependency of cattle body temperature with composite effects of many environmental variables is crucial in this respect. However, the applicability of many conventional statistical methods that depend on the assumptions of independent and identically distributed errors is severely restricted due to the fact that measurements taken over time are correlated. Commonly used multiple regression models can be misleading when predictor variables are non-stationary and issues of collinearity are ignored. Therefore, the purpose of this study is to make a comparative study of time series and multiple regression for modeling dependence of cattle body temperature on environmental variables during heat stress.

This paper is organized as follows. We describe data collection, statistical methods, statistical analysis and diagnostic procedures in section 2. In section 3.a, we discuss issue of multicollinearity among environment variables, inadequacy of classical regression models, and present results when variables are non-stationary and measurements are correlated over time. In section 3.b, we perform spectral analysis to identify dominant signals that governed underlying processes of several variables of interest and discuss coherency between responses with several predictor variables. Transfer function models (also called lagged regression, Shumway & Stoffer, 2006) with multiple environmental variables as input will be discussed thereafter. Finally, we present the conclusion in section 4., and in section 5., the summary of this study.

2. Materials and Methods

2. a. Response and Environmental Variables

The response variable for this study is tympanic temperature, an indicator of cattle core body temperature (Tb). The environmental variables included in the analysis are: air temperature (Ta), soil surface temperature (Tss), relative humidity (RH), temperature-humidity index (THI), wind speed (WS), net solar radiation (SNR), incoming shortwave solar radiation (SSWin), outgoing shortwave solar radiation (SSWout), incoming long wave solar radiation (SLWin), and outgoing long wave solar radiation (SLWout).

2. b. Data Collection

The dynamics of dependency of Tb on environmental variables is based on data collected during the summer of 2007. The experiment was conducted in the feedlot facilities of the Haskell Agricultural Laboratory in Concord, NE. One hundred and twelve steers (978.74± 83.1 lb) were used for this trial in order to assess the effects of the use of niacin when used as a depressant of heat stress in feedlot cattle. Steers were housed in two alleys (7head/pen). Two steers in each pen received a temperature data logger. Previous to the experimental period, cattle were implanted

(Revalor-S), vaccinated (Vision 7 and Vista Once), and ear tagged for individual identification. Tympanic temperature was collected for the period of July 5–12, 2007 at one-hour intervals. Environmental variables were collected at one-hour intervals for seven days from two weather stations, one for each alley. Both weather stations are located in the pens, thereby representing the micro-climate at the animal level. Soil surface temperature was recorded using a laser infrared gun located approximately at 1.85 m height, which was attached to the weather station. Incoming and outgoing shortwave solar radiation were collected using two precision spectral pyranometers (Eppley Lab. Inc., Newport, RI), whereas incoming and outgoing long wave radiation were collected using two precision infrared radiometers (Eppley Lab. Inc., Newport, RI). Simultaneously, net solar radiation was also collected using a REBS Net Radiometer model Q-7.1 (Radiation and Energy Balance Systems, Inc., Seattle, WA). All radiation measurements were collected hourly in an adjacent empty pen in the feedlot. For the purpose of this study, we present detailed analysis of one steer from each group (Drug and Control) and only general results were presented for other steers.

2. c. Selection of Environmental Variables

Regression analysis assumes no linear dependence among predictor variables. Serious collinearity among variables makes parameter estimates extremely unstable and strongly influences regression results. We examined redundancy among environmental predictor variables by performing collinearity analysis. Collinearity among the environmental variables was done by calculating the condition index (CI) and variation inflation factor (VIF) using SAS REG procedure. Environmental variables with mild collinearity (VIF < 30, CI < 15) were included in the reduced model. A scatter plot matrix of all ten variables was used to facilitate variable selection process.

2. d. Multiple Regression (MR) and Regression with Autocorrelated Errors (MRAE)

The classical multiple regression model (MR) is frequently used to characterize the dependency of a response variable on several predictor variables. Such models are based on the classical assumptions $\text{iidN}(0,\sigma^2)$. However, an uncorrelated error structure is not always plausible when the data are recorded over time. We need to modify the classical approach to take error structure into account. A multiple regression model with autocorrelated error (MRAE) is often used for this purpose. A MRAE model for p input variables $x_1, x_2, ..., x_p$ is expressed by the equation

$$y_t = \mu + \delta_1 x_{1,t} + ... + \delta_p x_{p,t} + \frac{\theta(B)}{\varphi(B)}\psi_t,\ \psi_t \sim \text{iidN}(0,\sigma_\psi^2)$$

where $\delta_1, ..., \delta_p$ are regression coefficients and $\xi_t = \frac{\theta(B)}{\varphi(B)}\psi_t$ is ARMA model for error term. In time series, we assume a stationary covariance structure for the error process that corresponds to a linear process and find an ARMA representation for the errors.

Predictor variables chosen after performing collinearity analysis were used to model Tb for the steer in each group. We checked the adequacy of MR model for time series data by examining the residuals for validity of model assumptions of iidN $(0, \sigma^2)$. Normality of residuals was checked by the Shapiro-Wilk test. We also inspected the sample autocorrelation function (ACF) of the MR residuals to check the validity of the assumptions of uncorrelated errors. We then proceeded to fit MRAE model for correlated error. The sample ACF and partial ACF were used to identify an appropriate ARMA model for residuals. The ARIMA commands in R were used to estimate the parameters for MRAE. Similar diagnostics of residuals of the MRAE models were performed to check model assumptions. The goodness of fit for MR and MRAE were compared using MSE and two information criterions (AIC and SBC).

2. e. Spectral and Cross-Spectral Analysis

Many time series are composed of periodic components**.** Any stationary time series that has periodic components is considered as the random superposition of sines and cosines oscillating at various frequencies (Wei, 1990). The spectral density function, interpreted as a variance of time series over given frequency bands, helps the researcher explain the physical and biological meaning of underlying process. The spectral density function is the analogue of the probability density function, which expresses information in terms of cycles. Such cycles can be discovered in a time series using the periodogram. To identify the dominant signal of the response as well as of each of environmental variables, we used the SPEC.PRGAM command in R to calculate and graph the raw periodogram. Given time series $x_1, x_2, ..., x_p$, the periodogram is defined to be

$$I(\omega_j) = |d(\omega_j)|^2, \; j = 0,1,2,..,n\text{-}1.$$

where discrete Fourier transform is $d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t e^{-2\pi i \omega_j t}$, and the fundamental frequency $\omega_j = j/n$.

When there is a wide spread around dominant peaks, the raw periodogram is a mediocre estimator of the spectral density function due to large variability. Smoothing the periodogram reduces the variance and highlights the dominant signals (Shumway & Stoffer, 2006). A smoothed periodogram is defined as an average of periodogram values over the frequency band B of L << n contiguous fundamental frequencies centered around $\omega_j = j/n$ that are close to frequency of interest ω (Shumway & Stoffer, 2006). The band B is given as

$$B = \left\{ \omega : \omega_j - \frac{m}{n} \le \omega \le \omega_j + \frac{m}{n} \right\}, \; L = 2m+1$$

The value of spectral density f (ω) is fairly constant in the band B and is estimated well by the smoothed spectra defined by

$$\bar{f}(\omega) = \frac{1}{L} \left(\sum_{-m}^{m} I(\omega_j + \frac{k}{n} \right) \text{ for } k = -m,...,0,...,m$$

Fluctuations and the power spread around the dominant peaks were smoothed using the Daniell kernel over the frequency band B. The 95% confidence interval for the spectrum $f(\omega)$

corresponding to peaks at the given frequency was calculated for response as well as environmental variables. Cross spectral analysis was then carried out between response and environmental variables to find associations among their periodic components. A measure of the strength of such an association between two time series is the square coherence function, which is defined as,

$$\rho^2_{y.x}(\omega) = \frac{|f_{yx}(\omega)|^2}{f_{xx}(\omega) f_{yy}(\omega)}$$

where f_{xx}, f_{yy}, f_{yx} are individual spectra of x_t and y_t series and the cross spectra of x_t and y_t, respectively. The coherency function is used as a tool for relating common periodic behaviors of time series. Strong coherence between series will help extend classical regression to the analysis of lagged regression or transfer function models (Shumway & Stoffer, 2006).

2. f. Fitting Transfer Function Model (TRF)

MRAE takes autocorrelated errors into account but is not a plausible approach if the predictor variables are non-stationary and stochastic. The classical regression approach assumes complete independence over time. We use regression analysis assuming that each input series is a fixed unknown function of time. But if an input series is a non-stationary stochastic process, we should consider a transfer function model (TRF). The transfer function model for p input series $x_1, x_2, ..., x_p$ can be written as

$$y_t = \sum_{j=0}^{\infty} \sum_{i=1}^{p} (\alpha_i)_j (x_i)_{t-j} + \eta_t \qquad (1)$$

We assume that each input process x_i, i = 1,.., p, and the noise series η_t are each stationary and mutually independent. The coefficients $(\alpha_i)_1, (\alpha_i)_2, ...$ given in (1) describe weights assigned to past values of the input variables that are used in predicting the response y_t. Often, we observe systematic patterns in these coefficients. Equation (1) can be written in the form of a rational function model which is given by

$$y_t = \mu + \frac{\delta_1(B)}{\omega_1(B)} x_{1,t-d_1} + ... + \frac{\delta_p(B)}{\omega_p(B)} x_{p,t-dp} + \frac{\theta(B)}{\varphi(B)} \zeta_t \qquad (2)$$

Each term in this rational function model includes a small number of coefficients and a specific delay or shift parameter d_i. The numerator and denominator of each term of the rational function are given by $\delta(B) = \delta_0 + \delta_1 B^2 + ... + \delta_\tau B^\tau$ and $\omega(B) = 1 - \omega_1 B - \omega_2 B^2 - ... - \omega_d B^s$ respectively. In addition, $\theta(B)$ and $\varphi(B)$ are autoregressive and moving average operators of the error process η_t. The goal of transfer function modeling is to determine a parsimonious model involving simple forms of δ (B) and ω (B) in equation (2) and estimate all parameters.

When identifying TRF model with multiple input variables, the cross-correlation functions may be misleading if the input series are autocorrelated. One solution to this problem is to prewhiten the input series. The SAS ARIMA procedure was used to prewhiten the input series and fit a TRF model as follows:

1. Detrend the response and all environmental variables to be used in the model
2. Fit ARMA model to each of the input series to obtain $\varphi_1....\varphi_p, \theta_1....\theta_q$ in $\varphi(B)x_t = \theta(B)\zeta_t$
3. Apply the operator from step 2 to get the transformed output series $\tilde{y}_t$, $\hat{\varphi}(B)\, y_t = \hat{\theta}(B)\, \tilde{y}_t$
4. Use the cross-correlation function between transformed output and prewhitened input to suggest form for components polynomial ratio and estimated time delay $\alpha(B) = \dfrac{\delta(B)B^d}{\omega(B)}$.
5. Obtain estimates of $\beta = (\omega_1,..,\omega_\tau, \delta_0,..,\delta_s)$ for each of input variables by fitting linear regression and retain estimates of residuals u_t: $y_t = \sum_{k=1}^{\tau} \omega_k y_{t-k} + \sum_{k=0}^{s} \delta_k x_{t-d-k} + u_t$.
6. Apply the MA transformation $u_t = \omega(B)\eta_t$ to the estimated residual in step 5 to estimate the noise η_t and fit ARMA model to the estimated noise $\tilde{\eta}_t$ to get estimates of the coefficients for $\tilde{\varphi}_n(B)$ and $\tilde{\theta}_n(B)$ where $\tilde{\varphi}_n(B)$ and $\tilde{\theta}_n(B)$ are autoregressive and moving average parts of noise.

2. g. Variable and Model Selection for TRF

First, all six environmental variables were included in the model. Insignificant variables were removed sequentially from the model. We repeated this process until all included variables were significant. After identifying the set of significant variables, we used MSE and information criteria (smallest AIC, SBC) to choose the best model. The final TRF model was the one that included significant variables and had the smallest AIC, SBC and MSE. In this way, we identified the most important variables to be included in the TRF model.

3. Results and Discussion

3. a. i. Multicollinearity among Environmental Variables

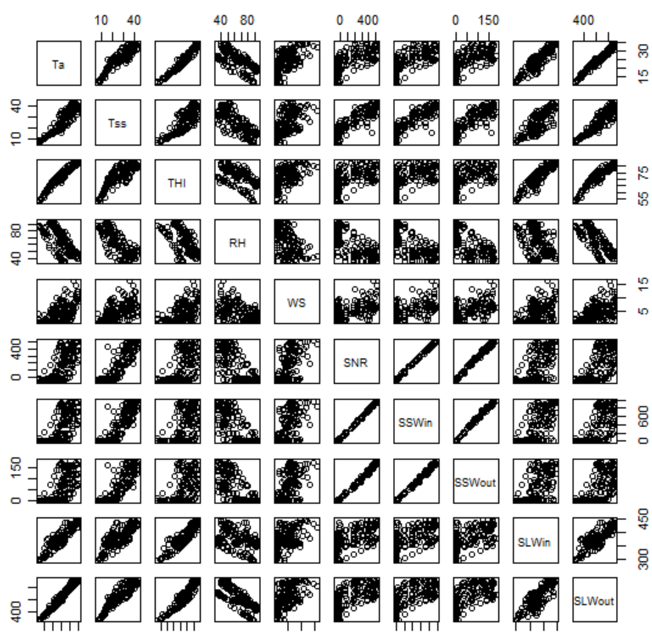
There was severe collinearity among the environmental variables (Table 1). The variance inflation factors indicated that the parameter estimates for Ta, THI, SNR, SSWin, SSWout, and SLWout were seriously affected by near-singularity in correlation matrix. Tests of collinearity by condition number also indicated that there was extreme collinearity among environment variables when all variables were included in the model since the maximum condition number > 100.

Table 1 Variation inflation factor and condition indices for the full model

Variation Inflation Factor			Condition Index		
Variable	Parameter Estimate	VIF		Eigenvalue	CI
Intercept	31.35	0	1	7.5	1
Ta	**-0.03**	**126.36**	2	1.4	2.32
RH	0.004	9.34	3	0.55	3.69
THI	**-0.06**	**85.89**	4	0.4	4.32
Tss	0.02	31.93	5	0.78	9.77
SNR	**-0.009**	**1384**	6	0.027	13.83
WS	0.00055	3.09	7	0.015	22.64
SSWin	**0.01**	**2151.9**	8	0.0067	32.71
SSWout	**-0.02**	**669.2**	9	0.001	85.04
SLWin	0.011	10.78	10	0.00029	158.03
SLWout	**0.016**	**71.23**			

Multiple regression models can be misleading when issues of collinearity are ignored. The scatter plot matrix shows that SSWin and SSWout are almost perfectly correlated with SNR, SLWin is highly correlated with SLWout and Ta has strong positive linear relationship with THI (Figure 1). Thus, we removed SSWin, SSWout, SLWout and THI from the analysis.

Figure 1 Scatter plot matrix showing plausible relation between environmental variables



There was little collinearity among the remaining six environmental variables Ta, Tss, RH, SNR,WS and SLWin as shown by the smaller variation inflation factors and condition indices (Table 2).

Table 2 Variation inflation factor and condition number among six environmental variables

Variation Inflation Factor				Condition Index		
Variable	Parameter	VIF			Eigenvalue	CI
Intercept	35.94	0		1	4.43	1
Ta	-0.02	19.63		2	0.69	2.53
Tss	0.06	24.54		3	0.53	2.9
RH	-0.007	4.04		4	0.28	3.98
SNR	-0.0007	9.06		5	0.05	9.23
WS	-0.03	2.45		6	0.02	13.56
SLWin	0.007	8.12				

3. a. ii. Multiple Regression Model (MR)

For the fitted MR model, the squared correlation coefficient (R^2) was higher for the steer in the drug group than the steer in the control group (0.82 vs. 0.67). The Ta was insignificant for both groups but SNR was significant only for the steer in the control group (Table 3).

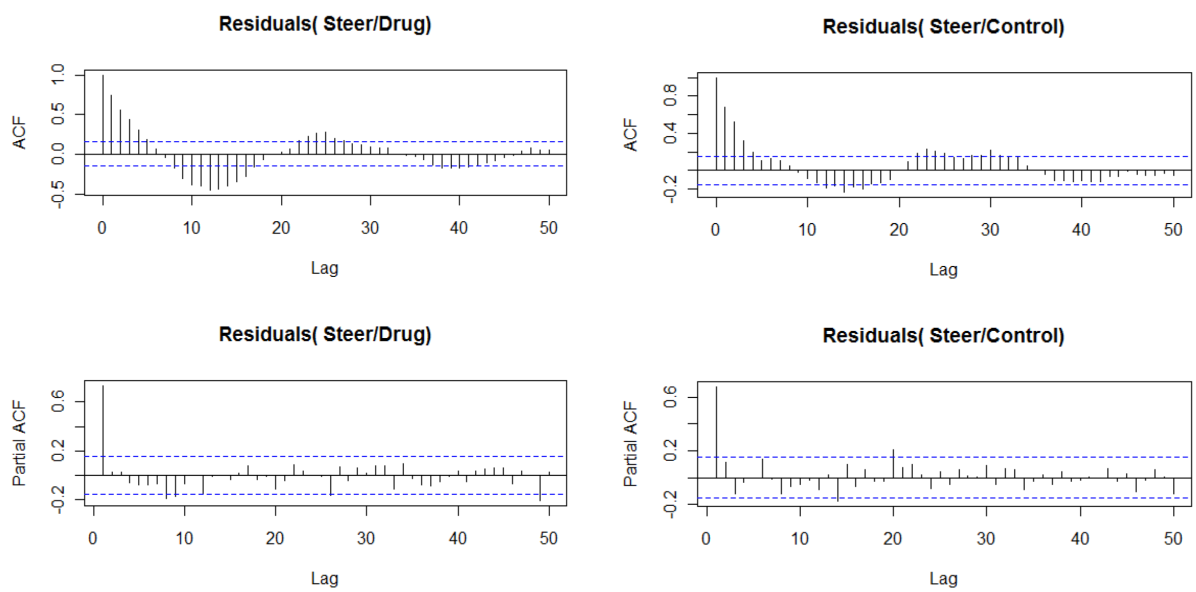
Table 3 Parameter estimates for multiple regression (MR) model of Tb vs. six environmental variables

Steer/Drug					Steer/Control			
Coefficient	Estimate	Std Error	P-value		Coefficient	Estimate	Std Error	P-value
Intercept	35.94	0.4	<0.0001		Intercept	38.29	0.44	<0.0001
Ta	-0.02	0.2	0.33		Ta	-0.009	0.02	0.65
Tss	0.06	0.01	<0.0001		Tss	0.04	0.01	0.003
RH	-0.007	0.003	0.01		RH	-0.02	0.003	<0.0001
WS	-0.03	0.011	0.014		WS	-0.09	0.012	<0.0001
SNR	-0.0006	0.0003	0.07		SNR	-0.0009	0.0004	0.02
SLwin	0.007	0.002	0.0002		SLwin	0.005	0.002	0.018

However, the applicability of this model was severely restricted due to the fact that this model violated the underlying assumptions of $iidN(0,\sigma^2)$ for error. Plots of the sample ACF and PACF (Figure 2) for the residuals for both steers indicated the AR (1) model would fit the residuals. The residuals of the models also deviated from the normality assumption for both the steers

(Shapiro-Wilk normality test, p=0.12 (Steer/Drug), p=0.0008 (Steer/Control). This suggested fitting MRAE model for Tb.

Figure 2 ACF of residuals from multiple regressions (MR) model



3. a. iii. Multiple Regression with Autocorrelated Errors (MRAE)

Even after removing the issue of collinearity from the models, the applicability of the MR model is restricted because the residuals of the fitted model do not satisfy the classical assumptions about independent residuals. Thus, we fitted the MRAE model. Table 4 shows that after taking AR (1) error structure into account, the significance of the predictor variables changed. Ta for the steer in both groups and SNR for the steer in the drug group were significant in the MRAE model whereas they were not significant in the MR model. On the other hand, WS and SLWin were not significant in the MRAE model for both steers. Therefore, Ta , Tss and SNR were the most important variables to model Tb during heat stress. These results are consistent with results from Rodrigo, 2008. Rodrigo reports that Ta, Tss and SNR are important predictor variables to model Tb during heat stress.

Inspection of the plot of the sample ACF for residuals showed no apparent departure from the model assumption of uncorrelated error structure. The Q-Q plot of residuals (Figure 3) indicates no departures from normality (Shapiro-wilk, p = 0.33 for Steer/Drug and p = 0.12 for Steer/Control).

A comparison between MR and MRAE (Table 5) shows that MRAE is a better approach to modeling Tb because AIC and MSE of MRAE are much smaller than for MR and the fitted residuals of MRAE satisfy the $iidN(0,\sigma^2)$ assumptions.

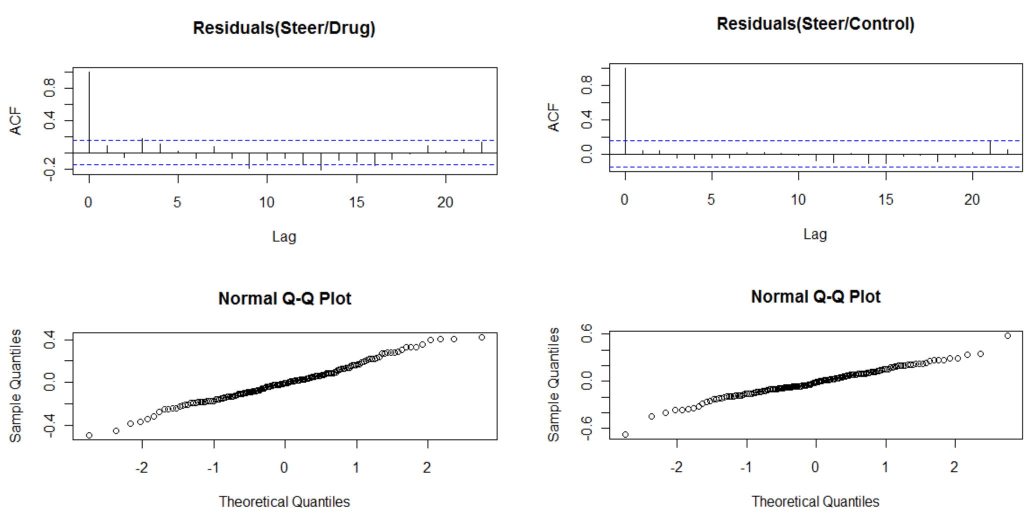
Table 4 Parameter estimates for multiple regression with autocorrelated errors (MRAE) model

Steer/Drug					Steer/Control			
Coefficient	Estimate	Std error	P > |t|		Coefficient	Estimate	Std error	P > |t|
Intercept	36.75	0.59	<0.0001		Intercept	37.53	0.62	<0.0001
Ta	0.05	0.016	0.002		Ta	0.07	0.015	<0.0001
Tss	0.023	0.005	0.0001		SST	0.006	0.006	0.02
SNR	0.0003	0.0003	0.03		SNR	-0.0005	0.0003	0.007
WS	0.002	0.007	0.77		WS	-0.0035	0.007	0.06
SLWin	0.001	0.002	0.44		SLWin	0.0003	0.002	0.07
AR(1)	0.87	0.04	<0.0001		AR(1)	0.94	0.028	<0.0001

Table 5 Comparison between multiple regressions (MR) and multiple regressions with autocorrelated errors (MRAR) models

	Steer/Drug				Steer/Control			
Model	AIC	MSE	Normality	Error Type	AIC	MSE	Normality	Error Type
MR	86	0.09	Normal	AR(1)	111	0.1	Non-Normal	AR(1)
MRAE	-98	0.03	Normal	White Noise	-110	0.03	Normal	White Noise

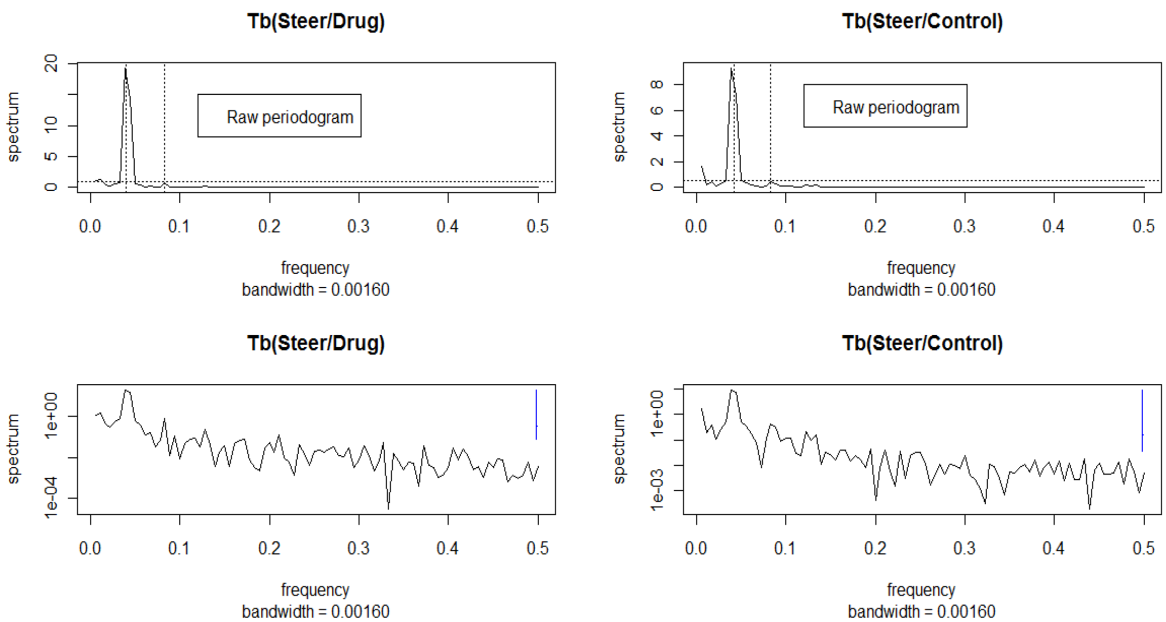
Figure 3 ACF and Q-Q plots of residuals for multiple regression with autocorrelated error (MRAE) model



3. b. i. Spectral Estimation with Nonparametric Smoothing

We noticed a major peak at the frequency ω ≈ 1/24 and a minor peak at ω ≈ 1/12 in the raw periodogram of Tb (Figure 4). Those two peaks suggest that dynamic of Tb is governed by periodic oscillations and cycles correspond to 24 and 12 hours periodic variation in Tb.

Figure 4 Raw periodgram of Tb, n = 165 showing peaks at ω ≈ 1/24 and ω ≈ 1/12 cycles/hour with 10% tapering (top). Average periodogram ordinates plotted on log10 scale displaying a generic 95% confidence interval in the upper right- hand corner (bottom)



The estimated periodogram of the Tb was $I_b(1/24) = 19.34$ for the Steer/Drug and 9.34 for the Steer/Control at the frequency of 24 hours per cycle (Also, $I_b(1/12) = 0.77$ for the Steer/Drug and 0.4 for the Steer/Control at the frequency of 12 hours per cycle). The approximate 95% confidence intervals for the spectrum $f_b(1/24)$ and $f_b(1/12)$ were too wide to be of much use for the steer in both groups (Table 6).

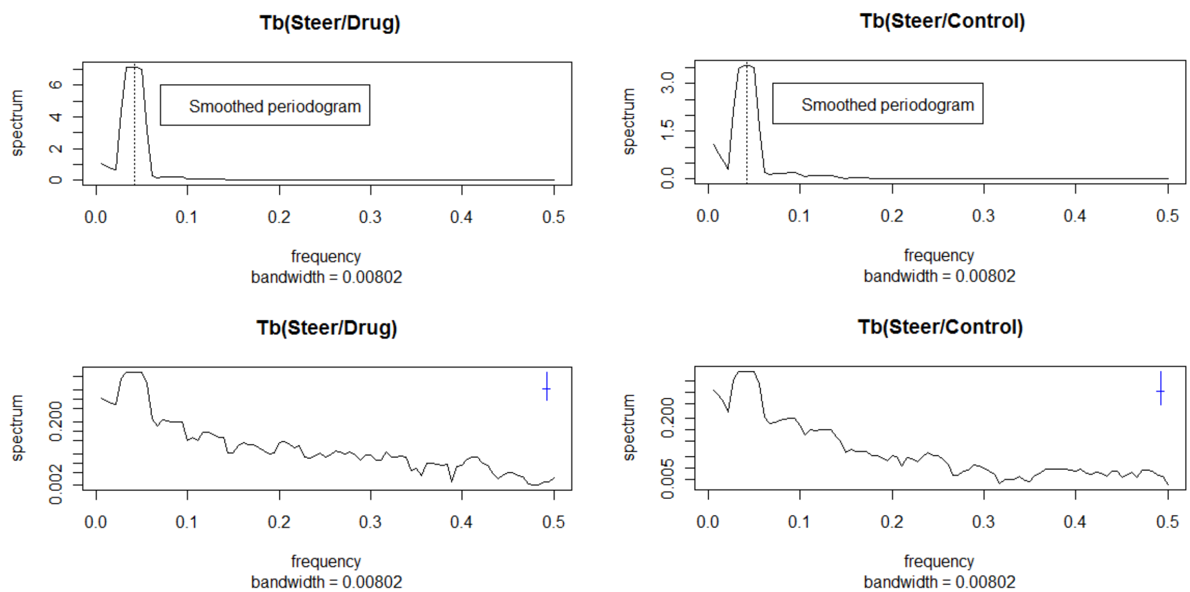
Table 6 Confidence intervals for the spectrum $f_b(\omega)$ of Tb

				Steer/Drug			Steer/Control		
	Series	ω ≈	Period ≈	I_b	Lower	Upper	I_b	Lower	Upper
	Tb	1/24	24 hr	19.3	5.25	765	9.3	2.53	369.12
Raw Periodogram	Tb	1/12	12 hr	0.77	2.1	30.3	0.4	0.11	16.36
Smooth Periodogram	Tb	1/24	24 hr	7.14	3.29	25.65	3.6	1.64	12.79

We noticed that there were fluctuations and some power spread around the dominant peak which indicates that the raw periodogram could be improved by smoothing the periodogram. Different trials lead to the choice L = 5 as a reasonable value to smooth the periodogram. Using Danielle nonparametric smoothing kernel (Shumway & Stoffer, 2006) with L = 5, the smoothed data had only one periodic component with an oscillation of roughly 24 hours (Figure 5). The peak at $\omega \approx 1/12$ disappeared. This indicated that $\omega \approx 1/24$ was the dominant frequency that governed the process.

The 95% confidence intervals for the smoothed periodogram were relatively narrow (Table 6) for the steers in both groups. In addition to the response Tb, a similar smoothing technique was applied to all six environmental variables. After smoothing using the Danielle kernel (Shumway & Stoffer, 2006), peaks at $\omega \approx 1/24$ remained but disappeared at $\omega \approx 1/12$ for Ta, Tss and RH. For SNR, WS and SLWin, there were peaks at $\omega \approx 1/24$ and 1/12.

Figure 5 Smoothed periodogram of Tb, n = 165 showing peaks at $\omega \approx 1/24$ and $\omega \approx 1/12$ cycles/hour with 10% tapering (top). Average periodogram ordinates plotted on log10 scale displaying a generic 95% confidence interval in the upper right- hand corner (bottom)



3. b. ii. Coherency

Spectral analysis showed that the response as well as all predictor variables, exhibited periodic oscillation behavior. In such cases, the classical approach of measuring association between variables via Pearson correlation is not appropriate. Coherency measures association when the variables have periodic components. The square coherency between Tb and Ta is illustrated in Figure 6.

The square coherency between Tb with Ta, Tss, SNR, WS and SLWin was significant but not significant for RH. Table 7 shows that there was coherency between Tb and Ta, Tss, WS, SNR and SLWin within a small neighborhood of the dominant frequency $\omega \approx 1/24$ except WS. For both groups, the coherence confidence intervals were wider for Ta, Tss, and SNR than for WS and SLWin. Significant coherency suggested an extension of classical regression to the transfer function model (Shumway & Stoffer, 2006).

Figure 6 Squared coherence between Tb and Ta, L = 5, n = 165, reference line at alpha = 0.01 (Steer/Drug (left); Steer/Control (right))

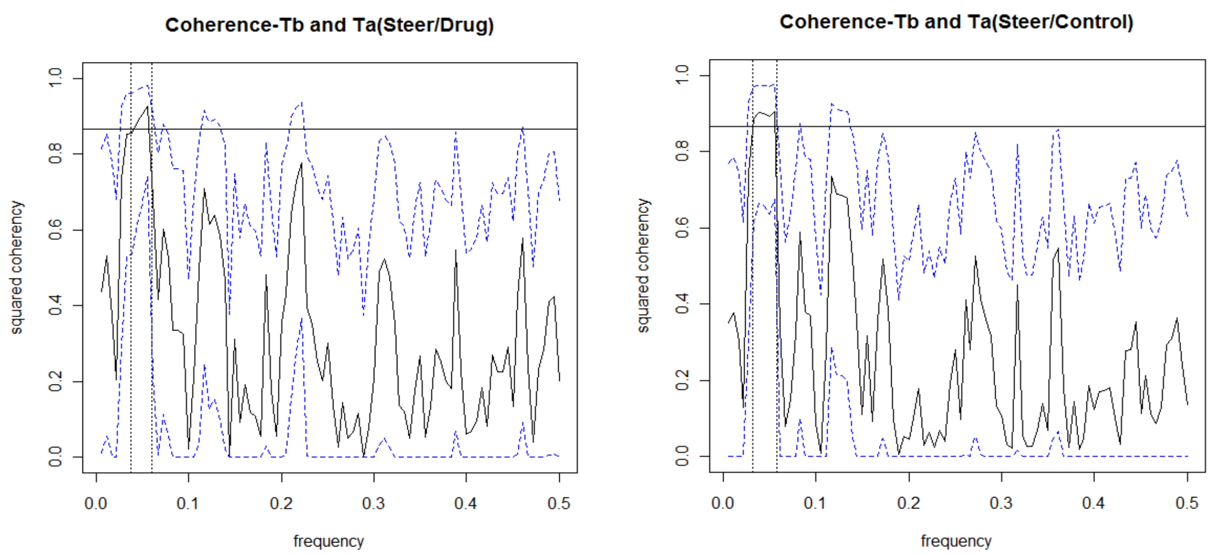


Table 7 Squared coherency confidence interval between Tb and other environmental variables, with Danielle smoothing L = 5, alpha = 0.01

Variable	Steer/Drug	Steer/Control
Ta	(0.038 0.58)	(0.032 0.06)
Tss	(0.029 0.058)	(0.026 0.058)
SNR	(0.02 0.056)	(0.026 0.058)
WS	(0.044 0.06)	(0.044 0.05)
SLWin	(0.04 0.056)	(0.034 0.056)

3. c. Transfer Function Model (TRF)

Regression of environmental variables with time (Table 8) showed that many predictor variables were non-stationary. The environment variables Ta, Tss and SLWin have a downward trend over time ($p < 0.0001$). This indicates that both MR and MRAE models that are based on static input is unrealistic although MRAE satisfied all model assumptions about error. When input series are stochastic and non-stationary, the transfer function model (i.e. lagged regression models) is more appropriate. The suitability of TRF was also suggested by spectral and coherency analysis.

Table 8 Stationarity check for the response and environmental variables

Raw Data				Detrended Data		
Variable	Slope	P-value		Variable	Slope	P-value
Tb/Drug	-0.05	<0.0001		Tb/Drug	-4E-08	0.99
Tb/Control	-0.04	<0.0001		Tb/Control	0.00003	0.48
Ta	-0.06	<0.0001		Ta	-0.0001	0.66
Tss	-0.07	<0.0001		Tss	-0.0008	0.88
RH	-0.01	0.64		RH	0.003	0.78
SNR	-0.4	0.23		SNR	0.01	0.9
WS	-0.008	0.12		WS	-0.001	0.71
SLwin	-0.4	<0.0001		SLwin	-0.007	0.64

All the non-stationary series were detrended before fitting the TRF model. Detrended series were stationary (Table 8) and the variance was stable over time for each variable ($p < 0.0001$, B-C transformation). Therefore, the detrended series were used to fit TRF.

Table 9 Comparison of transfer function (TRF) models

Seven Steer/Drug Models					Five Steer/Control Models			
Input Variable	AIC	SBC	MSE		Input Variable	AIC	SBC	MSE
Ta, Tss, SLWin	-110	-97	0.028		**Ta, Tss**	-118	-108	0.027
Ta, Tss	-105	-95	0.03		Tss, SNR	-111	-102	0.028
Tss, SLWin	-101	-92	0.033		Ta	-115	-109	0.028
Ta, SLWin	-89	-80	0.03		Tss	-81	-75	0.034
Ta	-85	-78	0.034		SNR	-99	-92	0.031
Tss	-86	-80	0.033					
SLWin	-34	-28	0.046					

Among several competing models, TRF with Ta, Tss and SLWin as input had smaller AIC, SBC and MSE for the Steer/Drug. The TRF with Ta and Tss as input had a smaller AIC, SBC and MSE for the Steer/Control (Table 9). In each case, the cross-correlation function between response and predictor variables contained only a finite numbers of impulses leading to TRF with only numerators terms in equation (3). Correlations between response and predictor variables were highest at zero lag for Ta and Tss and at lag 2 for SLWin for steer in the drug group (Appendix, Table 14, 15 & 16). For the steer in the control group, correlations were highest at lag one for Ta and at lag zero for Tss (Appendix, Table 18 & 19). The ACF plot of residuals of the fitted TRF for the steer in both groups is white noise (Appendix, Table 17 & 20). Parameter estimates for the final TRF models are given in Table 10 and 11.

Table 10 Parameter estimates for transfer regression function (TRF) model (Steer/Drug)

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
AR1,1	0.849	0.042	20.27	<.0001	1	Tb_detrend	0
NUM1	0.045	0.014	3.26	0.0011	0	Ta_detrend	0
NUM2	0.027	0.005	4.98	<.0001	0	Tss_detrend	0
NUM3	0.005	0.002	2.90	0.0038	0	SLWin_detrend	2

Table 11 Parameter estimates for transfer regression function (TRF) model (Steer/Control)

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
AR1,1	0.936	0.0275	34.10	<.0001	1	Tb_detrend	0
NUM1	0.068	0.010	6.59	<.0001	0	Ta_detrend	1
NUM2	0.008	0.004	1.96	0.0499	0	Tss_detrend	0

The Comparison of all three approaches to fit Tb is given in Tables 12 and 13. The superiority of TRF over MR and MRAE is consistent for all steers in both groups. However, input variables to be included in TRF are different for steer 3 from the drug group and the steers 2 and 4 from the control group. For the steer 3 in drug group, TRF with Ta and Tss has smaller AIC, SBC and MSE. The TRF with Ta, Tss and SLWin as input better fits data for steers 2 and 4 in the control group. This indicates that Ta, Tss and SLWin are the most important variables that affect Tb in heat stressed cattle.

Table 12 Model comparisons for the steers in the drug group

Model	Steer/Drug							
	Steer1**		Steer2		Steer3		Steer4	
	AIC	MSE	AIC	MSE	AIC	MSE	AIC	MSE
MR	86	0.09	50	0.07	149	0.13	88	0.1
MRAE	-98	0.03	-89	0.03	-70	0.04	-53	0.04
TRF	-110	0.03	-93	0.03	-79	0.03	-97	0.03

Table 13 Model comparisons for the steers in the control group

Model	Steer/Control							
	Steer1**		Steer2		Steer3		Steer 4	
	AIC	MSE	AIC	MSE	AIC	MSE	AIC	MSE
MR	111	0.1	95	0.09	96	0.1	89	0.09
MRAE	-101	0.03	-89	0.03	-12	0.1	-54	0.04
TRF	-118	0.03	-104	0.03	-16	0.1	-74	0.04

Steer1:** Detailed analysis of the steer 1 from both groups was presented in result section above**.**

Based on the parameter estimates (Tables and 10 and 11), we propose the following TRF models for the steers in each group.

TRF for the steer in the control group

$$\tilde{Tb}_t = 0.068*\tilde{Ta}_{t-1} + 0.008*\tilde{Tss}_t + \frac{\zeta_t}{(1-0.94B)}$$

which is equivalent to

$$\tilde{Tb}_t = 0.94*\tilde{Tb}_{t-1} + 0.068*\tilde{Ta}_{t-1} - 0.064*\tilde{Ta}_{t-2} + 0.008*\tilde{Tss}_t - 0.0075*\tilde{Tss}_{t-1} + \zeta_t$$

TRF for the steer in the drug group

$$\tilde{Tb}_t = 0.045*\tilde{Ta}_t + 0.027*\tilde{Tss}_t + 0.005*\tilde{SLWin}_{t-2} + \frac{\zeta_t}{(1-0.85B)}$$

which is equivalent to

$$\tilde{Tb}_t = 0.85*\tilde{Tb}_{t-1} + 0.045*\tilde{Ta}_t - 0.038*\tilde{Ta}_{t-1}$$
$$+ 0.027*\tilde{Tss}_t - 0.023*\tilde{Tss}_{t-1} + 0.005*\tilde{SLWin}_{t-2} - 0.004*\tilde{SLWin}_{t-3} + \zeta_t$$

4. Conclusion

Cattle body temperatures measured repeatedly over time are neither independent nor stationary. In addition to correlated errors, there was serious collinearity among the predictors, which were non-stationary and stochastic. Classical multiple regression models developed for the static case are inadequate for explaining all the interesting dynamics of cattle body temperature. Instead, time series analysis provides better insight of the underlying biological processes. Spectral analysis of the response, as well as, the predictor variables shows that all variables exhibit periodic oscillation that repeats roughly in every 24 hours and there is strong coherence between cattle body temperature with all environmental variables except relative humidity. In this case, transfer function (lagged regression) models fit the data better than the classical regression approach, even with adjustment for correlated errors. Among several environmental variables, air temperature, soil surface temperature and incoming long wave solar radiation and their lag variables are the most important predictor variables in modeling cattle body temperature during heat stress. For the steer in the control group, current cattle body temperature depends on its own previous value at lag1, air temperature at lag 1 and 2, current soil surface temperature and soil surface temperature at lag1. Similarly, for the steer in the drug group, current cattle body temperature depends on its own previous value at lag1, current air temperature, air temperature at lag1, current soil surface temperature, soil temperature at lag 1 and the incoming long wave solar radiation at lag 2 and 3.

5. Summary

In an attempt to model cattle body temperature, several points need to be considered. Among these, periodic and non-stationary behaviors over time are of fundamental importance that draws a line between classical regression and the time series approach. Contrary to the previous research in agriculture science where solar radiation, wind speed and relative humidity are important variables (Rodrigo, 2008), these variables do not show up in our final transfer function models. Study shows that air and soil surface temperature and incoming long wave solar radiation and their lags up to certain order (Model 1 and 2 above) are the most important variable for modeling summer cattle body temperature. These results are consistent with several other studies in this research area. However, there are certain limitations to the current study. First, we only analyzed data recorded in a summer of 2007. Studies show that solar radiation flux densities vary across the region and such changes may be due to season, year, time of day, and different geographical and environmental condition (Rodrigo, 2008). More research is needed to generalize the conclusions of this study in practice.

References

Arias R.A., 2008. Modeling the Effects of Environmental Factors on Finished Cattle. PhD Dissertation.

Davis M. S., Mader T. L., Holt S. M. & Parkhurst A. M., 2003. Strategies to reduce feedlot cattle heat stress: Effects on tympanic temperature. J. Anim. Sci. 81.

Da Silva R. G. & Minomo F. R., 1995. Circadian and seasonal variation of the body temperature of sheep in a tropical environment. International Journal of Biometeorology 39:69-73.

Feng S., Parkhurst A. M., Eskridge K., Spiers D., Travnicek D. A., Hahn G. L. & Mader T.L., 2000. Assessing refinements in modeling sinusoidal conditions used to drive cattle body temperature. Proc. 13th Annual Kansas State University Conference o Applied Statistics in Agriculture. 299-307. Kansas State Univ., Manhattan KS.

Finch V. A., 1986. Body temperature in beef cattle: - Its control and relevance to production in the tropics. J. Anim. Sci. 62: 531-542.

Gaughan J.B., Mader T.L. & Holt S.M., 2008. Cooling and feeding strategies to reduce heat load of grain-fed beef cattle in intensive housing. livestock science 113: 140-157.

Hahn G.L., 1999. Dynamic Responses of Cattle to Thermal Heat Loads, *J Anim Sci*. 77:10-20.

Hahn G. L. & Mader T. L., 1997. Heat waves in relation to thermo-regulation, feeding behavior and mortality of feedlot cattle. Proc. 5th Int. Livest. Environ. Symp. Amer. Soc. Agric. Eng., St. Joseph, MI. 563–567

Mader T. L., Holt S. M., Hahn G. L., Davis M. S. & Spiers D. E., 2002. Feeding strategies for managing heat load in feedlot cattle. J. Anim Sci. 80: 2373-2383.

Mader T. L. 2003. Environmental stress in confined beef cattle. J. Anim. Sci. 81:110-119.

Mader T. L., Davis M. S. & Kreikemeir W. M., 2005.Tympanic temperature and behavior associated with moving feedlot cattle. Prof. Anim. Sci. 21: 339-344.

Mader T. L., Davis M. S. & Brown-Brandl T., 2006. Environmental Factors Influencing Heat Stress in Feedlot Cattle. J. Anim. Sci. 84:712–719.

Lindley J.A. & Whitaker J.H., 1996. Agriculture buildings and structures. ASAE, St. Joseph, MI.

Parkhurst A. M. & Hahn G. L., 1987. Data dependent systems time series analysis of cattle tympanic temperature rhythms. ASAE Paper MCR 87-131. American Society of Agricultural Engineers, St. Joseph, MI.

Shumway H. R. & Stoffer S. D., 2006. Time Series Analysis and its Applications with R Examples. Second edition.Springer-Verlag, New-York.

Wei W.S., 1990. Time Series Analysis: Univariate and Multivariate Methods. Addison – Wesley.

Appendix

Table 14 Cross correlation between detrended Tb and Ta for the steer in drug group

```
Lag    Covariance    Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 -3     0.0098834       0.04575     |                  .  |* .                  |
 -2    -0.0062208       -.02880     |                  . *|  .                  |
 -1      0.033118       0.15331     |                  .  |***                  |
  0      0.044189       0.20456     |                  .  |****                 |
  1    -0.0030760       -.01424     |                  .  |  .                  |
  2    0.00031104       0.00144     |                  .  |  .                  |
  3     0.0010255       0.00475     |                  .  |  .                  |
```

Table 15 Cross correlation between detrended Tb and soil surface temperature, Tss, for the steer in drug group

```
Lag    Covariance    Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

 -3      0.017628        0.02648    |                 .  |* .                 |
 -2    -0.0065354        -.00982    |                 .  |  .                 |
 -1      0.023555        0.03538    |                 .  |* .                 |
  0      0.290235        0.43599    |                 .  |*********           |
  1      0.100911        0.15159    |                 .  |***                 |
  2     -0.012898        -.01938    |                 .  |  .                 |
  3      0.063293        0.09508    |                 .  |**.                 |
```

Table 16 Cross correlation between detrended Tb and incoming long wave solar radiation, SLWin, for the steer in drug group

```
Lag    Covariance    Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

 -3      0.196813        0.11511    |                 .  |**.                 |
 -2      0.079695        0.04661    |                 .  |* .                 |
 -1      0.288557        0.16876    |                 .  |***                 |
  0      0.202691        0.11855    |                 .  |**.                 |
  1      0.150468        0.08800    |                 .  |**.                 |
  2      0.383664        0.22439    |                 .  |****                |
  3     -0.122609        -.07171    |                 . *|  .                 |
```

Table 17 ACF plot of the residuals from proposed Transfer Function, TRF, model for the steer in drug group

```
Lag    Covariance    Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1      Std Error

  0      0.028848        1.00000    |                    |********************|             0
  1     0.0022078        0.07653    |                .   |**.                 |      0.078326
  2    -0.0015432        -.05349    |                . * |  .                 |      0.078783
  3     0.0040113        0.13905    |                .   |***                 |      0.079006
  4     0.0021390        0.07415    |                .   |* .                 |      0.080493
  5    -0.0002618        -.00908    |                .   |  .                 |      0.080911
  6    -0.0024915        -.08636    |                . **|  .                 |      0.080918
```

Table 18 Cross-correlation between detrended Tb and Ta for the steer in the control group

```
Lag    Covariance    Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 -3     -0.035119        -.16508    |                 ***|  .                 |
 -2    -0.0060400        -.02839    |                 . *|  .                 |
 -1      0.023200        0.10905    |                 .  |**.                 |
  0     0.0084161        0.03956    |                 .  |* .                 |
  1      0.034139        0.16047    |                 .  |***                 |
```

```
          2    -0.0038577         -.01813     |                 .  |  .                    |
          3      0.019166         0.09009     |                 .  |**.                    |
```

Table 19 Cross-correlation between detrended Tb and soil surface temperature, Tss, for the steer in control group

```
        Lag     Covariance    Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
         -3     -0.043676        -.07083     |                 . *|  .                    |
         -2     -0.072365        -.11736     |                 .**|  .                    |
         -1     -0.014178        -.02299     |                 .  |  .                    |
          0      0.126858        0.20574     |                 .  |****                   |
          1      0.071970        0.11672     |                 .  |**.                    |
          2      0.093208        0.15117     |                 .  |***                    |
          3      0.095981        0.15566     |                 .  |***                    |
```

Table 20 ACF of residuals from proposed Transfer Function, TRF, model for the steer in the control group

```
 Lag    Covariance    Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1      Std Error
   0      0.027727        1.00000    |                    |********************|            0
   1     0.0015072        0.05436    |                 .  |* .                 |     0.078087
   2    -0.0009629        -.03473    |                 . *|  .                 |     0.078317
   3    -0.0011249        -.04057    |                 . *|  .                 |     0.078411
   4    -0.0006450        -.02326    |                 .  |  .                 |     0.078539
   5    -0.0003751        -.01353    |                 .  |  .                 |     0.078581
   6    -0.0021323        -.07690    |                 .**|  .                 |     0.078595
   7    0.00021965        0.00792    |                 .  |  .                 |     0.079053
   8    0.00026925        0.00971    |                 .  |  .                 |     0.079058
```