Kansas State University Libraries New Prairie Press

Conference on Applied Statistics in Agriculture

2009 - 21st Annual Conference Proceedings

# USING TIME SERIES TO STUDY DYNAMICS OF SWEAT RATES OF HOLSTEIN COWS EXPOSED TO INITIAL AND PROLONGED SOLAR HEAT STRESS

B. Liang

A. M. Parkhurst

K. G. Gebremedhin

C. N. Lee

R. J. Collier

See next page for additional authors

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

 $igsimed{O}$  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**

Liang, B.; Parkhurst, A. M.; Gebremedhin, K. G.; Lee, C. N.; Collier, R. J.; and Hillman, P. E. (2009). "USING TIME SERIES TO STUDY DYNAMICS OF SWEAT RATES OF HOLSTEIN COWS EXPOSED TO INITIAL AND PROLONGED SOLAR HEAT STRESS," *Conference on Applied Statistics in Agriculture*. https://doi.org/ 10.4148/2475-7772.1084

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.

## Author Information B. Liang, A. M. Parkhurst, K. G. Gebremedhin, C. N. Lee, R. J. Collier, and P. E. Hillman

## USING TIME SERIES TO STUDY DYNAMICS OF SWEAT RATES OF HOLSTEIN COWS EXPOSED TO INITIAL AND PROLONGED SOLAR HEAT STRESS

B. Liang<sup>1</sup>, A.M. Parkhurst<sup>1</sup>, K.G. Gebremedhin<sup>2</sup>, C.N. Lee<sup>3</sup>, R.J. Collier<sup>4</sup>, P.E. Hillman<sup>2</sup>

Department of Statistics, University of Nebraska at Lincoln
 Department of Biological and Environmental Engineering, Cornell University
 Department of Animal Science, University of Hawaii at Hawaii
 Department of Animal Science, University of Arizona

## ABSTRACT

Sweating is a very important way for cows to cope with heat stress. We are interested in the ability of Holstein cows to sustain high sweat or evaporation rates when exposed to solar radiation. There were two solar heat stress treatments: onset and prolonged. The onset data provided an opportunity to examine the impact of sudden exposure to a solar thermal load. The prolonged data allowed us to examine the impact of exposure to solar heat stress for an expended period (5 hr). Two questions of interest were: Do cows sweat at a constant or cyclic rate? Is there a difference in the dynamics of the two treatments: onset and prolonged solar heat stress? The data were examined for stationarity. In the time domain, we fit ARIMA models and estimated the parameters. In the frequency domain, we used nonparametric spectral estimation to identify cyclic patterns in the sweat rates. The usefulness of each technique for analyzing the dynamics of sweat rates is discussed.

## **1. INTRODUCTION**

Many studies have been conducted to investigate the relationship between milk production and environmental conditions, especially air temperature, of high-producing cows. It is well known that a thermally challenging environment has detrimental effects in the characteristics of milk (Regan and Richardson 1938; West 2003). The initial purpose of this experiment was to study how a hot and humid environment affects the milk yields of high producing Holstein cows. However, in a pre-trial experiment, simulated hot and humid environment reflecting field conditions in many regions of the country, such as Florida, Arizona, Hawaii and southern California, proved too stressful for high producing dairy cows. After 3 hours of solar exposure in the pre-trial experiment, several cows had rectal temperatures above 40.0<sup>o</sup>C and a respiration rate above 125 /min. Thus, the focus of the study turned to how cows sweat during hot and humid environment. Sweating (or transpiration) is the production of water excreted by sweat glands in the skin. It is a means of thermoregulation. To stay cool, animals sweat to loose water by evaporation. The main purpose of this study was to learn whether the sweat (or evaporation) rate is constant or cyclical in nature. In order to do that, spectral analysis method was used to study the periodogram of evaporation rate in the frequency domain.

Water evaporation for a cow may be affected not only by the environmental situations, such as ambient temperature, relative humidity and air velocity, but also by the cow's own internal mechanisms. Since our data were collected as a time series, it is desirable to know how the evaporation rate of cow is affected by its own past values. Therefore, ARIMA models are used to explore the relationship between the current value of a cow's evaporation rate and its past values

in time. For both onset data and prolonged data, results from both the time and frequency domains are compared in an attempt to learn more about the dynamics of evaporation for both series. The objective of this study is to present a protocol for using time series to characterize sweat rates associated with initial and prolonged exposure to solar radiation.

## 2. MATERIALS AND METHODS

## 2.1. Data Collected and Experimental Design

The experiment was conducted in the environmental modules of the Parker Agricultural Research Complex, University of Arizona, Tucson AZ. Twelve high producing Holstein cows were subjected to a hot-humid environment with the solar load at 550 watts/m<sup>2</sup> for about 5 hours (10:00am-15:00pm). The initial room temperature was set to 30<sup>o</sup>C and relative humidity was 70%. After about 5 hours of exposure to solar radiation, the room temperature was increased to 35°C, and the relative humidity was decreased to 50%. In this experiment, measurements such as body temperature, relative humidity, air velocity (flow rate) and atmospheric pressure were measured by a portable calorimeter immediately before and after air passed through the sample area (76 mm x 102 mm) of the dorsal surface of a cow defined by the portable calorimeter. Data were collected for 20 minutes at 10-sec intervals. Measurements before the air passed through the sample area were defined as ambient measurements, while measurements after the air passed through the sample area were defined as sample measurements. Humidity Ratio was calculated as a function of all temperature, humidity and pressure. Unlike temperature, evaporation rate could not be measured directly. Instead, it is a calculated variable. It is the humidity ratio difference of ambient and sample (Hillman et al. 2001). For the onset treatment, data were collected right after the solar load was turned on. For the prolonged treatment, data were collected after the cows were exposed to solar radiation for at least 4 hours. Data for cow 5651 were made available to demonstrate the use of time series analysis for characterizing the two treatments.

#### **2.2. Statistical Methods**

#### 2.2.1. Spectral Analysis

In order to study the cyclical behavior of a time series, spectral analysis is often used. With the aid of a periodogram (an estimator of the spectral density function), strong periodic components in a time series can be identified by large periodogram values at or near those frequencies. In this study, the nonparametric spectral estimation method was used. The spectrogram was generated using the function *spec.prgam* in the statistical package  $R^{\text{(e)}}$  (Version 2.7.2). Smoothing was used to minimize the noise and make the signals more prominent for easier identification. But it requires a reasonable bandwidth of the L fundamental frequencies so that the real frequencies are not averaged out by taking too broad a band. Bandwidth, denoted by  $B_{\omega}$ ,

has the following relationship:  $B_{\omega} = \frac{L}{n}$ . A number of different L values were selected to get a set

of averaged periodograms and then the L value was selected by comparing the averaged periodograms. The values of L need to be kept small relative to n (the number of data points). In this study, three L values (3,5,7) were used. Also, the modified Daniell kernel was used to

improve the resolution of the peaks in the smoothed periodograms. In order to decide which band peaks are significant, we calculated 95% confidence intervals beginning with the largest peak. The lower limit of the confident interval was used as a baseline. Only those band peaks above the baseline were considered dominant peaks and the corresponding frequencies were used to define the periodic components.

## 2.2.2. ARIMA Model

## 2.2.2.a Stationarity

To fit an ARIMA model to a time series {  $x_t$ ;  $t = 0, \pm 1, \pm 2, ...$ }, the series must be stationary. In this study, weak stationarity was used.

Mean function of a series  $x_t$  is defined as:  $u_{xt} = E(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx$  (1)

Autocovariance function is defined as:  $\gamma_x(s,t) = E[(x_s - u_s)(x_t - u_t)]$  (2)

Weak stationary, means the mean function, defined in equation (1), is constant and does not depend on time t, and the covariance function, defined in equation (2), depends on s and t only through their difference |s-t|. Frequently, the trend and the variability of a series can be detected by looking at the time series plot. Therefore, a time series plot of each treatment was studied before proceeding with further analysis.

Differencing and detrending are two major methods that are often used to achieve stationarity of the data. Detrending provides more information about the dynamics of the series. After removing the trend from the data, the resulting data achieve stationarity in the mean. In this study, to achieve stationary data, the detrending method was used. Logarithm and power transformations are very useful to equalize the variability over the length of a single series. The Box-Cox power transformations *box.cox.powers* in the *car* library (R. 2008) are used to check the stability of the variance.

## 2.2.2.b ARIMA model

A time series {  $x_t$ ;  $t = 0, \pm 1, \pm 2, ...$ } will be ARMA(p,q) if it is stationary and

 $\phi(B)x_t = \theta(B)w_t$  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_n B^p$ 

 $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 \dots + \theta_q B^q$ 

Where,  $w_t$  is white noise, normally distributed with mean zero and variance  $\sigma_w^2$ , and  $\phi_p \neq 0, \theta_q \neq 0$ . The parameters p and q are called the autoregressive and the moving average orders respectively.

If  $x_t$  has a nonzero mean  $\mu$ , we set  $\alpha = \mu^* (1 - \phi_1 - \dots - \phi_p)$  and the model becomes  $x_t = \alpha + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q + w_{t-q}$ 

If differencing is used to get a stationary time series, the ARIMA model is used. A process  $x_t$  is said to be ARIMA (p,d,q) if  $\phi(B)^* (1-B)^d x_t = \theta(B) w_t$ .

Since only the detrending method was used in this study, we focused on ARMA(p,q) model for both treatments series.

## 2.2.2.b.i Causality and Invertibility

A valid ARMA(p,q) model  $x_t - \phi_1 x_{t-1} - \cdots \phi_p x_{t-p} = w_t + \theta_1 w_{t-1} + \cdots + \theta_q + w_{t-q}$  requires causality and invertibility for the corresponding AR(p) and MA(q) parts. Using the backshift operator, B, an ARMA(p,q) can be written as  $\phi(B)x_t = \theta(B)w_t$ , where  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ . We treat the backshift operator, B, as a complex number, z. This ARMA model is causal only when the roots of  $\phi(z)$  lie outside the unit circle and it is invertible only when the roots of  $\theta(z)$  lie outside the unit circle (Shumway and David, 2002). In this study, *polyroot* fn in R (2008) was used to check the values of roots to see whether or not they lie outside the unit circle and further determine the causality and invertibility of the model.

## 2.2.2.b.ii Model Selection

AIC was the information criterion used for model selection in this study. A grid search was used to find possible ARMA models for the series. The grid search method gave us the AIC values for all the models that were specified. The maximum order of p was specified as four and the maximum order of q was specified as three. Based on the rule that the model with smaller AIC value is better, the model with smallest AIC value was selected.

## 2.2.2.b.iii Diagnostics of Residuals

The fit of an ARMA model was checked for possible violations of the classical assumptions about the residuals after it was selected and fit. Deviations of the residuals from normality were checked with the Shapiro-Wilk test, Q-Q plots and histograms. Checks for uncorrelated errors were performed by inspecting time plots and ACF plots of the standardized residuals, as well as, the Portmanteau (Ljung-Box) statistical plot for up to twenty lags.

In addition, theoretical ACF plots were compared to the AR(p) model selected for each treatment series. Theoretical ACF plots of AR parts of the model compared to visually investigate how well the model fits the effect of past values of evaporation rate on the current value.

For an AR(1) model, ACF is known as  $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h$ , where  $\rho(h)$  is autocorrelation;  $\gamma(h)$  is

the autocovariance of h lags;  $\gamma(0)$  is autocovariance of 0 lag;  $\phi$  is the coefficient of AR(1) model; h is the number of lag.

For an AR(2) model, ACF is known as  $\rho(h) - \phi_1 \rho(h-1) - \phi_2 \rho(h-2) = 0$ , h = 1, 2, ...

Let  $z_1$  and  $z_2$  be the roots of the associated polynomial,  $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$ . For a causal model, when  $z_1 = \overline{z}_2$  are a complex conjugate pair, then  $\rho(h) = c_1 z_1^{-h} + \overline{c_1} \overline{z_1}^{-h}$ , where  $c_1$  and  $z_1$  are written in polar coordinates, for example,  $z_1 = |z_1| e^{i\theta}$ , where  $\theta$  is the angle whose tangent is the ratio of the imaginary part and the real part of  $z_1$ . Then, using the fact that  $e^{i\alpha} + e^{-i\alpha} = 2\cos(\alpha)$ , the solution has the form  $\rho(h) = a |z_1|^{-h} \cos(h\theta + b)$ , where *a* and *b* are determined by the initial conditions. In this study, theoretical ACF plots were generated by using *ARMAacf* in R (2008) by specifying the estimated AR coefficients from the models used to fit the treatment time series.

#### **3. RESULTS AND DISCUSSION**

#### **3.1. Frequency domain**

#### **3.1.a.** Onset Treatment

In the frequency domain, spectral analysis was applied to inspect the possible period components for both series. In Figure 1, three frequency bands were tried to smooth the original periodogram of onset series. Notice that using the bandwidth corresponding L=5 to smooth the periodogram, Figure 2 gives us one predominant peak at  $\omega$ =0.032. The second higher peak is lower than the 95% lower confidence limit of the highest peak. Therefore, the second peak was not considered a significant peak. When  $\omega$ =0.032, the cycle of evaporation was 0.032 cycles/10 seconds, which means the period of evaporation was about 313 seconds( about 5.2 minutes).



#### **3.1.b.** Prolonged Treatment

The same method was used to inspect the period components of the prolonged series. Three frequency bands were tried to smooth the original periodogram of the transformed stationary series in Figure 3. Finally L=7 with the corresponding bandwidth was chosen to smooth the

periodogram in Figure 4. The predominant peaks appeared at  $\omega$ =0.1167 and  $\omega$ =0.0417. When  $\omega$ =0.1167, the cycle of evaporation rate was 0.1167 cycles/10 seconds, which means the period of evaporation was about 86 seconds (about 1.4 minutes). When  $\omega$ =0.0417, the cycle of evaporation rate was 0.0417 cycles/10 seconds, which means the period of evaporation was about 240 seconds (4 minutes).



## 3.2. Time Domain

## 3.2.a. Onset Treatment

To analyze evaporation rate series of cow 5651 in time domain, we begin with the time series plot. Visual inspection of the onset time series plot, displayed in Figure 5, does not show an obvious trend. Simple linear regression of evaporation rate  $(x_i)$  on time (t) is not significant (p=0.63). Thus the onset series is stationary in the mean. The Box-Cox Power test for constant variance shows that the series is stationary in the variance. The original data without any power transformation (p=0.076) are used directly and fit with an ARMA model. A grid search is

performed and the AIC values of ARMA(p,q) models are shown in Table 1, where the maximum p=4 and the maximum q=3.

Figure 5. Onset Evaporation Rate Time Plot



Table 1.	Onset AIC	c values	for
	ARMA(p,	q)model	

AIC				
\q				
p	0	1	2	3
0	0	1403.63	1400.26	1402.22
1	<mark>1399.09</mark>	1399.47	1401.35	1403.16
2	1399.75	1401.52	1403.39	1404.4
3	1401.74	1403.33	1404.97	1405.76
4	1400.82	1402.78	1399.99	1401.52

The ARMA(1,0) model has the smallest AIC and its statistics are presented in Table 2. The AR1 coefficient is significant and the model is causal.

Table 2. Onset Parameter estimates for ARMA(1,0) model

Coefficient	Estimate	P-value	SE
ar1	0.372	2.08e-05	0.0838
Intercept	325.058	0.00e+00	20.7343

Shapiro-Wilk test indicates the residuals are normally distributed (p=0.263). Figure 6 shows a time plot of the standardized residuals (top), an ACF plot of the standardized residuals (middle), and a Portmanteau-statistic plot (bottom). Standardized residuals of the model are within 3







standard deviations; the ACF of residuals shows no apparent departure from model assumptions; the Portmanteau statistics are not significant except at lag 2. Figure 7 shows a normal Q-Q plot of the residuals (top), and a histogram of the residuals (bottom). The residuals are close to normality except for a few extreme values in the tails (Figure 7). Therefore, we fit an ARMA(1,0) model to evaporation rate as

$$y_t = 325.058 + 0.372 y_{t-1} + w_t \tag{3}$$

Where,  $y_t$  is the evaporation rate at time(every 10 seconds);  $w_t$  is the white noise and it follows a normal distribution with mean zero and variance  $\sigma_w^2$ .

## 3.2.b. Prolonged Treatment

For the prolonged series, unlike the onset series, we observe an upward trend in the time series plot shown in Figure 8. Simple linear regression of evaporation rate  $(y_t)$  on time (t) is significant (p=0.01). Therefore removing the trend to make the series stationary in the mean is required.

$$y_t = 308.26 + 0.459t + \varepsilon_t \tag{4}$$

Removing the trend from the evaporation series results in a residual series, call it  $\varepsilon_t$ .

$$\varepsilon_t = y_t - 308.26 - 0.459t$$
 (5)

In this case,  $\varepsilon_t$  is stationary in the mean. Stationarity of  $\varepsilon_t$  in the variance is also checked. Box-Cox power test shows that  $\varepsilon_t$  series is constant in variance (p=0.07). A table of the AIC values to identify possible ARMA models for  $\varepsilon_t$  series is shown in Table 3.

Figure 8. Prolonged Evaporation Rate Time Plot



Table 3. F	Prolonged AIC values	for
A	ARMA(p,q) model	

AIC				
$\sqrt{q}$				
p	0	1	2	3
0	0	1337.24	1331.31	1333.26
1	1336.12	1336.03	1333.23	1335.23
2	1332.35	1333.76	<mark>1331.10</mark>	1332.78
3	1333.28	1334.19	1332.69	1333.44
4	1334.17	1335.78	1334.14	1335.44

The ARMA(2,2) model gives the smallest AIC, Table 3, and this model has significant p-values for all AR and MA coefficients, Table 4. Since the p-value of intercept is not significant, the intercept is excluded from the model. The assumed ARMA(2,2) model for the detrended evaporation series is as follows:

$$\varepsilon_t = 1.07\varepsilon_{t-1} - 0.74\varepsilon_{t-2} + w_t - 0.93w_{t-1} + 0.81w_{t-2}$$
(6)

This model can also be written as

$$\mathcal{E}_{t}(1-1.07B+0.74B^{2}) = w_{t} - 0.93w_{t-1} + 0.81w_{t-2}$$
(7)

Coefficient	Estimate	P-value	SE	
ar1	1.06508	9.71E-11	0.1492	
ar2	-0.7395	2.48E-06	0.1490	
mal	-0.92543	1.98E-09	0.1418	
ma2	0.814642	7.81E-10	0.1213	
intercept	0.155992	9.92E-01	14.8257	

Table 4. Parameter estimates of ARMA(2,2) model

Diagnostics are performed on the residuals of the ARMA(2,2) model for the detrended prolonged series. The Shapiro-Wilk test indicates the residuals are normally distributed (p=0.08). Figure 9 shows a time plot of the standardized residuals (top), an ACF plot of the standardized residuals (middle), and a Portmanteau-statistic plot (bottom). Standardized residuals of the model are within 3 standard deviations; the ACF of residuals shows no apparent departure from model assumptions; the Portmanteau statistics are not significant. Figure 10 shows a normal Q-Q plot of the residuals (top), and a histogram of the residuals (bottom). The residuals are close to normality except for a few extreme values in the tails (Figure 10).







Therefore, for prolonged treatment, we fit an ARMA(2,2) model to the detrended evaporation series as equation (7).

Using equations (5) and (7), and simplifying yields  $y_t = 206.53 + 1.07 y_{t-1} - 0.74 y_{t-2} + 0.459t - 0.491(t-1) + 0.34(t-2) + w_t - 0.93 w_{t-1} + 0.81 w_{t-2}$  Where,  $y_t$  is the prolonged evaporation rate at time (every 10 seconds);  $w_t$  is the white noise, also called random noise.

#### 3.2.c Theoretical ACF's to Describe Dynamics of Model

Estimates from the ARMA model are used to generate theoretical ACF's. The theoretical models describe the dynamics of the series. Comparing the theoretical autocorrelations between onset series and prolonged series provides more information about how past values of evaporation rate influence the current value.

For onset data, the AR(1) model fit the series well using  $\phi = 0.372$ . For prolonged data, the AR part of the fitted ARMA model was investigated since the influence of past values on current value is of interest. Coefficients for AR part are  $\phi_1 = 1.07$ ,  $\phi_2 = -0.74$  and

 $\phi(z) = 1 - 1.07z + 0.74z^2$  has a pair of complex conjugate roots. The theoretical ACF of onset series (top) and theoretical ACF of AR part of prolonged series (bottom) are shown in Figure 11.



For the onset series, autocorrelation between past values and current value decreased to zero exponentially as  $h \rightarrow \infty$ . A specific past value had little impact on the current value after 3 lags; in this case, 30 seconds or 0.5 min.

For the prolonged series, autocorrelation dampened to zero exponentially as  $h \rightarrow \infty$ , but it did so in a sinusoidal fashion. Moreover, it took longer for the impact to disappear than in the onset series. After about 15 lags, the past values had little impact on the current value; in this case, it was 150 seconds or 2.5min.

## **4. CONCLUSIONS**

Cow 5651 sweats in a cyclic pattern rather than at a constant rate. In the frequency domain, the nonparametric spectral analysis for each treatment shows a cyclic pattern. For solar onset, the period is 5.2 min, while for the prolonged solar exposure there are two significant periods, 1.4 min or 4 min. The additional period for the prolonged solar exposure indicates an additional

physiological process is needed. Panting is an additional process which would help the cow dissipate body heat.

Cow 5651 responded differently to solar heat stress at the onset compared to prolonged exposure. At the onset, the solar effect decayed exponentially to zero in approximately  $\frac{1}{2}$  min. However, after prolonged exposure, the solar effect decayed sinusoidally to zero in approximately  $\frac{2}{2}$  min. The longer solar effect on the cow after prolonged solar exposure indicates possible physiological fatigue of the cow.

The sinusoidal nature of the decline in impact of a previous reading along with additional cycle suggests the cow needs to work harder to dissipate body heat during prolonged exposure.

Time series techniques in the frequency domain, such as calculating the nonparametric spectral analysis, are useful in identifying the existence of cyclic patterns of sweat rates in cows. While, time series techniques in the time domain, such as estimating ARMA models and exploring theoretical ACF's are useful in describing the differences in the dynamics of the onset and prolonged treatments.

## **5. REFERENCES**

Hillman, P.E., K.G. Gebremedhin, A. Parkhurst, J. Fuquay, and S. Willard. 2001. *Evaporative and convective cooling of cows in a hot and humid environment*. Livestock Environment VI. Proceedings of the Sixth International Symposium. Galt House Hotel, Louisville, KY. May 21-23, 2001. ASAE, 2950 Niles Rd., St. Joseph, MI 49085-9659, USA.

Hillman, P.E., C.N. Lee and S.T. Willard. 2005. *Thermoregulatory responses associated with lying and standing in heat-stressed dairy cows*. Trans. ASAE 48(2):795-801.

Gebremedhin, K.G., P.E. Hillman, C.N. Lee, R.J. Collier, S.T Willard, J. Arthington, and T.M. Brown-Brandl. 2009. *Sweating rates of dairy cows and beef heifers in hot conditions*. *Tranactions of ASABE* 51(6): 2167-2178.

R Development Core Team 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Regan W. M., G. A. Richardson. 1938. *Reactions of the dairy cow to changes in environmental temperature*. J. Dairy Sci. 21(2): 73-79.

Shumway Robert H., David S. Stoffer, 2006. *Time Series Analysis and Its Applications: With R Examples*, Second Edition.

West, J. W. 2003. Effects of heat-stress on production in dairy cattle. J. Dairy Sci. 86:2131-2144.