Kansas State University Libraries

New Prairie Press

USING TIME SERIES TO STUDY EFFECT OF AIR TEMPERATURE ON BODY TEMPERATURE OF COWS IN PUERTO RICO

Y. Zeng

A. M. Parkhurst

J. Pantoja

Recommended Citation

USING TIME SERIES TO STUDY EFFECT OF AIR TEMPERATURE ON BODY TEMPERATURE OF COWS IN PUERTO RICO

Y. Zeng[1], A.M. Parkhurst[2], and J. Pantoja[3]
1. Department of Industrial Engineering, University of Nebraska at Lincoln
2. Department of Statistics, University of Nebraska at Lincoln
3. Department of Animal Industry, University of Puerto Rico.

ABSTRACT

Body temperature is an important measure for monitoring the health status of cows. The objective of this study is to (1) determine if a cow's body temperature is related to ambient temperature, relative humidity, and/or temperature humidity index (THI); (2) look for signs of heat stress. The data are collected at five minute intervals during the summer months (December through February) in Puerto Rico. Regression analysis and a succession of time series analyses are conducted in time domains. Nonparametric spectral estimation and cross-spectra analysis are also performed in the frequency domain. A search for indications of heat stress is performed by characterizing the relationship between body temperature and environmental factors. Detailed approaches of regression with autocorrelated errors and transfer function model in time domain are presented, along with the comparison between two models.

1. INTRODUCTION

Body temperature (Tb) is a key measure for monitoring the health status of cows. In the dairy cattle industry, it is often important to detect factors affecting the body temperature. These factors may include the ambient temperature (Ta), relative humidity (RH) of surrounding environment, and the combined effect of the two which is often expressed in terms of a temperature and humidity index (THI) (Lefcourt and Schmidtmann, 1989).

Previous research on feedlot cattle indicates that under conditions of high Ta and THI there is a serious threat of heat stress which can affect physiological and behavioral adaptations of cows. Levels of heat stress have been identified by Hahn et al. (1997). In chamber experiments, acute heat stress occurred when there was a 4-hr delay in the effect of Ta, while a delay of 3 hours indicated chronic heat stress. In the naturally occurring environment, Ingraham and Wagner (1974) found that when THI was above 72 in sub tropical areas, the cows' breeding behavior was affected. West (2003) found that ambient temperature and humidity had significant impact on food intake and milk yield during hot periods when the average THI ranged from 72.1 to 83.6. Other researchers reported that THI of 72 was the threshold for initial heat stress on dairy cattle (Arnold and Dudzinski, 1978; Endres and Barberg, 2008).

The objective of this study is to investigate relationships among body temperature and environmental variables to detect signs of heat stress in the time domain; and to identify dominant frequencies for cow body temperature and the association between body

temperature and ambient temperature using time series analysis methods in the frequency domain.

2. MATERIALS AND METHODS

2. a. Response and Environmental Variables
The response variable for this study is the vaginal body temperature of a cow (Tb). The environmental variables included in this study are: ambient temperature (Ta), relative humidity (RH), and the temperature humidity index (THI).
THI is calculated by:
$$THI = 0.8 \times Ta + [(RH \div 100) \times Ta - 14.4] + 46.4 \dots \dots (1)$$

2. b. Data Collection and Selection
The data are collected every five minutes consecutively from one cow during three different periods of the same length (49 hours in 3 consecutive days) in Puerto Rico during the summer months (December, January, and February). The Tb and Ta are measured with a HOBO® temperature data logger. Relative humidity data are collected with a HOBO® relative humidity data logger. The average THI in February is 72.8 indicating potential risk of heat stress (Figure 1). This value is higher than the other two months and with a smaller standard deviation. So data collected in February are analyzed in this study.

2. d. Statistical Models in Time Domain

2. d. i Regression with Autocorrelated Errors
The classical regression model is frequently used to characterize the dependency of a response variable on several independent variables. The presumed assumptions of such model are normality and uncorrelated error structure. In this study, the error structure is taken into account to modify the classical regression approach. Regression model with autocorrelated errors is used for this purpose. A regression model with autocorrelated errors for p input variables $x_1, x_2, \dots, x_p$ is expressed by the equation:

$$y_t = \mu + \delta_1 x_{1,t} + \cdots + \delta_p x_{p,t} + \frac{\theta(B)}{\phi(B)} w_t, w_t \sim iidN(0, \sigma_w^2) \dots \dots (2)$$

where $\delta_1, \dots, \delta_p$ are regression coefficients and $\xi_t = \frac{\theta(B)}{\phi(B)} w_t$ is ARMA model for error terms. Under the assumption of constant covariance structure for the error process, all parameters are estimated simultaneously using maximum likelihood estimators.

2. d. ii Transfer Function Model
Classical regression approach assumes complete independence over time. Regression analysis can be used assuming that input series are fixed unknown functions of time. These considerations make sense for the time argument t, but are less satisfactory if input series are stochastic processes (Shumway and Stoffer, 2006). In this study when input series are stochastic and non-stationary, lagged regression model (transfer function) is considered with

environmental variables as input. The transfer function model for p input variables $x_1,$ $x_2,...,x_p$ can be written as

$$y_t = \sum_{j=0}^{\infty} \sum_{i=0}^{p} (\alpha_i)(x_i)_{t-j} + \eta_t \dots \dots (3)$$

assuming that each of input process x_t and noise η_t are stationary and mutually independent. The coefficients $\alpha_1, \alpha_2, ...$ given in equation (3) describe weights assigned to past values of input variables that are used in predicting response y_t.

2. e. Statistical Analysis in Frequency Domain

2. e. i Nonparametric Spectral Estimation

Time series are often composed of periodic components. Stationary time series that have periodic components can be considered as the random superposition of sines and cosines oscillating at different frequencies (Shumway and Stoffer, 2006). The spectral density function is the analogue of the probability density function which expresses information in terms of cycles. Such cycles can be detected using a periodogram. For time series $x_1,$ $x_2, ..., x_p,$ the periodogram can be defined as

$$I(\omega_j) = |d(\omega_j)|^2, j = 0, 1, 2, ..., n - 1 \dots\dots (4)$$

where $d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t e^{-2\pi i \omega_j t}$, fundamental frequency $\omega_j = \frac{j}{n}$. Considering the uncertainties and power spread around the dominant peak in the raw periodogram, the periodogram can be smoothed to reduce variance and to identify the dominant frequency. A frequency band B can be defined and periodogram values can be smoothed over the band. The frequency band B is defined as

$$B = \left\{ \omega: \omega_j - \frac{m}{n} \leq \omega \leq \omega_j + \frac{m}{n} \right\} where \; L = 2m + 1, \; \omega_j = \frac{j}{n} \dots\dots (5)$$

L is an odd number and should be small relative to the data size of the time series n. The averaged or smoothed periodogram over band B can now be defined as

$$\bar{f}(\omega) = \frac{1}{L} \sum_{k=-m}^{m} I(\omega_j + k/n), k = -m, ...,0, ..., m \dots\dots (6)$$

2. e. ii Cross-Spectra Analysis

Cross-spectra analysis allows one to determine the relationship between two time series as a function of frequency. To measure the strength of the relationship, coherence function, which is the correlation-squared between two stationary series indexed by frequency, can be used as a tool for relating common periodic behaviors of time series. The existence of coherence at a particular significance level α leads to transfer function modeling in time domain.

2. f. Model Building and Statistical Analysis

This study begins with input variable selection and classical regression analysis between body temperature Tb and other environmental input variables. The error structure is checked as required by the assumption of the classical regression model. As suggested by the error structure of the classical regression model, a regression model with autocorrelated errors is constructed. Then in frequency domain, the nonparametric spectral estimation is performed; after de-trending the body temperature data Tb and ambient temperature data Ta, cross-spectra analysis is conducted. Based on coherency, a transfer function model with lagged Ta and Tb values is developed. Diagnostics of the residuals are performed after each model is built. The regression model with correlated errors and the transfer function are compared based on the model fit and diagnostics.

Statistical package R$^®$ (Version 2.9.2) is used for classical regression analysis and performing regression with autocorrelated errors in the time domain; as well as non-parametric spectral estimation and cross-spectra analysis in the frequency domain. SAS$^®$ (version 9.2) PROC REG, PROC AUTOREG and PROC ARIMA procedures are used to fit the regression and transfer function models.

2. g. Model Diagnostics

Residuals are plotted and examined after model construction. Time plots and sample ACF of standardized residuals are inspected to check the validity of model assumptions. Normality is investigated by the normal Q-Q plot of the standardized residuals and the Shapiro-Wilk test.

3. RESULTS AND DISCUSSIONS

3. a. Multiple Regression Models

3. a. i Input Variables Selection

The scatter plot of all environmental variables (Figure 2) suggests a problem with RH. Many RH values above 100 also affect the validity of THI. Thus, only Ta is used as environmental predictor

3. a. ii Classical Regression Model

A multiple regression model with Tb as response value and Ta as the input variable is fit with a significant intercept (P<0.0001) and slope (P=0.0007)
The classical regression model is:

$$Tb_t = 38.62_{(0.042)} + 0.0057_{(0.0017)}Ta_t + \varepsilon_t \ldots\ldots (7)$$

The applicability of regression models is restricted because residuals of the fitted model are correlated with slowly decaying ACF as shown in Figure 3. Also Figure 4 shows that the normality of standardized residuals is not good.

3. b. Regression with Autocorrelated Errors

3. b. i Model Building

Figure 3 shows a slow decay in the ACF and the cutting off of PACF at lag of 2, which clearly indicates the autocorrelation of residuals. An ARIMA(2,0,0) model is constructed to fit a regression model with autocorrelated errors.

Both sides of original regression function with Ta as input variable are transformed with above ARIMA(2,0,0) model and all parameters are re-estimated. Table 1 shows that after taking the AR(2) correlated error structure into account, Ta is no longer significant (P=0.38). The final model of regression with autocorrelated errors is

$$Tb_t = 38.85_{(0.1001)} + 1.56_{(0.033)}Tb_{t-1} - 0.59_{(0.033)}Tb_{t-2} - 0.0034_{(0.0039)}Ta_t$$
$$+ w_t, w_t \sim iidN(0, \sigma_w^2) \ \dots\dots (8)$$

These results suggest the variation in Tb is primary due to its previous 5 and 10 minute readings. Ambient temperature (Ta) is not significant in this model.

3. b. ii Model Diagnostics

Inspection of time plot of standardized residuals in Figure 5 shows no obvious patterns. Only a few outliers exceed 3 standard deviations in magnitude. The ACF of standardized residuals shows no departure from the model assumption of uncorrelated error with zero mean. In Figure 6 Q-Q plot and histograms show the normality of residuals. Shapiro-Wilk test also indicates the normality (P = 0.95).

3. c. Statistical Analysis in Frequency Domain

3. c. i Nonparametric Spectral Estimation

The non-parametric spectral estimation is applied to Tb. The raw periodogram can be improved by smoothing the periodogram to reduce the variance. Different trials lead to the choice L=15 as a reasonable value to smooth the periodogram (Figure 7). Using Danielle smoothing with L=15, the smoothed data suggest that no dominant peak exist, which indicates that there is no detectable dominant frequency governing the process.

3. c. ii Cross-Spectra Analysis

De-trending of Ta and Tb is first performed. The coherence function between Tb and Ta is illustrated in Figure 8. In the plot the middle horizontal line is the reference line at α = 0.05; dotted lines represent confidence bands at α = 0.05. It can be seen that the squared coherency between Tb and Ta is significant at the 5% level.

3. d. Transfer Function Model

Based on the existence of coherency between Tb and Ta, a transfer function model is developed. Ta is used as the input variable. The cross-correlation function between

transformed Tb and de-trended Ta suggests a single impulse at a delay of 52 observations as shown in Table 2.   In order to whiten the residuals of the proposed transfer function model the ACF (Table 3) and PACF (Table 4) are examined.   An AR(3) model is used to get estimated coefficients for the residuals of the proposed transfer function model.

The correlation between response and predictor variables is highest at lag 52 (approximately 4.3 hours) as shown in Table 3. This suggests a time delay of 52 observations.   Parameter estimates for final transfer function model are given in Table 5.   Based on the parameter estimates, the following transfer function model is built:

$$Tb_t = 2.8Tb_{t-1} - 6.5Tb_{t-2} - 1.8Tb_{t-3} + 3.5Tb_{t-4} - 0.4Tb_{t-5} - Tb_{t-6} + Ta_{t-52}$$
$$- 2.8Ta_{t-53} + 2.45Ta_{t-54} + 3Ta_{t-55} - 1.1Ta_{t-56} + 50w_t - 85w_{t-2}$$
$$+ 35w_{t-4} \qquad w_t \sim iidN(0, \sigma_w^2) \dots \dots (9)$$

The transfer function model (9) suggests that current Tb is significantly dependent on previous Tb and previous Ta approximately 4.3 hours ago, which indicates the presence of heat stress.   Table 6 shows that the residuals of the final transfer function model are uncorrelated.   The histogram and Q-Q plot are shown in Figure 9, which illustrate that residuals are close to normality.   The Shapiro – Wilk test yields a p-value of 0.846 which indicates that the residuals are normal.

3. e. Model Comparison
Comparison between regression model with autocorrelated errors and transfer function model is performed.   The diagnostics of both models indicate the validity of models.   However, in the regression model with autocorrelated errors the ambient temperature (Ta) is not significant; the lagged effect of Ta is not detected.   By comparison, transfer function model depicts the lagged effect of Ta with an approximate lag of 4.3 hr, which further indicates the presence of heat stress. (Hahn and et al., 1997)   So the transfer function model in this study is more appropriate for exploring the relationship between the body temperature of cows and ambient temperature during the summer in Puerto Rico.

4.   CONCLUSION
Classical multiple regressions model developed for exploring the relationship between the body temperature of cows and ambient temperature during the summer in Puerto Rico are not useful since the error assumptions are violated.   Instead, regression with autocorrelated errors is conducted and the result suggests that the current Ta has no effect on current Tb. Spectral analysis in frequency domain suggests that there is no dominant periodic component in the body temperature of cows.   The existence of coherency between Tb and Ta via cross-spectra analysis in the frequency domain leads to developing a transfer function which better explains the dynamics of Tb and the underlying effect of Ta.   The transfer function indicates that current Tb is significantly dependent on previous Ta approximately 4.3 hours ago, which suggests the presence of heat stress.

5. SUMMARY

This study characterizes the relationship between time series data of the ambient temperature and body temperature of the cows during the summer in Puerto Rico.   Because of the non-stationarity of both data series, the assumptions of classical regression models that are widely used in agriculture are easily violated.   Specific time series analysis methods in both time domain and frequency domain are considered as the natural alternative.   The correlated residuals of the classical regression model suggest the use of the regression model with autocorrelated errors and the coherency suggests the construction of the transfer function model.   The transfer function model indicates that current body temperature of the cow is significantly dependent on the ambient temperature approximately 4.3 hours ago, indicating the presence of heat stress, which is consistent with several previous studies in this research area.

The limitations of this study are that there is a validity problem with the humidity data in the collection process.   The data cannot be used for analysis and since humidity also affects the calculation of THI, it also is not included in the model building process.

6. REFERENCES

Arnold, G. W. and Dudzinksi, M. L. (1978). *Ethology of Free-Ranging Domestic Animals*. Elsevier, Holland.

Endres, M. I. and Barberg, A. E. (2008). Behavior of Dairy Cows in an Alternative Bedded-Pack Housing System. *Proceeding of the 31 August – 4 September 2008 Conference (Iguassu Falls, Brazil).*

Hahn, G.L., Parkhurst,A.M., and Gaughan,J.B. (1997). Cattle Respiration Rate as a Function of Ambient Temperature. *ASAE Mid-Central Conf. Paper* , 97-121.

Ingraham, R. H., Gillette, D. D., and Wagner, W. D. (1974). Relationship of Temperature And Humidity To Conception Rate Of Holstein Cows In Subtbropical Climate. *Journal of Diary Science*, 57(4), 476-4481.

Lefcourt, A. M. and Schmidtmann, E. T. (1989). Body Temperature of Dry Cows on Pasture: Environmental and Behavior Effects. *Journal of Dairy Science*, 72, 3040-3049.

Shumway, R.H. and Stoffer.D.S. (2006). *Time Series Analysis and Its Applications with R Examples*. Second edition. Springer-Verlag, New York.

West, J. W., Mullinix, B. G., and Bernard,J.K. (2003). Effects of Hot, Humid Weather on Milk Temperature, Dry Matter Intake, and Milk Yield of Lactating Dairy Cows. *Journal of Diary Science,* 86(1), 232-242.

Figure 1.   Time plot of THI in February
(Middle horizontal line represents mean THI = 72.8; n=596)
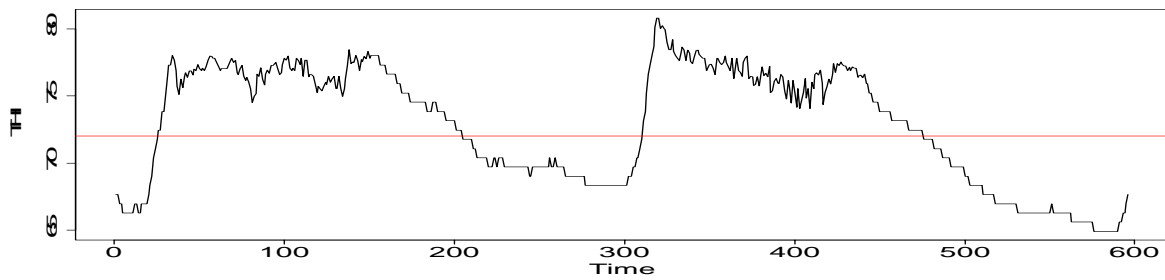
Figure 2.   Scatter plot matrix of variables
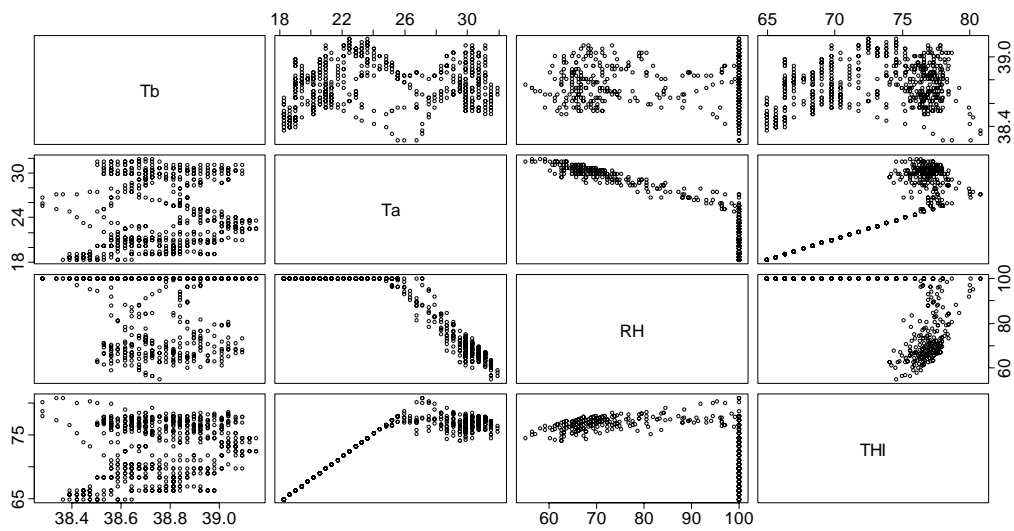
Figure 3.   The ACF and PACF of standardized residuals for
the classical regression model at α = 0.05
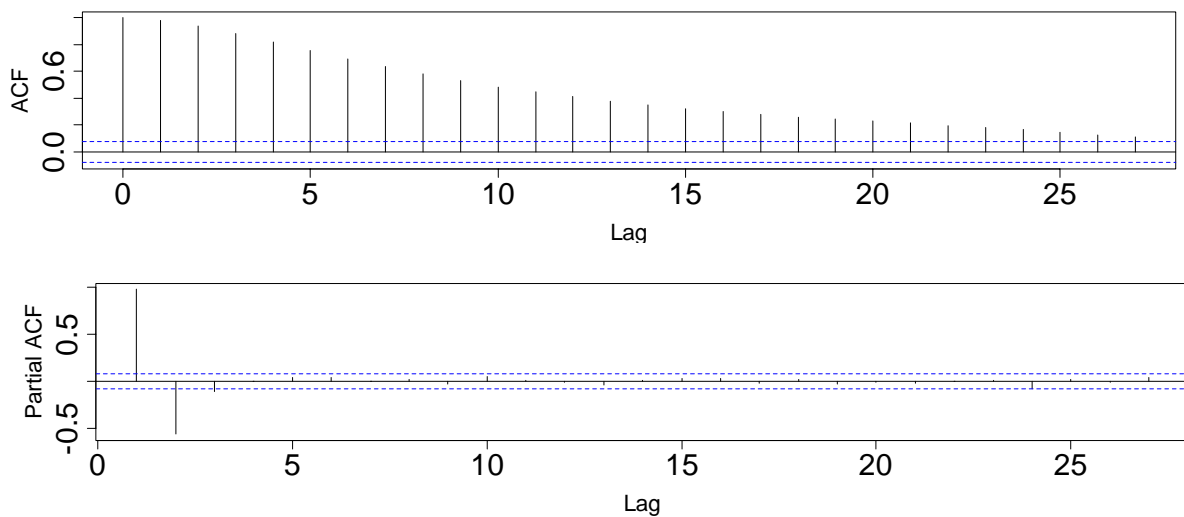
Figure 4.   The Q-Q plot and histogram of standardized residuals for
the classical regression model
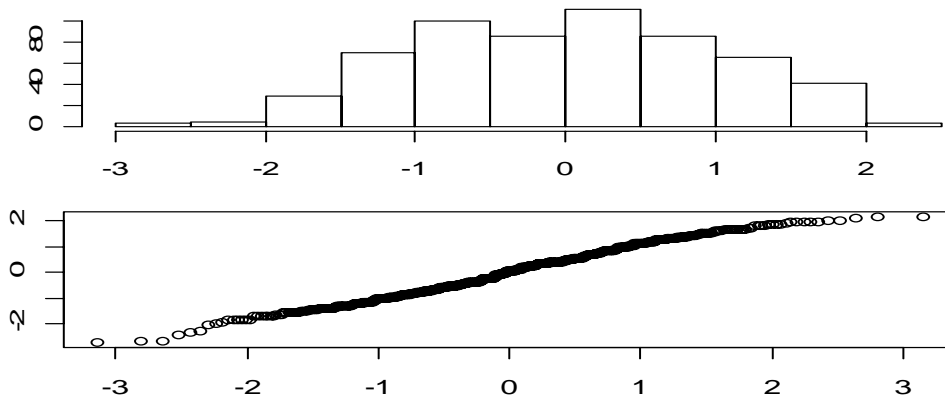
Figure 5.   Plot and ACF of standardized residuals of regression model with auto correlated errors
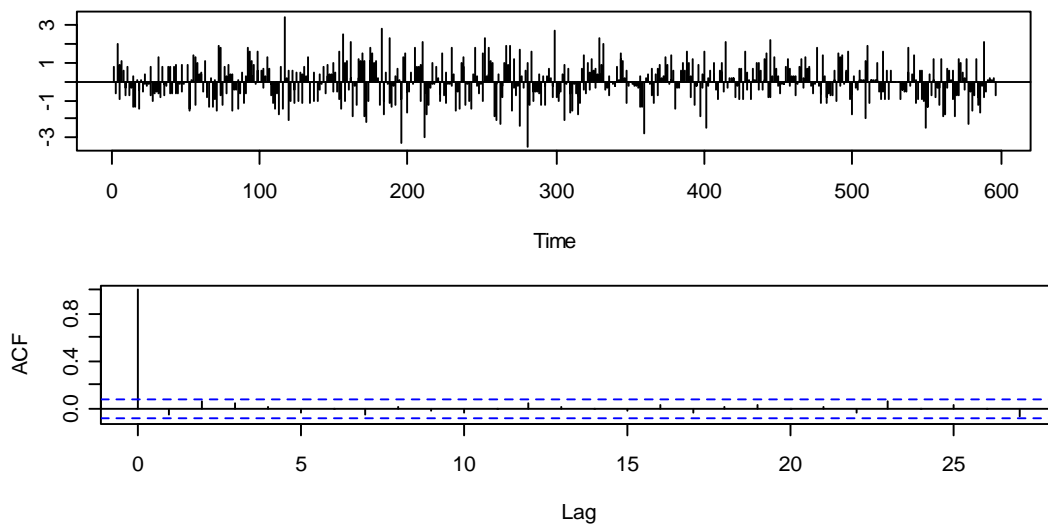
Figure 6.   Q-Q plot and histogram of standardized residuals
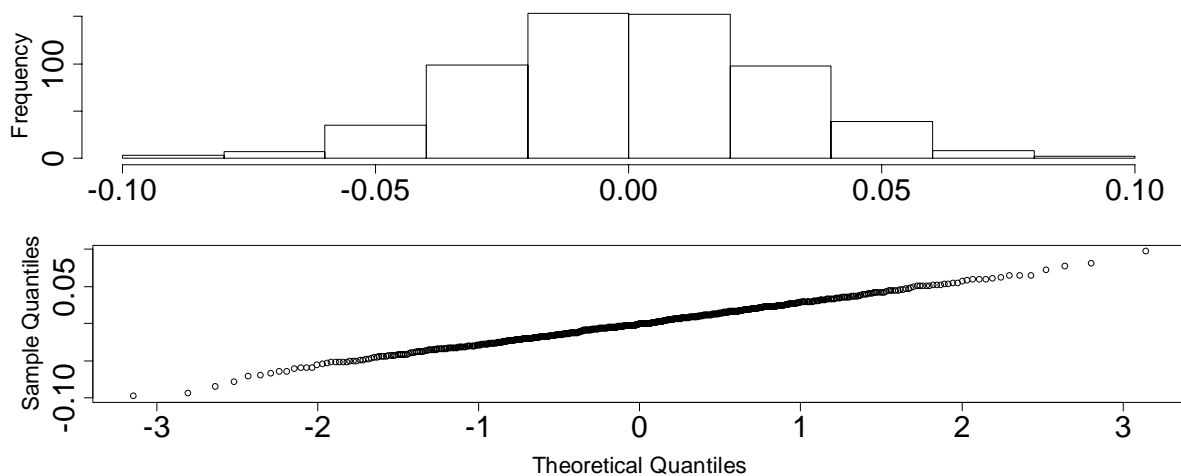
Figure 7. Smoothed periodogram of body temperature (n=596) showing no dominant peaks for L =3, 5, 7, 9, 11, 13, 15, 17, and 19 at 10% Tapering
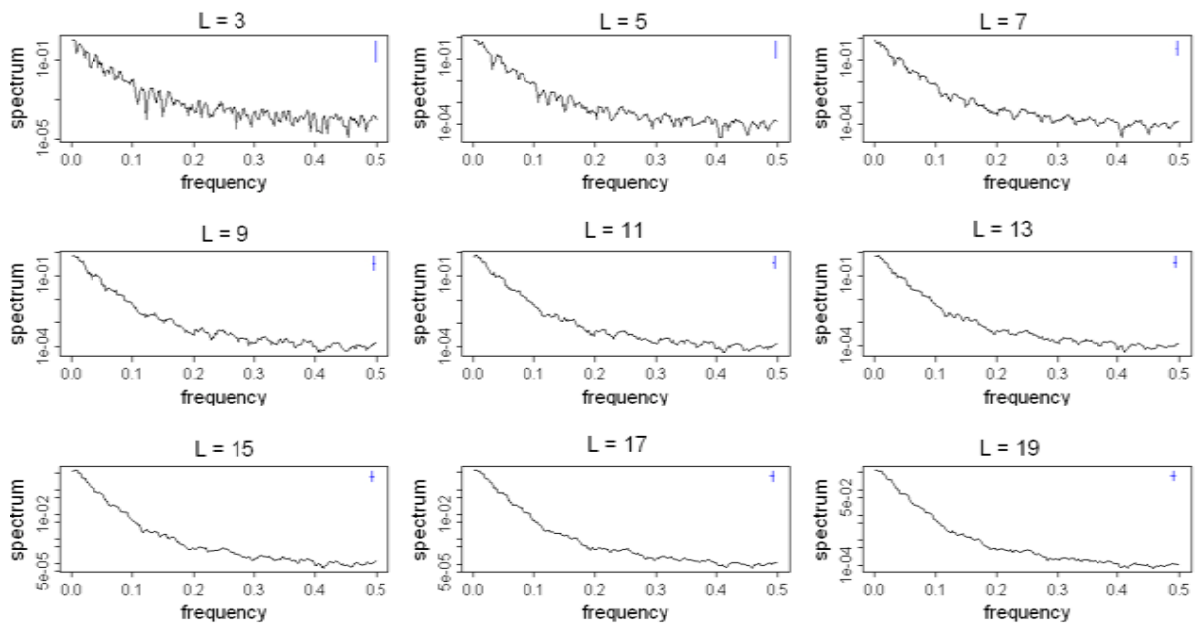
Figure 8. Coherence function between the Tb and Ta; L = 15, n = 596, and α = 0.05.
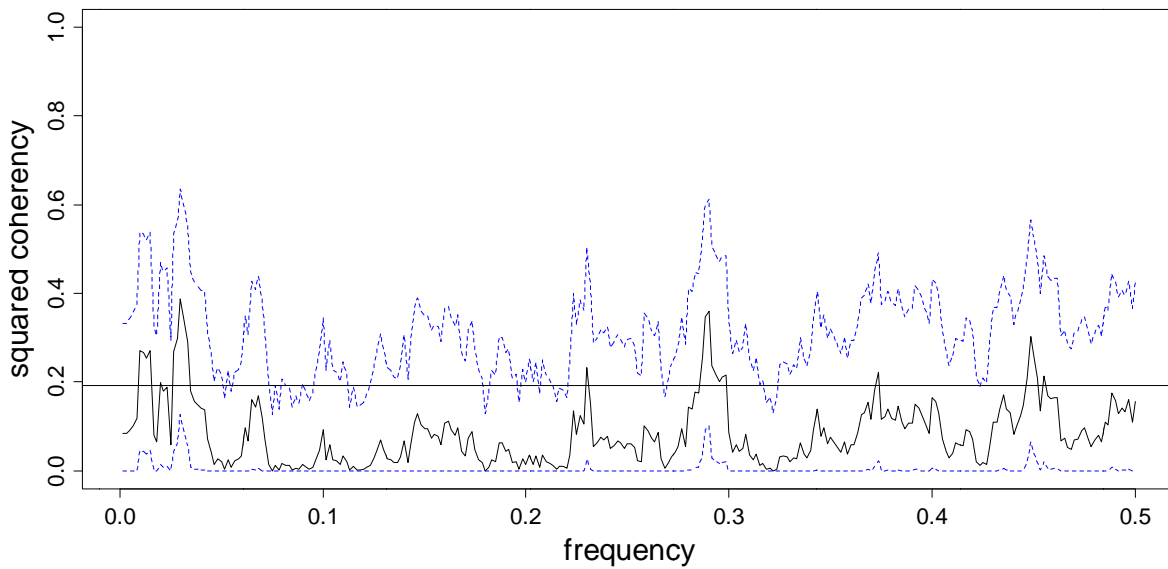
Figure 9. Histogram and Q-Q plot of residuals for Final Transfer Function Model
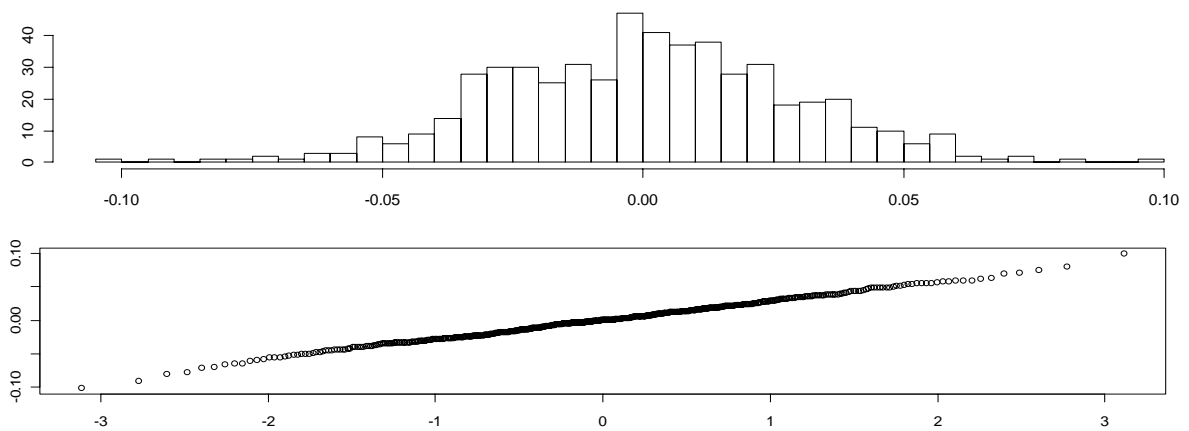
Table 1. Parameter estimates for regression with autocorrelated errors

Coefficient	ar1	ar2	Intercept	Ta
Estimate	1.5583	-0.5869	38.8464	-0.0034
S.E.	0.0331	0.0331	0.1024	0.0039
P-Value	0.0000	0.0000	0.0000	**0.38**

Table 2. Cross correlation between de-trended Tb and Ta

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

39   -0.0006304    -.04357      |              .*| .              |
40    0.00008158    0.00564     |              . | .              |
41   -0.0009526    -.06583      |              .*| .              |
42   -0.0003113    -.02151      |              . | .              |
43   -0.0008705    -.06016      |              .*| .              |
44   -0.0002159    -.01492      |              . | .              |
45   -0.0002057    -.01421      |              . | .              |
46   -0.0006942    -.04798      |              .*| .              |
47   -0.0001080    -.00747      |              . | .              |
48    0.00020610    0.01424     |              . | .              |
49    0.00040051    0.02768     |              . |*.              |
50   -0.0003572    -.02469      |              . | .              |
51    0.00075577    0.05223     |              . |*.              |
52    0.0012420     0.08583     |              . |**              |
53    0.00067019    0.04632     |              . |*.              |
54   -0.0001194    -.00825      |              . | .              |
55   -0.0005982    -.04134      |              .*| .              |
56    0.00014093    0.00974     |              . | .              |
57    0.00037065    0.02561     |              . |*.              |
58    0.00029564    0.02043     |              . | .              |
59    0.00007774    0.00537     |              . | .              |
60   -0.0001584    -.01095      |              . | .              |
```

Table 3. ACF of residuals of the proposed transfer function model

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	0.026734	1.00000	| |******************** |	0
1	0.025997	0.97242	| . |******************** |	0.042954
2	0.024616	0.92076	| . |****************** |	0.073036
3	0.022681	0.84837	| . |***************** |	0.091993
4	0.020554	0.76882	| . |*************** |	0.105445
5	0.018331	0.68568	| . |************** |	0.115324
6	0.016227	0.60698	| . |************ |	0.122616
7	0.014227	0.53217	| . |*********** |	0.128039
8	0.012446	0.46556	| . |********* |	0.132057
9	0.010812	0.40444	| . |******** |	0.135051
10	0.0094214	0.35241	| . |****** |	0.137268
11	0.0082018	0.30679	| . |****** |	0.138927
12	0.0071657	0.26803	| . |*****. |	0.140172
13	0.0061833	0.23129	| . |*****. |	0.141114
14	0.0053341	0.19952	| . |**** . |	0.141812
15	0.0045533	0.17032	| . |*** . |	0.142329
16	0.0039297	0.14699	| . |*** . |	0.142704
17	0.0033737	0.12619	| . |*** . |	0.142983
18	0.0029456	0.11018	| . |** . |	0.143189
19	0.0025372	0.09490	| . |** . |	0.143345
20	0.0021631	0.08091	| . |** . |	0.143461

Table 4. PACF of residuals of the proposed transfer function model

Lag	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
1	0.97242	| . |******************* |
2	-0.45649	| *********| . |
3	-0.27455	| *****| . |
4	0.07130	| . |*. |
5	-0.01011	| . | . |
6	0.06392	| . |*. |
7	-0.03921	| .*| . |
8	0.03425	| . |*. |
9	-0.03893	| .*| . |
10	0.05250	| . |*. |
11	-0.00716	| . | . |
12	-0.01152	| . | . |
13	-0.07565	| **| . |
14	0.04685	| . |*. |
15	0.00530	| . | . |
16	0.04875	| . |*. |
17	-0.03285	| .*| . |
18	0.01379	| . | . |
19	-0.04268	| .*| . |
20	-0.02329	| . | . |

Table 5.   Parameter estimates for Transfer Function Model

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
AR1,1	1.51017	0.04338	34.81	**<.0001**	1	bt_detrend	0
AR1,2	-0.46427	0.07603	-6.11	**<.0001**	2	bt_detrend	0
AR1,3	-0.08531	0.04336	-1.97	0.0497	3	bt_detrend	0
NUM1	0.0023747	0.0013038	1.82	0.0691	0	at_detrend	52
NUM1,1	0.0011943	0.0009570	1.25	0.2126	1	at_detrend	52
DEN1,1	0.93700	0.03982	23.53	**<.0001**	2	at_detrend	52

Table 6.   ACF of Residuals for Final Transfer Function Model

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	0.00083266	1.00000	| |******************** |	0
1	-2.47E-7	-.00030	| . | . |	0.042954
2	1.57829E-6	0.00190	| . | . |	0.042954
3	0.00002784	0.03343	| . |*. |	0.042954
4	-0.0000148	-.01783	| . | . |	0.043002
5	-0.0000204	-.02450	| . | . |	0.043015
6	-5.393E-6	-.00648	| . | . |	0.043041
7	-0.0000220	-.02648	| .*| . |	0.043043
8	0.00001711	0.02055	| . | . |	0.043073
9	-5.1548E-6	-.00619	| . | . |	0.043091