

# Long Range Dependency of Apparent Temperature in Birnin Kebbi

Babayemi A. W.<sup>1,\*</sup>, Onwuka G. I.<sup>1</sup>, Olorunpomi O. T.<sup>2</sup>

<sup>1</sup>Department of Mathematics, Kebbi State University of Science and Technology, Aliero, Nigeria

<sup>2</sup>Department of Mathematics, Nigeria Police Academy, Wudil, Nigeria

**Abstract** This research examined the long range dependency comportment of apparent temperature in BirninKebbi metropolis, using hourly data for a period of two years from 2013 to 2014. We found that the data points are highly correlated, there is a sense of return of a particular characteristic after some hours and this depicts seasonality traits and fractional integrated d series suspected. Autoregressive Fractional Integrated Moving Average (ARFIMA) modelrevealed that the series for the periods investigated is fractionally integrated with high-lag correlation structure,  $d=0.495$ . The long range dependency which describes the number of days for the return characteristics of persistence of the shock in a very long time period after it occurs was estimated to be three hundred and twelve days (312) days approximately a year; thus, once a shock is felt on apparent temperature, it will go on for about a year before the pattern or structure will be repeated.

**Keywords** Apparent temperature, ARFIMA, High-lag correlation structure, Long range dependency

## 1. Background to the Study

Apparent temperature allows for the quantitative assessment of heat stress and is used to determine the limit of heat exposure in order to prevent heat related illnesses. The Heat Index was originally developed by Steadman (2011) as an assessment for sultriness, but commonly referred to as the apparent temperature. Parameters involved in determining apparent temperature are water vapor pressure, surface area of skin, significant diameter of a human, clothing, core temperature, activity, effective wind speed, clothing resistance to heat transfer, radiation to and from the skin's surface, sweating rate, ventilation rate, skin resistance to heat transfer, and surface resistance to moisture transfer. The heat index is the measure of how hot it really feels when relative humidity is factored with the actual air temperature (Rothfus, 1990; Mendelsohn, 2003; Epstein, 2006; Robert, et al. 2010; Anderson et al. 2013; NOAA, 2013; Qu, et al. 2016). In estimation of ARIMA (p,d,q) model the researchers often peg the value of d as an integer value most especially between 0 and 2, however in estimating d in practice it is not always an integer value there are often at times that the value of d lies between -0.5 and 0.5. Many time series exhibit too much long range dependence to be classified as either I(0) or I(1) but do not belong to either of

them. In the recent, majority of these long range dependence models are often called long memory models are normally taken care of by the Autoregressive Fractional Integrated Moving Average (ARFIMA) model. The series that exhibits long memory process has autocorrelation function that damps hyperbolically more slowly than the geometric damping exhibited by "short memory" autoregressive moving-average, ARMA(p,q) process, which is predictable at long horizons(Granger, et al. 1980; Hosking, 1981; Ding, et al. 1993; Taqu, et al. 1995; Baillie, 2006; Robinson, 2003; Agostinelli, et al. 2003; Larranga, 2002; Nesrin, 2006; Haldrup, et al 2007; Baum, 2013; Contreras-Reyes, et al. 2013). Prior to this investigation on long memory of heat index in BIRNIN KEBBI metropolis, there have been previous studies involving heat in other fields such as Agric, Engineering, Physics, Chemistry or Biology in KEBBI State but little or no enquiry has been carried out on long memory effect of apparent temperature in KEBBI using time series approach. Over the years, there tends to be a severe heat during a particular period each year in the investigated area. This means we have higher values of heat index at a particular point in time every year. The research aims at investigating the long range dependency characteristics of apparent temperature in BIRNIN KEBBI metropolis with a view to determining: (i) the autocorrelation pattern for the long range dependency of the data points for the research; (ii) fractional value of d for the data points; and (iii) interpret the long memory characteristics.

\* Corresponding author:

whafsat@gmail.com (Babayemi A. W.)

Published online at <http://journal.sapub.org/statistics>

Copyright © 2017 Scientific & Academic Publishing. All Rights Reserved

## 2. Methodology

We adopted time series methodology to long memory characteristic of apparent temperature in BIRNIN KEBBI metropolis. It deals with hourly data of heat index collected over a period of two years from 2013 to 2014, and the set of data would be subjected to statistical testing procedures to confirm if the data are truly long memory, integrated, or stationary processes. It is expected to add to the existing literature on long memory behavior in heat index and will tend to help the metrology or practitioners in field of hydrology or dendrology. The data are secondary and are collected from Energy Research center BIRNIN KEBBI.

### 2.1. ARMA Process

An autoregressive moving average (ARMA) process consists of both autoregressive and moving average terms. If the process has terms from both an AR(p) and MA(q) process, then the process is called ARMA(p, q) and can be expressed as:

$$\phi(B)(y_t - \mu) = \theta(B)\varepsilon_t \quad (1)$$

where  $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$  is a white noise distribution term, so that for all  $t$ ,  $E(\varepsilon_t) = 0$ ;  $Var(\varepsilon_t) = \sigma_\varepsilon^2$  and  $\varepsilon_t$  are serially uncorrelated. All ARMA models are stationary. If an observed time series is non-stationary (e.g. upward trend), it must be converted to stationary time series (e.g. by differencing) (Anderson, et al 1979; Pelgrin, 2011).

### 2.2. ARFIMA Process

ARFIMA process allows for long memory series to be fractionally integrated. The general form of fractional ARMA model or the ARFIMA (p,d,q) model is:

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \quad (2)$$

Where  $d$  denotes non-integer fractional differencing parameter,  $y_t$  is the dependent variable at time  $t$ ,  $\varepsilon_t$  is distributed normally with mean 0 and variance  $\sigma_\varepsilon^2$ ,

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  and  $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$  represents AR and MA components respectively;  $\Phi(B)$  and  $\Theta(B)$  have no common roots,

$B$  is the backward shift operator and  $(1-B)^{-d}$  is the fractional differencing operator given by the binomial expansion:  $(1-B)^{-d} = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} B^j = \sum_{j=0}^{\infty} \eta_j B^j$  for

large  $j$  is  $\eta_j \approx \frac{j^{d-1}}{\Gamma(d)}$ ;  $d \in (-0.5/0.5)$  and  $(\varepsilon_t)$  is a white noise sequence with zero mean and Innovation variance  $\sigma^2$ . The properties of an ARFIMA process area bridged in the

following theorem. Let  $y_t$  be an ARFIMA (p,d,q) process then:

- $y_t$  is stationary and invertible if  $d < 0.5$  and all the root of  $\phi(B) = 0$  lie outside the unit circle.
- $y_t$  is non-stationary if  $d \geq 0.5$  and all the root of  $\theta(B) = 0$  lie outside the unit circle.
- if  $-0.5 < d < 0.5$ , the autocovariance of  $y_t; \gamma_k = E(y_t y_{t-k}) \sim B \cdot k^{2d-1}$ , as  $k \rightarrow \infty$ ; where  $B$  is a function of  $d$ .

The autocovariance function of ARFIMA process decay hyperbolically to zero as  $k \rightarrow \infty$  in contrast to the faster geometric decay of a stationary ARMA process.

### 2.3. Portmanteau Test

It is a test used for investigating the presence of autocorrelation in time series, number of lags  $\mu$  and  $h$  are predetermined.

#### Hypothesis

$H_0: \rho_{\mu,1} = \dots = \rho_{\mu,h} = 0$  (all the lags correlating are zero)

$H_1: \rho_{\mu,1} \neq 0$  i.e for at least one  $i=1, \dots, h$  is tested i.e at least one lag with non zero correlation.

#### Test Statistic

$LB_h = T^2 \sum_{j=1}^h \left( \frac{1}{T-j} \right) \hat{\rho}_{\mu,j}^2$  where

$\hat{\rho}_{\mu,j}^2 = T^{-1} \sum_{t=j}^T \hat{\mu}_{is} \hat{\mu}_{tj}$

## 3. Analysis of Data and Results

Figure 3.1 reveals the embedded long range dependency characteristics of the hourly data points in KEBBI between 2013/14. The data are highly correlated, there is a sense of return of a particular characteristic after some hours and this depicts seasonality traits and fractional integrated  $d$  series suspected. There was a drop after the first seven thousand hours of which it now picks and falls after fifteen hours, also the graph point shows the tendency of high correlation between the set of data point which amount to the fact the series is stable over time.

The autocorrelation function of the first two thousand five hundred (2500) observations in the data set displayed a trait of seasonality as it can be seen in the figure above that for every lag multiple of twelve (600) there appeared trough and picks, very strong correlation was noticed.

The figure 3.3 revealed that when the number of the lag is double we have a longer turning point, very strong correlation was noticed.

The figure 3.4 shows that there is a very strong correlation among the data points and the sinusoidal movement imitates a seasonal trait.

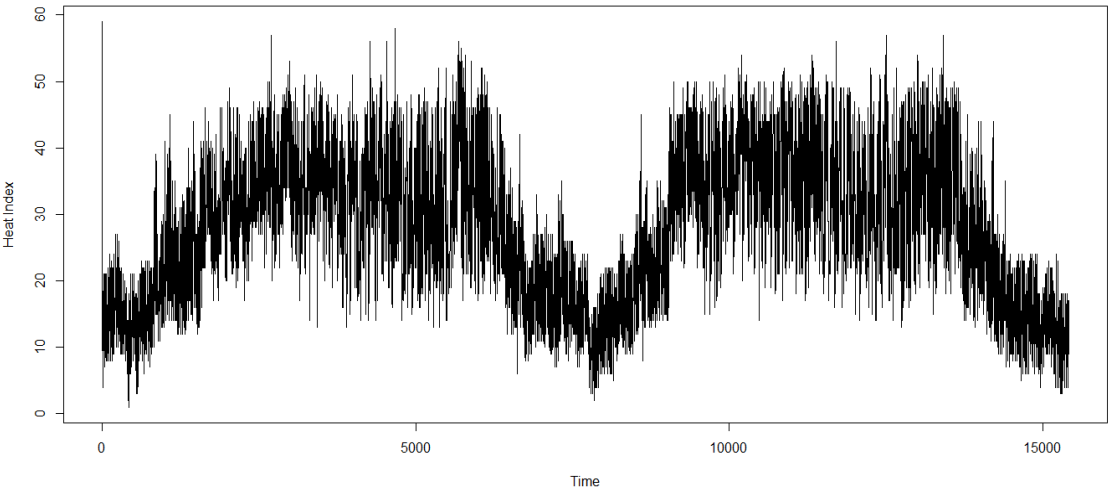


Figure 3.1. Visual Representation of Hourly Heat Index in KEBBI between 2013 and 2014

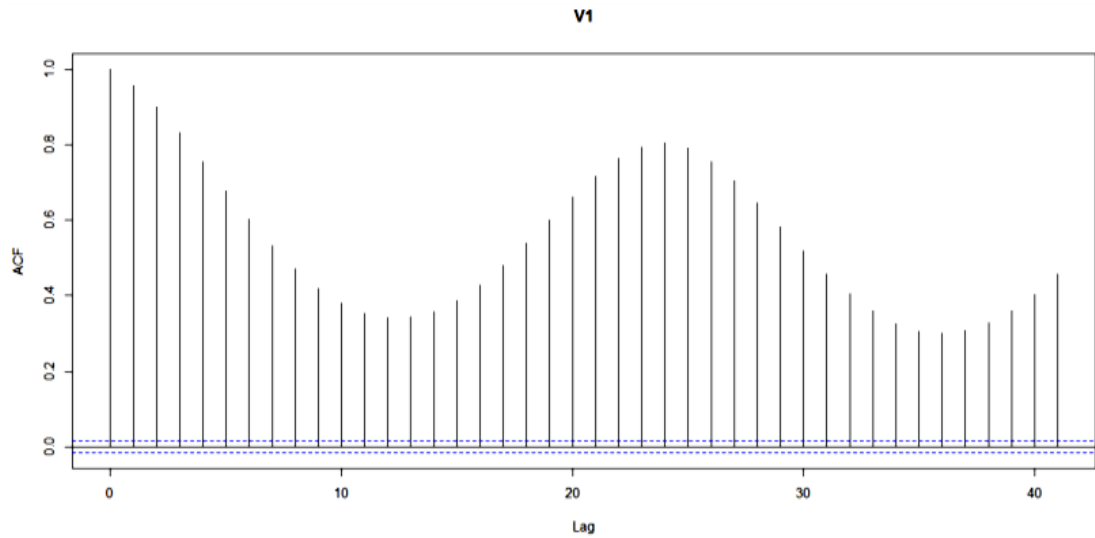


Figure 3.2. Visual Representation of First Two Thousand Five Hundred Hours Observation

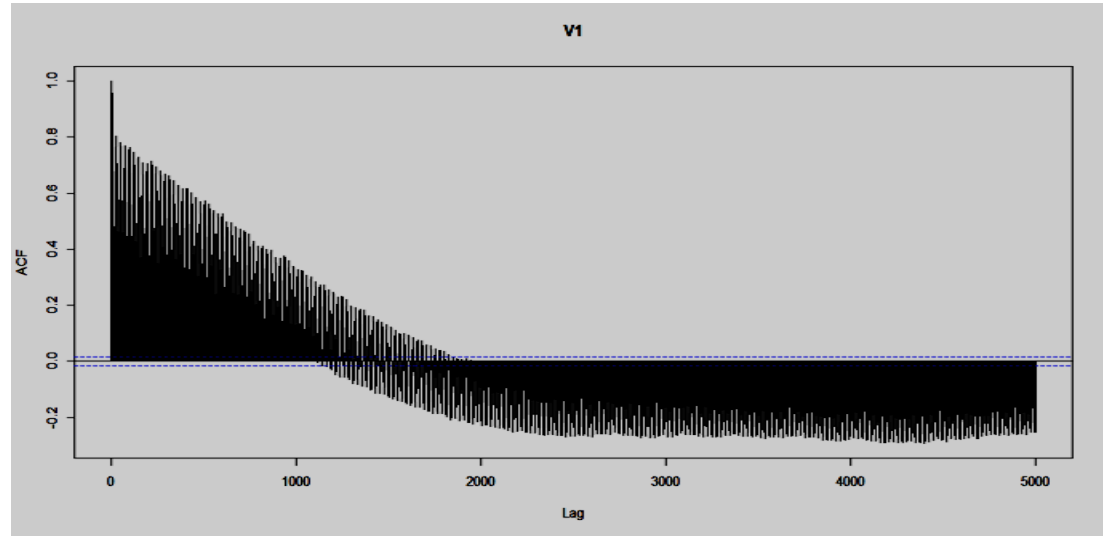
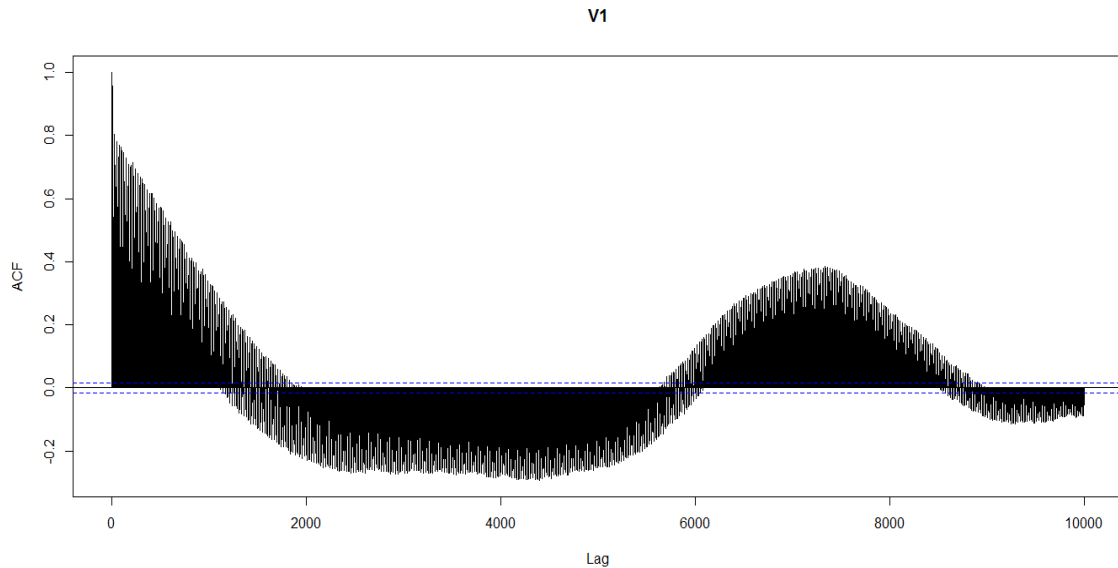


Figure 3.3. Visual Representation of Five Thousand Hours Observations



**Figure 3.4.** Visual Representation of Ten Thousand Hours Observations

**Table 3.1.** Descriptive Statistics for the Heat Index at Different Hours

NOB	Mean	Std	Median	Min.	Max.	Skewness	Kurtosis	Se
15418	27.72	11.31	28	1	59	0.12	0.87	0.09
2500	37.61	8.789	37	13.0	61	0.17	2.246	0.178
5000	42.59	11.99	43.00	5.0	55.6	15.6	67.2	15.6
10000	41.12	11.2	41.0	5.0	55.6	9.64	440.0	9.6

Table 3.1 shows the descriptive statistics of the heat index where the all samples are sub divided in three categories the first two thousand five hundred observations, five thousand observations and ten thousand observations; in order to examine the characteristics of the integrated order of at different sample hours. The minimum value of heat index for the hall sample is 1 and the maximum is 59. The average mean of the hall sample is 27.72.

16	0.045	0.026	828.96	0.000
17	0.050	0.030	867.43	0.000
18	0.054	0.032	912.44	0.000
19	0.059	0.035	967.04	0.000
20	0.065	0.038	1032.0	0.000
21	0.071	0.041	1108.9	0.000
...	...	...	...	...

**Table 3.2.** Portmanteau Test for Autocorrelation

S/N	ACF	PACF	Q-STAT	p-value
1	0.093	0.093	132.25	0.000
2	0.087	0.079	249.36	0.000
3	0.081	0.068	351.68	0.000
4	0.075	0.057	438.68	0.000
5	0.070	0.049	514.32	0.000
6	0.062	0.038	572.69	0.000
7	0.055	0.031	619.32	0.000
8	0.048	0.024	655.35	0.000
9	0.044	0.020	684.64	0.000
10	0.037	0.014	705.26	0.000
11	0.033	0.011	721.82	0.000
12	0.030	0.010	736.12	0.000
13	0.033	0.014	752.92	0.000
14	0.035	0.017	771.98	0.000
15	0.041	0.023	797.84	0.000

The table 3.2 above shows the results of the Portmanteau test at lag interval of 200 of which the first 21 of them are displayed. It reveals highly correlated set of data points which connotes non stationary level of the series.

**Table 3.3.** ADF Test for the Long Memory Process

Dickey-fuller Statistics	Lag-order	p-value
6.0258	24	0.01

The result from dickey fuller appears as if the series of apparent temperature is stationary unfortunately however, it is a long memory process where the value of the integrated order  $d$  in this case is not unit rather, a fractional value.

Having searched the ARFIMA models in an interval of  $[0,3]$  for both orders of  $p$  and  $q$  with fractional integrated  $d \in (-0.5,0.5)$  using Schwarz Information Criterion, the ARFIMA(3, 0.4950, 1) tend to be the most preferable model, the results are displayed in Table 3.4 below.

**Table 3.4.** Selection of ARFIMA Models

Model	Loglikelihood	d	Std error	p- value
ARFIMA(3,1)	13942.57	0.4950	0.000000172	0.00
ARFIMA(3,2)	13921.45	0.4991	0.000000492	0.00
ARFIMA(2,2)	13869.91	0.4998	0.000196000	0.00
ARFIMA(3,3)	12673.58	0.4999	0.011980000	0.00
ARFIMA(2,3)	12460.11	0.9960	0.000014500	0.00

Estimation of the ARFIMA(3, 0.4950, 1)

Variable	coefficient	Std Error	p-value
d	0.4950	0.000000172	0.0000
AR(3)	0.1858	0.008396000	0.0000
MA(1)	0.2536	0.008267000	0.0000

R-squared=0.864323; Loglikelihood=13942.57;  
 Inverted MA root=-0.25  
 Inverted AR Roots= 0.57, -0.29+50i, 0.29-50i

## 4. Discussion of Results

The research investigated the long range dependency characteristics of apparent temperature in Birnin Kebbi using ARFIMA model. The plots and the original autocorrelation structure are carefully examined and it is evident from the ADF test that the series under investigation does not contain unit root, neither can it be associated with a stationary to a series. It is noticed from the preliminary investigation the series possess long range dependency characteristics of  $d$  lying between 0 and 0.5 using Geweke, et al (2008) estimator and ARFIMA. ARFIMA model gives a better estimate. Long range dependency tend to reflect the persistence of shock in a very long time period after it occurs, meaning that; once a shocks is felt on the apparent temperature, it will go onto a very long time period for which the pattern or structure will be repeated. This can also be seen in the autocorrelation structure of all samples and when it is sub divided in to three categories, the average time of return is three hundred and twelve days which is approximately one year to all samples. Sample 2500 showed asinusoidal pattern in the structure, while samples 2500, 5000 and 10000 imitated the repetitive nature of shock persistence after a very long period of time. This shows that there is always a period where the apparent temperature is at pick or at lowest point. This research work prevent the misleading integrated of  $d = 1$  for ARIMA process.

## 5. Conclusions

The research was set out to investigate the long memory behaviour of apparent temperature in KEBBI metropolis, KEBBI State using hourly data for a period of two years (2013/14). Time series methodology of Autoregressive Fractional Integrated Moving Average model (ARFIMA) is used in the study. The results revealed that the series for the period investigated is highly correlated and fractionally

integrated with the value of  $d$  is 0.495 and the number of days for the return characteristics of the persistence of the shock is estimated to be 312 days approximately a year.

## REFERENCES

- [1] Agostinelli, C. and L. Bisaglia (2003). Robust estimation of ARFIMA process. Technical Report Università Cà Foscari di Venezia.
- [2] Anderson, O. D. and De Gooijer, J. G. (1979). \On discriminating between IMA(1,1) and ARMA(1,1) processes: Some extensions to a paper by Wichern," The Statistician, 28,(2), Pp. 119-133.
- [3] Baillie, R.T. (1996). Long memoery process and fractional integration in econometrics. Journal of Econometrics 73; Pp. 5-59.
- [4] Baum, C.F. (2013). ARFIMA (long memory) models, Boston college, www.bc.edu/Ec-c/s 2013/823/Ec823,s2013.nn08.sli de pgf.
- [5] Contreras-Reyes, J.E. and W. Palma (2013). Statistical Analysis of Autoregressive Fractionally Integrated Moving Average models in R. Springer-Verlag Berlin Heidelberg. arXiv: 1208.1728v1 [stat.CO] 8 Aug 2012; doi.10.1007/s00180-013-0408-7.
- [6] Dickey, D.A. and W.A Fuller (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root". Journal of the American Statistical Association, 74, Pp. 427-431.
- [7] Ding, Z., Granger, C.W.J. and R.F. Engle (1993). "A Long Memory Property of Stock Returns and a New Model," Journal of Empirical Finance, 1, Pp. 83-106.
- [8] Epstein, Y., and D. Moran (2006). Thermal comfort and the heat stress indices. Industrial Health, Pp. 388 – 398.
- [9] Pelgrin, F. (2011). Lecture 2: ARMA(p,q) model (part 3). University of Lausanne, \_Ecole des HEC Department of mathematics (IMEA-Nice).

- [10] Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*. 4: Pp.221-237.
- [11] Granger, C. W. J. and R. Joyeux (1980). "An introduction to long-memory time series models and fractional differencing". *Journal of Time Series Analysis*. 1: Pp.15–30. doi:10.1111/j.1467-9892.1980.tb00297.x.
- [12] Haldrup, N. and M. Nielsen (2007). Estimation of fractional integration in the presence of data noise. *Computational statistical and data analysis*, 51: Pp.3100-3114.
- [13] Hosking, J. R. M. (1981). "Fractional differencing". *Biometrika*. 68 (1): Pp.165–176. doi:10.1093/biomet/68.1.165.
- [14] Larranga, M. (2002). Thermal Standards and Measurement Techniques. In D. Anna, *The Occupational Environment: Its Evaluation, Control, and Management*. Fairfax: American Industrial Hygiene Association Pp. 918-951.
- [15] Mendelsohn, R. 2003. "Appendix XI: The impact of climate change on energy expenditures in California." *Global Climate Change and California: Potential Implications for Ecosystems, Health, and the Economy*, C. Thomas and R. Howard, Eds., Sacramento, California (2003).
- [16] National Oceanic Atmospheric Association (2013). National Weather Service Heat Index. Heat Index. NOAA <http://www.nws.noaa.gov/os/heat/images/heatindex.png>.
- [17] Nesrin, A. (2006). Long Memory Analysis of USD/TRL Exchange Rate, *World Academy of Science, Engineering and Technology International, Journal of Social, Human Science and Engineering*, 3(7): Pp.1-3.
- [18] Robert, H.S and D.S. Stoffer (2010). *Time series and its Application with R*. Third Edition. Springer New York Dordrecht Heidelberg London. doi 10.1007/978-1-4419-7865-3.
- [19] Robinson, P. M. (2003). *Time Series With Long Memory*. Oxford University Press. ISBN 0-19-925729-9.
- [20] Rothfus, L.P (1990). The Heat Index "Equation" or, More Than You Ever Wanted to Know About Heat Index". Technical Attachment, 1990. Available online: <http://www.srh.noaa.gov/images/ffc/pdf/tahtindx.PDF> (Accessed on 15 November 2015).
- [21] Steadman, R.G (2011), The assessment of sultriness. Part 1: A temperature-humidity index based on human physiology and clothing science. *J. Appl. Meteor.*
- [22] Taqqu, M. S.; Teverovsky, V. and W. Willinger (1995). "Estimators for Long-Range Dependence: An Empirical Study". *Fractals*. 3 (4): Pp.785–798. doi:10.1142/S0218348X95000692.
- [23] Qu, L., Chen, J., Dang, G, Jiang, S., Li, L., Guo, J. et al (2016). Heat Wave reduces Ecosystem Carbonsink Strength in a Eurasian Meadow Steppe. *Environ Res*. 144: 39-48. doi:10.1016/j.envres.2015.09.004.