Scientific Research Journal *Vol. 6 No. 2, 31-43, 2009*

A Study of Climate Data Using Least Square Method and the Fast Fourier Transform

Masriah Hj Awang[1,4], Zainazlan Md Zain[2] and Nur Sa'aidah Ismail[3]

[1]Faculty of Computer and Mathematical Sciences
[2]Faculty of Electrical Engineering
Universiti Teknologi MARA (UiTM), Malaysia
[3]Bahagian Matematik, Universiti Teknologi MARA (UiTM), Kedah
[4]Email: masriah@tmsk.uitm.edu.my

ABSTRACT

Energy utilization in buildings continues to increase as the quality of life increases. Buildings are built in an environment in which the climate surrounding a building is a factor influencing the energy requirements for the building services. The higher the thermal stress due to external conditions, the higher the energy required to provide consistent building services. This paper discusses the different types of climate analyses for Subang. The climate data has been calculated using averaged hourly values per month. The least squares method and fast Fourier transform have been used to explore the data further and elucidate climatic data. The climatic data collected and presented include temperature distribution, solar radiation, relative humidity distribution, rainfall distribution, wind-speed distribution and pressure distribution were presented. The least square polynomial of degree four and ten were chosen to represent the climate data. The least square error and the norm of the residual for these two polynomials were the smallest obtained amongst the other polynomials. The coefficients of determination were also calculated. The Fast Fourier Transform (FFT) from the MATLAB toolbox was used to evaluate patterns within the climate data. The FFT shows the Fourier coefficient on the complex plane. These studies reveal the climate patterns that need to be considered for optimum energy utilization in buildings.

ISSN 1675-7009

Keywords: *Climate, least squares method, Fast Fourier Transform, energy utilization*

Introduction

Generally, people are concerned about global climate change and its effect on their life. Due to this, there are an increasing number of studies and analyses of climate data. The analysts want to determine whether the changes indicated by climate data has any value and if it indicates the imminence of any changes to global life. Climate data can be in the form of temperature, solar radiation, rain, relative humidity, wind, pressure and other phenomena. In this study, analysis of the temperature, solar radiation, rainfall, relative humidity, wind-speed and pressure for Subang has been performed.

A major challenge when considering climate data is the approach (es) taken to analyze and interpret it thus elucidating any possible hidden information. Information visualization allows people to extricate important information and thus focus on items of greater interest. Approximation methods are helpful in that they provide a means to display as much of the necessary information as possible with respect to climate data distributions and patterns. There are patterns in the climate conditions and the ability to recognize these patterns and understand their significance will enable efficient management of the climate modification in a building, such as air conditioning. A building is a climate modifier and the energy utilized in a building is dependent on the climate conditions surrounding it. This work presented in this paper has three objectives. The first objective is to analyze a selection of climate parameters relating to Subang. The second objective is to perform time domain analysis using the least squares method. The third objective is to analyze the climate data using frequency domain analysis. All the aforementioned analysis has been performed using MATLAB®. Fourier transform enables efficient study of the repetitive or cyclic phenomena found in the climate data. It is hoped that this study will yield a general pattern by which to evaluate climate and improve efficiency in climate modifiers, as well as provide the opportunity for further studies on weather data and its derivatives.

Literature Review

The climate data parameters considered in this research are temperature, solar radiation, relative humidity, rainfall, wind-speed and pressure.

Temperature

Analyses of temperature data yields significant information regarding our life. Through a detailed study of temperature data, the climate change scenario (focusing primarily on temperature) and its implications on human, animal and plant life can be determined. The climate change scenario describes a future in which the gradual rising temperature accelerates the rate of warming, increasing the sea level, induces flooding in mountainous regions, causes droughts in coastal and agricultural areas and also contributes to the thinning of arctic ice [1]. Alnaser and Merzaa [2] stated that climate model should include monthly and yearly average temperature variations since air temperatures are an important factor in the estimation of the added energy ($E_{greenhouse}$) arising from the production of greenhouse gasses. The measurements of surface temperature were recorded daily at hundreds of locations for a prolonged period of time. The easiest way of obtaining daily scenario temperature data is to simply apply monthly scenario changes to a daily-observed station weather record [3]. For the purposes of synthesizing daily temperature data corresponding to a particular climate change scenario, a stochastic weather generator was applied [3]. This was done in order to observe and analyze the temperature data so that a warming pattern with respect to variations in the Earth’s temperature can be analyzed. It is proven that the year on year increase in average temperature is directly related to the increase in greenhouse gas emissions, primarily due to industrialization [4].

Historical climate change analysis is also based on temperature data and has been used as a point of reference for cyclic averaging of the Earth’s temperature variation [5]. The analysis of temperature data has been very useful in forecasting future temperature patterns as well as climate changes due to changes in temperature data. The climate model is capable of predicting and enabling summarization the potential consequences of climate change, such as the tendency for drought in some regions and higher rainfall in others, and also the potential for crop distribution to change due to hotter and drier conditions [5]. A historical weather record can be used to construct future analogue scenarios, which identify extreme conditions (drought or flood) and the consequent impacts [3].

Solar Radiation

The demand for solar radiation data is high since it is used in many aspects and life applications [6]. Information such as the availability, seasonal and geographical distribution of total, direct and diffuse radiation is important in the design and performance prediction of solar energy systems. In addition, it is pertinent to have performed atmospheric science studies and meteorological forecasting. There are several methods for analyzing and interpreting solar radiation data; one of the preferred methods uses statistics. An example of which is the work performed by Rehman *et al*. [7] who studied the spatial variation in Saudi Arabia using geostatistical techniques. Ianetz *et al*. [8] developed an analysis-based method that can be applied to characterize and differentiate between sites with regard to their potential for solar energy conversion systems. This statistical method was successfully applied to three sites located in southern Israel.

Al-Sanea *et al*. [9] applied the ASHRAE model to calculate the monthly average hourly global solar radiation on a horizontal surface in Riyadh. The ASHRAE model has also been used to determine the beam, diffuse and ground reflected solar radiation components independently. This solar model is beneficial in that it can be incorporated as an integral part of the entire heat-transfer analysis model, which contributes to the boundary conditions. Flor *et al*. [10] developed a method to calculate the incident total solar radiation for all types of surface both in open urban environments and inside buildings. This method is relevant for problems related to solar access, solar gains and day lighting. Solar radiation is the main contributor to heat gain in buildings, especially in residential buildings, where internal gains are very low. Solar Radiation data is essential in order to enable future solar radiation profile prediction. Solar radiation information is needed in a variety of applications and areas including agriculture, water resources, day lighting, architectural design and climate change studies [11]. Hence, solar radiation data plays an important role in many aspects of life.

Relative Humidity

Analysis of relative humidity data provides significant information with respect to our life. Alnaser and Merzaa [2] performed polynomial and sinusoidal regression on humidity data to yield a model capable of describing, with respect to inherent variations, average yearly and average monthly relative humidity.

Guan *et al*. [12] reported that there is a strong linear correlation between hourly variations in the dry bulb temperature and the relative humidity based on the statistical analysis of 10 years of historical hourly climatic data for Australia. It is observed that when dry bulb temperature increases, the relative humidity generally tends to decrease.

Rainfall

Several studies have been carried out focusing on rainfall data, specifically in terms of timing and volume. Ceballos *et al*. [13] analyzed the trends in rainfall and the behavior of dry spells in the Mediterranean. The work focused on the relationship between annual rainfall, the length and frequency of dry spells, and their annual variability. It was determined that there is no relationship between total annual rainfall, with an increment in intra-annual variability, and the pronounced occurrence of dry periods. In a study carried out in a typical Mediterranean semi-arid area in Spain, Lazaro *et al*. [14] analyzed a rainfall time-series (1967-1997) and evaluate the implications on vegetation using statistics. It was determined that all types of rainfall, in terms of volume, timing and intensity, have consequences on the composition, density and distribution of vegetation.

Wind-Speed

Wind data studies have been performed in various parts of the world. Nfaoui *et al*. [15] developed an autoregressive model to simulate hourly average wind speeds based on twelve years of hourly average wind speed data. The wind behavior was evaluated by comparing generated and real data. The validated model was used to generate reference monthly data to be used as a reference for yearly average wind speed data in Tangiers (Morocco). Blanchard and Desrochers [16] used the Box-Jenkins method to develop a new model to generate hourly wind speeds. It was shown that hourly wind speed data is sufficient to reproduce the main wind speed statistical characteristics such as monthly mean, standard deviation and high hourly autocorrelation.

Pressure

Several previous studies have been performed on pressure data. Guan *et al*. [12] reported that the final change in pressure due to changing temperatures is dependent on the product of new air densities and temperature. The study was based on the statistical analysis of 10 years of historical hourly climatic data in Australia. It has been observed that an increment in temperature will lead to a decrement in pressure and furthermore changes in pressure are comparatively small with respect to a large degree variation in temperature.

Methodology

The climate data for Subang was obtained from the Malaysian Meteorological Service Department. The climate profile data for Subang considered in this research had been recorded hourly for five years (1998-2002). Subang is located at a latitude of 3º07'N, a longitude of 101º33'E and an elevation of 16.5 m and represents a typical urban area in Malaysia. This research has used monthly averaged hourly data for the period 1998 to 2002.

The time domain analysis was performed using least squares method. The least squares approximation was used to evaluate the function, which best describes the climate data and is achieved by determining the error with respect to the sum of the square of the differences between the y-values on the approximating line and the actual climate y-values [17]. El-Metwally [18] used the coefficient of determination, r^2, as a measurement of the effectiveness of a new method to estimate global solar radiation based on meteorological data in Egypt. The higher the r^2 value better the fit and hence the better the method. Yuh-Shan *et al*. [19] analyzed the sorption isotherms of basic dyes on sugarcane dust by using the r^2 value as a means to perform error analysis.

Another analysis technique is the Fast Fourier Transform (FFT), which is a Discrete Fourier Transform (DFT) that reduces the number of computations needed from $2N^2$ to $2N \log_2 N$ [20]. The constants in the interpolating trigonometric polynomial can be calculated by using this technique. This method requires only O (m $\log_2$ m) multiplications and O (m $\log_2$ m) additions, provided m is chosen in an appropriate manner [17]. The FFT is an efficient way to compute the Discrete Fourier Transform if the number of points N is a multiple of two. If the number of points N is not a multiple of two, a transform can be performed on sets of points corresponding to the prime factors of N [20]. In general, there are two classifications of Fast Fourier Transform namely, decimation in frequency and decimation in time. The Fast Fourier Transform algorithm introduced by Cooley-Turkey constructs an output transform (decimation in time) after the input data has been rearranged in bit-reversed order [20]. This method allows the transformation from a time series of samples to a series of frequency domain samples. Other uses of Fast Fourier transform have been stated by Meirelles and Arruda [21] and include that Discrete Fourier Transform is suitable for the computation of the dynamic responses of damped linear systems to transient excitations with arbitrary initial conditions by using Fast Fourier Transform algorithms. Cheong and Haesun [22] stated that classification and identification of fingerprints can be determined by using Discrete Fourier transform.

Results and Discussions

The monthly averaged hourly temperature, solar radiation, relative humidity, rainfall, wind-speed and pressure data were calculated and plotted using MATLAB. The graphs were plotted for a 60 month, which started in January 1998 and ended in December 2002. Figures 1-3 present the monthly averaged hourly data for the six aforementioned climate parameters.

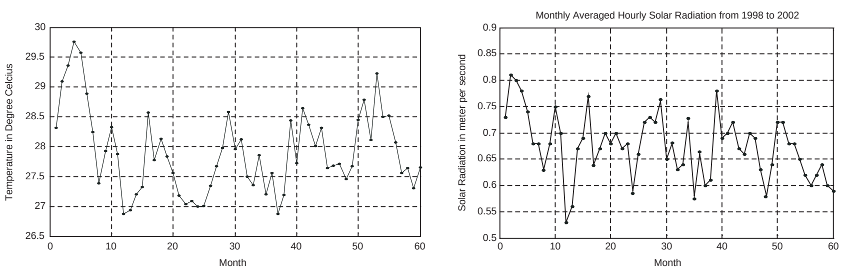

Figure 1: Monthly Averaged Hourly Temperature and Solar Radiation for Subang 1998-2002

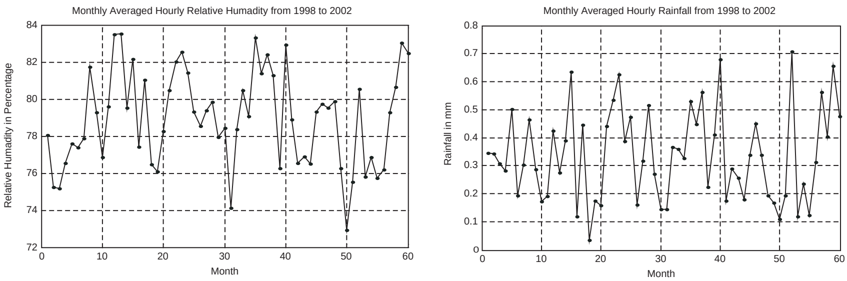

Figure 2: Monthly Averaged Hourly Relative Humidity and Rainfall for Subang 1998-2002

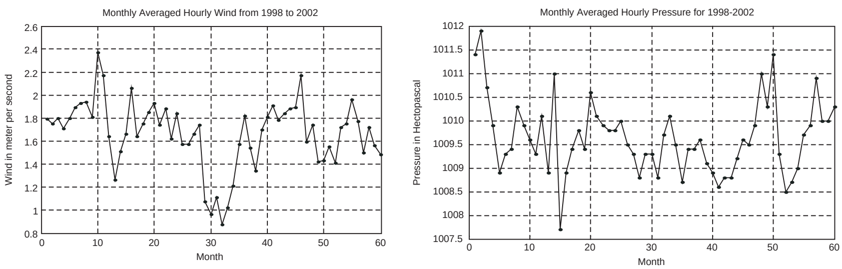

Figure 3: Monthly Averaged Hourly Wind-speed and Pressure for Subang 1998-2002

The time domain analysis was preformed using the least squares method. A number of polynomials of various degrees were tested and their quality of fit was evaluated with respect to the error, norm of residuals and coefficient of determination. The polynomials of degree four (P4) and ten (P10) have relatively smaller error than the other polynomials. Table 1 presents the results below:

Table 1: Least Squares Error (LSE), Norm of Residuals (NR) and Coefficient of Determination (CD) for Climate Parameters

	Temperature		Solar Radiation		Relative Humidity		Rainfall		Wind Speed		Pressure	
	P4	P10	P4	P10	P4	P10	P4	P10	P4	P10	P4	P10
LSE	16.7	12.0	0.2	0.1	319	254	1.4	1.4	3.7	2.1	28.3	23
NR	4.1	3.5	0.4	0.4	17.9	16	1.2	1.2	1.9	1.4	5.3	4.8
CD	0.4	0.6	0.3	0.4	0.2	0.4	0.1	0.1	0.3	0.6	0.2	0.4

The Fast Fourier Transform (FFT) of a sequence of input data results in two components that represent the complex frequency domain in rectangular form. The x-axis represents the real part of the transformed data (magnitude) and the y-axis represents the imaginary part of the transformed data phase). Figures 4-6 present the plots of the Fourier coefficients.

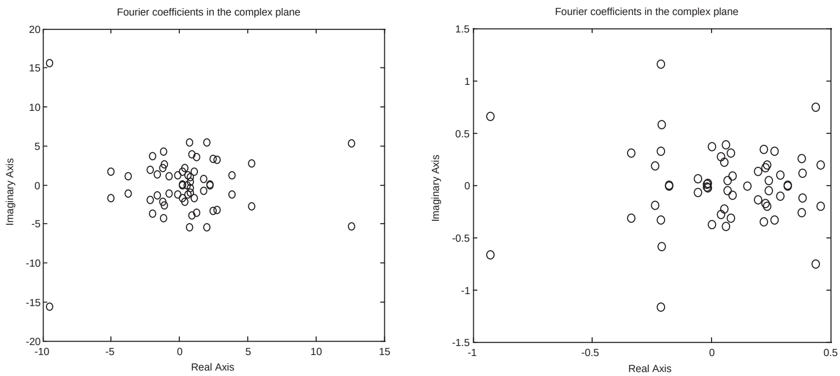

Figure 4: The Fourier Coefficients in Complex Plane for Temperature and Solar Radiation

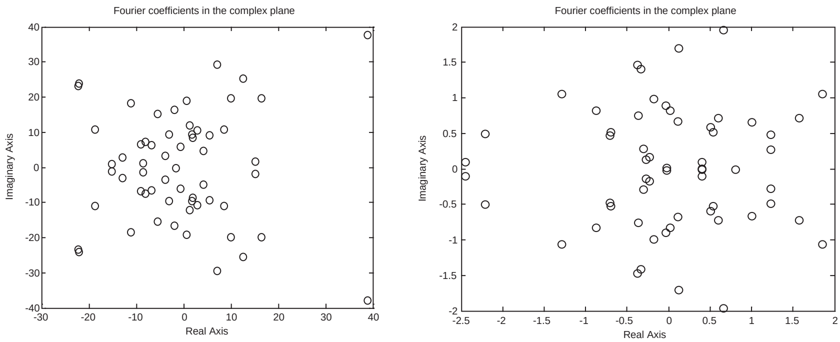

Figure 5: The Fourier Coefficients in Complex Plane for Relative Humidity and Rainfall

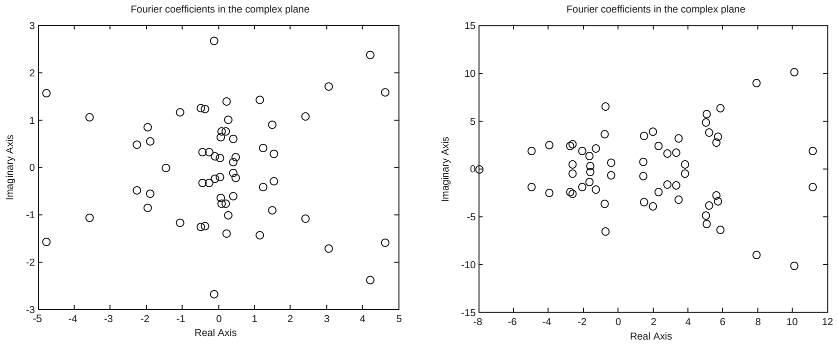

Figure 6: The Fourier Coefficients in Complex Plane for Wind-speed and Pressure

Conclusions

The objectives of this study have been achieved. All the parameters have been analyzed using fast Fourier Transform and it is evident that there are certain patterns in the climate parameters. Using the time domain, it can be observed that the period of the cycle is 11.8 months. The frequency domain has a cycle 0.9 cycles per month.

The longest cycle evident based upon the presented data is the annual cycle, if more data is available, i.e. extending beyond 5 years, it is possible that a cycle with longer period may be observed. Future work regarding the analysis of climate parameters includes the acquisition of more data from the Malaysian Meteorological Service Department and further in depth modeling and analysis. Data taken from higher altitudes and more frequently may elucidate more information and may help in evaluating wind harvesting potential, because the efficiency of wind is a function of height.

Acknowledgement

This research was supported by grant 600-IRDC/ST 5/3/1007 awarded by Research Management Institute. We thanked Universiti Teknologi MARA for the support.

References

[1] A. W. Shearer, 2005. Whether the weather: comments on 'An abrupt climate change scenario and its implications for United States national security', 38, pp. 445-463.

[2] W. E. Alnaser and M. K. Merzaa, 2005. Empirical Correlations for the climate data in the Kingdom of Bahrain, vol. 30, issue 15, pp. 2818-2830.

[3] E. M. Barrow and R. J. Lee, 2000. Climate Change and Environmental Assessment, *The Canadian Institute for Climate Studies*. Retrieved May 17, 2005.

[4] E. Wesker, 1996. Climate Change; a summary of some present knowledge and theories. Retrieved June11, 2005 from http://www.users/e_wesker/climate.html.

[5] Global Warming, 2000, United States Environmental Protection Agency, Retrieved August 18, 2005 from http://yosemite.epa.gov/oar/globalwarming.nsf/content/Climate.html

[6] D. G. S. Chuah and S. L. Lee, 1984. *Solar Radiation in Malaysia - A Study on Availability and Distribution of Solar Energy in Malaysia*. Oxford University Press, Oxford New York.

[7] S. Rehman and S. G. Ghori, 2000. Spatial estimation of global solar radiation using geo statistics, *Renewable Energy* 21, pp. 583-605.

[8] A. Ianetz, V. Lyubansky, I. Setter, G. E. Evseev and A. I. Kudish, 2000. A method for characterization and inter comparison of sites with regard to solar energy utilization by statistical analysis of heir solar radiation data as performed for three sites in the Israel Negev region, *Solar Energy* 69, no. 4, pp. 283-293.

[9] S. A. Al-Sanea, M. F. Zedan and S. A. Al-Ajlan, 2004. Adjustment factors for the ASHRAE clear-sky model based on solar-radiation measurements in Riyadh, *Applied Energy* 79, pp. 215-237.

[10] F. J. S. Flor, R. O. Cebolla, J. L. M. Fe'lix and S. A. Dominguez, 2005. Solar radiation calculation methodology for building exterior surfaces, *Solar Energy,* vol. 71, no. 4, pp. 217-224.

[11] H. C. Power, 2001. Estimating clear-sky irradiation from sunshine duration, *Solar Energy,* vol. 71, no. 4, pp. 217-224.

[12] L. Guan, J. Yang and J. M. Bell, 2006. Cross-correlations between weather variables in Australia, *Building and Environment,* pp. 1-22.

[13] A. Ceballos, J. Martinez-Fernandez and M. A. Luego-Ugidos, 2004. Analysis of rainfall trends and dry periods on a pluviometric gradient representative of Mediterranean climate in the Duero Basin, Spain, *Journal of Arid Environments*, 58, pp. 215-233.

[14] R. Lazaro, F. S. Rodrigo, F. Gutirrez and J. Puigdefabregas, 2001. Analysis of a 30-year rainfall record (1967-1997) in semi-arid SE Spain for implications on vegetation, *Journal of Arid Environments,* 48, pp. 373-395.

[15] H. Nfaoui, J. Buret and A. A. M. Sayigh, 1996. Stochastic simulation of hourly average wind speed sequences in Tangiers (Morocco), *Solar Energy*, 56, pp. 301-314.

[16] M. Blanchard and G. Desrochers, 1984. Generation of auto-correlated wind speeds for wind energy inversion system studies, *Journal of Solar Energy*, 33, pp. 571-579.

[17] J. D. Faires and R. Burden, 2003. *Numerical Methods*: 3rd Ed, Thomson Brooks/Cole, USA.

[18] M. El-Metwally, 2004. Simple new methods to estimate global solar radiation based on meteorological data in Egypt, *Atmospheric Research*, 69, pp. 217-239

[19] H. Yuh-Shan, *et al.,* 2005. Regression Analysis for the sorption isotherms of basic dyes on sugarcane dust, *Bioresource Technology*, 96, pp. 1285-1291.

[20] E. W. Weisstein, 1999. *Fast Fourier Transform*. Retrieved June 4, 2005.

[21] P. S. Meirelles and J. R. F. Arruda, 2005. Transient response with arbitrary initial conditions using Discrete Fourier Transform, *Journal of Sound and Vibration*, vol. 287, issue 3, pp. 525-543.

[22] H. P. Cheong and P. Haesun, 2005. Fingerprint classification using Fast Fourier Transform and nonlinear discriminant analysis, *Pattern Recognition*, 38, pp. 495-503.