International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

Temperature and Humidity Analysis using Data Logger of Data Acquisition System: An Approach

M.B. Waghmare[1], Dr. P. N. Chatur[2]
[1]*M.Tech Computer Science and Engineering-Head of Department Computer Science and Engineering, Government College of Engineering, Amravati, Maharashtra, India.*
[1]milind.btk@gmail.com
[2]chatur.prashant@gcoea.ac.in

***Abstract*-- Climate as the general weather conditions in the air above the earth such as wind, temperature, humidity and rain. Specially a particular area it said that there is a independence between those elements. In order to checked this connectivity a stastical analysis between relative humidity and temperature was carried out by analysing the correlation and linear regression as w ell as testing the independence between the relative humidity and temperature. Relative humidity, temperature measurement and monitoring are very useful for the purpose of its analysis about variation due to climate change and feature prediction of these values. This paper presents a simple technique in combining analog circuit and digital circuit theory together with programming technique to control the hardware. Data logger is a device used for storing measurement data. It can be used in wide range of embedded logging applications. Data loggers are computerized electronic devices that measure and record periodic environmental conditions. From this paper summarized the result of a statistical analysis of climate data by using this approach data collection, data processing, knowledge discovery.**

***Keywords*- Relative humidity and temperature, Remote sensing, Data Acquisition, Data logger.**

I. INTRODUCTION

The data logger is an invaluable tool to collect and analyze experimental data, having the ability to clearly present real time analysis with sensors and probes able to respond to parameters that are beyond the normal range available from the most traditional equipment [4]. The differences between various data loggers are based on the way that data is recorded and stored. Data logger is an electronic device that automatically records, scans and retrieves the data with high speed and greater efficiency during a test or measurement, at any part of the plant with time [4]. The type of information recorded is determined by the user i.e. whether temperature, relative humidity, light intensity, voltage, pressure or shock is to be recorded, therefore it can automatically measures electrical output from any type of transducer and log the value.

A data logger works with sensors to convert physical phenomena and stimuli into electronic signals such as voltage or current. These electronic signals are then converted into binary data. The binary data is then easily analyzed by software and stored on memory for post process analysis.

II. RELATED WORK

Many researchers have tried to use data mining technologies in areas related to meteorology and weather prediction. Kotsiantis et al. [12] predict daily average, maximum and minimum temperature for Patras city in Greek by using six different data mining methods: Feed-Forward Back Propagation (BP), k-Nearest Neighbor (KNN)[5], M5rules algorithm, linear least-squares regression (LR), Decision tree and instance based learning (IB3). They use four years period data [2002-2005] of temperature, relative humidity and rainfall. The results they obtained in this study were accurate in terms of Correlation Coefficient and Root Mean Square. The emphasis in [4] is on using DBSCAN (Density Based Spatial Clustering of Applications with Noise) clustering algorithm to categorize Turkey into regions according to climatic characteristics. They use the daily maximum and minimum temperature records between 1930 and 1996 from 258 stations. They draw that this type of data mining application can help meteorological to create faster forecast and decisions and provide more performance and reliability than any other methods. Data mining have been employed successfully to build a very important applications in the field of meteorology like predicting abnormal events like hurricanes, storms and river flood prediction [2][15]. These applications can maintain public safety and welfare. In this context, Zhang and Huang [17] propose a new framework to discover dynamic inter dimension association rules for local-scale weather prediction of Dallas City. The usefulness of applying association mining is to find a strong relation between severe conditions and the change tendencies of the measurements of the weather.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

The authors conclude with some predicates extracted from the obtained rules. Another contribution to detect severe events using data mining is by [14] and [16]. Peters at el. [16] used the volumetric radar data to detect storm events and classify them into four types: hail, heavy rain, tornadoes, and wind. Using data mining in meteorological application is not limited to prediction, but it also extends to participate in many important fields like water resource management [11] and air pollution management [13]. Mining techniques also can be applied to various types of data like weather images and radar maps extend to characteristic features extracted from this weather images can be used to represent various weather patterns [17].

III. Data Flow Diagram Of The Data Logger

The data flow diagram of embedded software that controls our data logger is showed in Figure 2.4. In this section, we briefly explain the operation of our data logger. In the first step, user needs to configure parameters for sensory data recording via USART. The parameters include time, date, sampling rate of data, and recording periods. Next, the MCU checks status of memory whether the memory is full or not connected. If there is no problem on the status of the memory, the MCU can continue to the next step. However, if there is a problem, the data logger will display error notification to its user on the LCD and will continue without saving data to the storage memory. Next, the MCU is put in standby mode and waits for measuring next data from sensor. The waiting period depends on predefined sampling rate. During the wait period, the MCU is put into power down mode to save its energy consumption. The MCU will be woken up to take the next measurement based on the interrupt from the built-in real time clock (RTC) of the MCU. After obtaining data from sensor, the MCU will put the sensing data into a buffer space. Next, it checks for the type of data whether it is a temperature or relative humidity data. Before storing the data to its corresponding sensor files using FAT16, the MCU attaches a timestamp to each data. Finally, the MCU rechecks the storage memory for its status before goes back to wait for the next measurement and changes its status to power down mode again.

A data logger (also data logger or data recorder) is an electronic device that records information over a period of time for later reference. You choose what information you need to gather and the data logger collects it. The data logger contains a sensor to receive the information and a computer chip to store it [10].

Then the information stored in the data logger can be transferred to computer for further analysis. This part is just like a central-processing unit in the computer. All the signals are sent to the data logger to be stored or transmitted to the computer for further analysis. The data logger has to be stand-alone so that remote sensing of relative humidity and temperature is possible. A micro controller is used for this purpose [12].

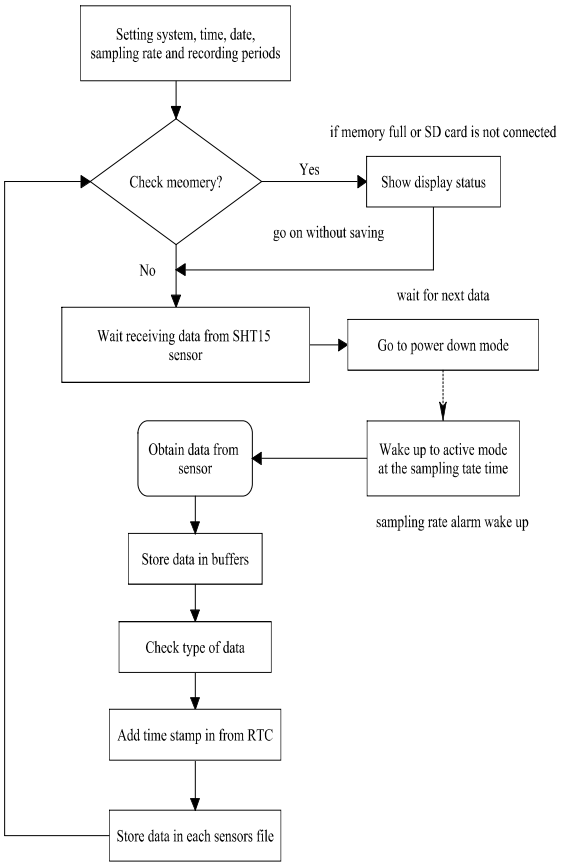

Fig1: Data Flow Diagram of the System

IV. System Structure

Inosculation between computers and instruments is an important direction of development in the field of instruments. Virtual instruments are computer-based apparatus, which are put into computers to implement various functions with common computer hardware and operating system.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

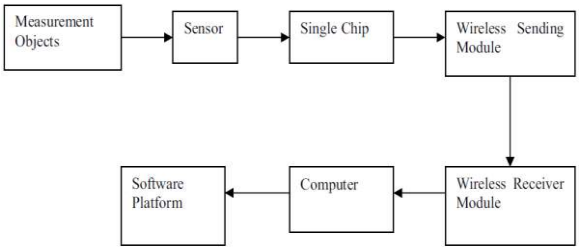

Fig2: Structure diagram of the system

Virtual instruments have several features

(1) Using common hardware as far as possible, the main differences between the instruments and software.
(2) It can give full play to the computer's capacity and have a powerful data-processing function so that we can create a powerful instrument.
(3) According to their own need, user can define and create all sorts of necessary instruments.

With the progressive development of meteorology, meteorological data measurement also enters the automated times. The traditional meteorological measurement system is huge in size and requires more hardware devices as well as influent factors like users, location, and space and so on. In addition, just like traditional instruments, when this kind of measurement systems are created, their function cannot be changed by users with their own need, because their functions have been defined by manufacturers. But virtual instruments are flexible, easy to develop and convenient to use, if we combine virtual instrument to traditional measurement system, use virtual instrument technology with wireless data transmission modules to implement Meteorological data acquisition and processing in addition that users can define and change the system function with their own need, it will save a lot of manpower and material, making the meteorological data acquisition and processing more easy and convenient.

A. Data Logging Versus Data Acquisition

The terms data logging and data acquisition are often used interchangeably. However, in a historical context they are quite different. A data logger is a data acquisition system, but a data acquisition system is not necessarily a data logger [9, 11].

Data loggers typically have slower sample rates. A maximum sample rate of 1 Hz may be considered to be very fast for a data logger, yet very slow for a typical data acquisition system [9, 11].

Data loggers are implicitly stand-alone devices, while typical data acquisition system must remain tethered to a computer to acquire data. This stand-alone aspect of data loggers implies on-board memory that is used to store acquired data. Sometimes this memory is very large to accommodate many days, or even months, of unattended recording. This memory may be battery-backed static random access memory, flash memory or EEPROM. Earlier data loggers used magnetic tape, punched paper tape, or directly viewable records such as "strip chart recorders" [9, 11].

Given the extended recording times of data loggers; they typically feature a time- and date-stamping mechanism to ensure that each recorded data value is associated with a date and time of acquisition. As such, data loggers typically employ built-in real-time clocks whose published drift can be an important consideration when choosing between data loggers [9, 11].

Data loggers range from simple single-channel input to complex multi-channel instruments. Typically, the simpler the device the less programming flexibility. Some more sophisticated instruments allow for cross-channel computations and alarms based on predetermined conditions. The newest of data loggers can serve web pages, allowing numerous people to monitor a system remotely [9, 11].

The unattended and remote nature of many data logger applications implies the need in some applications to operate from a DC power source, such as a battery. Solar power may be used to supplement these power sources. These constraints have generally led to ensure that the devices they market are extremely power efficient relative to computers. In many cases they are required to operate in harsh environmental conditions where computers will not function reliably [9, 11].

B. Data Acquisition Operations

After the MCU sent a request for data to SHT15 sensor, the sensor will measure the sensory data and signal a completion of measurement by transmitting four bits of low signal as shown in Figure 5 to the MCU. The actual measured data is transmitted next as 12 bits of data to MCU. The MCU will receive one bit at a time and continue to receive the rest of data until the whole 16-bit is collected. The 16-bit data will be calibrated and stored in float variable. Before saving the data into MicroSD card, the system will convert the float variable to ASCII string format and use this variable to write data into file and display it on LCD screen.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

V. Proposed Methodology Of Project

The project is completely designed in four different stages as shown in the fig. below:

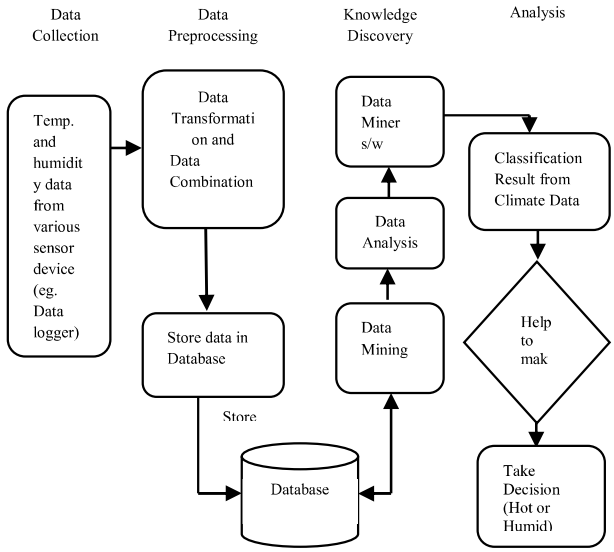

Fig3: Overall Design of Temperature and Humidity Data Analysis

The four main stages of the project are as follows:

1. Data Collection.
2. Data Preprocessing.
3. Knowledge Discovery.
4. Data Mining Technique.

A. Data Collection

The most important part while implementing any data related project is collection of proper data for the analysis using any technique (for eg. Data Mining). Thus, in this project we are collecting some amount of data using the 10 channel midi data logger system (GL220). As large amount of data is required for implementation of the project we are going to collect the requisite amount of data from various sources such as IMD (Indian Meteorological Department) or other sources.

B. Data Preprocessing

An important step in the data mining process is data preprocessing. One of the challenges that face the knowledge discovery process in climate data is poor data quality. For this reason we try to prepare our data carefully to obtain accurate and correct results. First we choose the most related attributes to our mining task.

For this purpose we neglect the year, temperature and humidity and time of the high temperature and humidity attributes. Then we try to fill the missing with appropriate values. In our data we have little missing, because we are working with climate data that is a form of time series, we must preserve the series smoothness and consistency. We use some effective method to fill missing values in the case of time series where the missed value is strongly related to its previous and next values. After filling the missing values we apply windowing operation on temperature attribute to create three lags (time frame) of temperature *lagt-1, lagt-2, lagt-3,* where lag is a past observations (days before) and ***t*** represent the day of the current class label (in the case of classification and prediction)**.**

C. Knowledge discovery

For knowledge extraction various data mining techniques are available such as Outlier Analysis, Clustering, Prediction and Classification and Association Rules. We are going to use Classification technique of data mining for our purpose, because Classification has been utilized in many meteorological applications; for example classifying to predict the weather on a particular day will be “sunny”, “rainy” or “cloudy”. Also it used widely in classifying geographical location based on its climate and classify weather conditions based on the agricultural crops suitable to cultivate on each climate. Classification includes methods that can produce useful results. We apply k-Nearest Neighbor classification techniques on our data to get the best results. Thus, we can classify the data class label into categorical. Classification task try to classify the data records into three classes hot (temperature is higher than 23 °C), warm (between 16 °C and 23 °C) or cold weather (bellow 16 °C).

D. Data Mining Technique

Technologies have been elaborated over the last few years, producing a huge amount of data. This huge raw data is difficult to analyze and understand. In this case clustering aim to improve the understanding of natural climate processes, to assess the quality of climate model results and to identify prevailing system features and their typical scales for specific atmospheric regimes. Clustering have been applied successfully in many meteorological application like determinate the precipitation weather type by finding the similarity between satellite cloud images, seasonal clustering [6] [7] and climatology [4].

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

In our project we use k-means clustering algorithm using k=4. K-means algorithm is the most popular clustering tool used in scientific and industrial applications. The name comes from representing each of k clusters by the mean (or weighted average) of its points, the so-called centroid. The centroid of a cluster is a point whose coordinates are the mean of the coordinates of all the points in the clusters [3]. Prediction is the most used data mining task in the field of meteorology. Data mining techniques provides with a level of confidence about the predicted solutions in terms of the consistency of prediction and in terms of the frequency of correct predictions [1]. Also it applied successfully to predict different weather elements like temperature and humidity. [8]. This is the most important part while implementing any of the data mining technique and for this purpose we are using midi data logger GL 220 system. This system provides temperature and humidity data in form of excel shit. The next important step in data mining is data preprocessing in climate data is poor quality. Thus data is to be pre-processed so as to remove the noisy data. For this project various data mining technique such as Outlier Analysis, Clustering Predication, Classification and Association rules are applied in statistical Data Miner software.

VI. Conclusion

In this different types algorithm is used for classification of time series climate data so as to predict climate as hot or humid based on temperature and humidity values. The result of the analyzed output would be compare with the actual values of temperature and humidity and accuracy of algorithm would be determined.

References

[1] Baboo S., and Shereef K., "*Applicability of Data Mining Techniques for Climate Prediction – A Survey Approach*," International Journal of Computer Science and Information Security, Vol. 8, No. 1, April 2010.

[2] Berkhin P., "*Survey of clustering data mining techniques*, Accrue Software, San Jose", CA, Tech. Rep., 2002.

[3] Bilgin T., and Çamurcu Y., "*A Data Mining Application on Air Temperature Database,*" Advances in Information Systems, Springer Berlin, Heidelberg, pp.68-76 .2004.

[4] Inniss, T. R., "*Seasonal clustering technique for time series data*". European Journal of Operational Research, 175, 376–384. 2006.

[5] Jan Z., Abrar M., Bashir S., Mirza A., "*Seasonal to Inter-annual Climate Prediction Using Data Mining KNN Technique*," Wireless Networks, Information Processing and Systems Communications in Computer and Information Science, pp. 40-51.2009.

[6] Kotsiantis S., Kostoulas A., Lykoudis S., Argiriou A., and Menagias K., "Using Data Mining Techniques for Estimating Minimum, Maximum and Average Daily Temperature Values

[7] Godfrey C. Onwubolu1, Petr Buryan, Sitaram Garimella, Visagaperuman Ramachandran,Viti Buadromo and Ajith Abraham "*Self-organizing data mining for weather Forecasting*" presented at IADIS European Conference Data Ming 2007.

[8] S. Kotsiantis, A. Kostoulas, S. Lykoudis, A. Argiriou, K. Menagias, "*A Hybrid Data Mining Technique for Estimating Mean Daily Temperature Values*".

[9] Dinelli, D., 1995: What weather stations can do. Landscape Manage., 34 (3), 6G.

[10] Peter Eredics "*Short-Term External Air Temperature Prediction for an Intelligent Greenhouse by Mining Climatic Time Series*" WISP 2009 • 6th IEEE International Symposium on Intelligent Signal Processing • 26–28 August, 2009 Budapest, Hungary

[11] Arman Kanooni, "*A survey of existing data mining techniques, methods and Guidelines within the context of enterprise data warehouse*" Master of Science Dissertation in Information Systems.

[12] Fix E, Hodges JL, Jr (1951) "*Discriminatory analysis, nonparametric discrimination*". USAF School of Aviation Medicine, Randolph Field, Tex., Project 21-49-004, Rept. 4, Contract AF41(128)-31, February 1951

[13] Tan P-N, Steinbach M, Kumar V (2006) Introduction to data mining. Pearson Addison-Wesley.

[14] Han E (1999) "*Text categorization using weight adjusted k-nearest neighbor classification*". PhD thesis, University of Minnesota, October 1999

[15] Cost S, Salzberg S (1993) A weighted nearest neighbor algorithm for learning with symbolic features. Mach Learn 10:57.78 (PEBLS: Parallel Examplar-Based Learning System)

[16] Cover T, Hart P (1967) "*Nearest neighbor pattern classification*". IEEE Trans Inform Theory 13(1):21–27

[17] Kuramochi M, Karypis G (2005) Gene Classification using Expression Profiles: A Feasibility Study. Int J Artif Intell Tools 14(4):641–660.