Tridimensional mapping model for the study of temperature and humidity behavior in data centers

Raúl D. Mejía [#1], Tatiana K. Montoya [#2], Raul J. Martelo [*3]

[#] Faculty of Engineering, University of La Guajira, Riohacha, Guajira, Colombia
[1] rdmeji@uniguajira.edu.co
[2] tmontoy@uniguajira.edu.co
[*] Faculty of Engineering, University of Cartagena, Cartagena, Bolivar, Colombia
[3] rmartelog1@unicartagena.edu.co

Abstract— **The objective of this study was to perform a three-dimensional mapping model for the study of temperature and humidity behavior in the data center. The above consists of the installation of hardware and software components that carry out processes to control the environmental conditions of the racks. In addition, a sensor positioning model was defined to design and implement a monitoring system taking into account the costs. It was verified that the temperature and humidity levels of the servers are below the 27° C threshold according to international standards. This is done through data mining techniques, to contribute to decision making and generate a strategic maintenance plan by the administrator.**

Keyword - Monitoring, sensor, rack, control system, data mining.

I. INTRODUCTION

Data centers are rooms that contain servers for storage and processing [1]. They are made up of buildings, facilities, telecommunications equipment, refrigeration and energy, which rapidly develop as an important part of the IT industry [2]. The growing multidisciplinary demand requires that this type of facilities that house high-performance computational machines have an appropriate balance and control of temperature and humidity levels, which is essential for operational efficiency in the protection of the equipment involved [3].

In the data centers an essential part to maintain stable environmental conditions are the cooling infrastructure, because it contributes significantly to energy consumption. With the above, the balance of temperature and humidity levels is achieved [4], however, in most cases, these high-performance computing facilities are not capable enough to handle additional cooling requirements as: recirculation or mixing of hot and cold air, in addition to little controlled humidity levels [5].

The coordination of the air conditioning units of the computer room (CRAC) and the supply of air temperature (SAT), represents a challenge of the logistics for the maintenance of the data centers; In aspects such as the speed of the fans, the energy consumption between the CRACs and the chillers decreases, while the temperature of the racks remains below the established thresholds. In addition, one of the biggest obstacles to overcome is the lack of simplified and computationally suitable models [6].

Among the works related to the development of the modeling of influential factors in the performance, maintenance and useful life of the data centers, is that of [7], who in their study used simulations to achieve the prediction of the behavior of heat flows in data centers with a sensitivity analysis, focused on the search for precision in the three-dimensional profile, and in which, with the help of a reduced-order modeling method, they manage to drastically reduce the execution time of the simulation (up to 600 times faster), maintaining the model's accuracy in an acceptable manner.

In addition, [6] proposed models that determine the efficiency of refrigeration processes, in order to improve the quality of these computing facilities and their thermal environment. A data center management and analysis scheme based on the dynamic model of rack entry temperature is used. The procedures for identifying the parameters of the model and its reduction are analyzed. An improved approach to thermal zone mapping is introduced through the grouping by the temperature of the model-based rack input. Another associated study is carried out by [5], in which they executed computer simulations based on a modeling of differential equations, in order to achieve a lower energy consumption and establish adequate flows of hot and cold air, to achieve greater use of refrigeration equipment.

Likewise, works such as that of [8] suggest the need to formulate administrative policies depending on the properties of each data center based on simple models that detail the thermal behavior.

Taking into account factors such as: The physical layout of the facilities, the electrical power consumption of the equipment and the cooling capacity. Another influential factor in the performance of data centers is humid, as denoted in [9], where a method is proposed to minimize energy consumption while maintaining humidity levels under acceptable ranges. Through a flexible control strategy that seeks the improvement of the ideal environmental conditions so that the equipment works properly. Taking into account the aforementioned investigations, a real-time system for temperature and humidity monitoring is required to obtain a modeling of the environmental conditions of the racks in a data center. The experimental development was based on a system of sensors that transmit the records to a computer that processes the data obtained and determines the appropriate model for the maintenance of the conditions in the data centers.

II. METHODOLOGY

A. Research instruments

To carry out this research, technical elements were defined as a network of sensors located in the data center racks, in order to perform the process of temperature and humidity measurements. In the software part, we used a web application based on a free license, such as PHP, JQuery, CSS and a MySql database; for the application of data mining techniques, SPSS (version 23) was used because it provides interpretations and graphs.

B. Scheme raised

For the development of this research, a model for the study of temperature and humidity behavior in the data center was proposed. In Figure 1, the phases and stages that make up the methodology are shown.
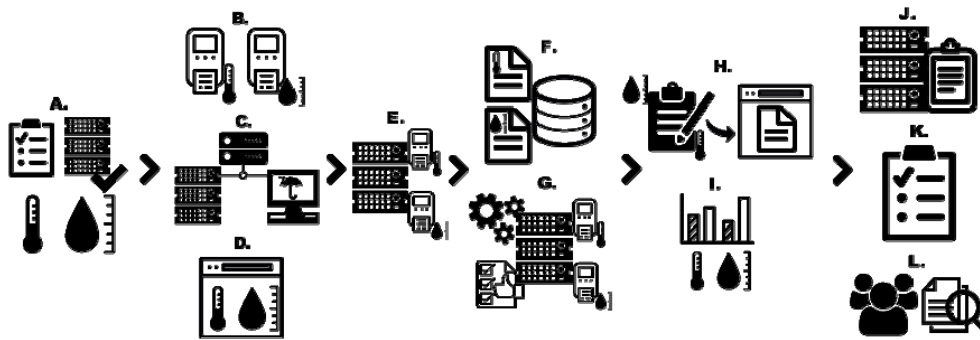

Fig. 1. Modeling methodology for three-dimensional mapping in the temperature and humidity analysis.

This model consists of six primordial phases:

1. Initial observation of data

It consisted of introductory steps prior to the beginning of the experimental process, for a recognition of the field of work and the object to be studied. It is identified with the letter A in Figure 1 and consists of: conducting technical visits to the data center, interviews with the personnel in charge to verify the configuration of the racks, finding out how the air conditioning works, ideal temperature and humidity, equipment prone to thermal faults and identify cooling failures in the system.

2. Selection of software and hardware tools for temperature and humidity measurement in the data center

It consists of the design of a hardware system for measuring the temperature and humidity of the data center at the entrance and exit of racks. This phase consists of stages B, C and D in Figure 1. In Stage B, sensors are selected to measure temperature and humidity based on observations to the data center, to verify thermal conditions. In Stage C, the acquisition, control and transmission of climate data hardware is chosen, in order to obtain data from each sensor, as well as allowing communication between the devices and the PC. In step D, the software is chosen for the recording and visualization of the data acquired by the sensors, in order to have a visualization of the records obtained in real time of the temperature and humidity variables.

3. Measurement of temperatures and humidity of each rack

This phase consists of real-time measurements in the data center. It is identified with the letter E in Figure 1, the sensors are placed in each rack in an equidistant manner, four temperature sensors in the front and four in the back to measure relative humidity, so that they are located in the central part of each rack. Initial tests are carried out with the proposed model and then based on the recommendations of the ASHRAE, to verify the behavior of the first.

4. Real-time recording of temperature and humidity

This phase is formed by steps F and G in Figure 1. In step F, continuous measurements are taken every 5 minutes for a week and recorded in the database, to know the thermal functioning of the data center. Stage G consists in carrying out a pilot test, in order to verify the correct functioning of the record in a continuous way of the temperatures and humidity entered in the database, for the subsequent analysis of the maximum and minimum ranges in each sensor.

5. Development of a three-dimensional dynamic thermal map with registered temperatures and humidity

This phase has as objective the elaboration of a thermal map; consists of stages H and I in Figure 1. Stage H consists of taking records associated with temperature and humidity in the database, so that a web application generates the necessary information for the creation of the mapping. In stage I, a thermal map with a graphical scale of the temperature in degrees centigrade and percentage of humidity associated with specific colors for the images is elaborated.

6. Analysis of resulting data on temperature and humidity measurements

It consists of stages J, K and L in Figure 1. In stage J, the temperature and humidity diagnosis of the data center was made, based on the measurements of the sensors to proceed with the evaluation, in accordance with the international standards of thermal conditions in the object of study. In stage K, verification processes were carried out to comply with the international standard of thermal conditions for the data center. In stage L, the information is analyzed through the use of data mining techniques (Correlation Matrix, Principal Components Analysis, Hierarchical Cluster Technique and Decision Tree), to identify patterns and trends based on the information recorded in the database. The information is classified according to the air conditioning that was in operation and initially with all the sensors.

III. RESULTS

Below is the execution of the proposed phases with their respective stages.

1. Initial observation of data

In stage A, technical visits were made to the place where the racks were located and specific points were analyzed for sensor installation, structured cabling assembly and conditions of the physical facilities that housed the data centers. In addition, the personnel in charge were interviewed to follow their recommendations and safety regulations.

2. Selection of software and hardware tools for temperature and humidity measurement in the data center

In stage B, the technical recommendations of the ASHRAE were taken into account, through observations to the data center for the measurement of thermal conditions in order to select suitable sensors. For temperature the DS18B20 was chosen because the signal is digital and in the case of humidity, the HIH-4030 sensor was chosen with a signal type similar to the previous one. In stage C, the appropriate hardware was chosen to obtain the data of each sensor, which was Arduino Leonardo, free hardware circuit suitable for connecting the sensors, achieving communication between the devices and a computer equipment. In stage D, the software environment for data recording is selected, which was the Processing development environment based on the library for communication with the Onewire protocol (OneWire.h), the library for the DS18B20 sensors (DallasTemperature.h), the library (DS2438) for the 1-Wore temperature-humidity sensors and the LCD Liquid Crystal.h display management library, in order for the Arduino controller to capture the temperature and humidity records. In the case of visualization of data in real time, free development such as HTML, PHP, JQuery, CSS and MySql are chosen for the development of a web application that shows the monitoring process in real time.

3. Measurement of temperatures and humidity of each rack

In stage E, the sensors are installed in each of the three racks to perform the appropriate measurements, the facilities have a cooling flow required for proper operation of the equipment, because the design allowed normal air flow. The above in order to control the humidity and temperature levels in an environment with absence of dust particles. In addition, it was required to study the distribution of the equipment in a data center, establishing cold zones of air intake to the equipment and hot zones for the exit of the superheated air. The three racks have sensors for measurements, this step is necessary to keep records of temperature and humidity that serve for modeling. Initially, the input measurement data for the proposed model (see Figure 2.a) were: HR (45%) and T (22.6° C) and are positioned horizontally. While in the ASHRAE model (see Figure 2.b), the sensors were positioned vertically to avoid possible anomalies obtained by applying the previous model.
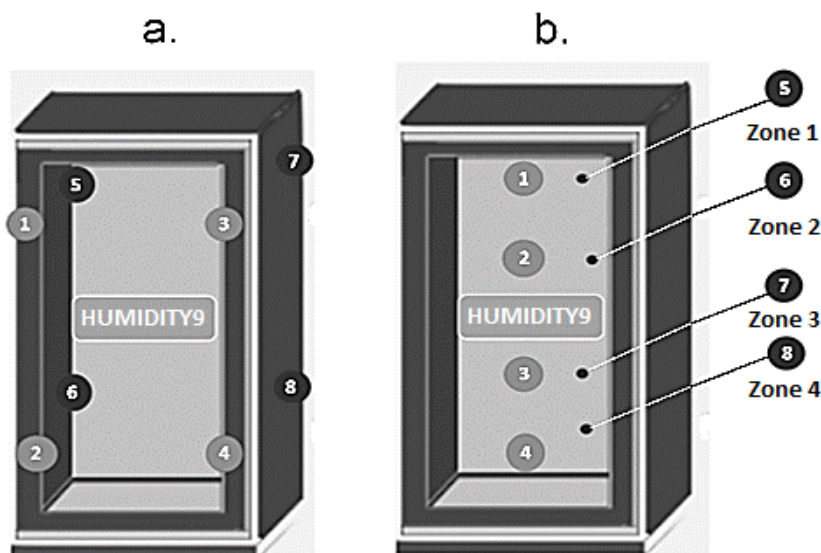
Raúl D. Mejía et al. / International Journal of Engineering and Technology (IJET)

Fig. 2: Positioning of the sensors in racks in model "a" (proposed) and "b" (ASHRAE)

4. Real-time recording of temperature and humidity

It consists of two stages. On stage F, measurements were taken from the moment the sensors began to work, this was done in order to know how each model behaves according to temperature and humidity with five minute intervals. In the model "a" of Figure 2, it was observed that in the first 4 hours of the dawn, the trend showed a variation in humidity between 34% and 42%, while the temperature after a small jump is maintained around 18° C. Later, similar ranges were observed from 6:00 a.m. to 9:00 a.m. and from 5:00 p.m. to 6:30 p.m., this mark is prone throughout the day, maintaining a variable behavior in the humidity and temperature data for all the racks. The sensors showed an unstable behavior in the environmental conditions of the data center and showed that temperature and humidity are quite similar, this indicates that one affects the other. The above can be seen in Figure 3.
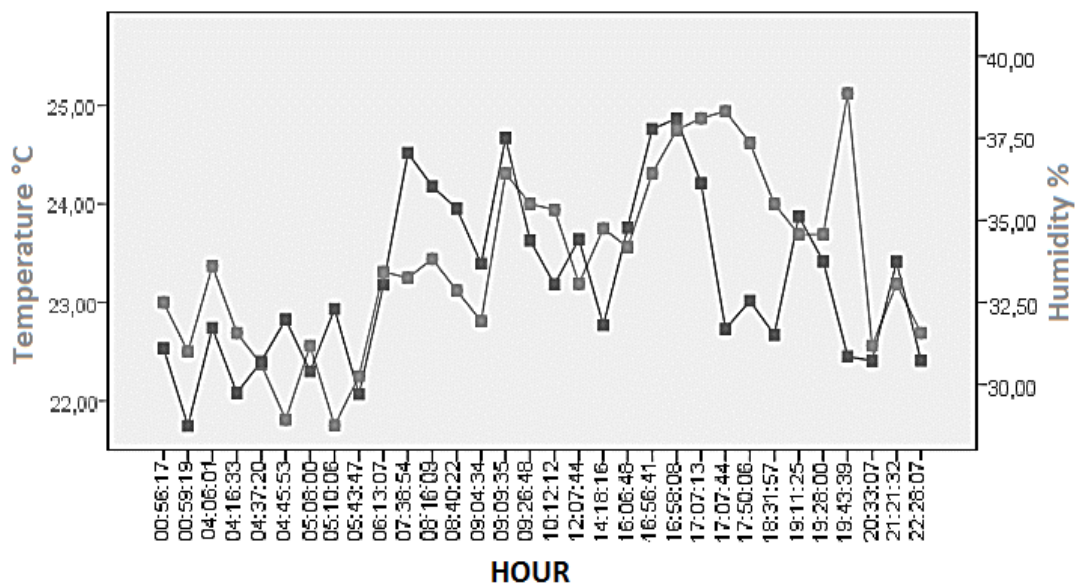

Fig. 3: Relative humidity and temperature rack 3

In the case of model "b" of Figure 2, the exit temperatures presented continuous changes and suffered a slight increase, because the servers transferred heat to the environment of the racks, except in rack 3 where the influence of the cooling of the center computation was given by the derivation of the air and produced a lower temperature at the outlet of the rack.

In stage G, a pilot test was carried out applying both models that allowed to obtain records of temperature and humidity, continuously for storage. The above procedure showed that model "a" of Figure 2 had humidity operating ranges between 40% and 55%; with respect to temperature, the range was from 17° C to 27° C. For the case studies in each rack, the maximum temperature level was exceeded. In the case of model "b" of Figure

2, the inlet and outlet temperatures were within the permissible operating ranges. These results were due to the behavior of the cooling unit, because the entry temperatures to the racks were between 17° C and 22° C, and the outputs with the influence of the servers were found temperatures between the ranges of 17° C and 28° C.

In addition, the studies of the previous models collaborated in the determination of temperature trends, for example, in the model "a" of Figure 2, there was a tendency to maintain the continuous changes over time, in the average, minimum temperature and maximum, in magnitudes 5.31° C; 3.7° C and 3.02° C; with a final maximum temperature of 30.19° C at 32.02° C. For maximum temperatures, peaks of behavior were estimated between 1:00 and 6:00 in the afternoon, while for the minimum temperatures occurred between 8:00 and 8:30 am. In the case of humidity, in racks 1 and 3, it increased around 7% during the study and in rack 2, the increase was around 16%.

In the case of model "b" of Figure 2, the behavior of the cooling unit showed that the entry temperatures to the racks were between 17° C and 22° C, for the outputs with the influence of the servers temperatures were found between the ranges of 17° C and 28° C. With respect to the higher temperatures in the two models the trend for hot spots was maintained.

Taking into account the previous results, it can be compared with the approach proposed by [7], which used two mathematical methods of modeling for the analysis of the behavior of temperature levels in a data center: Orthogonal Decomposition (POD)) and Simulation of Fluid Dynamics and Heat Transfer (CFD/HT), whose results coincided so that they could conceive a suitable modeling. Among the outstanding aspects that were highlighted were that the maximum temperature found was of 5.23 ℃, while the minimum of the average values were $3.2 \times 10 - 4°$ C and 0.325° C; in the case of modeling with parameter d $*$ = 0.01. Then, in the case of having assigned a value of 0.02, the accuracy improved considerably reaching values of 0.565° C, 0.165° C and $4.8 \times 10 - 5°$ C for the maximum, average and minimum values respectively.

In addition, in another study conducted by [1] showed that the average temperature of the interior and the surface of the servers improved in the study of a range of densities of electrical power, for example, inside the servers in a data center was from 22.5 ℃ to 19 ℃ for a density of 1898 W / m2 achieving an improvement of 11% due to the elimination of the hot air recirculation. In the same way, this work can be compared with a similar one carried out by [10], in which computer simulated modeling processes were carried out where the internal temperature of the CRACs was studied in 144 servers, with a time interval from zero to 358 seconds, it was observed that the maximum internal temperature reached was of 48° C in a t=264 s; At the end of the simulation, a point of convergence status of 40 ℃ was reached, so a heat peak close to 10 ℃ is lost that occurs in the CRAC while approaching cooling according to the model, demonstrating a temporary deficiency of the stationary thermal models and the need to achieve a conservative temperature estimate.

5. Development of a three-dimensional dynamic thermal map with registered temperatures and humidity

In stage H, the monitoring system is made up of a web application that receives information from sensors programmed in Arduino connected by means of a three-wire bus to the humidity and temperature sensors using the Onewire communication protocol. Additionally, a USB port was available to interconnect and send records to the PC for database storage, and the interconnection was implemented with a 20 x 4 LCD screen connected to one of the Arduino's outputs to show the measurements locally. The above is visualized in Figure 4.

1. Sensors

2. Hub rj11 (fixed telephone hub)

3. 4-wire telephone flat cable

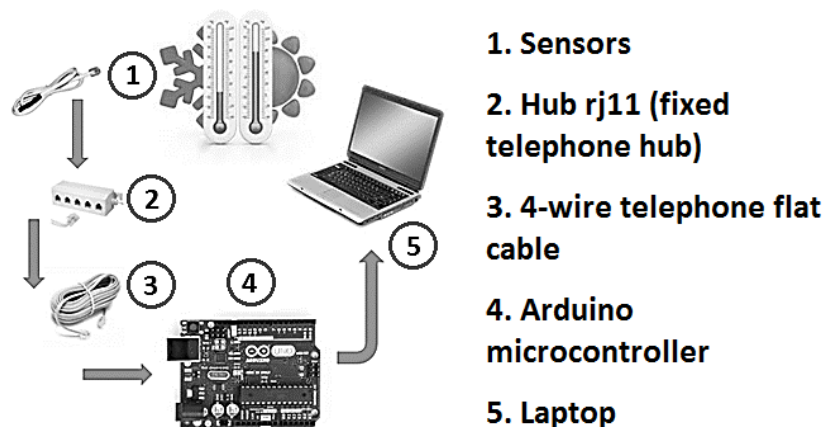4. Arduino microcontroller

5. Laptop

Fig. 4: Diagram of the data acquisition system

The software every three minutes captured records of the temperatures and humidities taken by the 25 sensors on the Onewire bus configured to work at 12 bits, grouped them into a specific structure and sent them through the USB port (emulated to a COM serial port on the PC) to be received in the destination computer, and showed in the LCD the eight temperature values of the input and output sensors of a specific Rack. Additionally, a

Raúl D. Mejía et al. / International Journal of Engineering and Technology (IJET)

security measure was programmed, with which the application activated an alert alarm when the external power supply of the sensors is lost.

In stage I, the software took the records and organized them for the respective processing, which could be viewed both on computers and mobile devices. The application showed each Rack with its respective sensors and at the same time a series of colors that inform immediately if the temperature and humidity data are outside the established limit, being the red color, the indicator of a very high value, as can be observed in Figure 5.
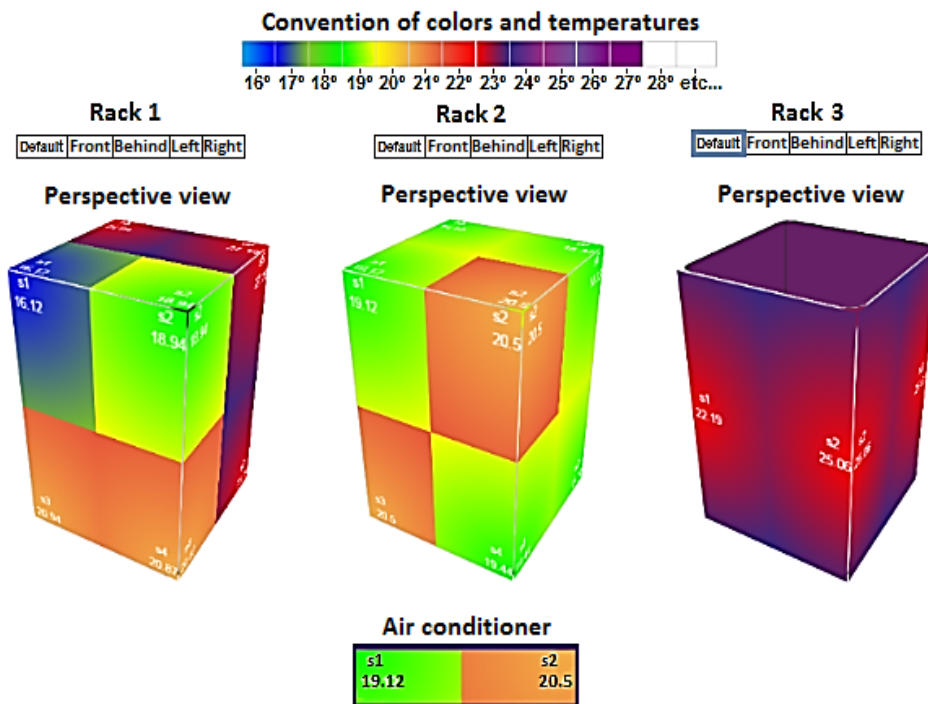

Fig. 5: Web visualization software (Temperature)

In the application, temperatures and humidity levels are shown in three-dimensional form, with the movement controls, the sides of the rack were observed, these rotate according to the order assigned to them. In the case of humidity could be observed by rotating the rack, this was shown in the center of them, which allowed to verify the relative humidity that was presented.

6. Analysis of resulting data on temperature and humidity measurements

In stage J, it could be concluded that the data center was found to be working normally and although the equipment was operating in what was assigned, they were exposed to possible damage due to temperature changes. In the K stage, verification of the international standard was verified with the application of model "b" of Figure 2 proposed by ASHRAE, Science Center and HP, taking temperature and humidity data as a reference for efficient estimation of model "a" of Figure 2. In stage L, the techniques of data mining to analyze information and determine a diagnosis were:

- Correlation matrix: With this technique we sought to estimate the relationship that exists between the temperature and humidity data on each of the sensors of the different racks and the cooling unit, this matrix was made with the SPSS software, which allowed apply data mining techniques through a graphical interface. In the resulting matrix there were relationship values that one sensor had with another. It was estimated that for all the racks and the cooling unit there were moderate correlations (above 0.8) and the rest of the correlations were considered low. We found quite high correlations in the tables (above 0.9), these were the ones that were taken into account. It is clarified that the correlation matrix was not used to find anomalies.

- Analysis of main components: Communities are represented by the variance of each of the variables explained by the main components, every element of the extraction is worth 1. For the studies, the two components with greater weight were taken. Therefore, the analysis of principal components (ACP) carried out, showed as results that the first two main components or dimensions, explained the total variability of the data, managing to accumulate between them a percentage of variance equal to 71.37% of the Total variance (inertia) of the data set.

- Hierarchical cluster technique: Following a complete link algorithm for the formation of clusters and allowing the grouping of data in five (5) clusters, they described which variables needed more attention. For humidity, Rack 2 suffered greater risks due to its high measurements and for temperature data, Racks 1 and 3 with the Sensors (5, 6, 8, 18, 19 and 20) were grouped together with greater risks of high heat, which had a maximum temperature.

- Decision tree: It was considered a supervised data technique and allowed to identify the occurrence of anomalies, in addition it contained a predictive investigative level. Due to the previous techniques of previously studied cases, the most significant clusters or groups to be studied were classified. For humidity there was the case of Rack 2 where a high level of measurements was found. For this variable, a decision tree was generated taking into account the time. With the above, it was possible to obtain a tree with nodes and branches that related anomalies of humidity levels, this served to make tabulations that led to show that the model is correct with 98.8% of the data in general. For each category of dependent variable there was a slightly lower success with the case of the normal state, with 96.2%.

The phases previously applied, led to a final comparison where the results of the two models that were presented in the research were taken, in this way it was shown that the data in the proposed experimental model may have gaps due to the distribution of the sensors, because it does not cover the entire rack space, unlike model b of Figure 6.
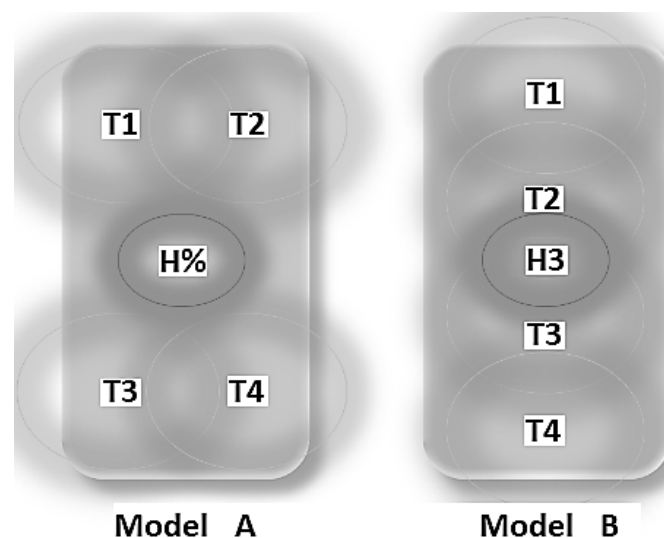

Fig.6: Temperatures in the front of the Rack

Model A of Figure 6 was not based on the study of the inlet and outlet temperature, nor was the load variable in the equipment contemplated. For the above, the behavior of each pair of sensors and their temperatures with respect to the established thresholds was taken as an initiative. In model B of Figure 6, three variables were considered: temperature, humidity and the distribution of workload; in the same way, thermal zones were established and the workload was established as a variable that can affect the state of the ambient temperature.

IV. CONCLUSION

The following conclusions were obtained:

a) A low-cost monitoring system can be designed to detect thermal anomalies of temperature and humidity data centers and thus be able to meet the standards established by ASHRAE TC9.9.

b) It was experimentally concluded that any failure reduces the reliability of the data center.

c) The monitoring system is effective when detecting thermal anomalies of temperature and humidity.

d) The system is suitable to be used in data centers of any type of organization, due to its low cost and reporting of anomalies in real time.

REFERENCES

[1]   S. Nada and K. Elfeky, "Experimental investigations of thermal managements solutions in data centers buildings for different arrangements of cold aisles containments," Journal of Building Engineering, vol. 5, pp. 41-49, 2016.
[2]   H. Zhang, S. Shao, H. Xu, H. Zou and C. Tian, "Free cooling of data centers: A review," Renewable and Sustainable Energy Reviews, vol. 35, pp. 171-182, 2014.
[3]   J. Cho, T. Lim and B. Kim, "Viability of datacenter cooling systems for energy efficiency in temperate or subtropical regions: Case study," Energy and buildings, vol. 55, pp. 189-197, 2012.
[4]   H. Sun and S. Lee, "Case study of data centers' energy performance," Energy and Buildings, vol. 38, no. 5, pp. 522-533, 2006.

[5]   M. Khan and N. Hassan, Thermofluid modeling for energy efficiency applications, 1st edition, pp. 31–58. Australia: Academic Press, 2016

[6]   R. Zhou, Z. Wang, C. Bash and A. McReynolds, "Data center cooling management and analysis - a model based approach," 28th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), pp. 98-103, 2012.

[7]   L. Phan, and C. Lin, "Reduced order modeling of a data center model with multi-Parameters," Energy and Buildings, vol. 136, pp. 86-99, 2017.

[8]   J. Moore, J.S. Chase and P. Ranganathan, "Weatherman: Automated, online and predictive thermal mapping and management for data centers," IEEE International Conference on Autonomic Computing (ICAC'06), pp. 155-164, 2006.

[9]   Y. Berezovskaya, A. Mousavi, V. Vyatkin, X. Zhang and T.B. Minde, "Improvement of energy efficiency in data centers via flexible humidity control," IEEE 42nd Annual Conference Industrial Electronics Society (IECON), pp. 5585-5590, 2016.

[10] M. Jonas, R. Gilbert, J. Ferguson, G. Varsamopoulos and S. Gupta, "A transient model for data center thermal prediction," IEEE International Green Computing Conference (IGCC), pp. 1-10, 2012.

AUTHOR PROFILE

Raúl D. Mejía works as full-time professor at the University of La Guajira (Colombia). Mr. Mejía completed his undergraduate in Systems Engineer at the University of La Guajira.

Tatiana K. Montoya works as full-time professor at the University of La Guajira (Colombia). Mrs. Montoya completed his Magister from University of La Guajira (Colombia). Mrs. Montoya completed his undergraduate in Industrial Engineering at the University of La Guajira.

Raul J. Martelo works as full-time professor at the University of Cartagena (Colombia). Mr. Martelo completed his magister from Industrial University of Santander (Colombia). Mr. Martelo completed his undergraduate in Systems Engineering at the Industrial University of Santander.