

Identification of Patterns of Consumption through the Daily Mean Outdoor Temperature

Aparicio-Ruiz Pablo, Guadix-Martín José, Onieva-Giménez Luis, Cortés-Achedad Pablo, Muñozuri-Sanz Jesús

Organization Engineering Group, School of Engineering, Universidad de Sevilla, Camino de los Descubrimientos, Seville, Spain
Email: pabloaparicio@us.es

How to cite this paper: Pablo, A.-R., José, G.-M., Luis, O.-G., Pablo, C.-A. and Jesús, M.-S. (2017) Identification of Patterns of Consumption through the Daily Mean Outdoor Temperature. *World Journal of Engineering and Technology*, 5, 20-28.
<https://doi.org/10.4236/wjet.2017.53B003>

Received: June 30, 2017

Accepted: August 8, 2017

Published: August 11, 2017

Abstract

The identification and recognition of patterns in the context of building is a necessary feedback to create intelligent buildings. In this context, the key is empowering the systems with learning elements to make decisions. The challenge is detected element to predict the human behavior in the building. Daily mean outdoor temperature is one of the variables with incidence in the human comfort due to the weather adaptation of the users. In this paper it analyzed the consumption in an office respect to the internal temperature and the daily mean temperature through cluster techniques. The cluster can be used as a forecasting of consumption.

Keywords

Pattern Recognition, Identification, Clustering, Energy

1. Introduction

The recent standards ISO 15251 [1] and other references, for example the guide CIBSE and ASHRAE remarking the dynamic behavior of the comfort temperatures, vary through the year as people adapt to changes in outside temperatures. This adaptation is different in buildings with the HVAC system in summer or in free running building's where the primary aim is to provide building occupants with a thermally comfortable indoor environment, using minimal energy consumption, through building techniques to maintain or improve performance. For this reason, it is decided to use daily mean temperature in the analysis of consumption.

The energy consumption data are influenced by the weather conditions. The temperature pattern is one of the primary factors of energy usage and electric

companies analyse the weather in the model that quantifies the relationship between energy sales and temperature. In [2] it analysed the relationship between electricity load and daily air temperature in Spain.

The problem of modelling and forecasting the energy use in a building can be complex. Data about the system behaviour in the building can be used in the forecasting and the behaviour patterns. It can help to know the consumption to make important decisions, including decisions on purchasing and generating electric power and load switching.

Load forecasts can be divided into three categories: short-term forecasts which are usually from one hour to one week, medium forecasts which are usually from a week to a year, and long-term forecasts which are longer than a year [3]. Short-term forecasts are important in day-to-day operation of the equipment. It is useful for building energy managers to schedule the operation of HVAC systems. In this paper, our interest focuses in analysing the demand in a short period based on the weather forecast and the information obtained in the pattern of consumption defined by clusters.

In [4], it is highlighted the impact of the occupant feedback can be useful for building diagnostics. On the one hand, the providing occupants with individualized controls can influence their thermal comfort. On the other hand, it is provided the occupants with an individual control can improve their subjective satisfaction towards their working environment.

A systematic review of research efforts and achievements mainly on methods of IA and expert systems is presented in [5]. The statistical techniques, artificial intelligence algorithms, neural networks, fuzzy logic, expert systems and other methods are used for forecasting. The consumption of HVAC systems is the principal issue. In [6] it is shown a comparison in energy consumption in offices by end use, between USA, UK and Spain. The HVAC energy end uses in Spain are 52%, followed by lighting 33% and equipment 10%.

The cluster analysis or clustering is the process of partitioning a set of data in characterizing subsets with similar characterized. Clustering can be used to organize data in groups and present the result in a concise and easy ways.

Clustering can be used as a data mining function, and the cluster analysis can be used as a tool to observe the characteristics of each cluster, and to maintain the focus on a set of clusters for the analysis and decision making, especially in the forecasting of consumption based on the weather prevision.

2. Methodology

The methodology is based on three variables, firstly is known the consumption of energy in the HVAC system for a period of 15 minutes, the indoor temperature in this period, and the daily mean outdoor temperature.

The data are saved in database; the outdoor temperatures are obtained from a weather station in the building. HVAC system data are obtained from a Building Automation System (BAS), HVAC mode (cooling or heating), ON-OFF mode,

indoor temperature and consumption is obtained from a network of sensors.

The first step was the definition of database and built the system. After this step, it is analyzed the conjunction of data in a year. It was filtered data into two groups, one for summer or cooling mode, and the other for winter or heating mode. Both groups are used for the clustering analysis.

On pattern classification, the k-nearest neighbor rule, first introduced by Fix and Hodges [7] is one of the most simple nonparametric techniques. The basic rationale for the KNN rule is such that: find KNNs for a query pattern, and then among these nearest neighbors, the most represented class is assigned to the query pattern by a majority voting. The 1-NN rule is a special form of KNN, when $k = 1$. In **Figure 1** is shown a centroid based on the nearest neighbor but based on c-means [8].

In this methodology it is used clusters, although not in the sense of cluster analysis as an unsupervised classification technique, but in the sense that whatever is done to apply in a data set where patterns want to be discovered. These clusters are used as a discriminant analysis to build a prediction model that is used to classify objects from a dataset depending on the features observed on them. In this case, the dependent variable is the grouping variable, which identifies which group and object belongs. This grouping variable is known at the beginning, the defined range of consumption, because there is a previous known classification of the elements of the dataset. In this sense, it is definitely the centroid of each range, and the prediction is analyzed by fuzzy C-means.

In the K-means algorithm, each vector is classified as belonging to a single cluster, and the centroids are updated based on the classified samples. In a variation of this approach known as fuzzy c-means [9], all vectors have a degree of membership for each cluster, and the respective centroids are calculated based

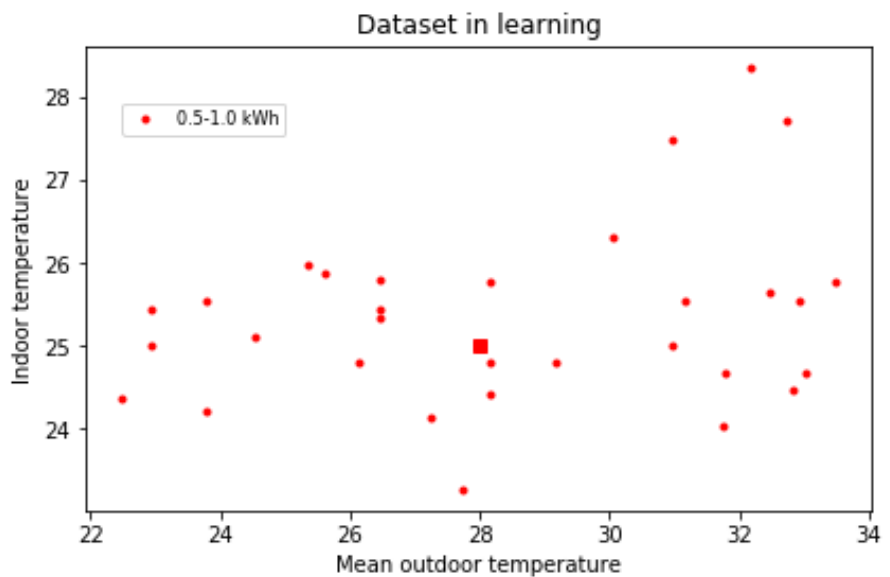


Figure 1. Example of centroid for 0.5 - 1.0 kWh based on the consumption in cooling mode with 15 min.

on these membership degrees. In K-means algorithm computes the average of the vectors in a cluster as the center. In this methodology are used the fuzzy c-means finds the center as a weighted average of all points, using the membership probabilities for each point as weights. Vectors with a high probability of belonging to the class have larger weights, and more influence on the centroid.

After the analysis of dataset in learning, **Figure 2**, the centroids used to classify are defined.

In **Figure 3** appear the centroid of the dataset in **Figure 2**. The data were clas-

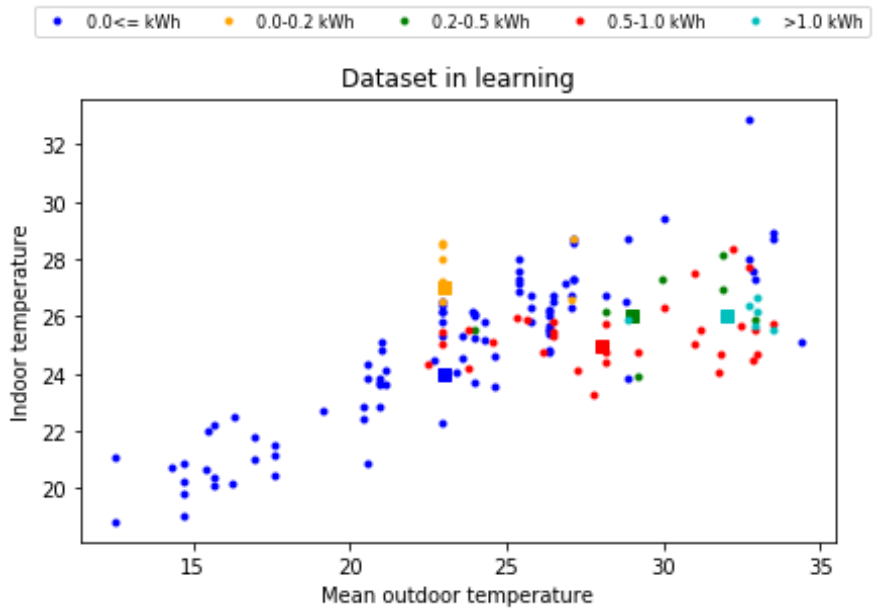


Figure 2. Example of dataset based on the consumption in cooling mode.

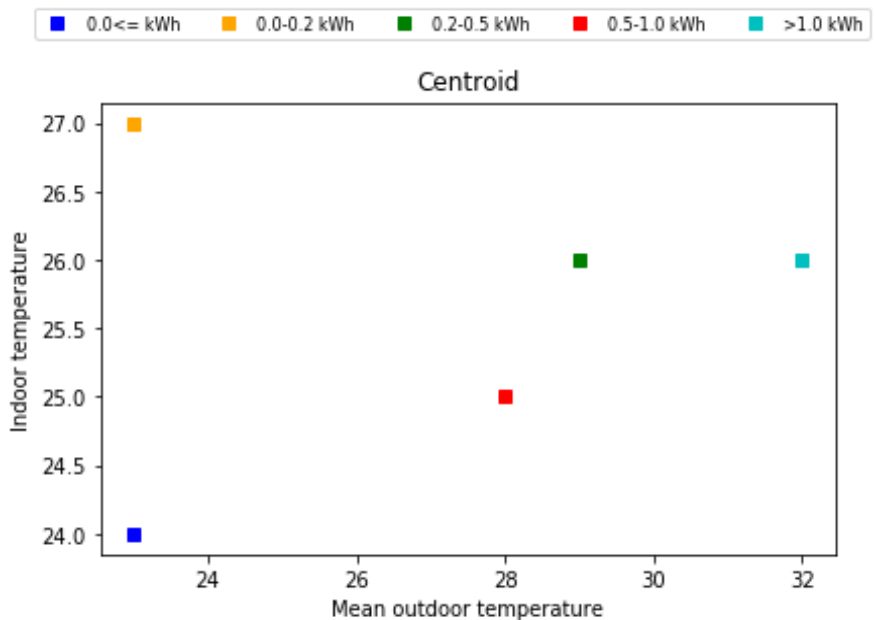


Figure 3. Example of centroids based on the consumption in cooling mode.

sified in 5 ranges of consumption in a period of 15 minutes. The first, no consumption 0 kWh, the second low consumption, a range between 0 to 0.2 kWh, then 0.2 - 0.5 kWh, in the penultimate 0.5 - 1 kWh and the last is more than 1 kWh.

In **Figure 4** it is shown the dataset uses to analyze the answer when the method is applied. At the same time the requirement of comfort was analyzed for this conjunction of data.

In **Figure 5** it is shown the cluster solution in cooling mode. Clearly, the

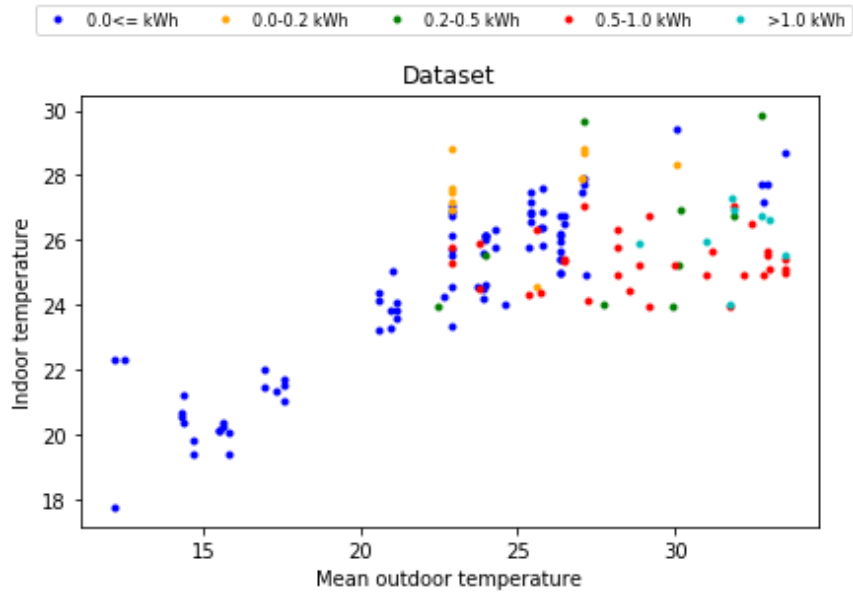


Figure 4. Example of dataset to analyse the detection based on the consumption in cooling mode.

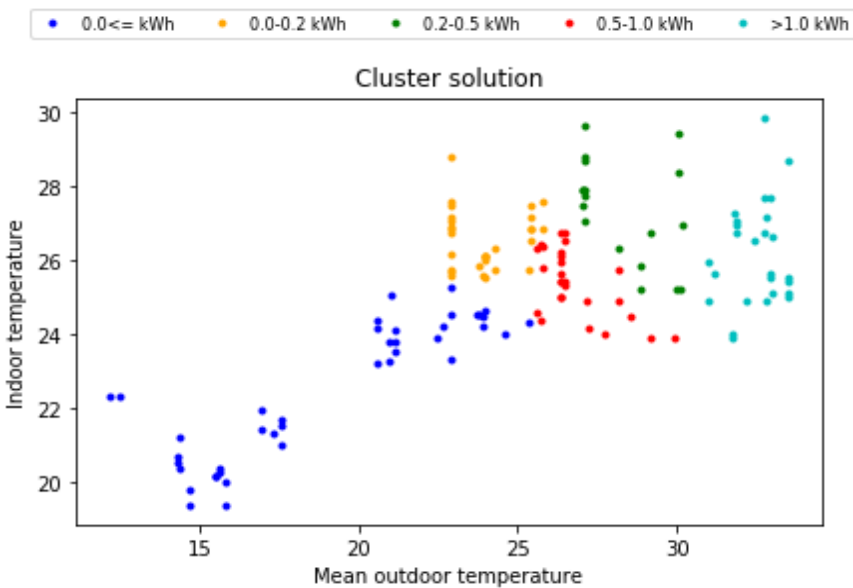


Figure 5. Example of forecast of consumption defined by the cluster solution of the **Figure 4** dataset, in a building in cooling mode.

system detects the five ranges. In this case, in summer mode, the system learned to detect consumption (kWh in 15 min.) when the indoor temperature (in °C) during this period was produced by a previous daily mean outdoor temperature.

The analysis of these data is joined to a survey answer about the user need in the period of study.

3. Case Study

A study of the application of the technique has been made based on a sample of 378 surveys in winter and 277 surveys in summer, and the physical data of the instant where those surveys were done. 838 surveys were discarded, because they represent situations without the use of HVAC system, since it was wanted to make forecasts with the system in heating or cooling mode.

The Figures in the previous section are based on those real data. These data contain situations in a period of one year.

The study was developed into room in an office in which four users sit. The office is not conditioned by external radiation. The study used half of the information for learning, and the other half to verify the ability to identify the pattern. The office was in Seville. Seville, at the south of Spain, it has a subtropical Mediterranean climate [10]. It has a dry summer and wet winter. The annual average temperature is 25.4 °C during the day and 13 °C at night, with an annual average of 19.2 °C according to the AEMET information (the State Meteorological Agency).

4. Results

The results of the application of the methodology were analyzed. In the case of the summer results (**Figures 2-5**), 43% of the cases were well classified. While 16% of the cases were classified in a higher range, 31% were classified in a range of consumption, while the real consumption in that period of 15 minutes was null, although the machine was on without action, finally in a 9% the consumption was higher than the prediction, although in 31% of that percentage corresponds to detections of non-consumption situations.

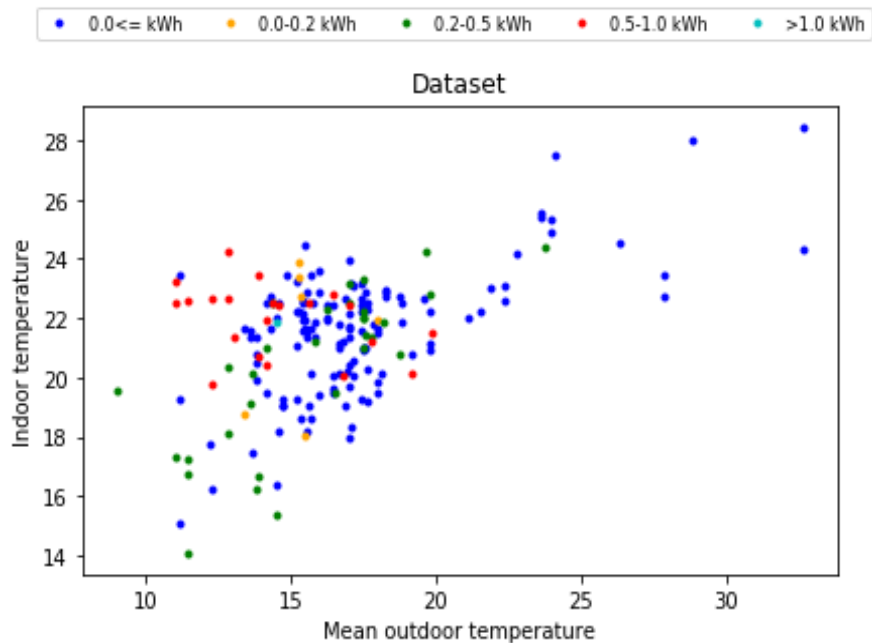
The operation of the system was analyzed with respect to what the users wanted, **Table 1**, the possible situations are described below:

- In A is shown the percentage of occasions where they responded that they wanted a lower or slightly lower temperature and exist HVAC actions with consumption to improve the temperature or users answered without changes and the system was active but without consuming.
- In B situations exists consumption to maintain the temperature and the user wanted to maintain that state of comfort.
- In C is shown the percentage of the situations in an overcooling condition.
- In D, the user wanted the temperature lower, although the system was active and no energy is consumed to decrease the temperature.

The results of the application of the methodology were analyzed. In the case of the winter results (**Figure 6** and **Figure 7**), 44% of the cases were well classified.

Table 1. Situations between consumption and the preference of users.

	Situations			
	A	B	C	D
Real	46%	22%	14%	18%
Forecasting	40%	40%	17%	4%

**Figure 6.** Example of dataset to analyse the detection based on the consumption in heating mode.

While 9% of the cases were classified in a higher range, 33% were classified in a range of consumption, while the real consumption in that period of 15 minutes was non-consumption, although the machine was on without action, finally in a 14% the consumption was higher than the prediction, although in 81% of that percentage corresponds to detections of non-consumption situations in which consumption was mostly less than 0.5 kWh.

The operation of the system was analyzed with respect to what the users wanted, **Table 2**, the possible situations are described below:

- In A is shown the percentage of occasions where they responded that they wanted a higher or slightly higher temperature and exist HVAC actions with consumption to improve the temperature or they answered without changes and the system was active but without consuming.
- In B situations exists consumption to maintain the temperature and the user wanted to maintain that state of comfort.
- In C is shown the percentage of the situations in an overheating condition.
- In D, the user wanted the temperature higher, although the system was active and no energy is consumed to increase the temperature.

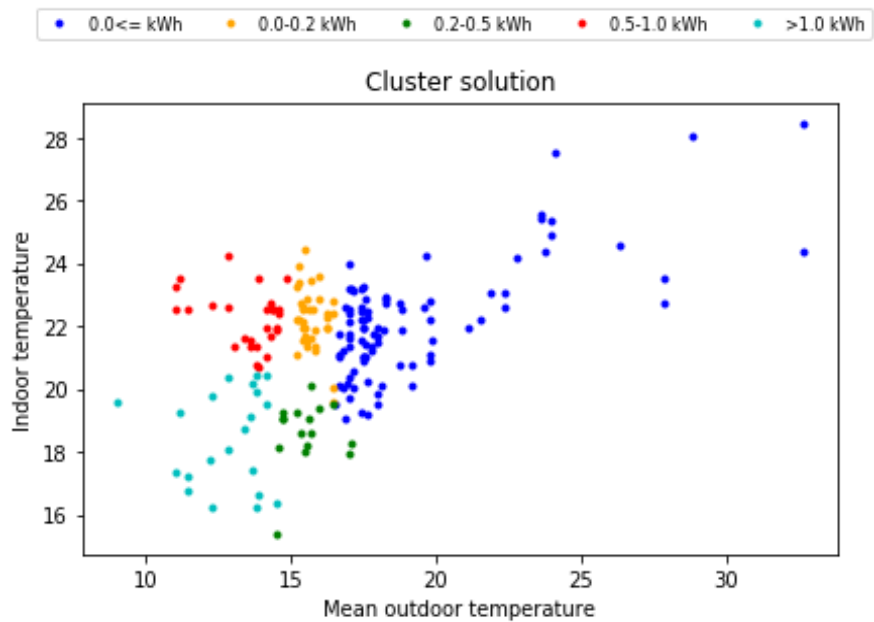


Figure 7. Example of forecast of consumption defined by a the cluster solution of the Figure 6 datasets, in a building in heating mode.

Table 2. Situations between consumption and the preference of users.

	Situations			
	A	B	C	D
Real	50%	15%	8%	27%
Forecasting	48%	25%	8%	19%

5. Conclusions

The cluster analysis can be used as a data mining and decision making. Specially in the forecasting of consumption based on the weather prevision, this technique can be joined with other techniques in forecasting and in artificial intelligence. This technique was used in data of a studio in an office in Seville. The results show that it can serve as a method to identify consumption patterns in the future, and the solutions are like the preference of users in the office. The results can be used in decisions making about store of energy, in the decisions about actions in building control and future intelligent systems.

References

- [1] AENOR, "AENOR: Norma UNE-EN 15251:2008 Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics". <http://www.aenor.es/>
- [2] Valor, E., Meneu, V., Caselles, V., Valor, E., Meneu, V. and Caselles, V. (2001) Daily Air Temperature and Electricity Load in Spain. *J. Appl. Meteorol.*, **40**, 1413–1421. [https://doi.org/10.1175/1520-0450\(2001\)040%3C1413:DATAEL%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040%3C1413:DATAEL%3E2.0.CO;2)

- [3] Feinberg, E.A. and Genethliou, D. (2006) Load Forecasting. *Appl. Math. Restructured Electr. Power Syst.*, 269-285.
- [4] Zhao, J., Lasternas, B., Lam, K.P., Yun, R. and Loftness, V. (2014) Occupant Behavior and Schedule Modeling for Building Energy Simulation through Office Appliance Power Consumption Data Mining. *Energy Build.*, **82**, 341-355. <https://doi.org/10.1016/j.enbuild.2014.07.033>
- [5] Dounis, A.I. (2010) Artificial Intelligence for Energy Conservation in Buildings. *Adv. Build. Energy Res.*, **4**, 267-299. <https://doi.org/10.3763/aber.2009.0408>
- [6] Pérez-Lombard, L., Ortiz, J., González, R. and Maestre, I.R. (2009) A Review of Benchmarking, Rating and Labelling Concepts within the Framework of Building Energy Certification Schemes. *Energy Build.*, **41**, 272-278. <https://doi.org/10.1016/j.enbuild.2008.10.004>
- [7] Cover, T. and Hart, P. (1967) Nearest Neighbor Pattern Classification. *IEEE Trans. Inf. Theory*, **13**, 21-27. <https://doi.org/10.1109/TIT.1967.1053964>
- [8] Bezdek, J.C., Ehrlich, R. and Full, W. (1984) FCM: The Fuzzy C-Means Clustering Algorithm. *Comput. Geosci.*, **10**, 191-203. [https://doi.org/10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)
- [9] Jain, A.K., Murty, M.N. and Flynn, P.J. (1999) Data Clustering: A Review. *ACM Comput. Surv.*, **31**, 264-323. <https://doi.org/10.1145/331499.331504>
- [10] Kottek, M. and Rubel, F. (2006) World Map of the Köppen-Geiger Climate Classification Updated. University of Veterinary Medicine Vienna.



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact wjet@scirp.org