Cluster analysis use for processing of the results of suspension's grain composition occurring in rainfall and snowfall

Ewa Burszta-Adamiak, Janusz Łomotowski

Institute of Building and Landscape Architecture
Agricultural University of Wroclaw
pl. Grunwaldzki 24
50-363 Wroclaw
burszta@wp.pl
lomot@ozi.ar.wroc.pl

Abstract: Grain composition of suspensions was determined in precipitation samples collected at five sites in the Upper and Lower Silesia region (southern and south-western Poland) over a period of three years (2003-2005). Laser diffraction particle size analyzer was used in this study.

Variety of probability density distribution functions of particles size in snow and rain waters obtained during analysis suggests the necessity of further research in order to systematize the data. Cluster analysis was used to this purpose.

1 Introduction

Snow and rain waters are characterized by quantitative and qualitative changeability that is dependent on atmosphere contamination extent, the season of a year, precipitation frequency (dry weather period) and sampling sites characteristic. Knowledge of suspended matter graining more and more arouses interest. It results from the need to determine the extent of suspensions removal for the storage reservoir design or other systems for rainwater pretreatment which are based at natural biological processes and water infiltration into the ground.

Samples of precipitation in the form of rain water and snow were collected from experimental sites situated in Bielawa, Strzelin, Wrocław and Racibórz towns. Samples in Bielawa and Strzelin were collected between November 2003 and July 2004, in Wrocław between July 2003 and July 2004 and in Racibórz from January to May 2004. Samples of a few hours' rainfall and snowfall were used for the analysis. Samples of snow collected within the winter sampling time were melted in the laboratory at ambient temperature. The total volume of snow or rain water sample collected within one sampling period varied from 600 to 900 ml.

Sampling site in Bielawa was located at housing estate area and in Strzelin at private promises which was situated 100 m from access road. Sampling site in Wrocław was located in the vicinity of the Szczytnicki park. There were a lot of trees, bushes and plants which blossom in April and at the beginning of May. During vegetation season pollen from these plants is washed away by rain water.

Two sampling sites in Upper Silesia region were located in Racibórz . One of them was situated in the neighborhood of chemical plant "Henkel" and the second measuring position was located in the centre of Racibórz at the park. The others sampling sites were situated 20 km from Racibórz.

The Mastersizer 2000 laser granulometer (diffraction analyzer) of Malvern Instruments Ltd. was used for the analysis of suspension's content and grain composition in the range of 0,02-2000 μm. A detailed description of the principle of this device operation can be found in [MS91, SMH04].

The results of the investigations that were carried out by authors [BŁ05, BŁS04] have shown large variety of particle's size and quantity occurring in rainfall and snowfall. Large quantities of the results encouraged authors to make an effort to systematize them. Cluster analysis was used to this purpose. Clustering techniques have received attention in many areas, such as computer science, medicine, biology, psychology and in business [AR96, MVW96].

Cluster analysis is an exploratory data analysis tool which aims at sorting different objects into groups in a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise.

Hierarchical tree and the k-means algorithms are the most widely used techniques for clustering. The goal of the tree clustering algorithm is to join together objects into successively larger clusters, using some measure of similarity or distance. A typical result of this clustering type is the hierarchical tree so-called dendrogram presenting distances between objects when forming the clusters. The basic operation of k-means algorithm relies on creating a number of k clusters, assign observations to these clusters so that the means across clusters are as different from each other as it is possible.

2 Results

Selected results of the analysis carried out by means of hierarchical cluster method are shown in Figures. 1 and 2.
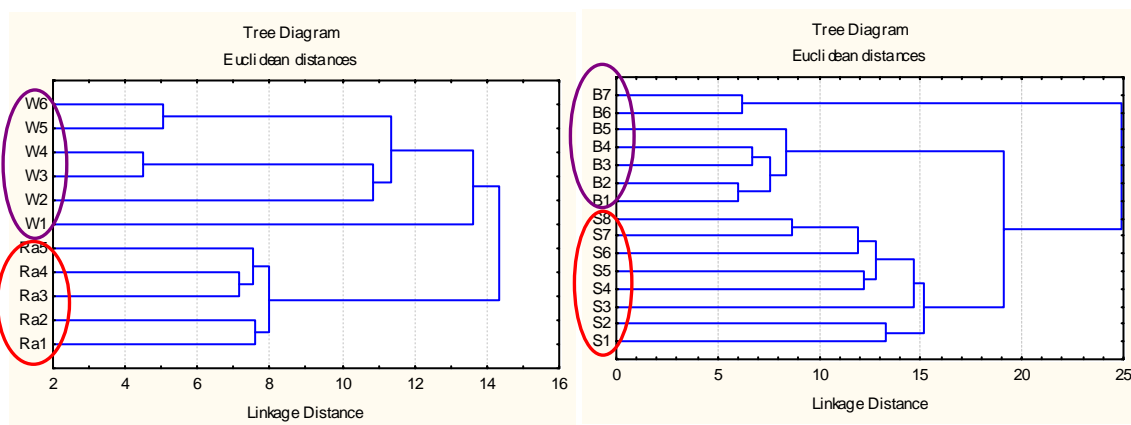

Figure 1: Dendrogram presenting clusters of rainfall and snowfall originated in Racibórz (Ra) and Wrocław (W) (on the left) as well as in Bielawa (B) and Strzelin (S) (on the right)
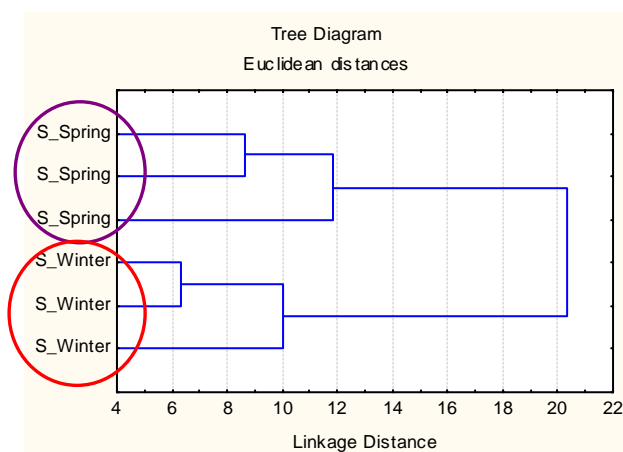

Figure 2: Dendrogram presenting clusters of rainfall and snowfall originated in Strzelin during spring and winter 2003/2004

Possibility of forming clusters proves the existence of similarity in frequency occurrence of particles in different precipitation events. Clusters of the research results depending on sampling site can be separated for the samples of rainfall and snowfall originated in Racibórz and Wrocław as well as in Bielawa and Strzelin (Figure 1). This relationship has not been confirmed for the research results from Bielawa and Racibórz.

Analysis, taking sampling season factor into a consideration, showed that only suspension's grain composition in precipitation originated in Strzelin was season dependent (Figure 2).

The research results showed that size of suspensions occurring in rainfall and snowfall is dependent on location of sampling site and in smaller extent on season of the year when precipitation is sampled.

References

[AR96]   Al.-Daoud, M. B; Roberts, S.A.: New methods for the initialization of clusters. Pattern Recognition Letters, No 17, 1996; pp. 451-455.

[BŁ05]   Burszta-Adamiak, E.; Łomotowski, J.: Grain composition studies of rainwater and the surface water by means of laser granulometer. (in Polish) Monografie Komitetu Inżynierii Środowiska PAN, II Kongres Inżynierii Środowiska, tom2, Vol.33 , Lublin 2005; pp. 331-338.

[BŁS04]   Burszta-Adamiak, E.; Łomotowski, J.; Stodolak, R.: Analysis of precipitation pollutions. (in Polish) Seria: Badania Systemowe. Wspomaganie informatyczne rozwoju społeczno-gospodarczego i ochrony środowiska., tom 36, Warszawa 2004; pp. 281-289.

[MVW96]   Mangiameli, P.; Vhen, S. K.; West, D.: A comparison of SOM neural network and hierarchical clustering methods. European Journal of Operational Research. No 93, 1996; pp. 402-417.

[MS91]   McCave, I. N.; Syvitski, J. P. M.: Principles and methods of geological particle size analysis. Principles, methods and application of particle size analysis. Cambridge University Press 1991.

[SMH04]   Sperazza, M.; Moore, J. N.; Hendrix; M. S.: High-resolution analysis of naturally occurring very fine-grained sediment through laser diffractometry. Journal of Sedimentary Research, Vol. 74, No.5, September 2004; pp. 736-743.