World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:14, No:12, 2020

Application of Combined Cluster and Discriminant Analysis to Make the Operation of Monitoring Networks More Economical

Authors : Norbert Magyar, Jozsef Kovacs, Peter Tanos, Balazs Trasy, Tamas Garamhegyi, Istvan Gabor Hatvani

Abstract : Water is one of the most important common resources, and as a result of urbanization, agriculture, and industry it is becoming more and more exposed to potential pollutants. The prevention of the deterioration of water quality is a crucial role for environmental scientist. To achieve this aim, the operation of monitoring networks is necessary. In general, these networks have to meet many important requirements, such as representativeness and cost efficiency. However, existing monitoring networks often include sampling sites which are unnecessary. With the elimination of these sites the monitoring network can be optimized, and it can operate more economically. The aim of this study is to illustrate the applicability of the CCDA (Combined Cluster and Discriminant Analysis) to the field of water quality monitoring and optimize the monitoring networks of a river (the Danube), a wetland-lake system (Kis-Balaton & Lake Balaton), and two surface-subsurface water systems on the watershed of Lake Neusiedl/Lake Fertő and on the Szigetköz area over a period of approximately two decades. CCDA combines two multivariate data analysis methods: hierarchical cluster analysis and linear discriminant analysis. Its goal is to determine homogeneous groups of observations, in our case sampling sites, by comparing the goodness of preconceived classifications obtained from hierarchical cluster analysis with random classifications. The main idea behind CCDA is that if the ratio of correctly classified cases for a grouping is higher than at least 95% of the ratios for the random classifications, then at the level of significance (α=0.05) the given sampling sites don't form a homogeneous group. Due to the fact that the sampling on the Lake Neusiedl/Lake Fertő was conducted at the same time at all sampling sites, it was possible to visualize the differences between the sampling sites belonging to the same or different groups on scatterplots. Based on the results, the monitoring network of the Danube yields redundant information over certain sections, so that of 12 sampling sites, 3 could be eliminated without loss of information. In the case of the wetland (Kis-Balaton) one pair of sampling sites out of 12, and in the case of Lake Balaton, 5 out of 10 could be discarded. For the groundwater system of the catchment area of Lake Neusiedl/Lake Fertő all 50 monitoring wells are necessary, there is no redundant information in the system. The number of the sampling sites on the Lake Neusiedl/Lake Fertő can decrease to approximately the half of the original number of the sites. Furthermore, neighbouring sampling sites were compared pairwise using CCDA and the results were plotted on diagrams or isoline maps showing the location of the greatest differences. These results can help researchers decide where to place new sampling sites. The application of CCDA proved to be a useful tool in the optimization of the monitoring networks regarding different types of water bodies. Based on the results obtained, the monitoring networks can be operated more economically.

Keywords : combined cluster and discriminant analysis, cost efficiency, monitoring network optimization, water quality
Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development
Conference Location : Chicago, United States
Conference Dates : December 12-13, 2020