Spatio-temporal modelling in the analysis of water quality data

Arminda Manuela Gonçalves[1] and Teresa Alpuim[2]

[1] Department of Mathematics for Science and Technology, University of Minho, Guimares, Portugal, Campus de Azurm 4800-058 Guimares.
e-mail: mneves@mct.uminho.pt
[2] Department of Statistics and Operations Research, University of Lisbon, Lisbon, Portugal, Campo Grande, Edifcio C5, 1148-016 Lisboa.
e-mail: mtalpuim@fc.ul.pt

Abstract: The purpose of this work is to develop new statistical methodologies to model the concentration of pollutants in the surface water of river basins. As a case study, we examine in detail the Ave River hydrological basin located in the north-west of Portugal (its main route is the Ave River which runs along 101 km, from its source to its mouth) and rigorously assess the performance of the Wastewater Treatment Plants (WTPs) installed in this basin at the end of the 1998 hydrological year.
The data set used refers to monthly values of 11 variables of water quality measured between 1988 and 2003 in a quality monitoring net of 20 sites. The measurements did not comply with the objectives of minimum quality for surface waters prescribed by the Portuguese legislation. Many untreated discharges are carried out in the waters, especially those related to industrial effluents from textile industries which result in extreme pollution of surface waters.
We first use descriptive statistics to perform a simple space-time analysis of the water quality variables to identify the most polluted segments of the Ave River and its adjacent streams. This analysis shows the quality variables with the best and worst performances and helps to detect both spatial and temporal patterns. This is a valuable step toward designing the more complex and rigorous procedure that we describe below. The proposed methodology starts using Cluster Analysis to classify the quality monitoring sites into homogeneous groups. The process is repeated, first using each quality variable separately and then using sets of similar types of variables together.
The spatial and temporal measurement of precipitation on a geographical area is essential when calculating hydrological balance for the indirect estimation of water flow and for the development of the study of the watershed charge. It is also very important for modelling many environmental phenomena, for example the variation of the water quantity of a hydrological basin of a river.
The dominant conditions in the source basin of that location, namely the hydrometeorological factors, affects the quality of the water. The variation of a space-time quality variable is associated with the variation of the flow (variable dilution effect) which in turn is related, in general, to the seasonal variation of rainfall.
In our study we need monthly measurements of rainfall averages which influence a

certain area of the River Ave basin and which represent the hydro-meteorological factor in the modelling of the quantity of water. We identify models which estimate monthly average rainfall in places where there are no observed values (in the monitoring points of quality of water) with the help of the spatial distribution of the rainfall and measurements in other locations.
Due to the large space and time variability of rainfall, the precise evaluation in real time of mean area estimates poses a difficult problem. These estimates are obtained from a set of rain gauges located at 19 points in the Ave River hydrological basin. We estimate the monthly values of rainfall per area flowing into each monitoring site. To accomplish this, we propose a methodology combining both deterministic (Thiessens polygons) and stochastic (Kringing) processes.
Finally, we adjust Linear Models for each variable to the average of each homogeneous group obtained by the Cluster Analysis. These models are selected taking into account the possible effects of trends, seasonality and the hydrometeorological factor estimated. Trends in water quality are studied in detail to evaluate the effective role of WTP's in the improvement of surface water quality in this basin.

Keywords: water quality variables; Cluster Analysis; area rainfall estimation; Kriging; Linear Models; trend; seasonality; WTP's assessment.