

# A Data Management System for Environmental Research

Ana Elena L. Conjares, Bobby D. Gerardo, Bartolome T. Tanguilig III

**Abstract**—Data generated from research projects, e.g. environmental data, could have vast potential for knowledge discovery other than that of the original purpose for what they were gathered. And further integration with other relevant datasets, historical or current, could provide for more knowledge discoveries, e.g. decision support, for forecasting and/or as inputs to other studies seeking better understanding of the environment. What is essential for this to happen is a database system that provides for proper data management with the capability for data integration, pre-processing, generate specified datasets for external data analysis and data mining tasks. This paper presents the design of a data warehousing system for environmental data generated from scientific research projects in a nuclear research and development institution as well as data coming from related environmental monitoring systems, e.g. meteorological and radioactivity. Data are loaded into the system through a parser module and metadata information is attached to each dataset. As initial data analysis capability, the system is designed to have the capability to generate and remember multivariate regression models.

**Keywords**— Data Warehouse, ETL, Data Mining, Metadata, Knowledge discovery