Geophysical Research Abstracts
Vol. 15, EGU2013-9322-1, 2013
EGU General Assembly 2013

Using Data Warehouses to extract knowledge from Agro-Hydrological simulations

Tassadit Bouadi (1), Chantal Gascuel-Odoux (2), Marie-Odile Cordier (1), René Quiniou (3), and Pierre Moreau (2)

(1) UMR IRISA, University of Rennes 1, Rennes, France (tbouadi@irisa.fr), (2) INRA, Soil Agro and hydroSystem, Rennes, France (Chantal.Gascuel@rennes.inra.fr), (3) INRIA, Rennes, France

In recent years, simulation models have been used more and more in hydrology to test the effect of scenarios and help stakeholders in decision making. Agro-hydrological models have oriented agricultural water management, by testing the effect of landscape structure and farming system changes on water and chemical emission in rivers. Such models generate a large amount of data while few of them, such as daily concentrations at the outlet of the catchment, or annual budgets regarding soil, water and atmosphere emissions, are stored and analyzed. Thus, a great amount of information is lost from the simulation process. This is due to the large volumes of simulated data, but also to the difficulties in analyzing and transforming the data in an usable information.

In this talk we illustrate a data warehouse which has been built to store and manage simulation data coming from the agro-hydrological model TNT (Topography-based nitrogen transfer and transformations, (Beaujouan et al., 2002)). This model simulates the transfer and transformation of nitrogen in agricultural catchments. TNT was used over 10 years on the Yar catchment (western France), a 50 km^2 square area which present a detailed data set and have to facing to environmental issue (coastal eutrophication). 44 output key simulated variables are stored at a daily time step, i.e, 8 GB of storage size, which allows the users to explore the N emission in space and time, to quantify all the processes of transfer and transformation regarding the cropping systems, their location within the catchment, the emission in water and atmosphere, and finally to get new knowledge and help in making specific and detailed decision in space and time.

We present the dimensional modeling process of the *Nitrogen in catchment data warehouse* (i.e. the snowflake model). After identifying the set of multileveled dimensions with complex hierarchical structures and relationships among related dimension levels, we chose the snowflake model to design our agri-environmental data warehouse. The snowflake schema is required for flexible querying complex dimension relationships. We have designed the *Nitrogen in catchment data warehouse* using the open source Business Intelligence Platform Pentaho Version 3.5.

We use the online analytical processing (OLAP) to access and exploit, intuitively and quickly, the multidimensional and aggregated data from the *Nitrogen in catchment data warehouse*. We illustrate how the data warehouse can be efficiently used to explore spatio-temporal dimensions and to discover new knowledge and enrich the exploitation level of simulations. We show how the OLAP tool can be used to provide the user with the ability to synthesize environmental information and to understand nitrates emission in surface water by using comparative, personalized views on historical data.

To perform advanced analyses that aim to find meaningful patterns and relationships in the data, the *Nitrogen in catchment data warehouse* should be extended with data mining or information retrieval methods as *Skyline queries* (Bouadi et al., 2012).

(Beaujouan et al., 2002) Beaujouan, V., Durand, P., Ruiz, L., Aurousseau, P., and Cotteret, G. (2002). A hydrological model dedicated to topography-based simulation of nitrogen transfer and transformation: rationale and application to the geomorphology denitrification relationship. Hydrological Processes, pages 493–507.
(Bouadi et al., 2012) Bouadi, T., Cordier, M., and Quiniou, R. (2012). Incremental computation of skyline queries with dynamic preferences. In DEXA (1), pages 219–233.