Geophysical Research Abstracts
Vol. 12, EGU2010-9506, 2010
EGU General Assembly 2010

A Data Model for Hydrologic Sensor Networks Monitoring River-Groundwater Interactions

Philipp Schneider (1) and Andreas Wombacher (2)

(1) Eawag - Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland (philipp.schneider@eawag.ch), (2) Data base group, Department of Computer Sciences, Univeristy of Twente, Enschede,The Netherlands

Real-time operated wireless sensor networks produce large amounts of data, so that typical eyeball based analysis of data comes to its limits. Consequently we have to adapt and automate our data handling and archiving procedures, as well as our data analysis tools. Management of sensor data requires metadata to understand the semantics of observations. While modelers have high demands on metadata, experimentalists prefer to minimize entering metadata, as this is an additional effort. Quite often this is done on subjective basis ("field notes") without following a strict and predefined structure with transparent criteria and consistent vocabulary. Nevertheless, data has to be semantically annotated. The claim of this presentation is to focus on the essentials, being described by location, time, owner, instrument and measurement. The applicability is demonstrated in a case study focussing on monitoring changes of river-groundwater interactions in the context of river restoration.

Fundamental steps are (i) a proper storage in a database, (ii) traceable link between data and meta-data and (iii) semantically annotation tagged to the data, e.g. concerning data quality and data interpretation. To some extend this can be done automatically (e.g. plausibility check, if values are in expected range). The scientific challenge lies in identifying periods (data strings) where high resolution data stresses expected system behavior and established process representations/conceptualizations used in well accepted and widely used models. When and where do we measure data which do not match our expectations? As the amount of data will increase dramatically, pre-aggregation and visualization have to be automated to focus on critical parts of time series which needs interpretation with further expert knowledge.