A Relational Model for Simulation Data

to Promote Interdisciplinary Collaboration

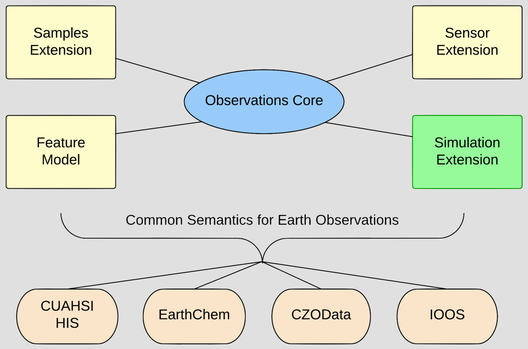
Anthony M. Castronova, Amber Spackman Jones, and Jeffery S. Horsburgh

Department of Civil and Environmental Engineering, Utah State University, Logan UT

H13H - 1212

Observations Data Model

The Observations Data Model (ODM2) is a community information model for representing spatially discrete, feature-based earth observations derived from sensors and samples. It provides a framework for representing a broad range of environmentally related data that can be used to archive many different types of naturally observed phenomena. The ODM2 has a central 'core' schema for information common to most types of data. The data model is extensible so that additional descriptive information can be provided for certain datasets. This design consideration enables the ODM2 to represent a broad spectrum of data types.



Model Simulations in ODM2

We have developed an extension to ODM2 to encapsulate simulation-specific metadata. The extension includes metadata to represent models, model simulations, and the relationships between simulations and input/output data. This will promote interdisciplinary collaborative modeling by standardizing the way we represent models, simulations, and their data.

Design Objectives

Geospatial Support

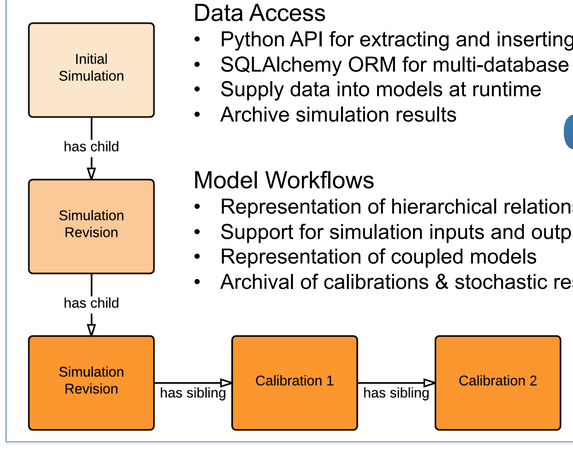
- Complete spatial data representation
- Enhance ODM2 with PostgreSQL + PostGIS
- Query by geometry, spatial transformations

Data Access

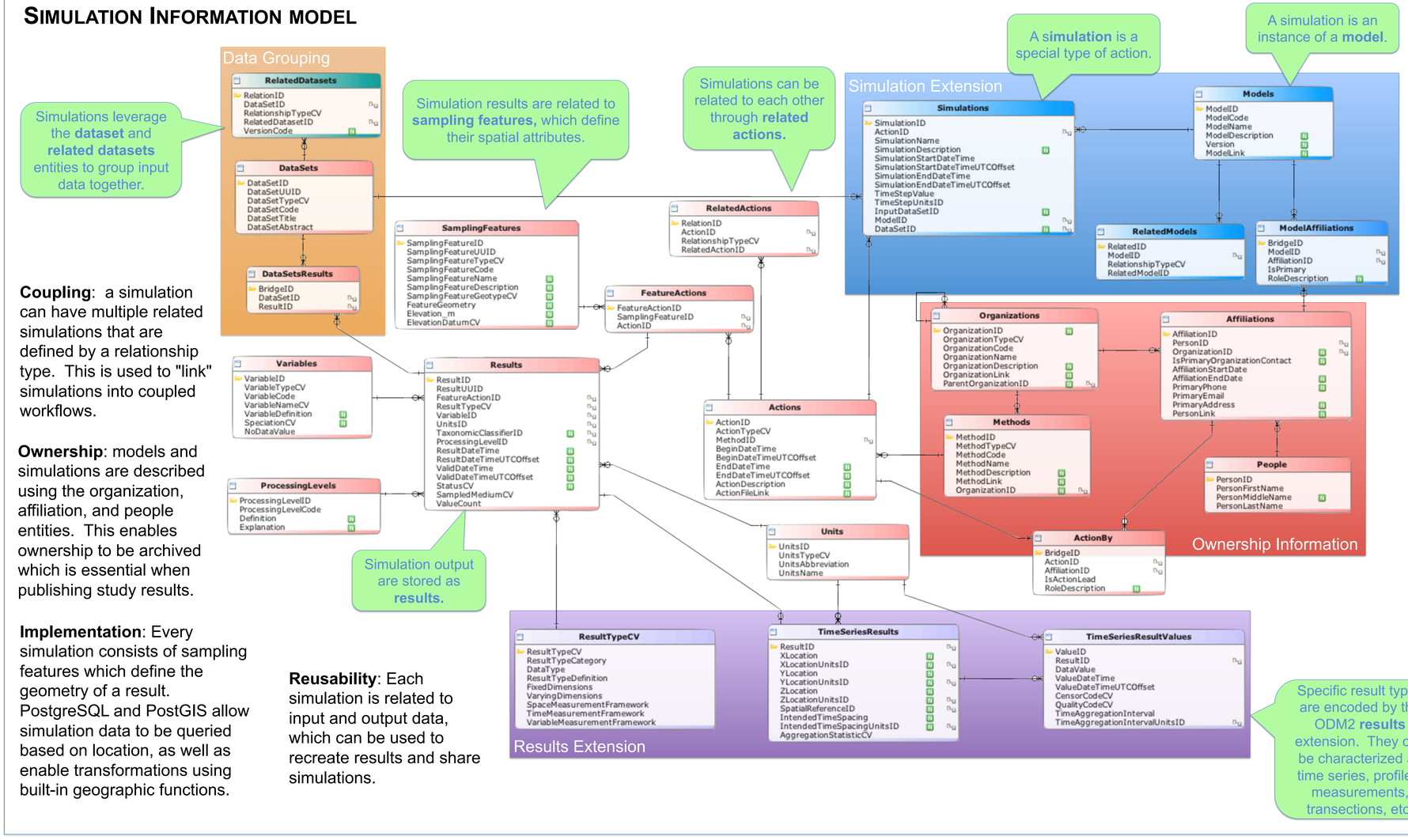
- Python API for extracting and inserting data
- SQLAlchemy ORM for multi-database support
- Supply data into models at runtime
- Archive simulation results

Model Workflows

- Representation of hierarchical relationships
- Support for simulation inputs and outputs
- Representation of coupled models
- Archival of calibrations & stochastic results



Simulation Information model



Coupling: a simulation can have multiple related simulations that are defined by a relationship type. This is used to "link" simulations into coupled workflows.

Ownership: models and simulations are described using the organization, affiliation, and people entities. This enables ownership to be archived which is essential when publishing study results.

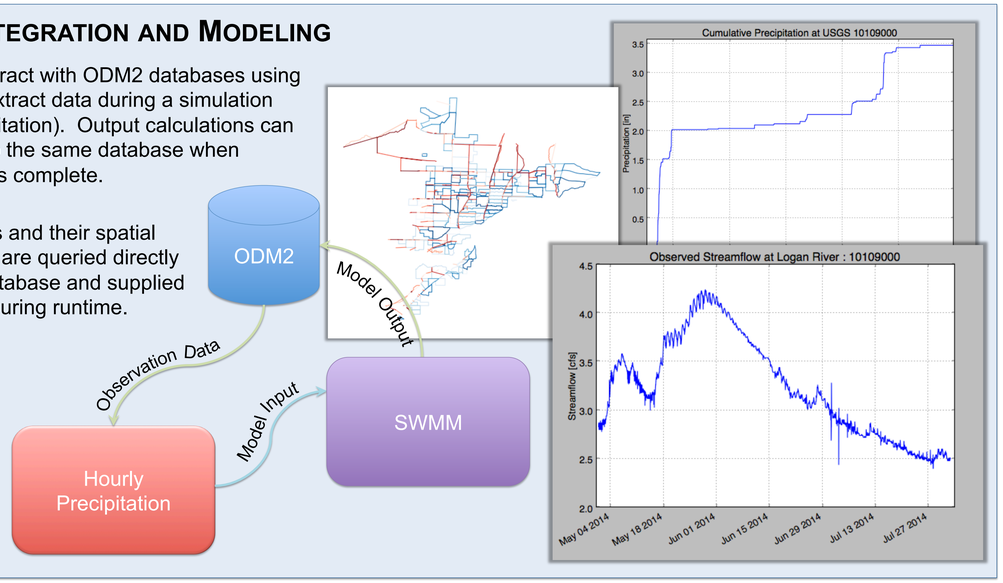
Implementation: Every simulation consists of sampling features which define the geometry of a result. PostgreSQL and PostGIS allow simulation data to be queried based on location, as well as enable transformations using built-in geographic functions.

Reusability: Each simulation is related to input and output data, which can be used to recreate results and share simulations.

Data Integration and Modeling

Models interact with ODM2 databases using an API to extract data during a simulation (e.g. precipitation). Output calculations can be saved to the same database when simulation is complete.

Data values and their spatial geometries are queried directly from the database and supplied to models during runtime.



Ongoing Work

1. Development of an integrated modeling framework that uses a data centric model coupling paradigm.
2. Archiving coupled model simulation results in a comprehensive manner (i.e. input parameters, calibrations, etc.)
3. Design of an API for inserting and querying simulation data during coupled model simulations.
4. Leveraging observation data in an ODM2 as input for an interdisciplinary water resources model as part of iUTAH modeling activities.

https://github.com/Castronova/EMIT.git

Acknowledgements

This material is supported by the National Science Foundation under grant number 1208732: **iUTAH-innovative Urban Transitions and Aridregion Hydro-sustainability.**