Workflows to accelerate preparing data for input to hydrologic models – DataNet Federation Consortium (DFC)

W. Christopher Lenhardt
RENCI
Spaghetti
“Anyone can cook in the French manner anywhere, with the right instruction. . . . Although you will perform with different ingredients for different dishes, the same general processes are repeated over and over again. . . . As you enlarge your repertoire, you will find that the seemingly endless babble of recipes begins to fall rather neatly into groups of theme and variations.” Forward, p xxiv, 2001 edition.
Did you know - Julia Child was a spy?
Mis en place
Floods

U.S. Drought Monitor

July 28, 2015
(Released Thursday, Jul. 30, 2015)
Valid 8 a.m. EDT

NOTE: To view regional drought conditions, click on map above. State maps can be accessed from regional maps.

http://droughtmonitor.unl.edu/
DataNet Federation Consortium (DFC)

Collaboration Environments for Data Driven Science

Major science and engineering initiatives are dependent upon massive data collections that comprise observational data, experimental data, simulation data, and engineering data. To support science and engineering collaborations, a policy-driven national data management infrastructure is being implemented. The prototype addresses both the life cycle of science and engineering data and the sustainability of data collections and repositories over time, across changes in the research communities.

NSF Grant Number OCI 0940841
Collaboration environments for data driven science.
VIC Hydrologic Model

- Variable Infiltration Capacity (VIC) Model
- Developed at University of Washington/Princeton University
- Widely used for macro-scale hydrologic modeling

Source: Gao et al. (2009)
Precipitation data processing

- Pre-processing data
  - Gridding, rescaling, etc.
- Run VIC model simulation
- Post processing of data
  - Summarize model output; create maps, figures, etc.
Data Pre-Processing for VIC

- Meteorological data are processed and combined into one dataset for each grid cell
- Soil and vegetation data are processed for each grid cell
Data post-processing for VIC

Model output is analyzed through data processing scripts that generate visualizations and publication ready figures.
Hydrologic Science Goals: RHESSys
(Regional Hydro-Ecological Simulation System)

- Model water, carbon, and nutrient cycling in headwater catchments
  - Forested
  - Urban
RHESSys Hydrologic Model

GIS Preprocessing
- Topography
- Vegetation
- Drainage Network
- Soil

Hierarchical and Distributed Elements
- Patch
- Hillslope
- Zone
- Watershed

Inputs
- Library of Vegetation and Soil Parameters
- Climate Time Series
- Disturbance History

RHESSys
- Meteorological Processes
- Hydrologic Processes
- Canopy Processes

Process Based Sub-Models

Output
- Timeseries
- Spatial

Creation of Worldfile and Flowtable
RHESSys data preparation: present and future

Manual workflows → Workflow framework
Ecohydrology data preparation software architecture
Bon apetit!

Thanks to Reagan Moore, Jon Goodall (co-PI), Mirza Billah (GRA), Larry Band (collaborator) & Brian Miles (GRA), and others for some of the material in this presentation.
Image Sources

- Slide #2
  - Julia Child (https://upload.wikimedia.org/wikipedia/commons/5/5c/Julia_Child.jpg)
  - Spaghetti (https://upload.wikimedia.org/wikipedia/commons/thumb/2/20/John%27s_Inc_Pizza_Spaghetti.jpg/800px-John%27s_Inc_Pizza_Spaghetti.jpg)
- Slide Julia Child (https://kpbs.media.clients.ellingtoncms.com/img/photos/2012/08/08/julia_turkey_tx800.jpg?aae402d4163f394116c3dd6e602f75682c526327)
- Mis en place (https://upload.wikimedia.org/wikipedia/commons/thumb/f/fa/Mise_en_place_for_hot_station.jpg/800px-Mise_en_place_for_hot_station.jpg)
- Hydrology
  - Drought (http://droughtmonitor.unl.edu)
  - Flood (http://floodobservatory.colorado.edu/GlobalFloodplains/090W040NCurrent.html)
- Book Cover (https://www.flickr.com/photos/george/4266821093/)
- Donuts (https://www.flickr.com/photos/eremita/12798304)