

An Intelligent Interface for Integrating Climate, Hydrology, Agriculture, and Socioeconomic Models

Daniel Garijo¹, Deborah Khider¹, Varun Ratnakar¹, Yolanda Gil¹, Kelly Cobourn², Ewa Deelman¹, Chris Duffy³, Rafael Ferreira da Silva¹, Armen Kemanian³, Craig Knoblock¹, Vipin Kumar⁴, Scott Peckham⁵, Yao-Yi Chiang¹, Ankush Khandelwal⁴, Minh Pham¹, Jay Pujara¹, Maria Stoica⁵, Kshitij Tayal⁴, Binh Vu¹, Dan Feldman¹, Lele Shu³, Rajiv Mayani¹, Anna Dabrowski⁶, Daniel Hardesty-Lewis⁶, Suzanne Pierce⁶

¹University of Southern California, ²Virginia Tech, ³The Pennsylvania State University, ⁴The University of Minnesota, ⁵University of Colorado Boulder, ⁶The University of Texas in Austin

ABSTRACT

Understanding the interactions between natural processes and human activities poses major challenges as it requires the integration of models and data across disparate disciplines. It typically takes many months and even years to create valid end-to-end simulations as different models need to be configured in consistent ways and generate data that is usable by other models. MINT is a novel framework for model integration that captures extensive knowledge about models and data and aims to automatically compose them together. MINT guides a user to pose a well-formed modeling question, select and configure appropriate models, find and prepare appropriate datasets, compose data and models into end-to-end workflows, run the simulations, and visualize the results. MINT currently includes hydrology, agriculture, and socioeconomic models.

Author Keywords

Intelligent workflow systems; model integration; environmental modeling; scientific discovery.

CCS CONCEPTS

H.5.m. Information interfaces and presentation (e.g., HCI); Miscellaneous.

ACM Reference format:

Daniel Garijo, Deborah Khider, Varun Ratnakar, Yolanda Gil, Kelly Cobourn, Ewa Deelman, Chris Duffy, Rafael Ferreira da Silva, Armen Kemanian, Craig Knoblock, Vipin Kumar, Scott Peckham, Yao-Yi Chiang, Ankush Khandelwal, Minh Pham, Jay Pujara, Maria Stoica, Kshitij Tayal, Binh Vu, Dan Feldman, Lele Shu, Rajiv Mayani, Anna Dabrowski, Daniel Hardesty-Lewis, Suzanne Pierce. 2018. An Intelligent Interface for Integrating Climate, Hydrology, Agriculture, and Socioeconomic Models. In Proceedings of IUI '19 Companion, March 17–20, 2019, Marina del Ray, CA, USA

<https://doi.org/10.1145/3308557.3308711>

1 Introduction

Understanding how human activities affect natural resources,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

IUI '19 Companion, March 17–20, 2019, Marina del Ray, CA, USA
© 2019 Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-6673-1/19/03...\$15.00

and how natural processes affect societies requires complex computational simulations that cut across disciplinary boundaries. For example, in order to understand how weather predictions and agriculture practices affect water availability or how flooding affects planting strategies and population migration, modelers integrate physics-based climate and hydrology models with biologically-informed agriculture models and market distribution economic models. While the questions are short fused in order to prepare for natural disasters or to make near-term policy decisions, integrating these diverse models may take many months or even years. Major challenges include finding relevant models to address a question and datasets with the necessary granularity and quality to run the models, developing sophisticated data transformations to set up and execute a model, and checking for models compatibility. Existing infrastructure supports some aspects of the modeling (e.g., [1]) but there are no comprehensive frameworks that tackle all the tasks involved in the model integration process. In addition, no frameworks have an integrated user interface, which allow to easily navigate among the various models and datasets, setting up scientific workflows, and visualizing the results.

2 The MINT Modeling Framework and UI

MINT is a novel framework for model integration that uses semantic representations to describe datasets and models to support users in data search and transformation, model selection and setup, combining them into end-to-end workflows, and execution and visualization of results [2].

The MINT user interface is designed for two types of users: 1) modelers, who can use MINT to reduce data friction across modeling approaches, and 2) users who lack a deep modeling expertise but need to interpret the results to implement policies. The MINT UI guides the user through a typical scientific workflow (Figure 1):

1) Question formulation and statement of the problem:

Users can set-up a question of interest and select relevant variables. Questions fall under three main categories: i) prognostic questions, which use forecasts to understand how some set of variables of interest will behave in the future; ii) diagnostic questions, where data about past situations is used to understand how some set of variables of interest behaved; and iii) counterfactual questions, where data about a situation is used to explore alternative timelines that are either diagnostic or prognostic in nature.

2) **Data Exploration:** Users can refine their question to a geographical area and a timeframe of interest. They may browse and visualize available data for that region.

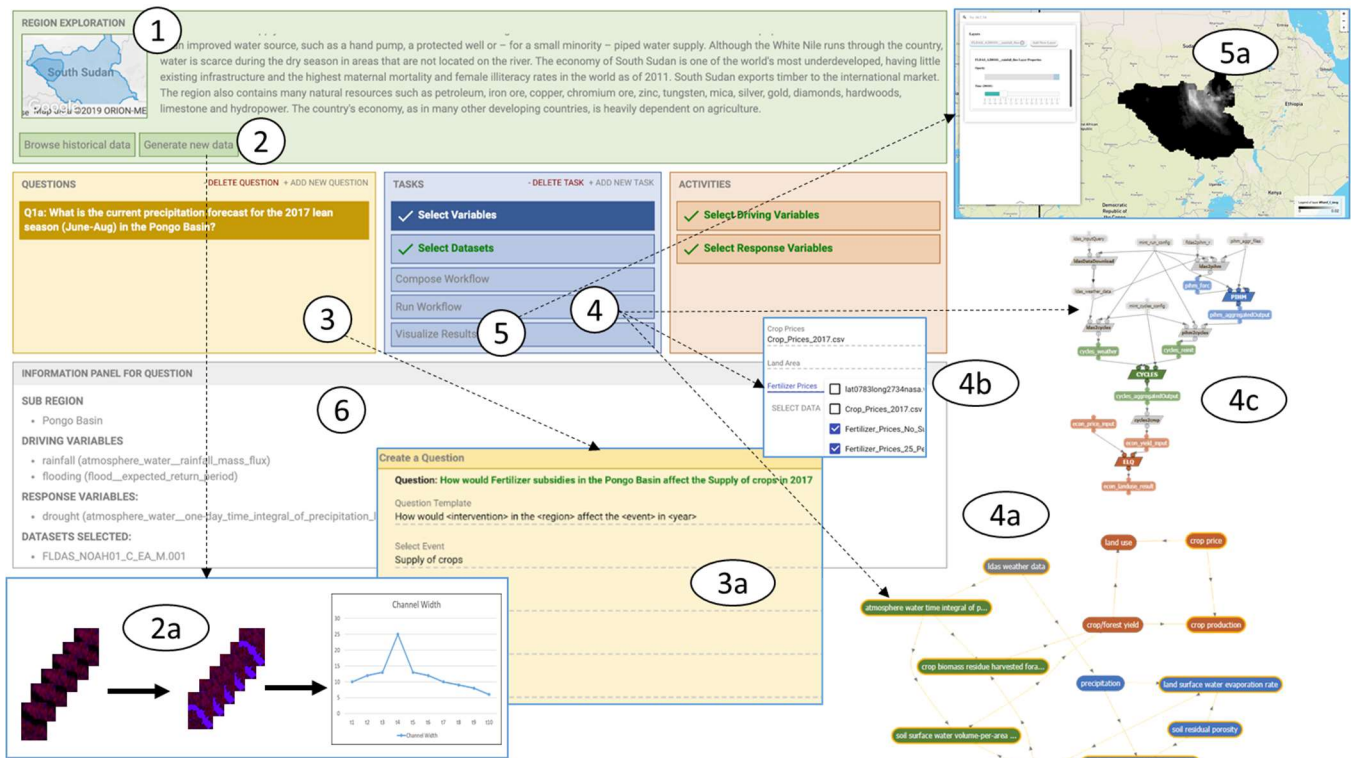


Figure 1: An overview of the main features of the MINT user interface. Once a region of interest has been selected (1), users may browse and generate additional data (2a), formulate modeling questions (3a), setup their models and data for execution (4a-c) and visualize and compare datasets and results (5). Provenance of the all user activities is captured to facilitate reproducibility (6).

3) Data Generation: Users may also generate new datasets by running pre-defined workflows. For example, a user can select a workflow to extract historical water depth for a river from satellite imagery of previous years by using machine learning techniques for remote sensing data (3a).

4) Model Setup, Data Preparation and Execution: One of the most challenging aspects of modeling is selecting the appropriate input data products and preparing them according to the models to run. In addition, if multiple models are integrated together, then additional transformation steps are required to convert the outputs of a model into the appropriate input formats for the next one. MINT helps assisting users to address this issue: users first specify the set of variables of interest for the modeling question at hand (4a) and the datasets they wish to use (4b). Then, MINT suggests appropriate models for the target variables, along with transformations to run them together as a workflow (4c). MINT executes the workflow efficiently in high-performance computing resources, and stores any results for future use.

5) Visualizing and Comparing Datasets and Results: MINT associates common visualizations to different types of data, and uses them to present results to the user (5a). Users may also compare results of different executions.

6) Capturing Provenance for Reproducibility: A critical requirement for modelers is to share and reproduce modeling products. MINT includes a history panel which records all the activities carried out by a user in the modeling interface. A provenance catalog stores these provenance traces, along with

workflow execution results, workflows and metadata that would allow other users to reproduce the results in the future.

3 Conclusions and future work

MINT is a model integration framework with a comprehensive user interface that uses knowledge about data and models to facilitate interdisciplinary modeling. Future work includes full automation of data transformations, running multiple workflows to reflect alternative scenarios, and exploring uncertainty of analysis results.

ACKNOWLEDGEMENTS

This work was funded by the Defense Advanced Research Projects Agency with award W911NF-18-1-0027, and the National Science Foundation with award ICER-1440323.

REFERENCES

[1] S. D. Peckham, E. W. H. Hutton, and B. Norris, "A component-based approach to integrated modeling in the geosciences: The design of CSDMS," *Comput. Geosci.*, vol. 53, pp. 3–12, 2013.

[2] Y. Gil et al., "MINT: Model Integration Through Knowledge-Powered Data and Process Composition," in *Proceedings of the Ninth International Congress on Environmental Modeling and Software*, 2018.