International Congress on Environmental Modelling and Software

9th International Congress on Environmental Modelling and Software - Ft. Collins, Colorado, USA - June 2018

Jun 27th, 3:40 PM - Jul 27th, 5:00 PM

Cyberinfrastructure for Enhancing Interdisciplinary Engagement in Coastal Risk Management Research

Ayse Karanci
North Carolina State University / Moffatt and Nichol, akaranc@ncsu.edu

Lisa Stillwell
University of North Carolina

Christopher Lenhardt
University of North Carolina

Casey Dietrich
North Carolina State University, jcdietrich@ncsu.edu

9th International Congress on Environmental Modelling and Software
Fort Collins, Colorado, USA, Mazdak Arabi, Olaf David, Jack Carlson, Daniel P. Ames (Eds.)
https://scholarsarchive.byu.edu/iemssconference/2018/

Cyberinfrastructure for Enhancing Interdisciplinary Engagement in Coastal Risk Management Research

Ayse Karanci[a,c], **Lisa Stillwell**[b], **Christopher Lenhardt**[b], **Casey Dietrich**[c]
[a] *Moffatt & Nichol (akaranci@moffattnichol.com)*
[b] *Renaissance Computing Institute (RENCI), University of North Carolina, Chapel Hill (lisa@renci.org, clenhardt@renci.org)*
[c] *Civil, Construction and Environmental Engineering Department, North Carolina State University (jcdietrich@ncsu.edu)*

Abstract: Tackling critical questions often requires the collaboration of researchers from different disciplines or institutions. Coastal hazards research is necessarily interdisciplinary and multi-methodological and often requires a team of researchers, due to its combination of storm-induced changes to the coastal environment, the effects of these changes on built infrastructure, and the combined effects on decision-making for individuals and communities. This paper introduces an interdisciplinary coastal hazard risk model that combines high resolution geospatial data, storm impact forecasts, and an agent-based model in the analysis, and then describes the model's implementation in a data science cyberinfrastructure. Lessons learned and limitations are also outlined.

Keywords: *open system architectures; component-based modelling; coastal risk management; interdisciplinary research*.

1 INTRODUCTION

The links and interactions between humans and their environment has intensified over the past century with increasing population, urbanization, and technological development (Hallagette et al., 2013). As systems become more interconnected, the effort to answer the critical questions has shifted toward interdisciplinary collaborations. This trend became especially dominant in research that aims to answer the critical questions at the interface of social and natural systems. Coastal systems encapsulate and demonstrate the interconnected nature of the anthropogenic activities and the natural systems. As such, in coastal hazard research, there is an increasing need for integration and re-use of models from different disciplines in the research collaborations. In their review of coastal disaster literature, Mileti and Noji (1999) stated that "research on disasters was dominated by physical scientists and engineers and that little attempt had been made to tap into social sciences to better understand the economic, social, and political dimensions of extreme natural events". Since then, a shift in scientific focus created a large body of research demonstrating interdisciplinary and inter-institutional efforts tackling critical questions in context of coastal systems (Prati et al., 2015; Lazarus et al., 2016; Van Dongeren et al., 2017).

However, interdisciplinary research presents challenges not only because of the complexity of the issues it addresses, but also because of the many challenges that come with interdisciplinary team efforts, some of which include interinstitutional collaborations via computer-mediated researchers across the globe. These challenges may include different scientific conventions and vocabulary used in different disciplines, the associated numerical and physical models that usually utilize different programming languages, and spatial and temporal grids used in the solution (Argent, 2004).

Despite providing the developer full control over the modeling process and optimizing the computational process, the traditional modeling approaches that use tightly coupled models (Sui and Maggio, 1999) require consistent internal conventions within the models (i.e. programming languages). In contrast, adopting loosely-coupled approaches facilitates engagement by different science fields by allowing different disciplines to construct autonomous modeling units that can be connected through shared boundary conditions during a simulation (Castranova and Goodall, 2010).

These loosely coupled models can be implemented in cyberinfrastructures to further enhance the collaborative research. Cyberinfrastructures consist of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked by high speed networks. The implementation of component-based models in cyberinfrastructure has the following advantages (Idaszak et al., 2017). First, it removes the interdependency of operating system and programming languages. Second, cyberinfrastructure allows the users to deploy the models without having to install the model or download the “big” data to their individual computers. Third, the cyberinfrastructure can adorn the user with the computing capabilities of the server infrastructure through a simple web browser.

Cyberinfrastructures with their data storage systems and data repositories also provides a secure way to manage small to large datasets. Employing data sharing rules, it can be leveraged to build a science community that is able to reuse applications and can utilize the datasets generated and published by the other researchers in the community. This is especially important for the dissemination and exchange of data produced by smaller scale projects, which might not have data repositories (Cragin, 2010)

To facilitate interdisciplinary research, a generalized data science cyberinfrastructure called xDCI, an acronym for cross-disciplinary data infrastructure, was developed by Renaissance Computing Institute (RENCI) (Krishnamurthy et al., 2017). xDCI provides researchers with a technology framework that enables interdisciplinary collaborations to rapidly deploy robust cyberinfrastructure that can easily ingest, move, share, analyze, and archive scientific data in all its varieties. In this paper, we describe our approach to implement an interdisciplinary coastal risk management research model using xDCI, as a pilot use case in the Risk Analytics Discovery Environment (RADE) project, as a part of the North Carolina Data Science and Analytics Data Science Initiative (NC DSAI). RADE uses xDCI with a goal to create a community platform for risk analytics research (Lenhardt et al., 2016). RADE defines risk analytics research using computationally and/or data intensive methods to assess and quantify risk and related impacts.

The RADE pilot use case combines high resolution geospatial data, storm impact forecasts, and an agent-based model in the analysis. In addition, the methodology requires access to analyze numerical storm model results, and an environment that can easily handle tens of thousands of discrete files.

In following sections, we briefly introduce the xDCI framework, describe the interdisciplinary coastal risk management model and outline the model’s application in xDCI. In addition to the descriptions, some preliminary lessons learned will also be outlined.

2 xDCI SYSTEM

xDCI is a technology framework designed to operate on shared, centralized computers. The main components of xDCI are:

- The integrated Rule Oriented Data System (iRODS): iRODS (http://irods.org) is an open source software that supplies a comprehensive set of tools to assist data management from initial collection through to archiving and re-use. Used by research organizations and government agencies worldwide, it is used to manage millions of files and tens of petabytes of data. The tools provided by iRODS enable a number of critical functions utilized in xDCI, namely, owner-specified access control for individual data sets, automatic transfer of data between computer systems, automatic backup and replication of data sets, generation of searchable metadata catalog that describes every file, every directory, and every storage resource in the database, and creation of workflows for data analysis. Additionally, the procedures that are constructed within iRODS are platform independent. Thus, they can be employed by end-users that are carrying their research on different operating systems - Windows operating systems, or Mac operating systems, or many Linux operating systems.

- The CyVerse Discovery Environment (DE): The DE (http://cyverse.com) provides a user-web interface for powerful computing, data management, and analysis application resources required for specialized scientific analyses. The DE has an Application Programming Interface that enables users to plug-in new analytical capabilities and can facilitate the implementation of resource-intense computations on remote resources.

  xDCI uses the DE both as a web interface, and to facilitate external analytical capabilities. Because the web interface provided by DE conceals the complexity from the user, carrying out analyses which are shared by computational savvy users are made easy for non-technical users of xDCI. Cyverse utilizes Docker container technology for integrating software applications to the CyVerse DE's Compute Cluster, which uses HTCondor for executing analyzes. Docker is a container technology that bundles the application of interest together with all its software dependencies (the code, system tools, system libraries, etc.) so it can function in a reproducible manner independent of operating platform. It also facilitates integration of more complicated and difficult to install software (e.g., software with many additional dependencies) in the DE. The docker images can easily be launched from the user's virtual desktop when they log into xDCI (Figure 1). Furthermore, the xDCI system will take care of any necessary data movement between xDCI and the compute resource, so that the analysis process itself will be a seamless experience for the user.

- xDCIShare: xDCIShare is derived from HydroShare (Tarbaton et al., 2014; Idaszak et al., 2016), an online collaborative system to enable users to share and publish data; to visualize and manipulate a comprehensive set of hydrologic data types and models; and to make this information available in a citable, shareable and discoverable manner. xDCIShare is the foundation for the Data Discovery and Data Persistence/Identifier functionality of xDCI by preserving these key capabilities of HydroShare in generalized form.

Using these components, xDCI enables researchers across disciplines and institutions to access data at different locations, choose applications from a comprehensive set of tools to manage and analyze their data, utilize computing resources at remote locations, run analyses, and share their results. Furthermore, the iRODS data grid secures the data and facilitates the re-use of it by providing a comprehensive metadata catalog to make data easier to find for replication studies and future research projects.

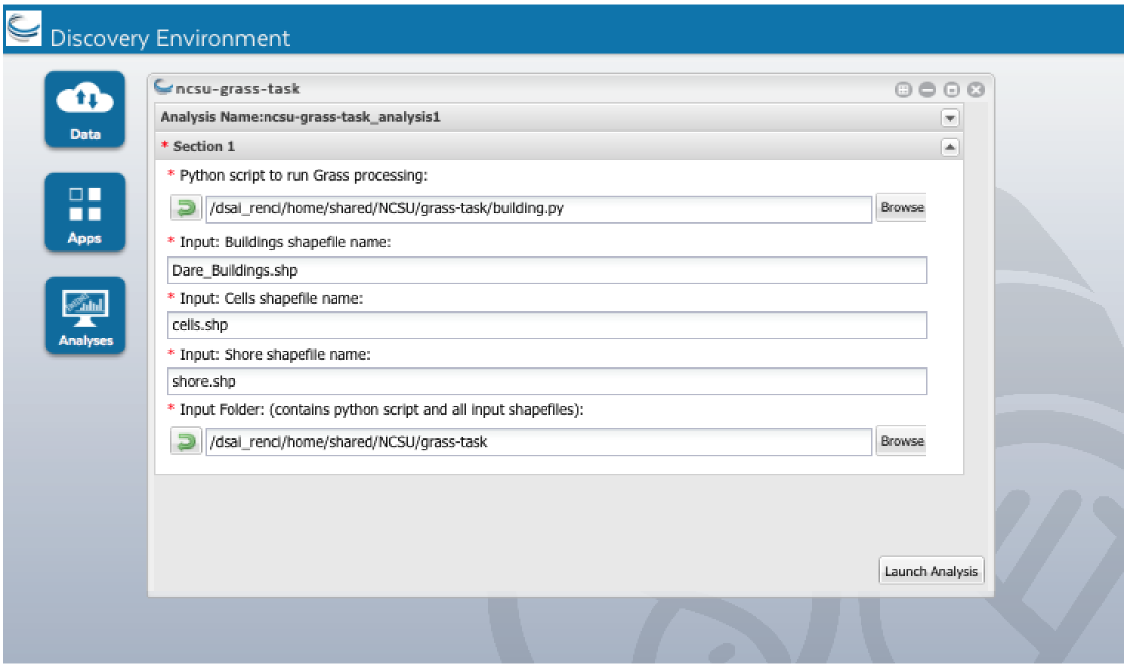

Figure 1. Screenshot of virtual desktop showing the Discovery Environment for an xDCI user. Note: The left side panel shows the navigation pane, which can be used to go between the available data in IRODS, the apps in the DE and the analyses results. The open window shows an example application configuration.

3 Coastal Risk Management Model Framework

Humans increasingly alter the coastal landscape and influence natural processes to reduce their vulnerability from hazards caused by tropical storms, hurricanes events and sea level rise (SLR) (Hapke et al., 2013; Halpern et al., 2015). Synchronously coastal landforms and natural processes modify the humans' environment and affect their decisions. Efficient and sustainable coastal management, therefore, requires the recognition of the two-way interaction between the human activities and natural processes. In an effort to investigate the reciprocal interaction of household decisions, coastal landform change and coastal hazard mitigation actions, and interdisciplinary agent-based model, referenced in this paper as Coastal Management and Occupation Dynamics model (CoMOD), was developed in Karanci et al., 2017. Using sea level rise and storms as exogenous drivers, CoMOD models the coupled evolution of coastal landscape and housing dynamics in a coastal town by simulating the responses of individual households' and town. However, CoMOD framework was implemented as a set of specialized models with varying input types, scales and ease of use. Therefore the framework required a sophisticated modeler to enable a full simulation including nearshore processes and agent-based decision-making. In this section, we describe the model components and how they were implemented in RADE to enable simulations by a wider range of researchers.

3.1 Coastal Management and Occupation Dynamics Model Overview

The model operates at the spatial scale of a town. The environment in the model upon which the agents interact has been designed to represent properties of the cadastral and physical coastal landform conditions. A topographic layer represents the coastal landform conditions (i.e. dune height beach width). A cadastral layer contains the areas that households inhabit and is generated from cadastral geographic data. Cadastral parcels can be either empty or occupied and have varying attributes that reflect their current physical and economic properties (i.e. parcel value, structure height from ground, structure distance to shore). The model environment is created using geospatial analysis using diverse data such as building databases, shoreline maps, and topography.

The main processes in the model are natural evolution of coastal landforms, implementation of soft engineered coastal projects, and trading of residential properties and cadastral parcels (Figure 2). For each time step, the coastal features are updated using user specified erosion and sea level rise rates. If triggered, the coastal features are further modified by coastal protection projects (beach nourishment or dune replenishment). The real estate decisions in each yearly cycle consists of several phases: relocation decision of homeowners and transformation of these agents into seller agents, decision to move into the coastal community by buyers, selection of the best affordable housing alternative by buyers, and determination of seller agent with whom to trade. Decision rules governing these processes include theories and parameters produced by engineering, social science, economy and planning disciplines. Details about these processes can be found in Karanci et al., 2017.

Each year, if a storm exists, CoMOD employs a storm-impact sub-model to (i) estimate the morphological change and inundation at buildings due to stochastic storms, and (ii) to determine the risk of flooding during feasibility analysis for possible coastal protection projects. To estimate the impact of storms, the model creates a Bayesian Network which predicts the change in morphological features and inundation at the structures using the pre-storm morphological conditions of the coast and hydraulic boundary conditions of the storm. The BN is trained using a suite of 1D XBeach simulations (Roelvink, 2009). Generating the 1D XBeach storm simulation database require coastal modeling expertise, computational time, and produces tens of thousands of discrete files.

The goal of CoMOD is to evaluate coastal hazard management strategies against historical and possible future conditions by exploring the evolution of the coastal town, typically over 50 years, under a diverse set of social and environmental conditions. Numerous possible scenarios can be generated and simulated by varying environmental (i.e. storm frequency, storm intensity, SLR) and socio-economic conditions (i.e. insurance rates, flooding risk perception, costs of prevention measures).

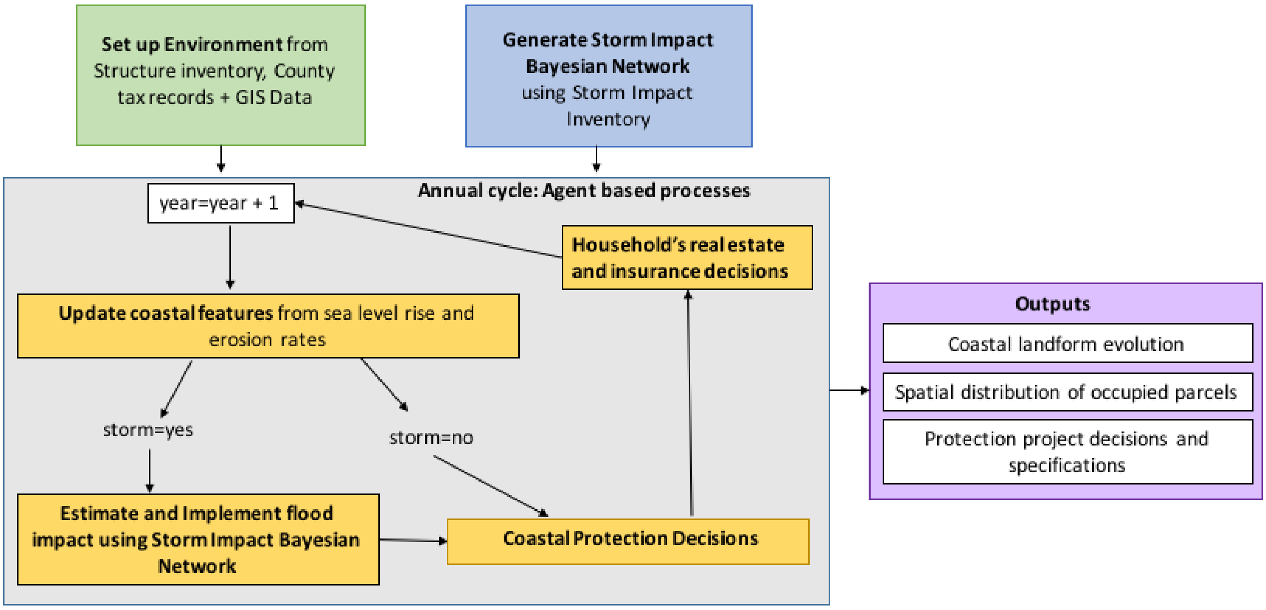

Figure 2. CoMOD Simplified Process Overview

3.2 Implementation of CoMOD in xDCI

CoMOD modelling process involves input from many disciplines such as coastal scientists, planners, social scientists, economists. Tightly coupling the model process can prevent or hamper the engagement of some fields. For example, a social scientist or an economist might not want to understand the underlying technology of creating the storm impact database to use/modify CoMOD. Whereas the partitioned process will let the social scientist work on the rules governing the interactions between the household agents without gaining the expertise to use a coastal modelling tool. Packaging the parts of the model into self-containing docker-based applications creates reusable tools which can be plugged into other workflows, thus enabling users to easily leverage the existing efforts, creating a stronger research community. By isolating the applications from one another, the researchers will be provided the ability to modify the existing tools and test new theories without considering the compatibility with the underlying technologies in the other docker-based applications. Furthermore, the researchers can plug their separate docker-based applications into the workflow, enabling test of new or modified apps without disrupting the usage of other researchers.

For the pilot study, we partitioned the CoMOD model process into three reusable docker containers (Figure 3):

- CoMOD Environment Creation: This docker-based application takes the geospatial building, topography, shoreline databases and creates input files for generation of topographic and cadastral layer for the coastal town environment. An open source software, GRASS GIS (Neteler et al., 2012), is utilized in this process to determine the properties of the coastal landforms and the spatial relationships between the structures and shoreline.

- Storm Impact Database Generation: This docker-based application shifts through tens of thousands of binary XBeach simulation results and determines the morphological change and inundation depth for each simulation. Then the results of the analysis are stored in a text file relating the pre-storm morphological properties of the coast and hydraulic boundary conditions of the storm to the morphological change in coast and inundation depth. Open source software, GNU Octave, is employed to handle the processes in this docker-based application.

- Coastal Town Agent-Based Model: The rules governing the interactions in-between agents and their environment, as well as, the rules describing the change of environment is encapsulated in this docker-based application. The outputs from the other two docker-based applications as well as some user defined inputs (i.e. nourishment dimensions) are used for initialization of this model. Then using Netlogo, the docker-based application simulates evolution of the coastal town through time, following the aforementioned processes and specified rules.

The analysis supported by the workflow could easily be used against different sets of input data. For example, through the docker-based applications, researchers and planners can generate location specific model environments enabling expanding the user base of the model, boosting interinstitutional use. The output data is also easily staged for further processing for additional research insights. Batches of simulations can be performed to observe household occupation and mitigation decisions, and evaluate the community viability under varying initial community, environment, and global climate scenarios. These results allow for scenario-based analyses, which could be useful for communities as they plan for the effects of climate change and sea level rise. More details about CoMOD research results can be found in Karanci et al., 2017.

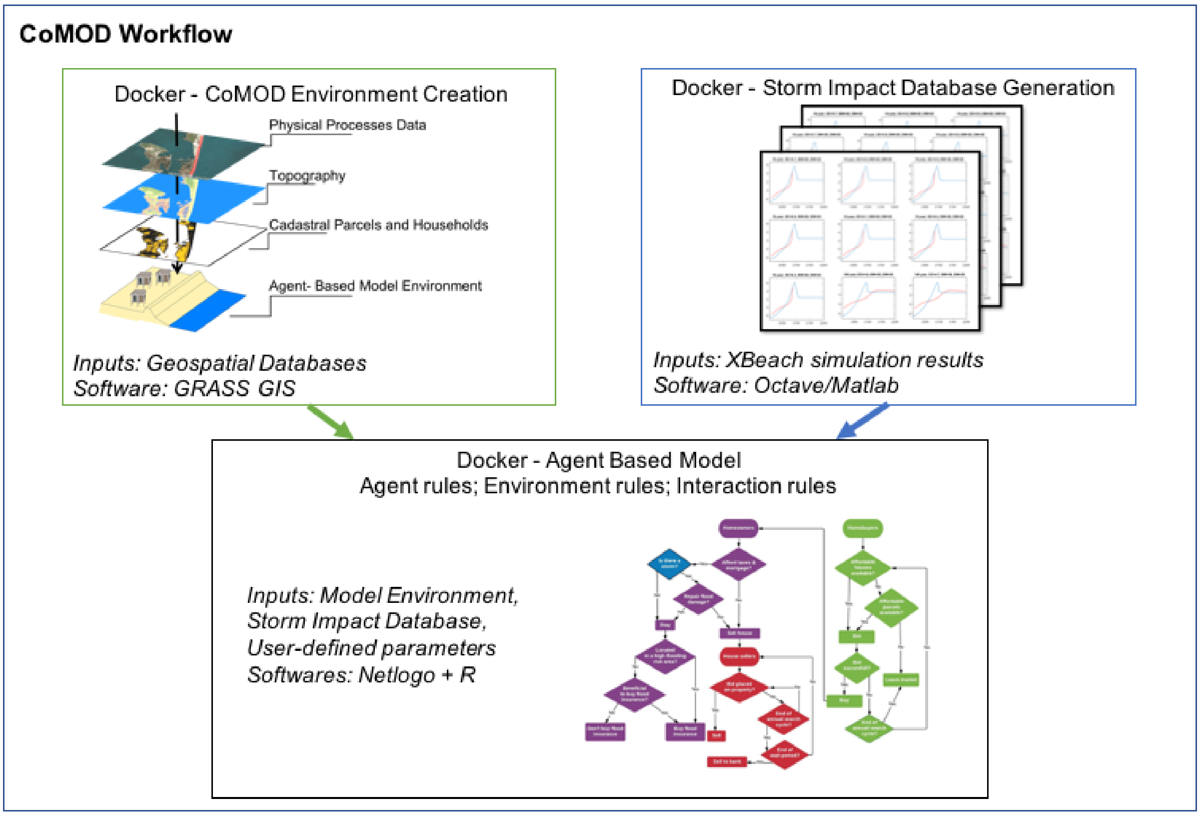

Figure 3. CoMOD Workflow and Docker Apps

4 CONCLUSION

This paper outlined the implementation of the CoMOD as a pilot use case in the xDCI cyberinfrastructure. The use case integrated various open source components to provide access to robust storage, data management tools, access to compute resources, and the ability to integrate their particular science models and algorithms as applications that are much more easily manipulated.

Utilizing the xDCI-enabled cyberinfrastructure for this pilot case facilitated 1) the development of reusable tools and apps that allow for customization by different disciplines, 2) connectable workflows to further enhance the research, 3) straightforward data input and output management, 4) collaborative research across institutions, and 5) a platform for community-specific analytical capabilities, data and vocabulary.

The data share capabilities within the research community was also recognized as a significant benefit for this use case which involves tens of thousands of XBeach simulations. Through the cyberinfrastructure, this database can be reused, which will lower the computation time and resources required immensely. Furthermore, the database can be expanded by supplementing more simulations using different morphological and hydraulic boundary conditions. The community can leverage the database as a valuable asset for numerous future coastal hazard risk and mitigation research.

The cyberinfrastructure will increase the productivity of collaborations permitting specialists in the application domain to create components incorporating their expertise, and providing a focus on discourse in development at a level far more higher-level than a programming language. Establishment of a cyberinfrastructure for component-based software development can also realize significant benefits through reduced software project costs, enhanced software quality, and expanded applicability of less expensive technology. Additionally, the automation will increase the productivity in software development by improving software quality characteristics, and reducing the time to develop, test, and certify software.

Through the iRODS grid, the cyberinfrastructure also provides the option to more easily integrate metadata, data management and curation workflows. It is anticipated that these types of approaches will become more prevalent as the need to conduct interdisciplinary research grows to address existing and future grand challenge science questions.

Because components can be developed and maintained by different groups, yet can still be coupled within the cyberinfrastructure, interdisciplinary teams can be more effective in tackling the big long-term questions in research. This could shift the focus of the researchers in the community from a focus on individual goals, which are generally of a short-term character, and provide incentives leading to a greater emphasis on effective collaboration in interdisciplinary, interinstitutional settings.

However even if all the strictly technical challenges of reuse are resolved, social, behavioral and programmatic barriers may still prevent realization of the full potential of reuse (Lenhardt et al., 2016). Recognizing the constraints imposed by governance, reuse can only succeed through shaping changes in stakeholder behavior and decision making regarding reuse.

REFERENCES

Lenhardt, W.C., Conway, M., Scott, E., Blanton, B., Krishnamurthy, A., Hadzikadic, M., Vouk, M., Wilson, A., 2016. Cross-institutional research cyberinfrastructure for data intensive science. In: High Performance Extreme Computing Conference (HPEC), 2016 IEEE, pp. 1-6.

Karanci, A., Velásquez-Montoya, L., Paniagua-Arroyave, J.F., Adams, P.N., Overton, M.F., 2018. Beach Management Practices and Occupation Dynamics: An Agent-Based Modeling Study for the Coastal Town of Nags Head, NC, USA. In: Beach Management Tools-Concepts, Methodologies and Case Studies. pp. 373-395. Springer, Cham.

Castronova, T.A., Goodall, J.L., 2010. A generic approach for developing process-level hydrologic modeling components. Environ. Modell. Softw. 25, 819-825.

Hallegatte, S., Green, C., Nicholls, R.J., Corfee-Morlot, J., 2013. Future flood losses in major coastal cities. Nature climate change 3(9):802

Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S., Rockwood, R.C., Selig, E.R., Selkoe, K.A., 2015. Spatial and temporal changes in cumulative human impacts on the world/'s ocean. Nature communications 6.

Hapke, C.J., Kratzmann, M.G., Himmelstoss, E.A., 2013. Geomorphic and human influence on large-scale coastal change. Geomorphology 199, pp. 160-170.

Idaszak, R., Tarboton, D., Yi, H., Christopherson, L., Stealey, M., Miles, B., Dash, P., Couch, A., Spealman, C., Ames, D., Horsburgh, J., 2016. Software Engineering for Science. HydroShare – A case study of the application of modern software engineering to a large distributed federally-funded scientific software development project. J. Carver, N. Chue Hong, and G. Thiruvathukal, Eds. CRC Press, Chap. 10, pp. 219–236.

Idaszak, R., Ahalt, S., Bradford, K., Calloway, C., Castillo, C., Coposky, J., Crabtree, J., Davis, S., Guo, Y., Jiang, F., Krishnamurthy, A., Lander, H., Lenhardt, W.C., Rajasekar, A., Robasky, K., Russell, T., Scott, E., Sizemore, D., Stealey, M., Xu, H., Yi, H., Zhang, W., 2017. xDCI - Accelerating Data Cyberinfrastructure and Research for Community Science Gateways. In: Proceedings of Gateways 2017, University of Michigan, Ann Arbor, MI. October 23-25, 2017.

Krishnamurthy, A., Bradford, K., Calloway, C., Castillo, C., Conway, M., Coposky, J., Guo, Y., Idaszak, R., Lenhardt, W.C., Robasky, K., Russell, T., 2017. xDCI, a data science cyberinfrastructure for interdisciplinary research. In: High Performance Extreme Computing Conference (HPEC), 2017 IEEE . pp. 1-7.

Lazarus, E.D., Ellis, M.A., Murray, A.B. and Hall, D.M., 2016. An evolving research agenda for human–coastal systems. Geomorphology, 256, pp.81-90.

Mileti, D., Noji, E.K., 1999. Disasters by Design. Joseph Henry Press, Washington.

Neteler, M., Bowman, M.H., Landa, M., Metz, M., 2012. GRASS GIS: A multi-purpose open source GIS. Environ Model Soft 31, 124–130.

Prati, G., Albanesi, C., Pietrantoni, L. and Airoldi, L., 2016. Public perceptions of beach nourishment and conflict management strategies: A case study of Portonovo Bay in the Adriatic Italian Coast. *Land Use Policy*, *50*, pp.422-428.

Tarboton, D., Idaszak, R., Horsburgh, J., Heard, J., Ames, D., Goodall, J., Band, L., Merwade, V., 2014. A resource centric approach for advancing collaboration through hydrologic data and model sharing. In: Proceedings of the 11th International Conference on Hydroinformatics, New York City, USA, 2014.

Van Dongeren, A., Ciavola, P., Martinez, G., Viavattene, C., Bogaard, T., Ferreira, O., Higgins, R. and McCall, R., 2017. Introduction to RISC-KIT: Resilience-increasing strategies for coasts. Coastal Engineering.