

OGC Testbed-13

*NA001 Climate Data Accessibility for Adaptation  
Planning*

# Table of Contents

1. Summary .....	4
1.1. Requirements .....	5
1.1.1. Deliverables .....	5
1.2. Key Findings and Prior-After Comparison .....	7
1.2.1. Key Findings Relevant to DSI Modeling Thread in Testbed-13 .....	7
1.2.2. Other Examples of Advancing in Academia and Science .....	7
1.3. What does this ER mean for the Working Group and OGC in general .....	8
1.4. Document contributor contact points .....	9
1.5. Future Work .....	9
1.6. Foreword .....	9
2. Terms and definitions .....	10
2.1. Abbreviated terms .....	10
3. Overview .....	11
3.1. Summary Chapter .....	11
3.2. Components Chapter .....	11
3.3. Future Work and Recommendations Chapter .....	11
4. Components & Component Scenario .....	12
4.1. Setting the story .....	12
4.2. Component Overview Chart .....	12
4.3. Individual components .....	13
4.3.1. NA101 - Agricultural Scientist Client .....	13
4.3.2. NA102 Non-Scientist or Analyst Client .....	16
4.3.3. NA101b, NA102b Combined Scientist, Non-Scientist Client .....	19
4.3.4. NA103 Prediction WPS .....	24
4.3.5. NA104 WCS access to climate data .....	37
4.4. Component Implementation & Scenario for the Modeling Package .....	38
4.4.1. TIE Component Implementation .....	38
4.4.2. Scenario "Blueprint" for the Result Demonstration Video .....	39
5. Recommendations and Future Work .....	40
5.1. Observations and caveats .....	40
5.2. Change requests .....	40
Appendix A: Revision History .....	42

Publication Date: 2018-01-11

Approval Date: 2017-12-07

Posted Date: 2017-11-08

Reference number of this document: OGC 17-022

Reference URL for this document: <http://www.opengis.net/doc/PER/t13-NA001>

Category: Public Engineering Report

Editor: Guy Schumann

Title: OGC Testbed-13: NA001 Climate Data Accessibility for Adaptation Planning

---

## **OGC Engineering Report**

### **COPYRIGHT**

Copyright © 2018 Open Geospatial Consortium. To obtain additional rights of use, visit <http://www.opengeospatial.org/>

### **WARNING**

This document is not an OGC Standard. This document is an OGC Public Engineering Report created as a deliverable in an OGC Interoperability Initiative and is not an official position of the OGC membership. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an OGC Standard. Further, any OGC Engineering Report should not be referenced as required or mandatory technology in procurements. However, the discussions in this document could very well lead to the definition of an OGC Standard.

## LICENSE AGREEMENT

Permission is hereby granted by the Open Geospatial Consortium, ("Licensor"), free of charge and subject to the terms set forth below, to any person obtaining a copy of this Intellectual Property and any associated documentation, to deal in the Intellectual Property without restriction (except as set forth below), including without limitation the rights to implement, use, copy, modify, merge, publish, distribute, and/or sublicense copies of the Intellectual Property, and to permit persons to whom the Intellectual Property is furnished to do so, provided that all copyright notices on the intellectual property are retained intact and that each person to whom the Intellectual Property is furnished agrees to the terms of this Agreement.

If you modify the Intellectual Property, all copies of the modified Intellectual Property must include, in addition to the above copyright notice, a notice that the Intellectual Property includes modifications that have not been approved or adopted by LICENSOR.

THIS LICENSE IS A COPYRIGHT LICENSE ONLY, AND DOES NOT CONVEY ANY RIGHTS UNDER ANY PATENTS THAT MAY BE IN FORCE ANYWHERE IN THE WORLD. THE INTELLECTUAL PROPERTY IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NONINFRINGEMENT OF THIRD PARTY RIGHTS. THE COPYRIGHT HOLDER OR HOLDERS INCLUDED IN THIS NOTICE DO NOT WARRANT THAT THE FUNCTIONS CONTAINED IN THE INTELLECTUAL PROPERTY WILL MEET YOUR REQUIREMENTS OR THAT THE OPERATION OF THE INTELLECTUAL PROPERTY WILL BE UNINTERRUPTED OR ERROR FREE. ANY USE OF THE INTELLECTUAL PROPERTY SHALL BE MADE ENTIRELY AT THE USER'S OWN RISK. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR ANY CONTRIBUTOR OF INTELLECTUAL PROPERTY RIGHTS TO THE INTELLECTUAL PROPERTY BE LIABLE FOR ANY CLAIM, OR ANY DIRECT, SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES, OR ANY DAMAGES WHATSOEVER RESULTING FROM ANY ALLEGED INFRINGEMENT OR ANY LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR UNDER ANY OTHER LEGAL THEORY, ARISING OUT OF OR IN CONNECTION WITH THE IMPLEMENTATION, USE, COMMERCIALIZATION OR PERFORMANCE OF THIS INTELLECTUAL PROPERTY.

This license is effective until terminated. You may terminate it at any time by destroying the Intellectual Property together with all copies in any form. The license will also terminate if you fail to comply with any term or condition of this Agreement. Except as provided in the following sentence, no such termination of this license shall require the termination of any third party end-user sublicense to the Intellectual Property which is in force as of the date of notice of such termination. In addition, should the Intellectual Property, or the operation of the Intellectual Property, infringe, or in LICENSOR's sole opinion be likely to infringe, any patent, copyright, trademark or other right of a third party, you agree that LICENSOR, in its sole discretion, may terminate this license without any compensation or liability to you, your licensees or any other party. You agree upon termination of any kind to destroy or cause to be destroyed the Intellectual Property together with all copies in any form, whether held by you or by any third party.

Except as contained in this notice, the name of LICENSOR or of any other holder of a copyright in all or part of the Intellectual Property shall not be used in advertising or otherwise to promote the sale, use or other dealings in this Intellectual Property without prior written authorization of LICENSOR or such copyright holder. LICENSOR is and shall at all times be the sole entity that may authorize you or any third party to use certification marks, trademarks or other special designations to

indicate compliance with any LICENSOR standards or specifications.

This Agreement is governed by the laws of the Commonwealth of Massachusetts. The application to this Agreement of the United Nations Convention on Contracts for the International Sale of Goods is hereby expressly excluded. In the event any provision of this Agreement shall be deemed unenforceable, void or invalid, such provision shall be modified so as to make it valid and enforceable, and as so modified the entire Agreement shall remain in full force and effect. No decision, action or inaction by LICENSOR shall be construed to be a waiver of any rights or remedies available to it.

None of the Intellectual Property or underlying information or technology may be downloaded or otherwise exported or reexported in violation of U.S. export laws and regulations. In addition, you are responsible for complying with any local laws in your jurisdiction which may impact your right to import, export or use the Intellectual Property, and you represent that you have complied with any regulations or registration procedures required by applicable law to make this license enforceable.

# Chapter 1. Summary

This Engineering Report describes all Testbed-13 activities relating to the Climate Data Accessibility for Adaptation Planning requirements of the National Aeronautics and Space Administration (NASA). It discusses relevant experiences made during implementation including recommendations to the sponsor, and provides resulting standards change requests to the appropriate working groups. Additionally, it develops best practices for data and model integration and serves as a guidance document to work with NASA Earth Science Data System (ESDS) working groups and externally provided data. The added value of this Engineering Report is to improve interoperability and to advance location-based technologies and realize innovations with regards to NASA Climate Data and NASA ESDS objectives.

Testbed 13 supports the ESDS Vision [<https://earthdata.nasa.gov/earth-science-data-systems-program/about-the-esds-program/esds-mission>] stated as:

"Make NASA's free and open earth science data interactive, interoperable and accessible for research and societal benefit today and tomorrow".

The ESDS Program oversees the lifecycle of earth science data with the principal goal of maximizing the scientific return from NASA's missions and experiments for research and applied scientists, decision makers and society at large. Testbed-13 results are applicable to this ESDS Goal:

"Ensure access to data and services that are useful and usable by a wide community of user."

To meet this goal, the Climate Data Accessibility thread in Testbed 13 sought to:

- Improve climate data accessibility for the scientist and non-scientist.
- Broaden climate adaptation essentials.
- Collaborate with the Federation of Earth Science Information Partners (ESIP) [<http://www.esipfed.org>].
- Work with the sponsor to continually improve components so that they are fit-for-purpose throughout Testbed-13 lifecycle.

A pre-Testbed-13 concept development study executed by OGC identified a number of datasets, portals, data centers, simulation models, and other web services suitable for addressing the stated goals of the Testbed-13 modeling thread. OGC have issued a Request for Information (RFI) to its membership to solicit interest, experiences, and data and model availability. In this context, OGC closely cooperates with ESIP to ensure knowledge transfer with ESIP partners. ESIP is an open, networked community that brings together science, data and information technology practitioners. ESIP's mission is to support the networking and data dissemination needs of its members and the global Earth science data community by linking the functional sectors of observation, research, application, education and use of earth science. The OGC and ESIP have a Memorandum of Understanding (MoU) that promotes coordination between the two organizations on topics of common interest to facilitate effective knowledge transfer partnerships.

Testbed-13 participants were requested to help complete the set of available material. The overall goal was to gain experiences from the integration process, communicate the best practices and recommendations back to the Sponsor on potential improvements that will further simplify the

integration process, understand interoperability issues arising from different formats, interfaces, protocols, and access policies. The Climate Data Accessibility for Adaptation Planning work package was demonstrated in a scenario defined in this document.

## 1.1. Requirements

Figure 1 illustrates the work items and requirements in the context of Climate Data Accessibility for Adaptation Planning.

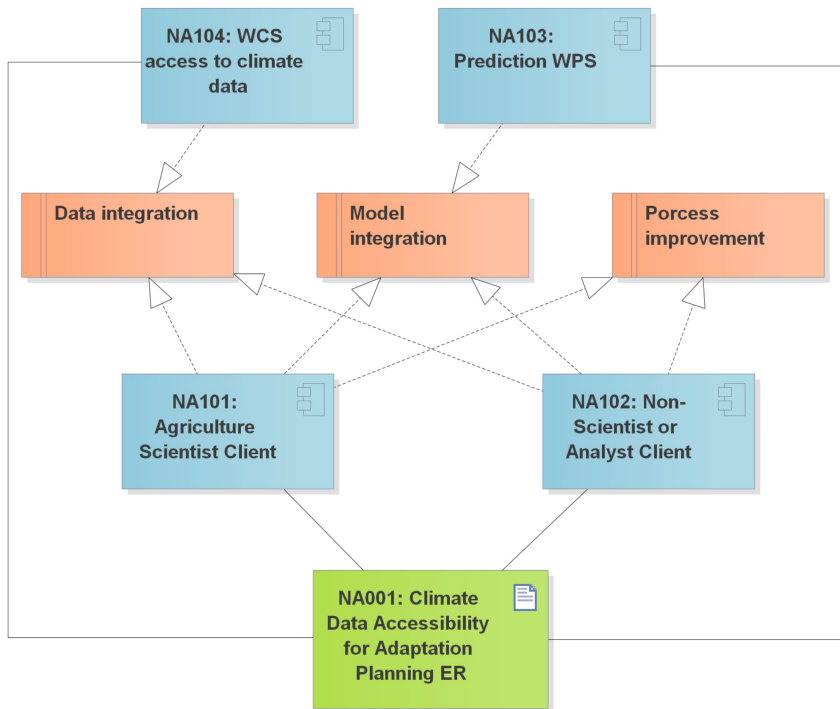


Figure 1. Climate data accessibility for adaptation planning requirements and work items.

The Testbed-13 deliverables described in this document meet the following requirements:

1. Data integration: Demonstrate the integration and analysis of earth observation data made available at various portals (e.g. <http://www.prepdata.org>) and data serving systems in the context of use cases one and two described in the Deliverables section.
2. Model integration: Integrate simulation models into the demonstration that can be parameterized and executed on demand through standardized Web Application Programming Interfaces (API) such as the OGC Web Processing Service (WPS) interface standard.
3. Process improvements: All demonstrations implement the context described above. The work captures all experiences made during the integration process and derives best practices and recommendations for improvements that simplify workflows for scientists and non-scientists. This is done whilst considering account data, data access and data formats, models and model parameterization and execution, portals, access policies, and other elements that influence the integration experience.

### 1.1.1. Deliverables

The following list identifies the work items assigned to the requirements, note that work items are

not exclusive to the work package described in this Engineering Report (ER).

- NA101: Agriculture Scientist Client – This client component provides access to climate data, agriculture predictions and other data relevant for an analyst to assess the effects of drought on mass migration. The client supports the integration of on demand models through a WPS 2.0 interface.
- NA102: Non-Scientist or Analyst Client – This client component provides access to climate data and other data relevant for an agriculture scientist to predict the effects of drought on crop production. This client also provides access to on demand models through the WPS 2.0 interface.
- NA103: Prediction WPS – This component is made accessible through a WPS and enables access and control of predictive models relevant to drought and agriculture prediction based on climate and other data. The service supports different parameterization options to serve both scientists and non-scientists made accessible through the clients in NA101 and NA102.
- NA104: WCS access to climate data – This Web Coverage Service (WCS) will provide access to NASA Climate data for use by the other components in Testbed-13. The climate data shall be made available in formats and resolutions appropriate for scientists and non-scientists.

The following use-cases shall be considered:

- Use Case 1: Agriculture Researcher considering the effects of climate on crop production needs easy access to climate prediction/data for ingest into crop forecasting models, potentially requiring temporal and spatial sub-setting of climate model variables relevant to rain and soil conditions. Access to the right datasets, with effective system-to-system API, GIS based data formats (EG Shape, GeoTIFF) in addition to providing option to ingest subset HDF, NetCDF formats.
- Use Case 2: Use by non-scientist or analysts to better understand climate impacts to populations by accessing climate model prediction data for ingest into local GIS based systems for population and critical infrastructure overlay.

These deliverables are designed to broaden distribution access and formats of climate reanalysis, climate model data, and climate based observational data for ease of system-to-system ingest, access (EG: API, WCS, other), and data delivery formats (File: HDF, NetCDF, Shapefile, GeoTIFF) for ingest by the scientist and non-climate scientist.

The two use cases for the system are represented by the two client work items, as discussed earlier. The main differences between the clients are as follows:

- The amount of configuration required to produce a successful output.
- The ability to upload manually configured files for running APSim, an Agricultural Production Systems sIMulator.
- The constraints associated with the non-scientist client (i.e. the interface only offers a subset of the functionality to the user).
- Contrarily, the freedom offered with the expert client to configure and execute the model without constraint.

The basis for the requirements is that many people needing climate reanalysis and climate model



data (some cases observational data) do not have effective system-to-system access to climate model data and do not necessarily have geospatial expertise, i.e. understanding of file formats such as HDF, NetCDF and GRIB. They are also not necessarily climate scientists yet crop yield predictions based upon climate change are required for their work.

## 1.2. Key Findings and Prior-After Comparison

This section describes the status of discussion in the OGC Emergency & Disaster Management/Law Enforcement And Public Safety (EDM/LEAPS) Domain Working Group (DWG), identified as being most relevant for the addressed topic. This document is reviewed by the EDM/LEAPS DWG to ensure that the latest developments have been considered.

### 1.2.1. Key Findings Relevant to DSI Modeling Thread in Testbed-13

The flowchart of the DSI Modeling work package as illustrated in figures 1 and 2 develops full interoperability between geospatial data, models and their parameterization as well as visualization of data and model outputs. More specifically, with reference to NASA Climate Data and Accessibility, this Testbed-13 thread shows proper use of NASA data, primarily the Modern-Era Retrospective Analysis for Research and Applications (MERRA) version 2 dataset, for crop yield prediction. It also introduces geographic intelligence to the crop modeling process, something it is lacking at the moment for the particular model used.

Specifically, for the MERRA-2 data used, a key finding is that, although NASA data is freely and openly accessible to all, the MERRA-2 data, as also other NASA data, are scientific data that are often included in large structured files. Finding data in the large number of collections of files NASA holds proved to be very time-consuming. If the NASA (or other) cataloging systems supported searching for *variables* within datasets/collections that have specific *units of measure*, data location would have been less of a manual chore. Once found, the data files typically contain many variables and have their own variable naming which are oftentimes ambiguous. Moreover, when served in a web service setting using WCS 2.0, temporal subsetting proves difficult (WCS 2.1 addresses this, but the server will have to implement that optional feature). In this testbed, under the DSI Modeling work, the spatial subsetting of the MERRA-2 NetCDF file has been tested successfully for the Bay area. However, the temporal subsetting was not tested and such a test is recommended under the future work section. Other, less significant caveats, and also advantages of NASA HDF5-, NetCDF-, and GeoTIFF-formatted data are reported throughout this ER.

One of the key goals for this ER was to make valuable but complex datasets such as NASA's MERRA-2 more accessible. This has been achieved by exploring the ease of access and performance of common clients. However, it remains a fact that datasets like the MERRA-2 are high-level scientific data and therefore require operations performed within a scientist client context. In fact, one of the recommendations for future work is that it should be explored how a scientist client could be used to predefine or preset the data visualization and simple parameterization for a non-scientist client and/or non-expert user.

### 1.2.2. Other Examples of Advancing in Academia and Science

The DSI Modeling work package for the Testbed-13 requirements NA001, NA101, NA102, NA103 and NA104 looked into how to efficiently couple data and numerical models such that they operate

through web protocols and results as well as data visualized by clients. For instance, the National Science Foundation is already supporting similar activities in academia and science. The Community of Surface Dynamics Modeling Systems (CSDMS; <http://csdms.colorado.edu/wiki>) is establishing a web interface that lets users configure and operate a model. Although CSDMS mostly uses a web interface to run single or coupled models, they have experimented in the past with coupling data as input to models, and will explore this more in the coming 5 years. Setting standards in how models and data should be configured is of key importance to make coupling work and the DSI Modeling work package of Testbed-13 helps realize this step.

### **1.3. What does this ER mean for the Working Group and OGC in general**

The purpose of the Emergency & Disaster Management (EDM) DWG is to promote and support the establishment of requirements and best practices for web service interfaces, models and schemas to enable the discovery, access, sharing, analysis, visualization and processing of information to the forecasting, prevention, response to and recovery from emergency and disaster situations. The mission lies in improving interoperability of geospatial products and other information consumables that can be shared across these communities. The following are the two main objectives for this ER:

- Identify interoperability standards gaps and opportunities to support improved EMDR information sharing, collaboration and decision making.
- Propose or encourage initiation of Interoperability Program studies, experiments, pilot initiatives, testbed threads or demonstrations to address technical, institutional and policy related interoperability challenges, and identify and engage the interest of potential sponsors for these activities.

The Law Enforcement And Public Safety (LEAPS) DWG promotes and supports the establishment of local, national, regional and international requirements and best practices for web service interfaces, data models and schemas for enabling the discovery, access, sharing, analysis, visualization and processing of information. This geospatial and temporal information is used comprehensively to address crime, terrorist activities and public safety incidents in an operationally effective way.

These two groups now form one DWG (EDM/LEAPS) given that their objectives and general purpose are very similar and overlapping for many applications, especially during disaster situations.

The objectives of this DWG are synergistic with the requirement and deliverables in the Testbed-13 Modeling Thread, i.e. to facilitate access and formats of NASA climate reanalysis, climate model data, and climate based observational data for ease of system-to-system ingest, access, and data delivery formats for ingest by the scientist and non-climate scientist. Many needing observations and model data do not have effective system-to-system access to models or data, however crop model predictions are still required for their work.

This situation is very common in the emergency response and disaster management sector, and hence the importance of this ER to the EDM/LEAPS DWG.

## 1.4. Document contributor contact points

All questions regarding this document should be directed to the editor or the contributors:

*Table 1. Contacts*

Name	Organization
Guy Schumann	Remote Sensing Solutions, Inc.
Albert Kettner	UC Boulder, DFO, Observer through Remote Sensing Solutions, Inc.
James Gallagher/Nathan Potter	OPeNDAP
Sam Meek	Helyx SIS
Lei Hu	George Mason University
Dean Hintz	Safe Software, Inc.

## 1.5. Future Work

Details on Future Work are given in the final chapter of this ER. Below is just a short list of the most significant items recommended for future work:

- WCS and/or clients should support some method of getting data at a position or set of positions over a time series without having to download the entire dataset or subset.
- In light of large-scale dynamic process-based computer simulations, spatial and temporal subsetting of a NetCDF output from such models that produce a temporally and spatially dynamic result would be needed.
- Test implementation of a transactional WPS (WPS-T), which would relieve burden on implementers to second guess requirements from users by providing a platform for them to execute their own processes at will.
- In the context of complex scientific data and data formats, explore how a scientist client could be used to predefine data visualization and simple parameterization for a non-scientist client and/or non-expert user.

## 1.6. Foreword

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

# Chapter 2. Terms and definitions

For the purposes of this report, the definitions specified in Clause 4 of the OWS Common Implementation Standard [OGC 06-121r9](https://portal.opengeospatial.org/files/?artifact_id=38867&version=2) [https://portal.opengeospatial.org/files/?artifact\_id=38867&version=2] shall apply. In addition, the following terms and definitions apply.

## 2.1. Abbreviated terms

- API Application Programming Interface
- APSim Agricultural Production Systems sIMulator
- CSDMS Community of Surface Dynamics Modeling Systems
- CU University of Colorado
- EDM Environmental and Disaster Management
- ER Engineering Report
- EPA Environmental Protection Agency
- ESDS Earth Science Data System
- ESRI Environmental Sciences Research Institute
- FME Feature Manipulation Engine
- GIS Geographic Information System
- GML Geographic Markup Language
- GMU George Mason University
- HDF Hierarchical Data Format
- LEAPS Law Enforcement And Public Safety
- MERRA-2 Modern-Era Retrospective analysis for Research and Applications, Version 2
- MoU Memorandum of Understanding
- NASA National Aeronautics and Space Administration
- NetCDF Network Common Data Form
- RFP Request For Proposals
- RSS Remote Sensing Solutions, Inc.
- TIE Test Interoperability Experiment
- WCS Web Coverage Service
- WPS Web Processing Service
- WMS Web Map Service
- WMTS Web Map Tile Service

# Chapter 3. Overview

This ER includes the following major chapters:

## 3.1. Summary Chapter

This chapter serves as introduction and background material to this ER and links to the LEAPS DWG and also the Testbed-13 Request for Proposals (RFP) document contents. It focuses on the DSI Modeling Component work, outlined in detail in the Components Chapter.

## 3.2. Components Chapter

This is the main chapter of this ER and includes detailed sections on the following Components and describes the connections between them as well as the respective TIEs. It also includes a brief scenario setting and framework of the demonstration video.

The different Components of the Modeling subthread are:

- WCS from OPeNDAP: Providing NASA (MERRA-2) data to the WPS and all clients (scientist expert & non-scientist clients)
- WPS from Helyx: APSim crop yield (agricultural) model process and implementation description
- Scientist Clients:—Agriculture expert client from George Mason University (GMU)—FME expert client for Safe Software
- Non-Scientist Clients: includes the following three clients:—Web GIS-style client from Remote Sensing Solutions (RSS/CU)—QGIS open-source desktop client from RSS/CU—FME client from Safe Software

## 3.3. Future Work and Recommendations Chapter

This chapter outlines general observations and caveats encountered and gives recommendations and suggestions for future work.

# Chapter 4. Components & Component Scenario

## 4.1. Setting the story

Agricultural yield prediction is a complicated process reliant on a variety of factors such as soil, relief, weather, crop type, and disease. Configuration of each of these variables and sub-variables requires knowledge beyond that of parties who could utilise the outputs from crop-yield predictions.

Rice crop/fields or grapes are interesting to simulate since these crops provide a compelling story about water usage and/or droughts. Rice was used as the main crop of interest for this testbed scenario since grapes require an extension module in the APSim prediction model which is beyond the scope of this sub-thread. Additionally, in the wider Bay area, Sacramento, Colusa, or Glenn counties have the bulk of California's rice crop, and this crop is already included in the APSim model.

As an introductory part to the demonstration video, a short video (<https://youtu.be/PHGRdovNGmM>) and high-res satellite imagery of the area that includes rice fields (<https://landsat.gsfc.nasa.gov/mapping-rice-managing-water/>) have been selected.

The NASA video shown in the link below has also been used since it tells a compelling story about how satellite data and rice fields are being linked to support migration bird habitat.

<https://phys.org/news/2015-09-satellite-migrating-birds-survive.html>

For the story building, the United States Environmental Protection Agency (EPA) document accessible from the link below has been used to set the story. This is all very general and the focus is on rice crop, and the impact of drought/climate change.

<https://www.epa.gov/sfbay-delta/what-are-challenges>

Also, a lot of interesting content material that has been used can be found here: <http://www.energy.ca.gov/2012publications/CEC-500-2012-033/CEC-500-2012-033.pdf>.

## 4.2. Component Overview Chart

Figure 2 illustrates the connections and interoperability between the various component parts discussed in this ER.

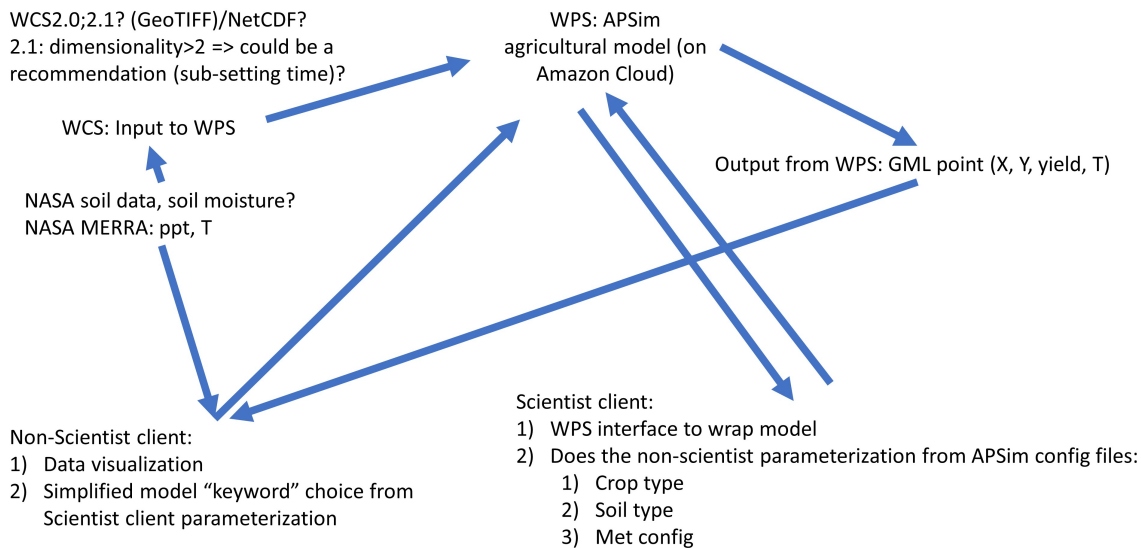


Figure 2. Overview flowchart

## 4.3. Individual components

The components that make up the scenario are described in this section.

### 4.3.1. NA101 - Agricultural Scientist Client

The Modeling thread contains two implementations of the scientist or analyst client: the agricultural scientist client by GMU, and the FME scientist client.

#### Agricultural Scientist Client by GMU

This client visualizes the result and the process steps of NA104 WCS and the NA103 Prediction WPS. For the agricultural scientist client scenario of the OGC Testbed-13 project, the agricultural researchers considering the effects of climate on crop production require access to climate and prediction data for ingest into crop forecasting models, potentially including temporal and spatial sub-sets of climate model variables relevant to rain and soil conditions. In this case, the focus is on providing the data access of NASA MERRA-2, exposed through the NA104 WCS 2.0, moreover, the client also supports the integration of on-demand crop yield prediction outputs from the NA103 Prediction WPS running the APSim crop model for the expert user. This client will greatly help researchers to access the correct datasets, with an effective API, and Geographic Information Systems (GIS) data formats including GeoTIFF and Geography Markup Language (GML).

The agricultural scientist client developed is based on the GMU Geobrain framework (<http://www3.csiss.gmu.edu/OnAS/>) , the client portal is located here: <http://www3.csiss.gmu.edu/OnAS/>

The area of interest in the scenario (described later) is the San Francisco bay area.

The client supports the following functions:

Support WCS MERRA-2 data as the crop yield prediction model input:

1. WCS interface parameterization (GetCapabilities, DescribeCoverage, and GetCoverage).
2. Input data visualization (GeoTIFF).

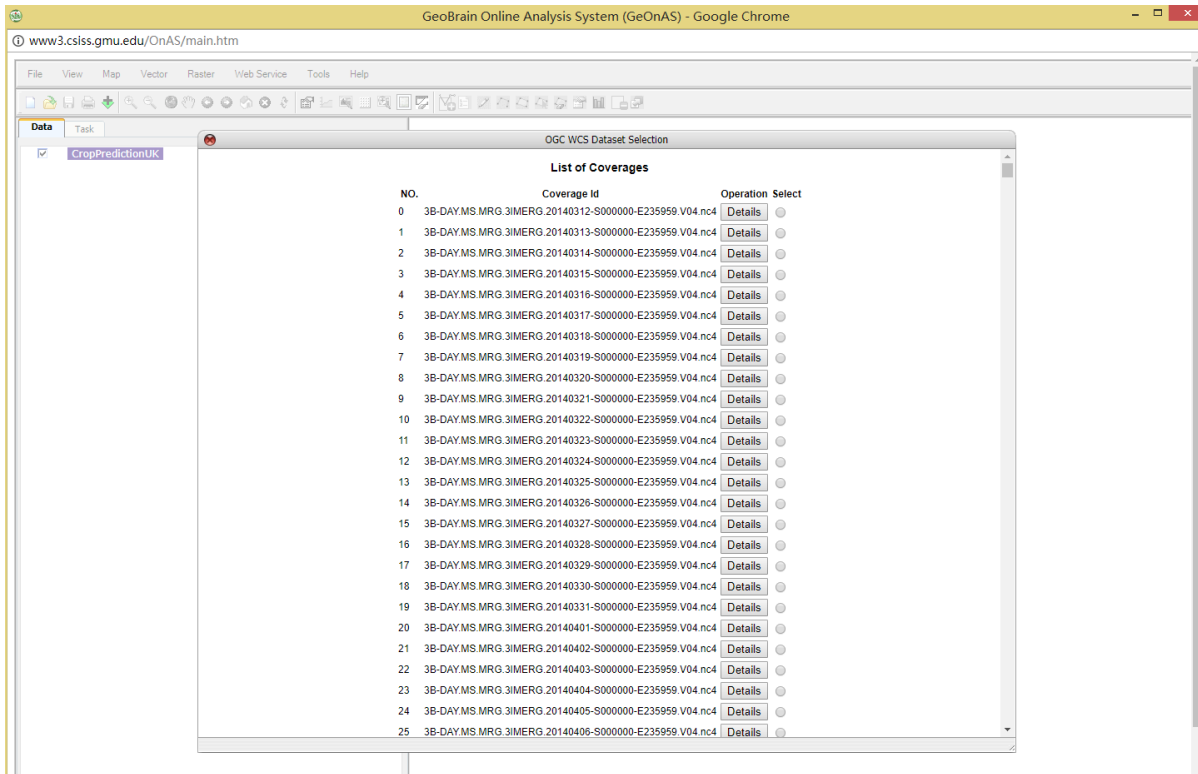


Figure 3. Visualization of WCS process steps

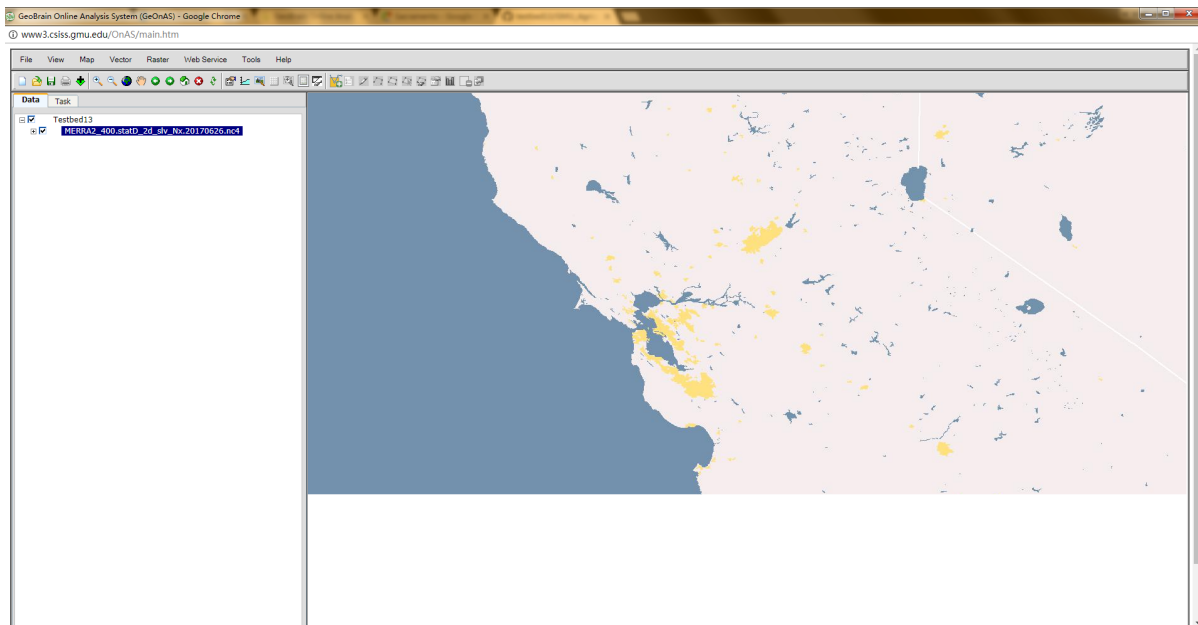


Figure 4. Visualization of the WCS climate data

Support APSim WPS service as the crop yield prediction model

1. WPS interface parameterization (GetCapabilities, DescribeProcess, Execute, GetStatus, and GetResult).
2. Support synchronous/asynchronous execution of WPS service.



3. Support the customization of latitude, longitude, the relevant soil, and meteorological data.
4. Output data visualization (GML point and bar chart).

The expert user could interact with the WPS through the Agricultural Scientist Client as it provides the ability to produce predictions based upon finely grained parameterization of the service.

1. The functionality of APSim is manually configured via direct input into the client, or by uploading a series of files that are required by the crop model to run correctly. The user enters parameters including soil variables (moisture at depth, temperature, nitrogen content), meteorological file (minimum and maximum temperatures, irradiance, precipitation, evapotranspiration), the beginning and end dates of the simulation, and the crop parameters (soil and water condition).
2. There is very little error catching available, as the user guidance is minimal beyond the process descriptions.
3. Additionally, the shared functionality is included to enable the user to present the system with a latitude and longitude to produce climate data for a location over a given date range as this maybe an easier method of determining input climate data.

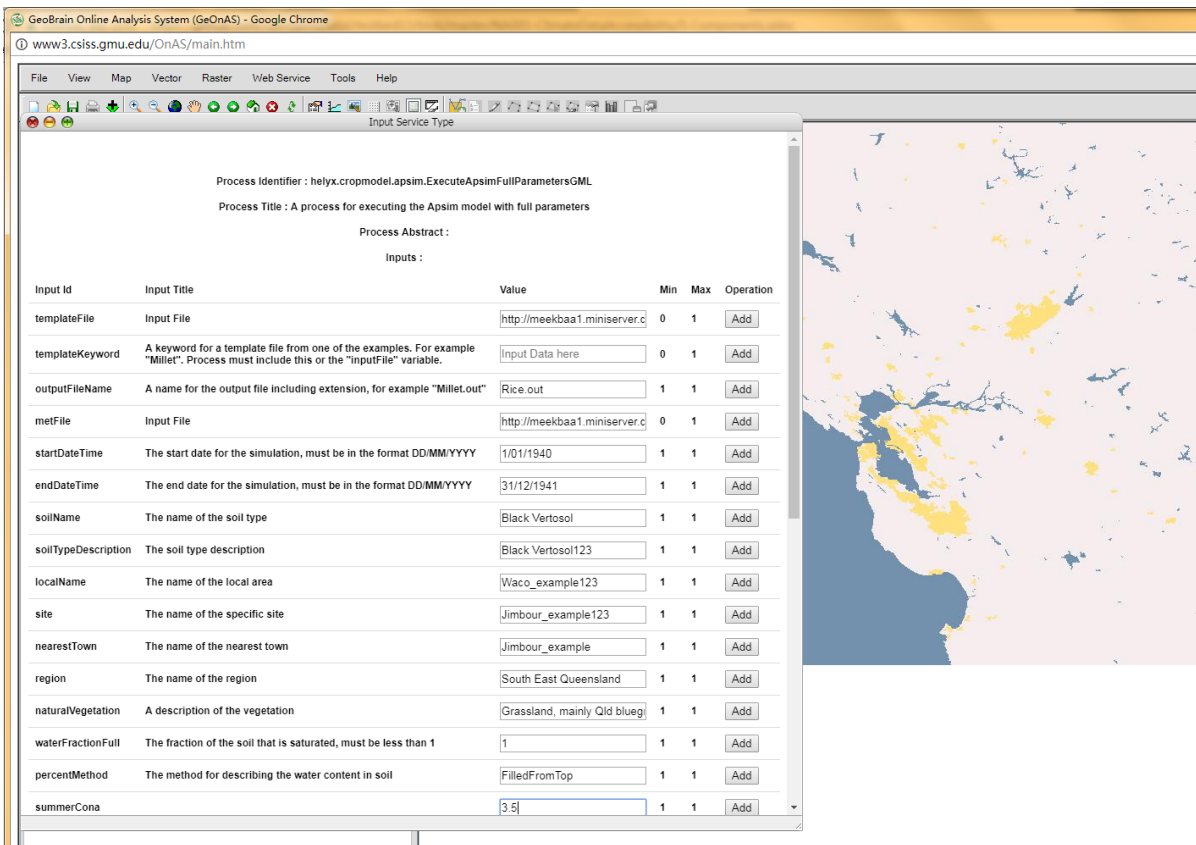


Figure 5. Visualization of the detailed configuration of WPS service

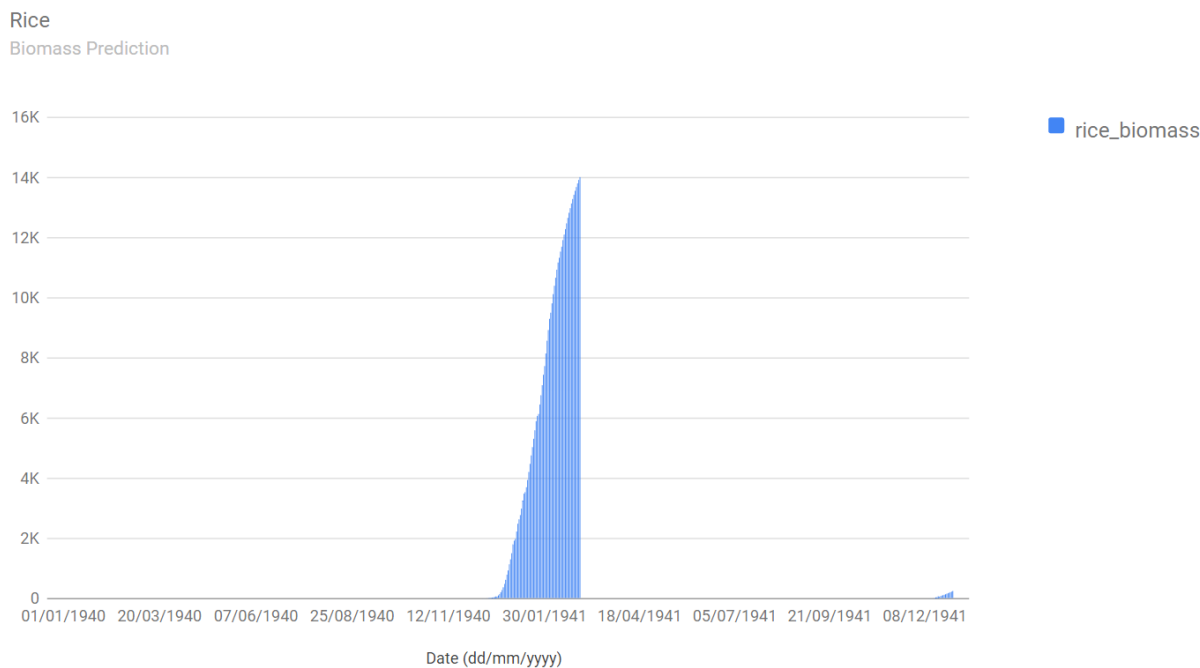


Figure 6. Visualization of the WPS result

### 4.3.2. NA102 Non-Scientist or Analyst Client

The Modeling thread contains two implementations of the non-scientist or analyst client, a variation on the QGIS client and a client in the Feature Manipulation Engine (FME).

#### Visualizing NASA data (input to APSim model)

Visualization of the MERRA-2 data in the Modeling thread is accomplished through utilization of Portal for ArcGIS' (<http://server.arcgis.com/en/portal/>) and GeoServer (<http://docs.geoserver.org>). WCS is simply imported into GIS desktop software such as QGIS or ArcMap and the provided properties can then be used in standard analysis. However, visualizing WCS services through Esri Portal is not straightforward, as is the case with WMS services where images (or maps) are simply "displayed". Portal for ArcGIS does not include methods to include WCS services and GeoServer 2.9.1 does not have cascading service capability for external provided WCS services. A workaround is to display the NASA data directly in its native raster format, GeoTIFF or NetCDF, or use the WMS endpoint for each MERRA data coverage. Visualization of WCS data was tested using two methods:

1. Direct incorporation of the WMS endpoint into Portal for ArcGIS.
2. Cascade the WMS endpoint before displaying in Portal for ArcGIS.

An Esri web map has been setup in Portal for ArcGIS for testing purposes, "Testbed-13 Non-scientist" displays both the WMS's mentioned above (see also Figure 7, where daily max precipitation is displayed for 1 day): <https://diluvium.colorado.edu/arcgis/home/webmap/viewer.html?webmap=0856d14f047b4db3aca9f045241aaa62>

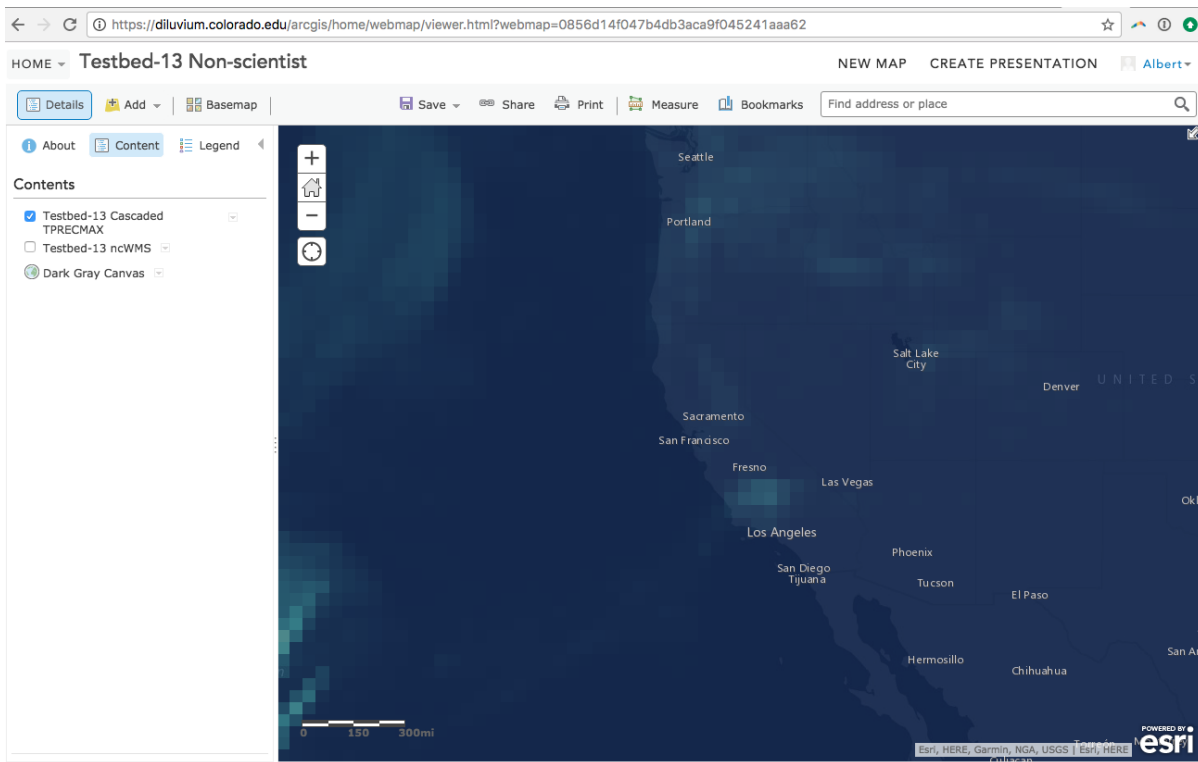


Figure 7. Display of Merra Maximum precipitation input data over California.

Both the cascading and the direct incorporation of the WMS endpoint have the disadvantage of only showing a single time-slice of the input data. Input data is currently not hosted with WMTS capability, therefore no time dimension visualization is available.

The second use case for the non-scientist client is to have the user select a simple pre-configured APSim model parameterization file that can then be submitted as a WPS service that executes the model and returns the APSim model output as a point GML, the methodology for this is described in the following sub-sections.

### Executing simple APSim setup in open-source GIS software QGIS

The latest stable version of QGIS (version 2.18.13 <http://qgis.org/en/site/>) is used to test the connection and execution the non-scientist APSim WPS process. The plugin 'WPS Client' needs to be installed to enable WPS client connection capability, and can be installed using the standard QGIS plugin repository. Once installed, the plugin can be activated or deactivate under the 'web' tab, by selecting 'WPS-client'. All that needs to be done to connect to a WPS service, is to provide the URL, in this case: <http://54.149.145.158/wps/WebProcessingService>. The 'non-scientist client' was selected from all available WPS services that will display. Once a specific WPS service is selected a WPS input parameters window will appear (Figure 8).

The output format can be chosen, after populating all required input parameters, following the provided standard input parameters from the WPS test client URL. Given output formats from the WPS are currently as shapefile or as text, the shapefile option is suitable for GIS software and the text output provides a table with APSim yields for given dates.

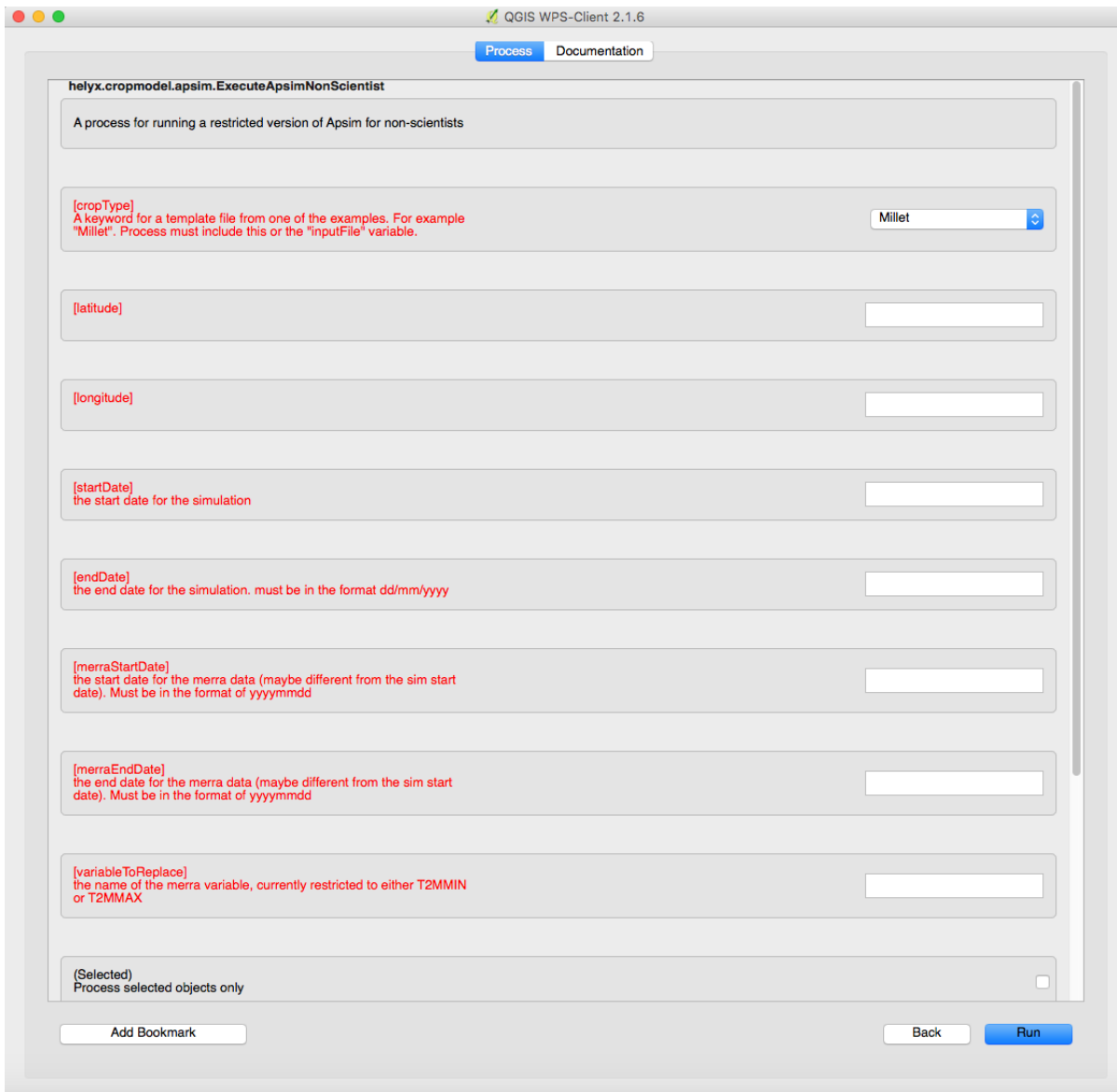


Figure 8. Input window when connecting to a WPS service in QGIS.

After executing the WPS service, the status of running the WPS is provided, in this case 'Execution of APSim Non Scientist finished successfully', and the output is returned as shapefile and automatically read in and displayed by QGIS. Results of the APSim simulations can be viewed by opening the attribute table of the shapefile (Figure 9).

	field_0	field_1
437	24/05/1981	1.623
438	26/05/1981	1.871
439	28/05/1981	2.466
440	30/05/1981	3.677
441	01/06/1981	7.132
442	03/06/1981	15.914
443	05/06/1981	29.287
444	07/06/1981	60.660
445	09/06/1981	116.952
446	11/06/1981	198.026
447	13/06/1981	329.732
448	15/06/1981	509.809
449	17/06/1981	623.498
450	19/06/1981	784.664
451	21/06/1981	931.098
452	23/06/1981	1059.810
453	25/06/1981	1175.943
454	27/06/1981	0.000
455	29/06/1981	0.000
456	01/07/1981	0.000

Advanced Filter (Expression) | 05,906,907,908,909,910,911 | Apply

Figure 9. Attribute table of the shapefile

The WPS shapefile can be saved locally. Using the 'time manager' plugin of QGIS, the changing yield values of rice can be displayed as for example changes in size of the point shapefile.

### 4.3.3. NA101b, NA102b Combined Scientist, Non-Scientist Client

The latest release of FME 2017.1 (Feature Manipulation Engine by Safe Software) was used to visualize results from the NA103 APSim Prediction WPS as both a scientist and non-scientist client. While in practice these modes would typically be accessed by different users, this test serves to evaluate the ease of integration of these OGC services with a COTS data conversion tool like FME which is often used to support data loading to common GIS applications and systems. FME was also used to visualize results from the NA104 ODAP MERRA-2 WCS 2.0 server.

#### Executing simple APSim setup in FME

The FME client is configured using an FME data transfer model called a workspace that submits requests and consumes responses from the APSIM WPS. This configuration has a choice to select either the scientist or non-scientist access mode. The scientist option employs the full parameter query and the non-scientist mode employs the abbreviated, non-expert query. The user also specifies the location in degrees for which APSim results are required. In both cases, when the client receives the WPS response, it parses the GML contained into features. See [Figure 10](#) and [Figure 11](#) from the reader workspace and [Figure 12](#) showing the client display.

One of the challenges with consuming the response GML was that it was difficult to parse. Every time a query is submitted, the feature member container name changes. The typical feature container in GML is just <featureMember>. In this case, the WPS response after one query had a feature container of: <n52:Feature-fa8574c1-6bb8-4c21-84e0-e50cda71eb2a gml:id="ID21">. Then the next time the query was run it became: <n52:Feature-02cac30c-f5ac-44a0-9d2d-cbac72adf5b2

gml:id="ID21">. So the feature container object had a unique id embedded in the tag. This may not be fully GML compliant because the application schema (xsd) would have to change with every result in order to read this by the schema. However, FME has an 'ignore schema' mode in the GML reader so this was used with the query 'featureMembers/\*'. This is essentially an XQuery that extracts all children of 'featureMembers'.

The other challenge was the time delay involved with using the non-scientist access mode. As discussed in more detail under the NA103 and NA104 sections, the non-scientist service mode involves longer run times on the server side due to the amount of data that is downloaded by the server to support each request. Also, the request process times appear to vary significantly, perhaps partially depending on the number of competing requests. Initially, access using the FME non-scientist client resulted in time-out errors. Increasing the timeout tolerance on the client resolved this, but testing is limited as every request / response iteration may involve a delay of several minutes.

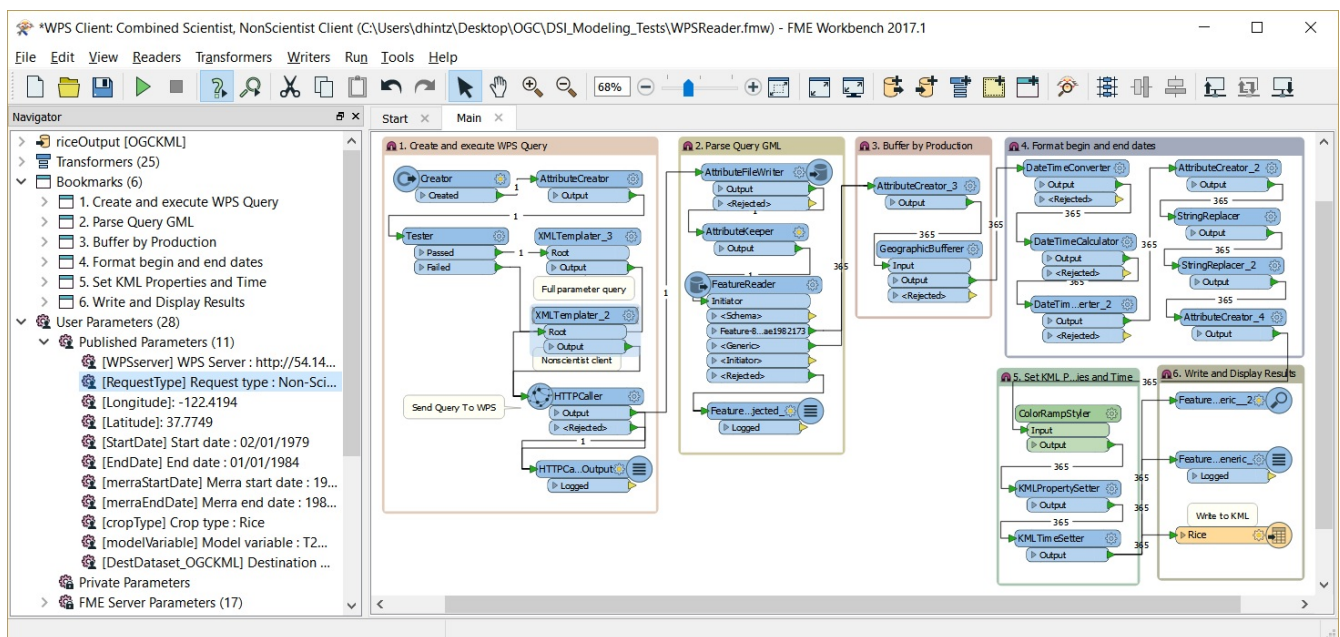


Figure 10. APSim WPS FME combined scientist and non-scientist client implemented using FME workspace

Parameter Name	Value
WPS Server	http://54.149.145.158/wps/WebProcessingService
Request type	Non-Scientist
Longitude	-122.4194
Latitude	37.7749
Start date	02/01/1979
End date	01/01/1984
Merra start date	19800101
Merra end date	19800102
Crop type	Rice
Model variable	T2MMAX
Destination Google KML File	z:\Desktop\OGC\DSI_Modeling_Tests\riceOutput.kml

Figure 11. APSim WPS FME combined client workspace parameters



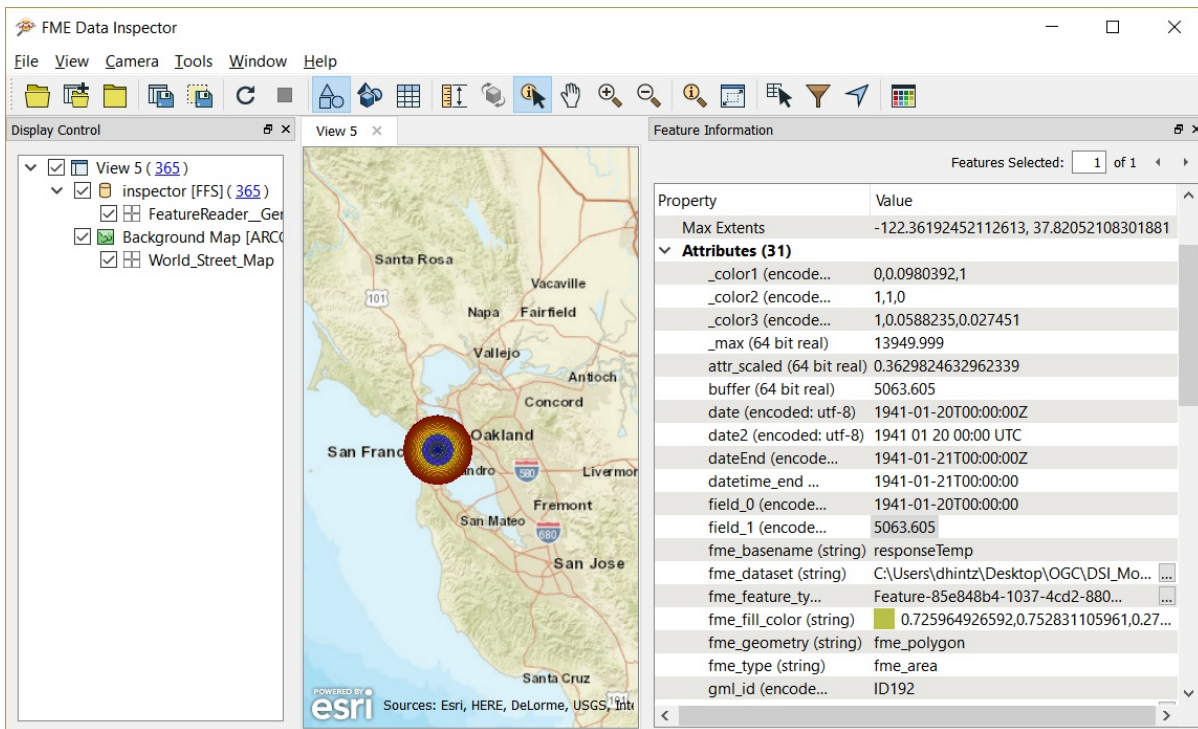


Figure 12. APSim WPS combined scientist and non-scientist client results displayed in FME Data Inspector

### Time-slider of Model Outputs

To make this output available in a more accessible format, after GML feature extraction, the FME WPS client workspace was then used to generate a time step OGC KML file with a georeferenced, buffered circle where the circle is colored according to a gradient from blue = low to red = high. The buffer amount is also set to the value of field1 in meters + 1000 - in the case shown below this is the value of T2MMAX. This enables the functionality to hit play on the time slider and watch the values of T2MMAX change over time. This was then visualized in NASA's WorldWind software - see [Figure 13](#) below.

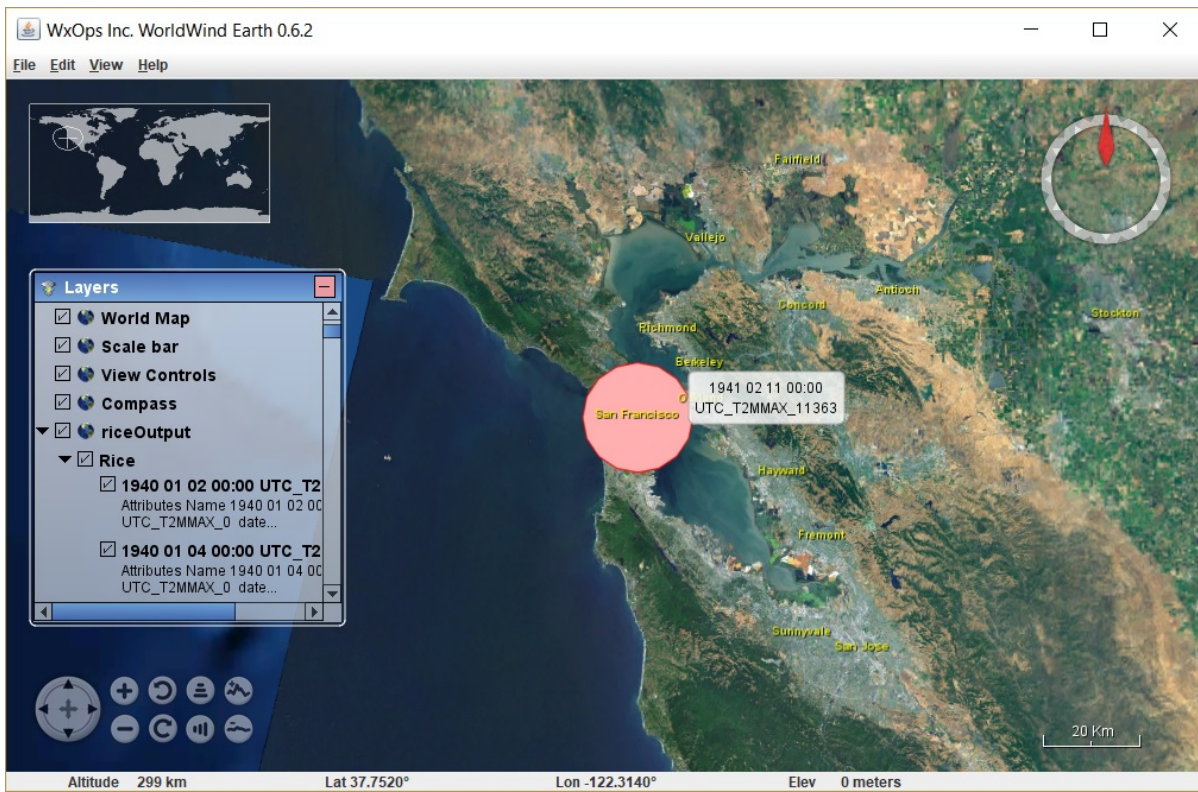


Figure 13. APSim WPS combined scientist and non-scientist client showing time series result converted to KML by FME for animated display in NASA Worldwind.

### FME WCS Client for NA104 ODAP MERRA-2 WCS

A custom WCS reader was implemented in FME to query the WCS that the WPS depends on directly. The goal was to test sub-setting the NetCDF time series from the WCS client to minimize the data required for transmission by defining the minimum x,y and t range required for the analysis. This sub-setting appeared to offer more responsive performance. For example, to retrieve a 4x4 grid subset of NetCDF data only took 2.5 seconds. This sub-setting approach should be considered in future testbeds in order to reduce the transmission and processing footprint for WCS client requests and downstream processing from any dependent WPS's. Below are screenshots of the FME WCS custom reader workspace, parameters and output as displayed in Data Inspector.



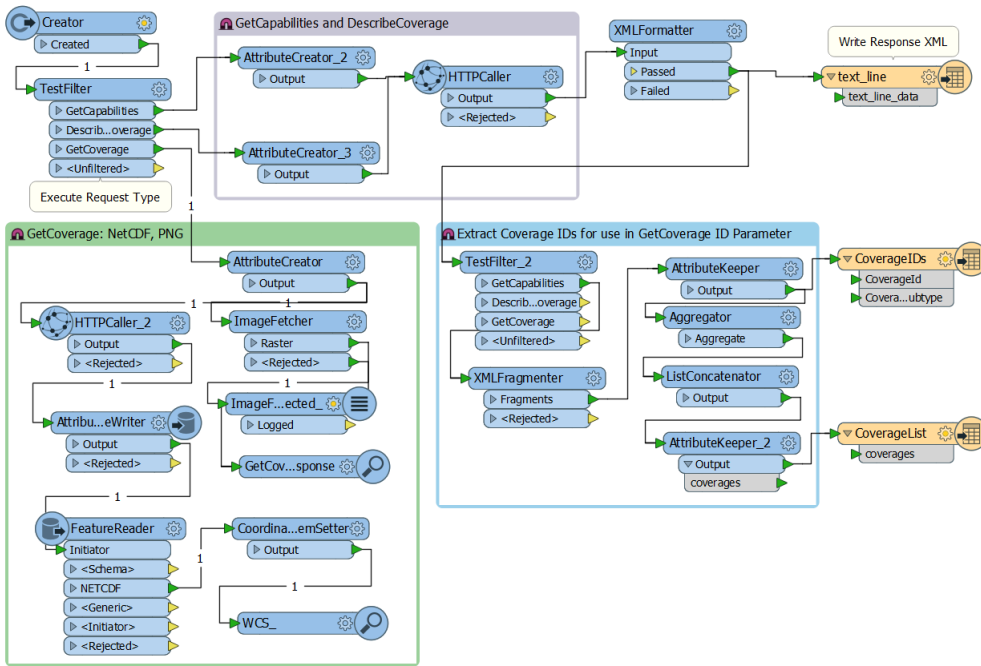


Figure 14. WCS subset client implemented using FME workspace

Figure 15. WCS subset client parameters showing query long, lat and range (subset width)

This is a sample WCS request generated by the above FME WCS client to request a subset of the MERRA-2 data: [http://testbed-13.opendap.org:8080/WCS-2.0?service=WCS&version=2.0.1&request=GetCoverage&CoverageId=MERRA2\\_100.statD\\_2d\\_slv\\_Nx.19800101.SUB.nc4&format=netcdf&subset=latitude\(37.2,40.2\)&subset=longitude\(-123.1,-120.1\)&rangesubset=T2MMAX&format=netcdf](http://testbed-13.opendap.org:8080/WCS-2.0?service=WCS&version=2.0.1&request=GetCoverage&CoverageId=MERRA2_100.statD_2d_slv_Nx.19800101.SUB.nc4&format=netcdf&subset=latitude(37.2,40.2)&subset=longitude(-123.1,-120.1)&rangesubset=T2MMAX&format=netcdf)

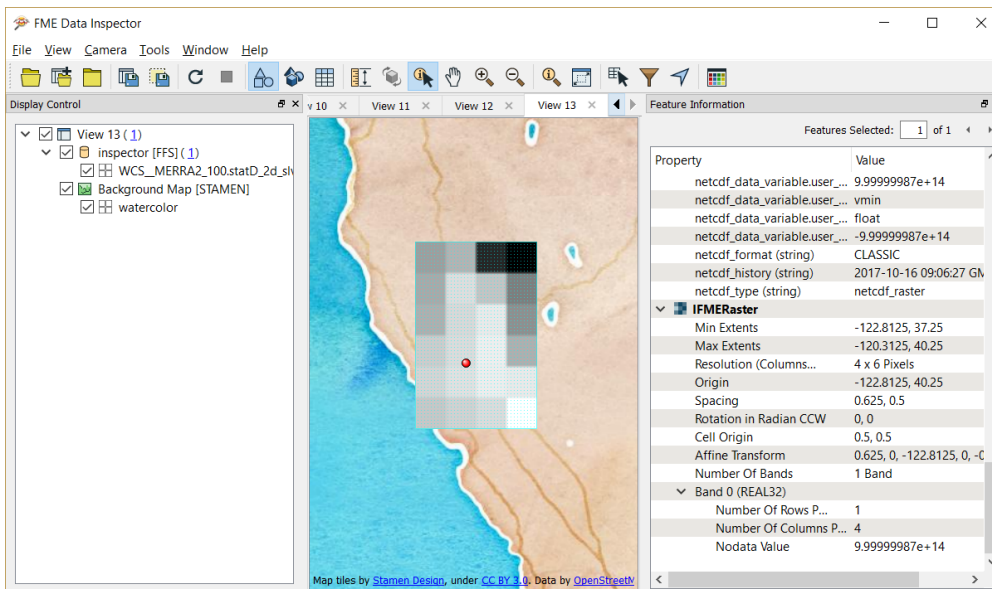


Figure 16. WCS subset client results displayed in FME Data Inspector

Note that the custom client evaluated did involve configuration setup of transformers within the FME workspace which define the data transformation workflow. The client requests are implemented by feeding values from published parameters into XML templates of WPS or WCS requests in the XMLTemplater transformer. These requests are then sent to the WPS or WCS server using HTTPCaller. The response is then parsed into features using FeatureReader or ImageFetcher depending on the response type. Although some configuration was needed, no coding was required, and the flexibility of this approach means that the client can easily be customized to generate a range of query types including spatial and temporal ranges, with results that can be readily transformed to meet a variety of output or external service requirements.

#### 4.3.4. NA103 Prediction WPS

The Prediction WPS acts as an interface to the APSim crop model. The user is able to configure the crop model via the thread clients (NA101, NA102) and execute a simulation, additionally there is a simple test client included with the WPS to create execute documents and instruct the WPS ([http://54.149.145.158/wps/test\\_client](http://54.149.145.158/wps/test_client)). The system returns a point of the simulation location with daily predictions for crop yield. The WPS also allows the user to specify a location, a date range and the meteorological data for use with the crop model. The data is provided by NASA sources in NetCDF and therefore contains multiple variables, and has a temporal element that is queried in the WPS back-end.

The test client drop-down contains an execute document called ExecuteAPSimFullParametersGML\_Rice.xml, for executing the crop model using a set of default parameters, it returns a spatially enabled point file with daily crop yield predictions. This process is designed for the *Expert Scientist* user client as it provides a configurable interface to execute the APSim crop model. A second process, ExecuteApsimNonScientistGML\_Rice.xml, is designed to be used with the non-science client. This process enables execution of APSim with the minimum parameters required as many of the required variables are inferred by crop type and location. This process interacts with NA104 to get the MERRA-2 climate data for incorporation into the meteorological aspects of the crop yield prediction model. This demonstrator includes a subset of the full model parameters to provide an example demonstration of what can be achieved. Additionally, the processes only include the base model of APSim without the numerous extensions

available from the originators.

### **The APSim crop model**

The Testbed-13 RFP requires: a "component accessible by Web Processing Service (WPS) will enable access and control of predictive models relevant to drought and agriculture prediction based on climate and other data."

This requirement has been clarified to mean a flexible, configurable crop model with academic credentials that is suitable for augmenting with external data. APSim (Agricultural Production Systems Simulator) is the crop model of choice as it is:

1. Flexible
2. Configurable
3. Powerful
4. Academically sound
5. Open source
6. Otherwise fits the requirements

APSim is modular and is able to be extended to support different aspects of crop yield prediction, for example prediction of grape crop yield requires metrics on water retention for grape species, which is implemented as a module in the APSim and requires a large set of parameters in addition to the existing set. For Testbed-13 the base APSim model is used to demonstrate execution via WPS and interoperability with the WCS (work item NA104) provided by OPeNDAP.

The APSim crop model is usually accessed by a GUI interface provided with the software, it also has options for execution on a cluster for high performance computing, presumably for complex modeling. Behind the scenes, the crop model is configured using an XML (.apsim) file to store all of the values to execute and return results. This use of XML lends itself to external configuration without the use of the UI. The .apsim file also provides paths to other files for use in the simulation. The most important of these is the meteorological (Met) file that provides daily measures to simulate crop growth including:

- Maximum temperature
- Minimum temperature
- Precipitation
- Radiation
- Evapo-transpiration rate

The variables in the Met file can be augmented by extractions from the MERRA-2 data provided by the OPeNDAP WCS.

### **MERRA-2 Data**

Modern Era Retrospective analysis for Research and Applications version 2 (MERRA-2) is a product released by NASA for climatic research and development projects. The datasets provide 1 degree

squared rasters of the world with the following variables encoded:

- HOURNORAIN - hours without any rain for the time period.
- T2MAX - the maximum temperature for the time period.
- T2MIN - the minimum temperature for the time period.
- T2MEAN - the mean temperature for the time period.
- T2PRECMAX - the maximum precipitation for the time period.

Additionally, the data are available at different frequencies ranging from hourly to monthly, with the low frequency data derived from the high frequency data. As the APSim model creates a meteorological file, the data provided can be incorporated into the APSim model by updating the values in the meteorological file using MERRA-2 data. The lookups between the required data and the MERRA-2 data are as follows, please note that not all of the required variables within the crop model have a suitable corresponding variable in the MERRA-2 data, also, unit conversion is required in the available MERRA-2 data (for example Kelvin to Celsius):

*Table 2. Look up table between MERRA-2 and APSim data parameters*

Description	MERRA-2	APSim
Maximum temperature	T2MAX	maxt
Minimum temperature	T2MIN	mint
Radiation	none	radn
Precipitation in the form of rain	T2PRECMAX	rain
Evapotranspiration	none	evap

### **Detailed WPS description**

Figure 17 outlines the use case diagram for the APSim WPS, which forms the software component to the Prediction WPS requirement outlined in the RFQ. There are three main actors in the use case diagram, these conform to the systems that will interact with the WPS. The Expert User interacts with the system via the Agricultural Scientist Client (NA102) and has the ability to produce predictions based upon finely grained parameterisation of the WPS. Essentially, all of the functionality of APSim is exposed to the user and can be manually configured via direct input into the client, or by uploading a series of files that are required by the crop model to run correctly. There is very little error catching available, as the user guidance is minimal beyond the process descriptions. Additional, shared functionality is included to enable the user to present the system with a latitude and longitude to produce climate data for a location over a given date range as this maybe an easier method of determining input climate data.

The second actor is the Non-Agricultural Scientist, that is, an individual or team who wants to run crop simulations using climate data, but may not have the expertise to produce the required files for the crop model. This actor is reliant on the WPS to configure the model using the information provided by the client at a high level. For example, the user may not have the knowledge to fully create a crop model, but they are likely to understand parameters such as the location, crop and date range of the predictions. Therefore, the WPS will attempt to take these parameters from the user and infer the required parameters, either using external sources of data or internal logic.

These predictions are unlikely to be as robust as those for the agricultural scientist, but will still offer value to the users.

The third actor is the WCS (NA104) holding NASA data. This acts as a temporally enabled data repository for NASA data. The Prediction WPS makes a call to this service to query point values for temperature, precipitation and irradiance variables. Other data, such as soil water at different depths may be made available through this service, although this is marked as an extension to the current functionality.

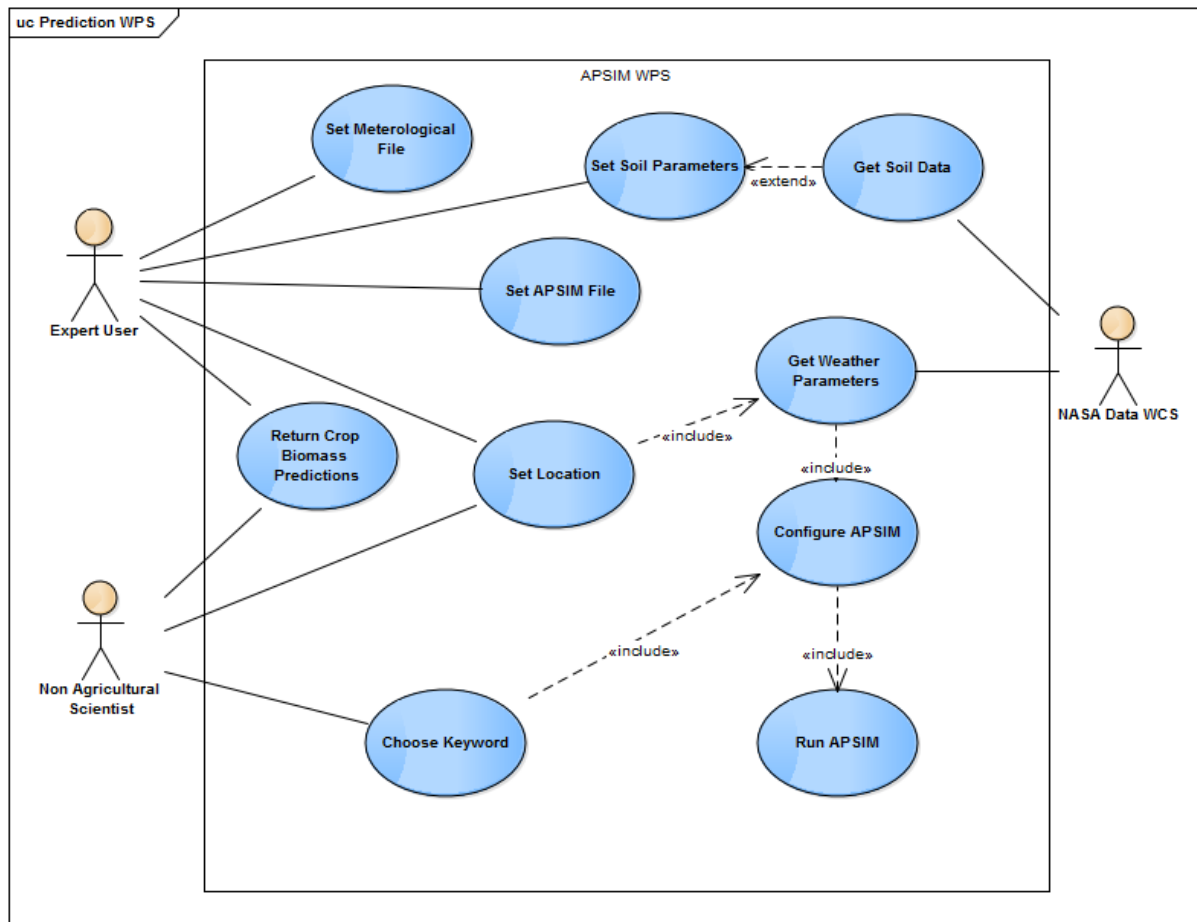


Figure 17. The use case diagram for the APSim WPS.

The workflow for the WPS can be found in Figure 18. Essentially, the process workflow is determined according to what process is executed in the WPS. The scientist process is simplistic in that it translates the input variables into an APSim execution document, executes the process and then returns the result. The non-scientist process executes using few parameters and makes decisions on the type of template APSim file to use from the crop keyword (a restricted WPS string input). The meteorological file is specified in the scientist client, whereas the most suitable file is identified in the non-scientist client and then augmented with MERRA-2 data for a provided date range. Note that the non-scientist process contains two separate date ranges, one for the prediction and one for MERRA-2 meteorological data augmentation. This is for testing purposes to manage the responses of the WCS and avoid client time out, an improvement for future iterations is to produce a set of asynchronous processes.

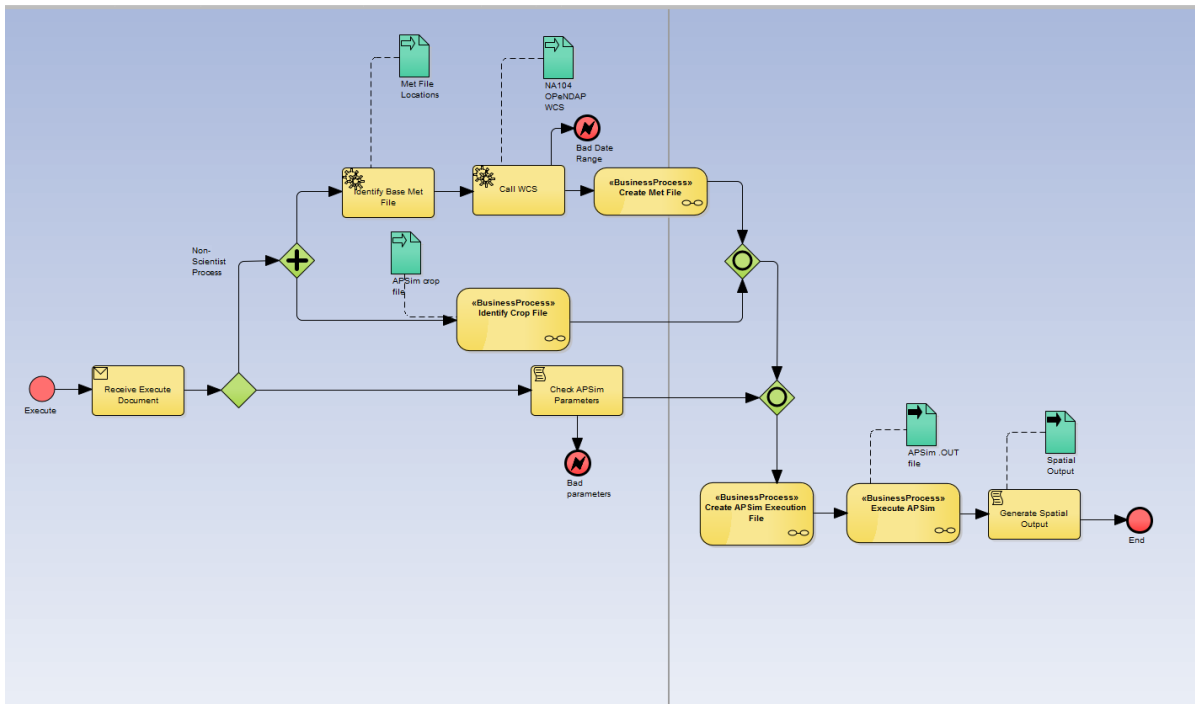


Figure 18. A BPMN diagram describing the workflow for the APSim WPS.

## Description of WPS Processes

Table 3. WPS processes

Identifier	Title	Description
ExecuteAPSimFullParametersGML_Rice	Execute APSim model using all parameters for the Rice scenario	This process requires all parameters to be entered including an APSim file. This process is designed for the Agricultural Scientist client.
ExecuteAPSimNonScientistGML_Rice	Execute APSim model with simple parameters	This process just requires a latitude and longitude and some information about the dates of the simulation to run. Data are gathered using a series of backend processes including calls to the WCS. The different date ranges refer to the total length of the simulation, and the number of values to augment with WCS data. Note that the larger the WCS call range, the longer the process takes to execute.

### Agricultural Scientist Process

This process is designed to be used with the Agricultural Scientist and requires many inputs to function. The process also uses several template files that are uploaded along with the required variables. The process returns a point file with the predictions attached, the point file can be requested as GML, a zipped shapefile or JSON.

The full input parameters are described here:

Table 4. Agricultural scientist process input parameters

Identifier	Title	Description
templateFile	The template APSim file	An input file that forms the basis of the simulation
outputFileName	The name of the output file	This is what is returned as the _outputResultFile
metFile	The meteorological file to simulate against	A MET file in APSim format
startDateTime	The start date of the simulation	The date that the simulation starts, must be included in the Met File. Needs to be in the following format; dd/mm/yyyy.
endDateTime	The end date of the simulation	The date that the simulation ends, must be included in the Met File. Needs to be in the following format; dd/mm/yyyy.
soilTypeDescription	A description of the soil type	This is for descriptive purposes only.
localName	The name of the simulation area	This is for descriptive purposes only.
site	The specific site of the simulation area	This is for descriptive purposes only.
nearestTown	The nearest town to the simulation area	This is for descriptive purposes only.
region	The region that this simulation sits	This is for descriptive purposes only.
naturalVegetation	The type of natural vegetation in the area	This is for descriptive purposes only.
waterFractionFull	The amount of water available for the simulation	Provided as a fraction of water available.
percentMethod	The method of measuring the water	Describes whether the water profile is filled from the top or the bottom, refers to the precipitation and water profile in the area.
summerCona	Stage 2 drying slope	This represents the slope of the stage 2 drying curve (PERFECT model)
summerU	Stage 1 drying limit	This represents the upper limit of the stage 1 drying curve (PERFECT model)

summerDate	Summer start date	The start date for summer in the simulation
winterCona	Stage 2 drying slope	This represents the stage 2 drying curve (PERFECT model)
winterU	Stage 1 drying limit	This represents the upper limit of the stage 1 drying curve (PERFECT model)
winterDate	Winter start date	The start date for winter in the simulation
diffusConst	The diffuse constant	The soil water diffuse constant
diffusSlope	Capillary flux	A measure of the capillary flux and flow for sandy soils
salb	Albedo	The albedo of the bare soil
cn2Bare	CN2 measure	The CN2 adjusted for bare soil
cnRed	Runoff reduction	Soil runoff reduction for the corresponding curve number
cnCov	Soil surface threshold	The effect of runoff mitigated by the threshold surface cover
Latitude	The latitude of the site	Note that this variable defines the WCS call and the base meteorological file.
Longitude	The longitude of the site	Note that this variable defines the WCS call and the base meteorological file.

The output parameters are as follows:

*Table 5. Agricultural scientist process output parameters*

Identifier	Title	Description
outputResultFile	The APSim .out file	The raw output from APSim, no spatial component.
outputResultPoint	A spatial output from APSim	The output spatial file, options include GML, Shapefile and JSON.

An example WPS execute POST request is as follows:

```
<?xml version="1.0" encoding="UTF-8"?>
<wps:Execute
  xmlns:wps="http://www.opengis.net/wps/2.0"
  xmlns:ows="http://www.opengis.net/ows/2.0"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.opengis.net/wps/2.0
http://schemas.opengis.net/wps/2.0/wps.xsd"
```



```

service="WPS" version="2.0.0" response="document" mode="sync">
<ows:Identifier>helyx.cropmodel.APSim.ExecuteAPSimFullParametersGML</ows:Identifier>
<wps:Input id="templateFile">
  <wps:Reference mimeType="text/plain" xlink:href=
"http://meekbaa1.miniserver.com/dl/Millet.APSim"/>
</wps:Input>
<wps:Input id="outputFileName">
  <wps:Data>
    <wps:LiteralValue>Millet.out</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="metFile">
  <wps:Reference mimeType="text/plain"
xlink:href="http://meekbaa1.miniserver.com/dl/Goond.met"/>
</wps:Input>
<wps:Input id="startDateTime">
  <wps:Data>
    <wps:LiteralValue>1/01/1940</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="endDateTime">
  <wps:Data>
    <wps:LiteralValue>31/12/1941</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="soilTypeDescription">
  <wps:Data>
    <wps:LiteralValue>Black Vertosol123</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="localName">The output parameters are as follows:
  <wps:Data>
    <wps:LiteralValue>Waco_example123</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="site">
  <wps:Data>
    <wps:LiteralValue>Jimbour_example123</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="nearestTown">
  <wps:Data>
    <wps:LiteralValue>Jimbour_example</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="region">
  <wps:Data>
    <wps:LiteralValue>South East Queensland</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="naturalVegetation">

```

```

<wps:Data>
  <wps:LiteralValue>Grassland, mainly Qld bluegrass</wps:LiteralValue>
</wps:Data>
</wps:Input>
<wps:Input id="waterFractionFull">
  <wps:Data>
    <wps:LiteralValue>1</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="percentMethod">
  <wps:Data>
    <wps:LiteralValue>FilledFromTop</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="summerCona">
  <wps:Data>
    <wps:LiteralValue>3.5</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="summerU">
  <wps:Data>
    <wps:LiteralValue>6</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="summerDate">
  <wps:Data>
    <wps:LiteralValue>1-Nov</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="winterCona">
  <wps:Data>
    <wps:LiteralValue>3.5</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="winterU">
  <wps:Data>
    <wps:LiteralValue>6</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="winterDate">
  <wps:Data>
    <wps:LiteralValue>1-Apr</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="diffusConst">
  <wps:Data>
    <wps:LiteralValue>40</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="diffusSlope">
  <wps:Data>

```

```

    <wps:LiteralValue>16</wps:LiteralValue>
  </wps>Data>
</wps:Input>
<wps:Input id="salb">
  <wps>Data>
    <wps:LiteralValue>0.13</wps:LiteralValue>
outputResultFile  </wps>Data>
  </wps:Input>
<wps:Input id="cn2Bare">
  <wps>Data>
    <wps:LiteralValue>70</wps:LiteralValue>
  </wps>Data>
</wps:Input>
<wps:Input id="cnRed">
  <wps>Data>
    <wps:LiteralValue>23</wps:LiteralValue>
  </wps>Data>
</wps:Input>
<wps:Input id="cnCov">
  <wps>Data>
    <wps:LiteralValue>0.8</wps:LiteralValue>
  </wps>Data>
</wps:Input>
<wps:Input id="latitude">
  <wps>Data>
    <wps:LiteralValue>52.0315</wps:LiteralValue>
  </wps>Data>
</wps:Input>
<wps:Input id="longitude">
  <wps>Data>
    <wps:LiteralValue>-2.1498</wps:LiteralValue>
  </wps>Data>
</wps:Input>
<wps:Output id="outputResultFile" transmission="value"/>
<wps:Output id="outputResultPoint" transmission="value"/>
</wps:Execute>

```

## Non-Scientist Process

The non-scientist process is designed to be used with the non-scientist client. Unlike the agricultural scientist process, it has few variables and gathers much of the required data from geospatial elements such as where the site is located. This is meant as a demonstrator of what is possible to automate, however there are shortcomings with the approach, which will be described later in the document.

The full input parameters are described here:

Table 6. Non-scientist process input parameters

Identifier	Title	Description
------------	-------	-------------

cropType	The type of for prediction	This pre-configures the data associated with a crop from the provided example files, it is restricted by the availability of examples but could be extended in the future.
merraStartDate	The start date for the MERRA data	This configures the OpenDAP WCS to download the correct NASA NetCDFs.
merraEndDate	The end date for the MERRA data	This configures the OpenDAP WCS to download the correct NASA NetCDFs.
startDate	The start date for the simulation	This differs from the MERRA dates as the MERRA data is used to augment the existing MET data that may contain a wider range of dates than offered by the MERRA WCS.
endDate	The end date for the simulation	This differs from the MERRA dates as the MERRA data is used to augment the existing MET data that may contain a wider range of dates than offered by the MERRA WCS.
variableToReplace	The variable to edit in the MET file	This is currently restricted to maximum daily temperature and minimum daily temperature, but could be extended to other variables.
latitude	The latitude of the site	The latitude defines where the site is located and also how the process calls the WCS and chooses the best base MET file.
longitude	The longitude of the site	The longitude defines where the site is located and also how the process calls the WCS and chooses the best base MET file.

The output parameters are as follows:

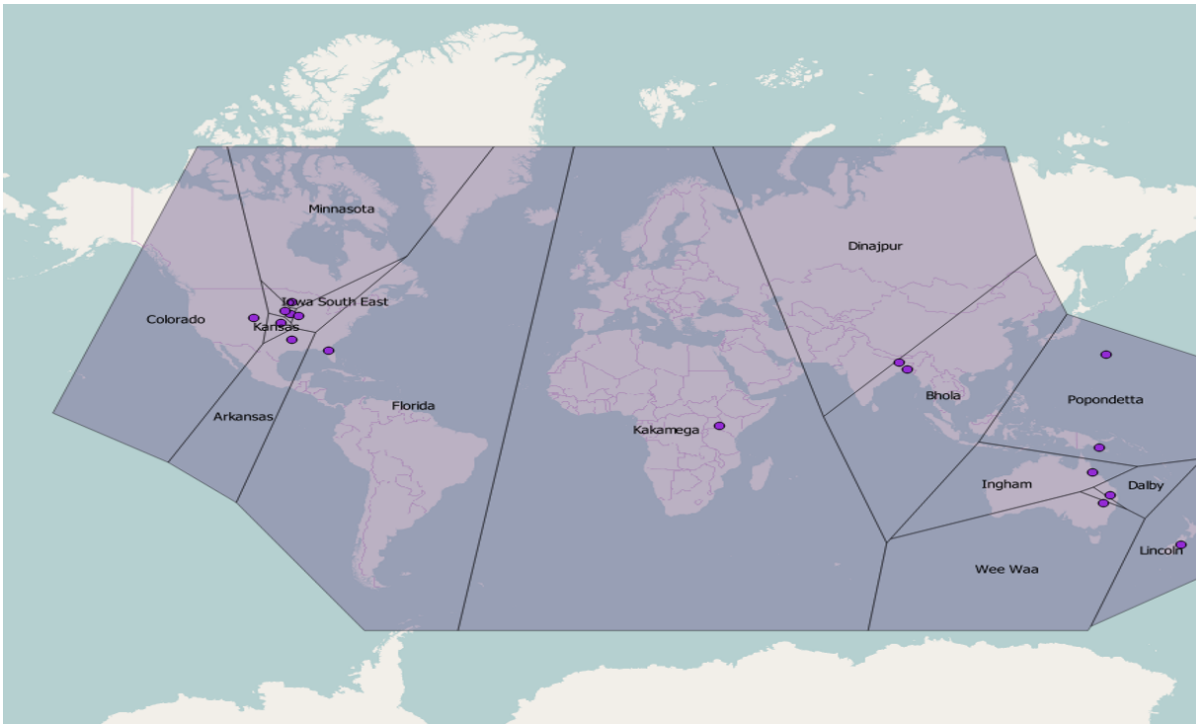
*Table 7. Non-scientist process output parameters*

Identifier	Title	Description
outputResultFile	The APSim .out file	The raw output from APSim, no spatial component
outputResultPoint	A spatial output from APSim	The output spatial file, options include GML, Shapefile and JSON

## Establishing the correct meteorological base file

The MERRA-2 data available in the WCS only covers a few months' worth of daily data, the meteorological data from the APSim files contain many more data points than the available MERRA-2 data and needs to cover the entire duration of the required prediction. However, the provided data is geographically coarse, the scenario for this Testbed is based in San Francisco, but the closest data point is Colorado. Therefore, the Colorado data is a suitable base, but is much more accurate if augmented with MERRA data over the available dates.

The process for selecting the correct base file is to use a simple point in polygon operation, the point location is created from the latitude and longitude inputs. The base polygon file is shown below.



An example truncated meteorological file is here:

```
[weather.met.weather]
```

```
Location = Colorado example
```

```
latitude = 40.56 (DECIMAL DEGREES)
```

```
longitude = -104.89 (DECIMAL DEGREES)
```

```
Elevation = 1540 (meters above sea level)
```

```
tav = 9.697991 (oC) ! annual average ambient temperature
```

```
amp = 24.27206 (oC) ! annual amplitude in mean monthly temperature
```

```
! source: http://daymet.ornl.gov/ (verified - November 2014)
```

```
! Note: The year 1979 indicates average climatic values from 1980 to 2013
```

```
! The met file generated on November 23, 2014, Department of Agronomy, Iowa State University,  
Ames, IA
```

```
year day radn maxt mint rain snow vp dayL
```

```
0 0 (MJ/m^2) (oC) (oC) (mm) (mm) (kPa) (hours)
```

```
1979 2 8.118852 5.794117 -10.32353 2.9411760 9.882300 0.28352 9.216111
```

```
1979 3 7.750381 6.029412 -9.161765 0.2058824 9.647058 0.2917647 9.216111
```

```
1979 4 7.641088 5.220588 -9.470589 0.3823529 10.00000 0.2870588 9.216111
```

```
1979 5 7.750381 5.529412 -9.470589 0.5294118 10.47059 0.2800000 9.216111
```

1979 6 7.972088 5.838235 -9.617647 0.2941177 10.58823 0.2776470 9.216111  
 1979 7 7.859673 4.970588 -10.17647 0.4117647 10.82353 0.2647059 9.216111  
 1979 8 7.797220 6.352941 -8.676471 0.2352941 10.70588 0.2929412 9.216111  
 1979 9 7.783646 6.308824 -8.220589 0.2941177 10.82353 0.2894118 9.311944  
 1979 10 8.165414 6.205883 -9.485294 0.3529412 11.05882 0.2670588 9.311944  
 1979 11 8.272688 6.794117 -8.985294 0.2647059 11.17647 0.2741176 9.311944  
 1979 12 8.121243 5.514706 -9.970589 0.4705882 11.52941 0.2635294 9.311944  
 1979 13 8.549333 5.970588 -9.867647 2.9411760 11.52941 0.2576471 9.408055  
 1979 14 8.667276 7.000000 -8.852942 0.2352941 11.76471 0.2635294 9.408055  
 1979 15 8.871287 6.852941 -9.250000 5.8823530 11.76471 0.2564706 9.408055

### Augmenting the file with MERRA data

When the correct base file has been selected, the MERRA date inputs are used to edit a WCS query and download all of the required daily files. The query is constrained by the variable to augment (controlled by the variableToReplace input) and the location given by the latitude and longitude inputs. An example query is as follows:

[http://testbed-13.opendap.org:8080/WCS-2.0?service=WCS&version=2.0.1&request=GetCoverage&coverageId=MERRA2\\_100.statD\\_2d\\_slv\\_Nx.19800101.SUB.nc4&subset=latitude\(35.0,45.0\)&subset=longitude\(-126.0,-120.0\)&rangesubset=T2MMEAN&format=netcdf](http://testbed-13.opendap.org:8080/WCS-2.0?service=WCS&version=2.0.1&request=GetCoverage&coverageId=MERRA2_100.statD_2d_slv_Nx.19800101.SUB.nc4&subset=latitude(35.0,45.0)&subset=longitude(-126.0,-120.0)&rangesubset=T2MMEAN&format=netcdf)

The variables highlighted in **bold** are changed according to the input parameters:

- MERRA2\_100.statD\_2d\_slv\_Nx.19800101.SUB.nc4 - This refers to the coverage that is requested, the date is 01/01/1980. The date parameter is changed in the iteration.
- latitude, longitude - these are the minimum and maximum latitudes and longitudes of the requested image, the requested image is 2 degrees by 2 degrees from the center point of the image.
- T2MMEAN - this is the variable that the pixels refer to in the image, it also defines the variable that will be augmented in the meteorological file. Currently, this is restricted to T2MMAX and T2MMIN, the maximum and minimum temperatures recorded at the defined location. Note that the MERRA data supplies temperatures in Kelvin and APSim requires data in Celsius, the conversion is done on the back-end. Additionally, the meteorological file is parsed out to be edited, the two aforementioned variables refer to *maxt* and *mint*, conversion is also done on the backend.

This is an example WPS execute document for the non-scientist client.

```
<?xml version="1.0" encoding="UTF-8"?>
<wps:Execute
  xmlns:wps="http://www.opengis.net/wps/2.0"
  xmlns:ows="http://www.opengis.net/ows/2.0"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.opengis.net/wps/2.0
  http://schemas.opengis.net/wps/2.0/wps.xsd"
  service="WPS" version="2.0.0" response="document" mode="sync">
  <ows:Identifier>helyx.cropmodel.APSim.ExecuteAPSimNonScientist</ows:Identifier>
```

```

<wps:Input id="cropType">
  <wps:Data>
    <wps:LiteralValue>Millet</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="merraStartDate">
  <wps:Data>
    <wps:LiteralValue>19800101</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="merraEndDate">
  <wps:Data>
    <wps:LiteralValue>19800130</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="startDate">
  <wps:Data>
    <wps:LiteralValue>02/01/1979</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="endDate">
  <wps:Data>
    <wps:LiteralValue>01/01/2005</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="variableToReplace">
  <wps:Data>
    <wps:LiteralValue>T2MMAX</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="longitude">
  <wps:Data>
    <wps:LiteralValue>-122.4194</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Input id="latitude">
  <wps:Data>
    <wps:LiteralValue>37.7749</wps:LiteralValue>
  </wps:Data>
</wps:Input>
<wps:Output id="outputResultFile" transmission="value"/>
</wps:Execute>

```

### 4.3.5. NA104 WCS access to climate data

OPeNDAP provides the WCS 2.0 service offering data for both the WPS service and the WCS clients in this thread. The service offers a substantial amount of data, primarily taken from the NASA MERRA 2 service.

The WCS 2.0 endpoint can return both NetCDF and GeoTIFF responses and supports scaling,

interpolation and trimming operations. The WCS 2.0 service also supports the Earth Observation Profile for WCS 2.0, but is not utilized in the thread clients.

In an effort to get datasets up and available for client ingest and testing, several datasets have been made available from the larger MERRA-2 collection at NASA on the [OpenDAP Testbed-13 WCS-2.0 service](http://testbed-13.opendap.org:8080/WCS-2.0/) [http://testbed-13.opendap.org:8080/WCS-2.0/]. The coverages are identifiable through a standard prefix: "MERRA2\_400.tavgM\_2d\_int\_Nx" .

**NOTE** Some of the other coverages on that server are part of our work on the WCS EO Profile for Testbed-12. Many of these are 4D data and cannot be served through the basic WCS-2.0 service profile.

## 4.4. Component Implementation & Scenario for the Modeling Package

The component overview graphic at the beginning of this section illustrates how the different components described above fit together. The following scenario will be used to demonstrate this in a real-life example demonstration video.

### 4.4.1. TIE Component Implementation

Table 8. Testing and Integration Experiment (TIE)

Component	Description	Link Request	Tested
NA101: Expert Client	Agricultural analyst client	WPS (configure APSim execute parameters), WCS (GeoTIFF visualization)	Y: TIE on 10/03/17
NA101b: (Other) Scientist Client	Additional Scientist Client: FME client form Safe Software	WCS (spatial subsetting of NASA data), WPS (APSim run with parameter selection/execute & output display over time)	Y: TIE on 10/03/17
NA102: Non-Scientist Client	Simple GIS-style web client/Open-source desktop client (QGIS)	WCS (direct visualization of NASA data)/WPS execution and output visualization	Y: TIE on 10/03/17
NA102b: (Other) Non-Scientist Client	Additional Non-Scientist Client: FME client form Safe Software	WCS (spatial subsetting of NASA data), WPS (simple APSim run selection/execute & output display over time)	Y: TIE on 10/03/17
NA103: WPS	APSim: Agricultural yield prediction model	WCS & both clients	Y: TIE on 09/19/17



Component	Description	Link Request	Tested
NA104: WCS	Serving NASA data	Linking to WPS & both clients	Y: TIE on 09/19/17 & 10/03/17

#### 4.4.2. Scenario "Blueprint" for the Result Demonstration Video

After setting the story for the demonstration video including the material and short video(s) described at the start of this chapter, the TIE showcasing follows the sequence:

Step 1: Expert and non-expert use cases. Expert use case - selects area based on latitude longitude, determines model parameterization, executes model, gets model output, visualizes model output. Non-expert user - selects area based on latitude longitude, selects a pre-determined model setup, gets model output, visualizes model output.

Step 2: Actual working component implementation. Show how everything works together:

- 1) WCS to grab NASA MERRA-2
- 2) WPS with APSim on Amazon Cloud with parameterization file and pre-determined selections
- 3) Non-expert and expert clients in operation with WPS/APSim model
- 4) Visualization of model results in GML format
- 5) WCS also directly accessed by non-expert client for direct data visualization. Also non-expert client can access NASA data directly for display.

# Chapter 5. Recommendations and Future Work

## 5.1. Observations and caveats

There is somewhat of a reliance on the concept of keywords for the non-scientist client. Ideally, the non-scientist user will be able to enter plain text as variables such as crop type (for example, Rice) and other derived variables will come from data inputs such as the NA104 WCS. However, this may prove difficult in practice and another approach may have to be adopted if the management of this plain text input becomes difficult. Additionally, using keywords is restrictive as the user can only execute the crop model for crops that already have example information, this can be mitigated by offering the expert user client as an alternative in this scenario.

## 5.2. Change requests

A significant issue with WCS is its reliance on a fixed set of metadata fields, precluding WCS from exposing metadata information encoded in the dataset that fall outside what can be encoded in those fixed fields. In practice, it seems implausible that any data system will completely foresee all possible uses of the information it contains, and failing to leverage users and their innate ability to evaluate the fitness of data for a given purpose seems to be a needless limitation. Note that this idea does not preclude the notion of a metadata standard, but argues for a representation flexible enough to include other information *in addition*. Had we been able to do this, we could have spent more time using data and less time hunting for it!

WCS and/or clients should support some method of getting data at a position or set of positions over a time series without having to download the entire dataset or subset. This is particularly relevant when attempting to get point samples of time series data in a single location, a use case in this work item requires daily rainfall and temperature data over a number of years, therefore one may end up making 100s of WCS calls to get the required data. Currently, the methodology for extracting multiple time slices of a dataset is to make multiple calls to the WCS and download the requested data. This can be made efficient within confines of the standard by restricting the bounding box to a very small area to retrieve many small-kilobyte tiles and extract value from those. Overall this is an inefficient way of working as there is the added complications of latency and possible outages from the WCS server.

During this testbed activity, Safe Software has successfully demonstrated spatial subsetting of the MERRA-2 data in NetCDF using their FME client and has undergone TIE as part of this testbed. This capability was not part of the original RFP tasks but it shows great interoperability. In addition to this, it is recommended/suggested that in future work (during Testbed-14), temporal subsetting is also demonstrated. It would be interesting and beneficial to do this using a physical process or model that produces a more spatially dynamic output, for instance a simple flood-fill operation on a digital elevation model. The APSim model used in this testbed produces a temporally dynamic output at a point coordinate location and thus lacks a spatially dynamic output.

The WPS parameters required to execute the crop model are variable depending on the amount of fine-tuning required and the modules included in the calculation. For example, an idea for the

scenario was to use grapes, due to the economic importance of California's wine regions. However, creating WPS processes to accept inputs for every module in every configuration scenario is unrealistic. Therefore, this is another scenario where a transactional WPS (WPS-T) would relieve the burden on implementers to second guess requirements from users by providing a platform for them to execute their own processes at will.

One of the key goals for this ER was to make valuable but complex datasets such as NASA's MERRA-2 more accessible. "These deliverables are designed to broaden distribution access and formats of climate reanalysis, climate model data, and climate based observational data for ease of system-to-system ingest, access and data delivery" (Section 1.1.1). In order to further this goal, methods of presenting results in more widely accessible forms should be explored, including those accessible by common web clients and tools (GeoJSON, MapML, PNG etc). Also, consideration should be given to access scalability, in terms of both request and user volume, given the performance challenges encountered with some of the WCS related processes described above.

Finally, it remains a fact that datasets like the MERRA-2 are high-level scientific data and therefore require operations performed within a scientist client context. A recommendation to make this a more seamless and user-friendly process for a non-scientist or non-expert client is to explore how a scientist client could be used to predefine simple parameterization and/or preset data visualization for a non-scientist client and/or non-expert user. For instance, a "hybrid" form of client as presented in this ER using the FME client by Safe Software could be used to define the settings for creation and visualization of simple but informative KML files for NASA's WorldWind or indeed a WMS-type image format for straightforward visualization in a more typical GIS client, like QGIS or ESRI ArcMap.

# Appendix A: Revision History

Table 9. Revision History

<b>Date Release</b>	<b>Editor</b>	<b>Primary clauses modified</b>	<b>Descriptions</b>
May/June, 2017	G. Schumann	all	initial version
July/August, 2017	S. Meek	various	additions integrate
August, 2017	OpenDAP	components	additions integrate
September, 2017	D. Hintz	components	additions integrate
September, 2017	A. Kettner	components	additions integrate
September, 2017	G. Schumann	various	additions integrate
October 6, 2017	G. Schumann	various	preparation for publication
October 17, 2017	S. Meek	various	preparation for publication
October 17, 2017	G. Schumann	various	preparation for publication
October 17-19, 2017	All	various	preparation for publication
October 19/20, 2017	D. Hintz	various	preparation for publication
October 20, 2017	G. Schumann	various	final edits and proof-reading for internal review
October 30, 2017	G. Schumann	various	implemented first round of comments received by Gobe
November 7, 2017	G. Schumann	various	implemented comments received by Terry