Data Management Plan

The primary caretakers of project data will be the University of Puerto Rico, University of Alabama in Huntsville (UAHuntsville), and University of Florida (UF). The data management plans for the three institutions are provided below.

Types of data

The types of data in the proposed work include: model formalism, developed algorithms, and software components (such as web applications and user interfaces), as well as experimental measurements for prototypes that demonstrate privacy preservation on web applications and in general in the Clouds. This will also include educational materials and sample data. The algorithms will be developed using Java Web Services Development Pack (JWSDP) technology as the tool to implement the software components. Using JWSDP allows the team to leverage on many tools (such as, ASP, ASP.NET, CGI, JSP/Java, PHP, or Python) for application programming, metadata representation, and security infrastructure. As the database systems, the PI plans to utilize MySQL and PostgreSQL, in different prototypes, as they are among the most common open-source databases. The tools might change during the course of the research as needed in order to facilitate development and acquisition of results. The expected development platforms for the algorithms and software are Linux and Windows. The algorithms and software developed will be available to the public upon request.

During the course of this proposed research, several types of data will be collected. GOES data for the creation of the solar insolation archive will be obtained from either the NASA Marshall Space Flight Center (at the National Space Science Technology Center on the UAHuntsville campus), or from NOAA NESDIS, when real-time feeds are required, and through the Comprehensive Large Array-data Stewardship System (CLASS) when historical data are needed. MODIS, VIIRS and LandSat imagery will be applied to the DisALEXI approach for evaluating soil moisture and evapotranspiration at the farm scale. Numerical weather prediction model data will be obtained from NOAA. For selected cases, numerical weather prediction will be conducted using the mesoscale Weather Research and Forecasting (WRF) model for atmospheric conditions.

With respect to the solar insolation used by GOES-WEB and ALEXI algorithms, the new GOES-R (and perhaps GOES-S) datasets will become available within this project's timeframe, by 2016. Hence, one research component will involve developing and tailoring these algorithms for using 500 m resolution visible and 2 km resolution infrared data from these new systems. The project's goals can be met with current GOES data, yet we expect significant improvements to (especially in cases of small convective clouds) with the higher resolution observations.

Data and metadata standards

Data and metadata standards have not been selected for the experimental measurements and simulation results. As the work progresses and the experimental data generated, formats will be chosen and the appropriate documentation will be developed.

The GOES data (in McIDAS AREA format) will be used directly to produce the solar insolation product. The WRF model will be used for simulating operation weather data over the study areas. The model output will be archived as netcdf format.

Policies for access and sharing and provisions for appropriate protection/privacy
The GOES and NWP are under Dr. John Mecikalski's management (Atmospheric Science, UAHuntsville). The archived raw data for specific cases can be accessed per request from the above people. These data may be used free of charge, and without restriction for non-profit research activities. GOES-WEB image results are provided to the public at http://pragwater.com free of charge to the public. GOES-WEB maps and time-series data at selected map locations will be available on line using a data visualization tool currently being developed under the NOAA CREST project.

Field data collected by UPR and UF for model calibration will include LAS RET and other weather parameters. Additional field data will be collected for testing the irrigation scheduling tools including irrigation amounts, weather parameters, crop growth, and crop yield. Data will be stored electronically after review and quality control assessment. Data will be stored on UPR and UF servers with remote nightly backup (Clouds). Data will be shared with the project team and available upon request to the public.

Policies for access and sharing
The data obtained in the experiments will be available through the project's site and by direct requests to the PI. The data formats and the date on which the data will be available immediately after publication. The experimental data is not considered 'personal data' as defined by the Data Protection Act 1998. The PI will follow the policies of the University of Puerto Rico at Mayaguez (UPRM) on patents, intellectual property and commercialization as described by Certification 132 (2002-2003) on any intellectual property that results from this research.

The new irrigation scheduling tools developed from this proposal will be available free of charge to the public which will access data or data products created by the proposed work. Tools developed as apps would use the native programming languages for iOS (Objective C) and/or android (Java).

Policies and provisions for reuse, redistribution
There will be no permission restrictions on the data obtained through our experimental measures. The developed software components and algorithms will be available to members of the scientific community by direct requests to the PD through the project's site. The obtained data will be of interest to students, scientists, and industry members who are currently working on database security, assurance techniques, and user privacy. The data can be used for further works.

Plans for archiving and preservation of access

All raw datasets collected from this proposed research will be stored on computers within a data storage infrastructure that allows for ready access. A RAID storage array at NSSTC in UAHuntsville (under Dr. Mecikalski's supervision) will be used to archive the observational datasets and numerical model output. The obtained data will be archived in a repository for at least 5 years after the completion of the project.

Funds are requested for one server type computer totaling $11,245.00, which will be the backbone computational resource supporting the research for this AFRI effort, and that will host data distribution capabilities between the UAHuntsville, the University of Maryland, the USDA (Beltsville, MD), the University of Florida, the University of Puerto Rico–Mayaguez, and The University of the Virgin Islands. We anticipate that the PC computer will be used for the Ph.D. students, as well as made available to all other Investigators on this project (and their students as needed). A quote is provided.

The details of the server-type workstation are below, and a full quote is included.

ACE Powerworks 4-Way   **$11,245.00**
Workstation with 4x AMD 6320 8C
2.8GHz, 64GB Memory, 20 x 4TB
Enterprise drives, RAID 5
Supermicro Motherboard, 2 PCI-e x16, 1 PCI-e x8(x4 electrical), 32 Dimms,
4x AMD Abu Dhabi 8C 6320 2.8G 16M 6400MT with 4 Actve Heatsinks
64GB memory, 16x 4GB DDR3 1600MHz ECC Registered
(1)LSI 9260-8i RAID 5 Controller
4U Chassis, 24-HS Bay, Rack-mountable with Rail Kit, 1620 Watt Redundant PS
Optical, 24X SATA DVD
20x Internal 4TB SAS Enterprise Drives setup as RAID 5 - WD4001FYYG
Assembly, Built and tested in ISO-Certified Facility
Warranty 3 Years Parts and Labor
Shipping Fees Included