EO Open Science and FutureEO
12–16 November 2018 | ESA–ESRIN | Frascati (Rome), Italy

NASA’s AIST Program – Fueling Innovation

Michael M Little

18/10/2016
Earth Science Technology Program Elements

ESTO manages, on average, 120 active technology development projects. Most are funded through the primary program lines below. Nearly 800 projects have completed since 1998.

<table>
<thead>
<tr>
<th>Program Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Space Validation of Earth Science Technologies (InVEST)</td>
<td>on-orbit technology validation and risk reduction for small instruments and instrument systems that could not otherwise be fully tested on the ground or airborne systems</td>
</tr>
<tr>
<td>Advanced Component Technologies (ACT)</td>
<td>critical components and subsystems for advanced instruments and observing systems</td>
</tr>
<tr>
<td>Instrument Incubator Program (IIP)</td>
<td>innovative remote sensing instrument development from concept through breadboard and demonstration</td>
</tr>
<tr>
<td>Advanced Information Systems Technology (AIST)</td>
<td>innovative on-orbit and ground capabilities for communication, processing, and management of remotely sensed data and the efficient generation of data products</td>
</tr>
</tbody>
</table>
Technology Development:
A Flexible, Science-driven Strategy

- Competitive, peer-reviewed proposals enable selection of best-of-class technology investments
- Risks are retired before major dollars are invested: a cost-effective approach to technology development and validation
- Successful partnering establishes leveraging opportunities
- This approach has resulted in:
  - A portfolio of emerging technologies that will enhance and/or enable future science measurements
  - A growing number of infusion successes into science campaigns, instruments, applications, ground systems, and missions

ESTO Enables – Science Selects
ESTO’s all-time infusion success, drawn from 775 completed projects through the end of FY17. In this fiscal year, at least 10 ESTO projects achieved infusion into science measurements, airborne campaigns, data systems, or follow-on development activities.

57% of ESTO technology projects funded during FY17 advanced one or more TRLs over the course of the fiscal year (16 advanced more than one TRL). The average TRL advancement for all years is 41%.
AIST Goals

1) Reduce the risk, cost, size, and development time of NASA Earth Science Division (ESD) space-based and ground-based information systems;
2) Increase the accessibility and utility of science data; and
3) Enable new observation measurements and information products.

AIST 18 Objectives

Solicit AIST18 Competed Research

Understand National Academy of Science Decadal Survey guidance
Acquire Center/HQ science community input and partnership
Develop acquisition strategy and limit scope to key focus areas
Draft and approve for release

Improve partnerships with Research and Applied Sciences to develop useful advanced info technology
Use of modern commercial technology: cloud computing, GIS, analytic tools
Focus advanced analytic techniques on supporting NASA ESD R&A and Applied Science
Assess the needs for multi-vantage point Sensor Web-based Observing Strategy

Accelerate Community Acceptance of advanced information technology
Develop technology useful to the Earth Science community
Establish credibility in the community
Sensor Web Observing Strategy

- Constellations of sensors from different vantage points
- Designing a complete architecture to create a unified picture of a phenomenon
  - Forecast models as a measure of quality of understanding
  - Non-NASA sources of data or relevant services

Analytic Center Framework to determine needs for tool integration

- Focus on supporting Science Investigations
  - Allow maximum discretion on part of Science PI in using data and tools
  - Reduce the repetitive work in data access and

Related Science support questions

- How do we make an objective and quantitative comparison of multi-dimensional data
- How do we measure science value?
Program Strategy for AIST18

Solicit competed research
  Public competition with peer review and programmatic alignment

Optimize technology infusions identified in proposals
  Increase collaboration with R&A and AS Program Managers
  Solicit R&A and Applied Sciences partners for input
  Augment Awards for unexpected technology infusion opportunities

Assess commercial services to support Research and Applied Sciences
Technology advances have created an opportunity to make new measurements and to continue others less costly

- Smallsat equipped with science-quality instruments
- Machine Learning techniques permit handling large volumes of data

New Observing Strategies

- Sensor webs producing data integrated from multiple vantage points
- A unified picture of the physical process or natural phenomenon
Analytic Center as a Framework for OpenScience
Focus on the Science User

User
- Project Definition
- Plan for Investigation

Tools
- Discovery & Catalog
- Work Management
- Data Interfaces
- Analytic Tools
- Modeling
- Collaboration
- Visualization
- Sharing/Publication
- Local/custom

Data
- Catalog
- NASA DAAC
- Other US Govt
- Non-US
- Local or non-public

Computational Infrastructure
- Computing
  - Capacity
  - Capability
- Storage
- Communications

Storage
- Data Containers
- Thematic model
- Metadata/Ontology
- Resulting Products
- Published data
- Provenance

Computing
- Local systems
- High End Computing
- Cloud Computing Capability
- Quantum Computing
- Neuromorphic Computing

Focus on the Science User
Focus technology projects on solving science problems

- Clear requirement in Solicitation provisions
- Augmentations to demonstrate proof of concept to science adopters
- Machine Learning (Analytic Center) Workshops
- Sensor Web (New Observing Strategy) Workshops
- Regular conversation with HQ Program Scientists and Applied Sciences PM’s
- What science problems are out of reach today?

Increase awareness/acceptance by science community

- Public Cloud Computing (AMCE)
- Machine Learning tools for analyzing data
- Geographic Information System (GIS)
- Commercial Analysis Services (i.e., Decartes Labs, Radiant.Earth, Esri, Digital Globe)
PI: Ved Chirayath, NASA ARC
Co-I: Alan Li, NASA ARC

Objective
- Develop novel machine learning methodology to fuse low resolution (meter-scale) NASA EOS remote sensing data from satellites with extremely high-resolution (mm and cm-scale) airborne NASA Fluid Lensing & MiDAR instruments for augmentation.
- Use augmentation algorithm to significantly enhance classification accuracy for coral percent cover and morphology breakdown by reducing classification error.
- Release open-source machine learning augmentation toolboxes and methodology.

Accomplishments
- Developed a robust machine learning augmentation algorithm for fusing mm-scale to m-scale multispectral data.
- Demonstrated that augmented classification results show significantly increased classification accuracy for coral percent cover and morphology assessment.
- Released the publicly available toolbox online at www.nasa.gov/ames/las/ml-augmentation.

<table>
<thead>
<tr>
<th>Method</th>
<th>0.3 m</th>
<th>0.5 m</th>
<th>2 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-means</td>
<td>67%</td>
<td>71%</td>
<td>66%</td>
</tr>
<tr>
<td>SVM</td>
<td>74%</td>
<td>74%</td>
<td>63%</td>
</tr>
<tr>
<td>Augmented</td>
<td>83%</td>
<td>77%</td>
<td>69%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>0.3 m</th>
<th>0.5 m</th>
<th>2 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>59%</td>
<td>61%</td>
<td>38%</td>
</tr>
<tr>
<td>Augmented</td>
<td>70%</td>
<td>68%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Left to right: satellite image, predicted coral coverage and morphology.
Objective

- Improve the ability of scientists to discover new surface deformation phenomena
  - Develop enhanced environments to enable computer-aided discovery and analysis for diverse scientific datasets
  - Allow cross-comparison of multiple datasets
- Demonstrate the value of this approach in 3 scenarios:
  - Volcanology Case Study
  - Groundwater Case Study
  - Study on interaction of standing atmospheric waves with Sierra Nevada topography distorting GPS locations (Lee Waves)

Accomplishments

- Developed a cloud-based platform for analysis to facilitate new scientific discoveries using high-performance parallel computing and NASA data; approach led to several discoveries in volcanology (previously undiscovered inflation events)
- Tested and refined platform tools in Jupyter/Python around volcanic, groundwater, atmospheric data sets; created Jupyter extension for transparent cloud offloading of data processing pipelines (e.g., Amazon Cloud)
- Developed components for Principal Component Analysis (PCA), neural networks/deep learning of patterns in MODIS, data fusion of GPS + MODIS.
- Authored publications on all objective case studies, and developed workflow warehouse for processing code reuse (Open-source code available through MIT and github)
“A sensor web is a distributed system of sensing nodes that are interconnected by a communications fabric and that functions as a single, highly coordinated, virtual instrument. It autonomously detects and dynamically reacts to events, measurements, and other information from constituent sensing nodes and from external nodes (e.g., predictive models) by modifying its observing state so as to optimize science information return.”

- Steve Talabac et al, 2003
Earth Science Use Cases

- Sensor webs for improving weather prediction
  - Integrate models with in situ, airborne and orbital instruments
- Real-time targeting of transient and transitional phenomena
  - In situ triggering of observing system
  - Train configuration prolongs observations of an event
  - Viewing an event from multiple angles
- Integrated sensor webs for phased arrays
  - Beat down error with statistics
  - Improve resolution with multi-node instruments
2006: Technology project launched to determine if sensor webs could provide a “revolutionary” improvement in the skill of numerical weather forecasts

Could 7-day skill in 2005 improve to 14-day skill by...

from ESTO Weather Architecture Study, 2003

from F. Yang, 2013
Real-Time Data for Adaptive Targeting

1. Observations are collected at specified locations in space & time

2. Data analysis is performed

3. Numerical forecast is executed

4. Forecast error is estimated

Autonomous and On-Demand Targeting to Collect “Best” Observations

Adaptive Targeting
Automated / Manual

Sensor Web Feedback Loop
Movie depicts mission CONOPS -- wind lidar working with an operational atmospheric model’s first-guess field that identifies regions that are sensitive to forecast error

Spacecraft will slew toward sensitive regions and lidar is placed in a high data rate collection mode

“Leader” spacecraft is included to depict how optimization for cloud-free lines of sight could be performed

(Operational CONOPS would be more complex)
SWOS Candidate Science Customers

- Hydrology
  - River flow and Flooding
  - Snow fall in 3D
  - Aquifer degradation

Precipitation
  - Extreme precipitation events

- Cryosphere
  - Glaciers changes
  - Sea Ice changes

Urban Air Quality Events
  - At a resolution (vertical and horizontal)

- Biodiversity
  - Migrations
  - Invasive species
  - Transient spring phenomena

- Solid Earth and Interior
  - Landslides
  - Plate movement
  - Volcanic activity
  - Interior magma movement
Use cases are executed based on the proposed architecture to determine technology gaps and to estimate the feasibility of a “generic” simulator (i.e., high level of reuse).
Analytic Center Concept

An environment for data analysis in a Science investigation

Tailored to the individual study
Stand it up when needed, archive when complete
Collects publication submission materials in background (data, source code, version tracking)

Harmonizes data, tools and computational resources to permit the research community to focus on the investigation

Reduce the data preparation time to something tolerable through re-use
Catalog of optional resources (think HomeDepot shopping or AppStore)
Collect relevant publications
Provide established training data sets of varying resolution
Provide effective project confidentiality, integrity and availability
Seamless integration of new and user-supplied components and data
Analytic Center Strategy

Socialize the concepts with the science community
   Does it resonate with Program Scientists and Managers
Conduct a workshop to assess needs and technology gaps
   If it works, do more of them
Solicit competed research
   Accelerate the process where appropriate
   Leverage other solicitations (NESSF, NIP)
Analytic Center Features

Low infrastructure cost
- Can be stood up on short notice for an individual investigation
- Only maintained if a reason to maintain continuous availability

Seamless integration of new components
- User provided tools
- User provided data
- Local computing environment

Comprehensive catalog of data sources and tools
- Clear applicability (or not)
- Shopping for data, not searching
- Expert system as an operator aid in selecting accepted tools

Help in using them
- Video examples and training
- Expert System Support
Define fundamental or threshold characteristics
  Inputs from Science community
  ConOps and Common elements

Inventory Existing Examples
  Lessons Learned, weaknesses, corrective action
  Inventory technologies for each tool type
  Inventory viable data storage containers and models

ID Technology deficiencies
  From experiments
  Solicit AIST18

Socialize inside the science community
  Workshop Examples
    Atmospheric Composition
    Hydrology
    Cryo
    Biodiversity

  Opportunistic
    GIS
    AWS Experiment as an AC
    Commercial: Descartes Labs, Radiant Earth, Digital Globe, Esri

  Public Awareness
Candidate Useful ML Tools

Implementations
- TensorFlow
- TensorFire (ML in the browser using jess)
- arc-GIS server (Esri)
- MathLab (Mathworks)

Algorithms
- Logistic regression
- Symbolic regression
- Random forest
- Convolutional Neural Networks (CNN)
- Deep Neural Networks (DNN)
- Case-based reasoning
- Restricted Boltzman Machines
Tropical Cyclone Research Program

Data
- Integrated web portal
- Hurricane List & Tracks
- HS3 Field Campaign
- SMAP
- SST
- TPW
- WIND
- ECMWF output
- HAMSR
- GOES-East
- GOES-West
- Microwave Rain Signature

Tools
- Integrated through web portal
- Re-use requires TCIS Team
- NEOS3 instrument simulator
- Statistical evaluation
- Storm Structure
- Environment vertical structure
- Visualization
- Models

Computational Infrastructure
Storage
- Local Storage

Computing
- Local systems
- AWS through AMCE

POC: svetla.m.hristova-veleva@nasa.gov
https://tropicalcyclone.jpl.nasa.gov/
Earth Data Analytics System (EDAS)

Data
- MERRA
- ECMWF ERA
- NOAA NCEP CFSR
- JMA JRA
- UA ERA

Tools
- Earth System Grid Federation Framework
- Discovery & Catalog
- Work Management – Jupyter Notebook
- Data Interfaces - OpenDAP
- Analytic Tools: Min, max, sum, diff, average, rms, anomaly, std deviation
- Visualization using time series or spaceplot plot routines

Computational Infrastructure
- Local Sparc
- SGI Data Migration Facility
- High End Computing (DISCOVER)
- GSFC Science Cloud

Storage
- Local systems

Climate Impact
- Project Definition
- Plan for Investigation

POC: laura.carriere@nasa.gov

https://www.nccs.nasa.gov/services/Analytics
NASA Earth Exchange (NEX) as Analytic Center

Land Change/Use Community

- Project Work Environment
- NEX

Data
- Built into website
- Dataset Sources
  - Landsat
  - Sentinel 1A
  - Modis
  - ASTER
  - TRMM
  - AVHRR
- Climate Datasets
- Land cover
- Digital Elevation Map
- STATSOO Soils
- USDA Aerial NAIP
- And others

Tools
- Models
- Tops
- Biome-BGC
- LPJ Dynamic Global Model
- Sandbox for small scale experiments
- Analysis
  - R and python based tools
  - Matlab VIIRS HDF5 swath conversion

Computational Infrastructure
- NASA Advanced Supercomputing (NAS)
- Amazon Web Services (AWS) Public Cloud
- Ames Quantum Computer (D-wave)

Storage
- Data Containers
- Thematic model
- Metadata/Ontology
- Resulting Products
- Published data
- Provenance

POC: rama.nemani@nasa.gov
https://nex.nasa.gov/nex
<table>
<thead>
<tr>
<th>PI Name</th>
<th>Org</th>
<th>Prop #</th>
<th>Project Title</th>
<th>ESTO PM</th>
<th>End Date</th>
<th>Machine Learning Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamalika Das</td>
<td>UC Santa Cruz</td>
<td>115</td>
<td>Uncovering Effects of Climate Variables on Global Vegetation</td>
<td>Oza</td>
<td>5/31/17</td>
<td>Symbolic regression</td>
</tr>
<tr>
<td>Hook Hua</td>
<td>JPL</td>
<td>109</td>
<td>Agile Big Data Analytics of High-Volume Geodetic Data Products for Improving Sc</td>
<td>Norton</td>
<td>5/31/17</td>
<td>Fault recognition and Science processing redirection</td>
</tr>
<tr>
<td>Thomas Huang</td>
<td>JPL</td>
<td>28</td>
<td>OceanXtremes: Oceanographic Data-Intensive Anomaly Detection and Analysis Portal</td>
<td>Norton</td>
<td>8/29/17</td>
<td>NEXUS as a data delivery tool - oceanography</td>
</tr>
<tr>
<td>Kristine Larson</td>
<td>U of CO</td>
<td>4</td>
<td>AMBIGHO: Automated Metadata Ingest for GNSS Hydrology within OODT</td>
<td>Hines</td>
<td>5/31/17</td>
<td>Sensor characterization</td>
</tr>
<tr>
<td>Victor Pankratius</td>
<td>MIT</td>
<td>36</td>
<td>Computer-Aided Discovery of Earth Surface Deformation Phenomena</td>
<td>Little</td>
<td>5/31/17</td>
<td>Computer aided discovery</td>
</tr>
<tr>
<td>Chaowei Yang</td>
<td>GMU</td>
<td>82</td>
<td>Mining and Utilizing Dataset Relevancy from Oceanographic Dataset (MUDROD) Meta</td>
<td>Cole</td>
<td>5/31/17</td>
<td>Natural language processing, CNN, SVM, deep learning</td>
</tr>
<tr>
<td>Tomasz Stepinski</td>
<td>Cincinnati</td>
<td>27</td>
<td>Pattern-based GIS for Understanding Content of very large Earth Science datasets</td>
<td>Quam</td>
<td>6/31/17</td>
<td>Classification and similarity</td>
</tr>
<tr>
<td>Jonathan Gleason</td>
<td>LaRC</td>
<td>95</td>
<td>NASA Information And Data System (NAIADS) for Earth Science Data Fusion and A</td>
<td>Oza</td>
<td>1/31/17</td>
<td>Precision ontology foundation</td>
</tr>
<tr>
<td>Constantine Lukashin</td>
<td>LaRC</td>
<td>14</td>
<td>NASA Information And Data System (NAIADS) for Earth Science Data Fusion and An</td>
<td>Murray</td>
<td>1/31/17</td>
<td>Scientific enterprise service bus</td>
</tr>
<tr>
<td>Anshish Chaudhary</td>
<td>Kitware</td>
<td>65</td>
<td>Visualization Pipelines for big-data on the NASA Earth Exchange (NEX)</td>
<td>Hines</td>
<td>3/31/17</td>
<td>Workflow</td>
</tr>
<tr>
<td>Martyn Clark</td>
<td>UCAR</td>
<td>88</td>
<td>Development of Computational Infrastructure to Support Hyper-resolution Large-ensemble Hydrology Simulations from Local-to-Continental Scales</td>
<td>Hines</td>
<td>4/30/17</td>
<td>Assimilation and ensembles</td>
</tr>
<tr>
<td>Kwe-Sen Kuo</td>
<td>Bayesics</td>
<td>56</td>
<td>DERECHOS: Data Environment for Rapid Exploration and Characterization of Organized Systems</td>
<td>Little</td>
<td>4/30/17</td>
<td>Foundation for data delivery</td>
</tr>
<tr>
<td>Seunghwon Lee</td>
<td>JPL</td>
<td>32</td>
<td>Climate Model Diagnostic Analyzer</td>
<td>Norton</td>
<td>4/30/17</td>
<td>Foundation for data delivery</td>
</tr>
<tr>
<td>Christian Mattman</td>
<td>JPL</td>
<td>34</td>
<td>SciSpark: Highly Interactive and Scalable Model Evaluation and Climate Metrics for Scientific Data and Analysis</td>
<td>Norton</td>
<td>4/30/17</td>
<td>Foundation for data delivery</td>
</tr>
<tr>
<td>PI Name</td>
<td>Org</td>
<td>Prop #</td>
<td>Project Title</td>
<td>End Date</td>
<td>Machine Learning Elements</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>---------</td>
<td>---------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Dan Duffy</td>
<td>GSFC</td>
<td>AIST-QRS-16-0002</td>
<td>Demo of VR Technology with live data (Phase 2) &amp; GIS (Phase 3)</td>
<td>12/31/16</td>
<td>VR</td>
<td></td>
</tr>
<tr>
<td>Sreeja Nag</td>
<td>ARC</td>
<td>AIST-QRS-16-0003</td>
<td>Scheduling Satellite Pointing within Constellations</td>
<td>12/31/16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ved Chirayeth</td>
<td>ARC</td>
<td>AIST-QRS-16-0004</td>
<td>MiDAR-fused Supervised Machine Learning (SML)</td>
<td>1/31/17</td>
<td>Use of high resolution training sets to improve global scale moderate resolution data</td>
<td></td>
</tr>
<tr>
<td>Chris Mattman</td>
<td>JPL</td>
<td>AIST-QRS-16-0007</td>
<td>Deep Web Search Analytics</td>
<td>10/31/16</td>
<td>Text analytics</td>
<td></td>
</tr>
<tr>
<td>John Readey</td>
<td>HDF Group</td>
<td>ACCESS15 -0031</td>
<td>Object Store-based Data Service for Earth System Science</td>
<td>5/31/17</td>
<td>Foundation for data storage to improve access</td>
<td></td>
</tr>
<tr>
<td>Yehuda Bock</td>
<td>Scripps</td>
<td>AIST-QRS-16-0010</td>
<td>Latency test of realtime warning systems in AMCE Cloud Computing</td>
<td>9/30/17</td>
<td>Feature Detection in near real time</td>
<td></td>
</tr>
<tr>
<td>Amy Braverman</td>
<td>JPL</td>
<td>AIST-QRS-16-0005</td>
<td>Probabilistic Climate Model Evaluation</td>
<td>9/20/16</td>
<td>What do you mean by similar?</td>
<td></td>
</tr>
<tr>
<td>Chris Lynnes</td>
<td>GSFC</td>
<td>AIST-QRS-16-0001</td>
<td>Experiment with Data Containers in ESDIS Context</td>
<td>12/20/16</td>
<td>Evaluation of alternative storage models for data to enable analysis</td>
<td></td>
</tr>
<tr>
<td>PI Name</td>
<td>Org</td>
<td>Project Title</td>
<td>End Date</td>
<td>Machine Learning Elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-------</td>
<td>-------------------------------------------------------------------------------</td>
<td>----------</td>
<td>--------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victor Pankratius</td>
<td>MIT</td>
<td>Computer Aided Discovery and Algorithmic Synthesis for Spatio-Temporal Phenomena in InSAR</td>
<td>7/30/19</td>
<td>Computer Aided Discovery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jacqueline LeMoigne-Stewart</td>
<td>GSFC</td>
<td>TAT-C ML</td>
<td>7/30/19</td>
<td>Design space iterator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branko Kosovic</td>
<td>NCAR</td>
<td>Fuel moisture content for improved fire prediction</td>
<td>5/31/17</td>
<td>Data assimilation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andrew Michaelis</td>
<td>ARC</td>
<td>Framework for Mining and Analysis of Petabyte-size Time-series on the NASA Earth Exchange (NEX)</td>
<td>7/31/17</td>
<td>Time series analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barton Forman</td>
<td>UMD</td>
<td>A mission planning tool for next generation remote sensing of snow</td>
<td>8/30/19</td>
<td>OSSE and assimilation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ved Chirayath</td>
<td>ARC</td>
<td>NeMO-Net - The Neural Multi-Modal Observation &amp; Training Network for Global Coral Reef Assessment</td>
<td>8/31/19</td>
<td>High resolution training sets improve moderate resolution imagery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dara Entekhabi</td>
<td>GSFC</td>
<td>Autonomous Moisture Continuum Sensing Network</td>
<td>8/31/19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milton Halem</td>
<td>UMBC</td>
<td>Hybrid Quantum Annealing Approaches for Inferring &amp;Assimilating Satellite Surface Flux Data into Global Land Surface Models.</td>
<td>8/31/19</td>
<td>Assimilation, registration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeffrey Morisette</td>
<td>USGS</td>
<td>Advanced Phenological Information System</td>
<td>8/31/19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walter Jetz</td>
<td>Yale</td>
<td>Software workflows for remote sensing-based biodiversity change monitoring</td>
<td>9/14/19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jonathon Hobbs</td>
<td>JPL</td>
<td>Simulation-based Uncertainty Quantification</td>
<td>9/30/19</td>
<td>Statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martyn Clark</td>
<td>NCAR</td>
<td>Climate risks in water sector: Advancing the readiness of emerging technologies in climate downscaling &amp; hydrologic modeling</td>
<td>9/30/19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
• An **Observing System Simulation Experiment (OSSE)** is an experiment designed to assess the potential impact of planned missions on Numerical Weather Prediction. OSSEs, now widely used, were pioneered at NASA by Dr. Robert Atlas.

• OSSEs help quantify the potential benefits of an observing system before it is designed, built and launched into orbit.

• Trade-offs in instrument or orbital configurations and methods of assimilating a new type of observing system can be determined by an OSSE and ultimately result in both time and cost savings.

• A **Nature Run (NR)** is a high resolution long integration from a state-of-the-art numerical weather prediction model.

• It acts as a proxy atmosphere for OSSE's from which synthetic observations from existing and future observing systems are derived.

• It is also used as the verification or truth data set when evaluating assimilations and forecasts which use the synthetic observations.

• Available Nature Runs include: fvGCM, T511 and WRF-ARW.
OSSE Assessment

Recent and current projects

Weather (GMAO/McCarty, Bob Atlas)
Hydrology (GSFC/Peters-Lidard)
Snow (UMd/Forman)
General purpose (Johnson/OSU)

Need for improved OSSE technologies

Reduce cost and turnaround time
Feed into optimizing model-based engineering tools
Ensure simulation includes all data sources

Approach for improving effectiveness of OSSEs
Comparing multi-dimensional datasets

Motivation

How do you validate models with observations?
How do you fuse data?

Technologies needed to compare multi-dimensional datasets

Needs a better definition of the problem in order to develop the technologies

What are the requirements for integrating or comparing

Introduced discussion at Theory of Data Systems workshop at CalTech

Sponsored by Statistical and Applied Mathematics Sciences Institute (SAMSI)

FY17 and FY18 studies to define the problem enough to compete research in AIST-18