→ THE ESA EARTH OBSERVATION Φ-WEEK

EO Open Science and FutureEO

12–16 November 2018 | ESA–ESRIN | Frascati (Rome), Italy

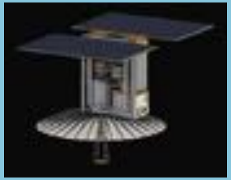
NASA's AIST Program – Fueling Innovation

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.



In-Space Validation of Earth Science Technologies (InVEST)
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

Advanced Component Technologies (ACT)
critical components and subsystems for advanced instruments and observing systems

Instrument Incubator Program (IIP)
innovative remote sensing instrument development from concept through breadboard and demonstration

Advanced Information Systems Technology (AIST)
innovative on-orbit and ground capabilities for communication, processing, and management of remotely sensed data and the efficient generation of data products

Technology Development : A Flexible, Science-driven Strategy

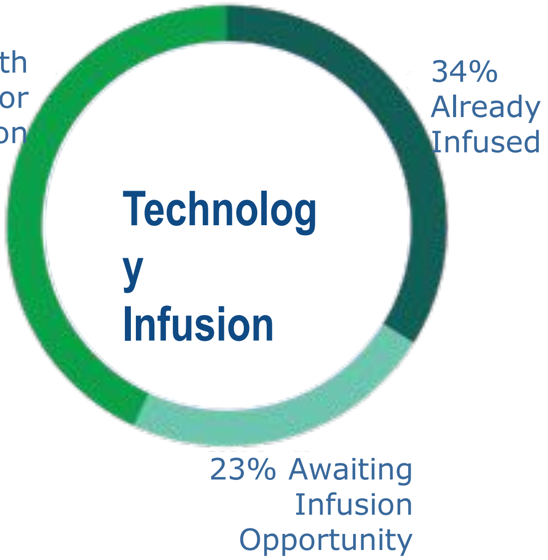
- Competitive, peer-reviewed proposals enable selection of best-of-class technology investments
- Risks are retired before major dollars are invested: a cost-effective approach to technology development and validation
- Successful partnering establishes leveraging opportunities
- This approach has resulted 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

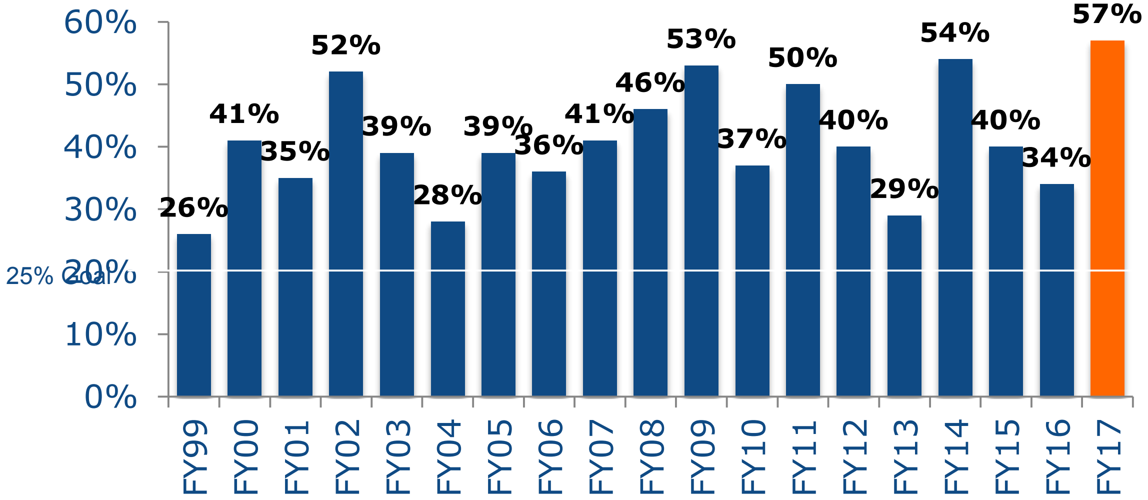
ESTO Enables – Science Selects

FY17 Program Metrics



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.

TRL Advancement



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%.**

NASA AIST Objectives

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

FY18 Primary Technology Thrusts

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
- What tools are needed and how to accelerate development through AIST18

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

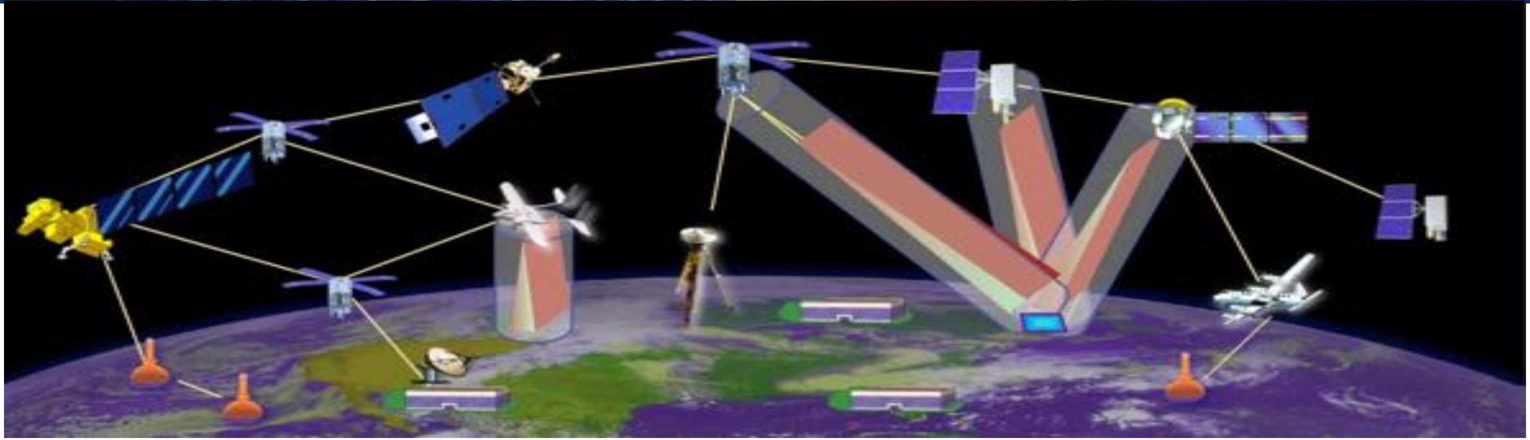
- Public competition with peer review and programmatic alignment

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

Sensor Web Observing System



Technology advances have created an opportunity to make new measurements and to continue others less costly

- Smallsats 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

Project Work Environment

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

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

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

Fostering Technology Infusion

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)

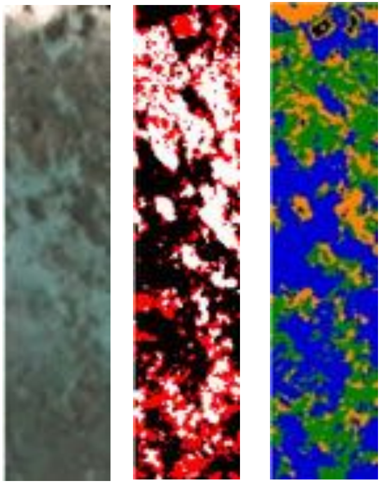
Machine Learning Augmentation & Fluid Lensing for Coral Reef Assessment

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



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

Coral Cover Prediction Accuracy

Method	0.3 m	0.5 m	2 m
K-means	67%	71%	66%
SVM	74%	74%	63%
Augmented	**83%**	**77%**	**69%**

Morphology Prediction Accuracy

Method	0.3 m	0.5 m	2 m
SVM	59%	61%	38%
Augmented	**70%**	**68%**	**60%**

Organic Inorganic Error

Rock Branching Mounding Sand

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

Computer Aided Discovery of Surface Deformation

Objective

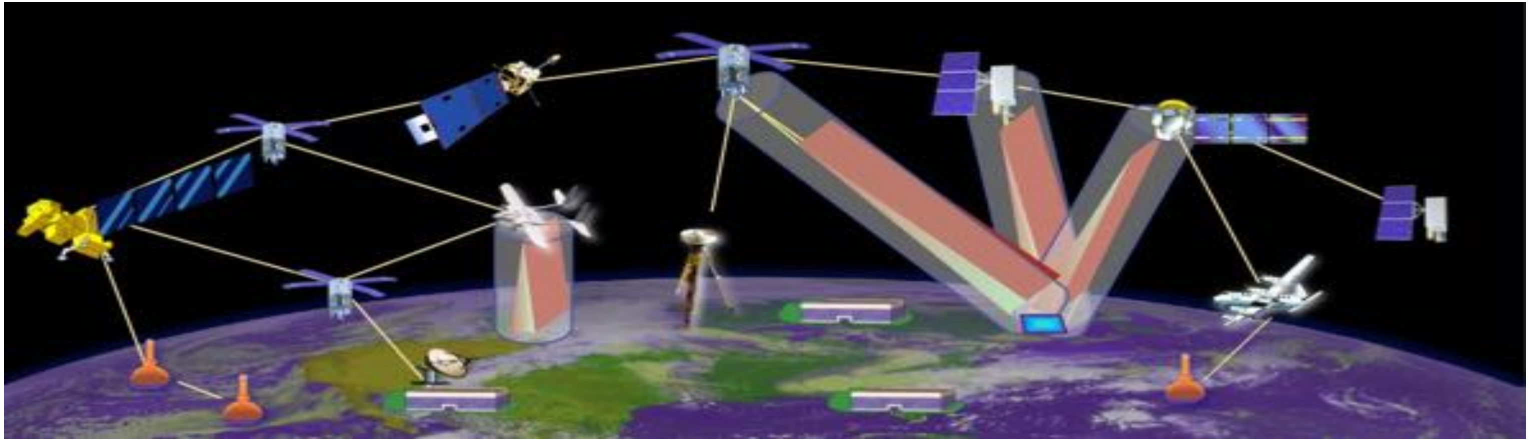
- Improve the ability of scientists to discover new surface deformation phenomena
  - Develop enhanced environments to enable computer-aided discovery and analysis for diverse scientific datasets
  - Allow cross-comparison of multiple datasets
- Demonstrate the value of this approach in 3 scenarios:
  - 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)

Backup

So What is a sensor web?



"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."

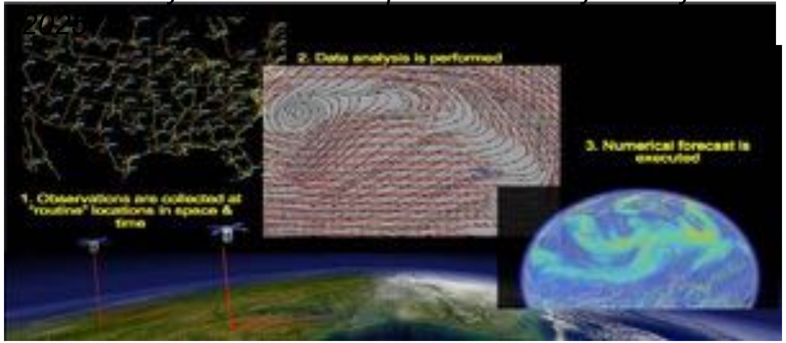
- Steve Talabac et al, 2003

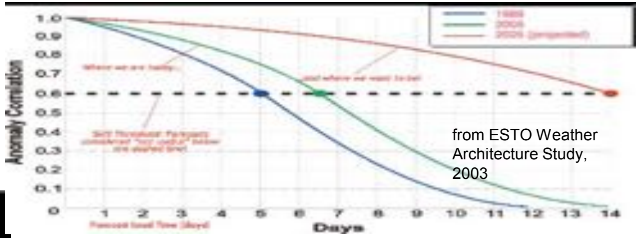
Earth Science Use Cases

- 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

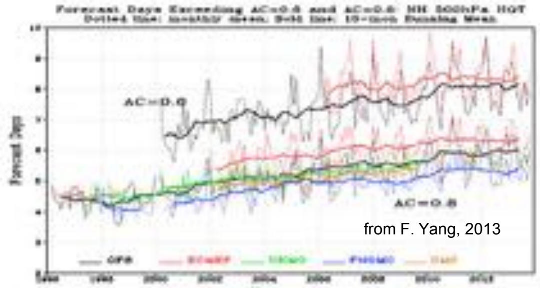
2006: Technology project launched to determine if sensor webs could provide a "revolutionary" improvement in the skill of numerical weather forecasts

Could 7-day skill in 2005 improve to 14-day skill by 2020?



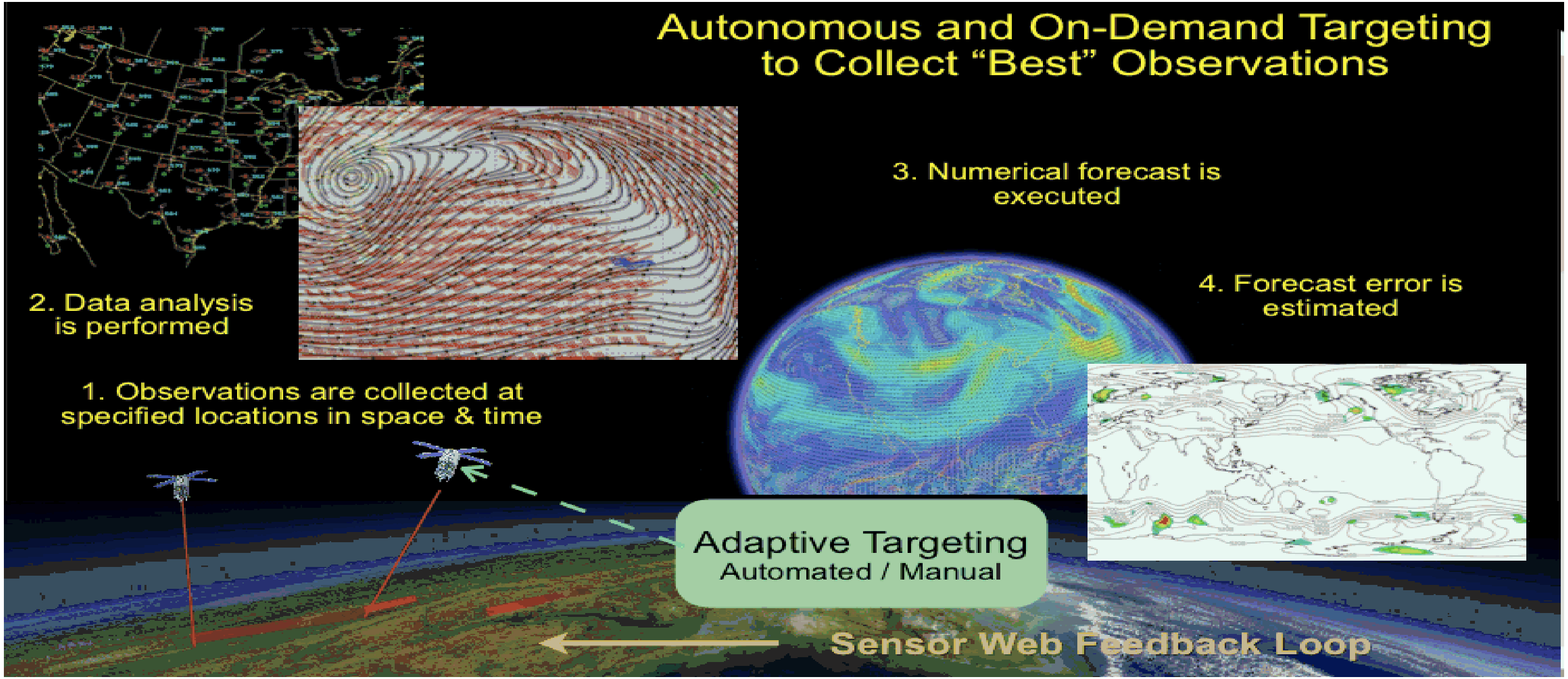


from ESTO Weather Architecture Study, 2003



from F. Yang, 2013

Real-Time Data for Adaptive Targeting

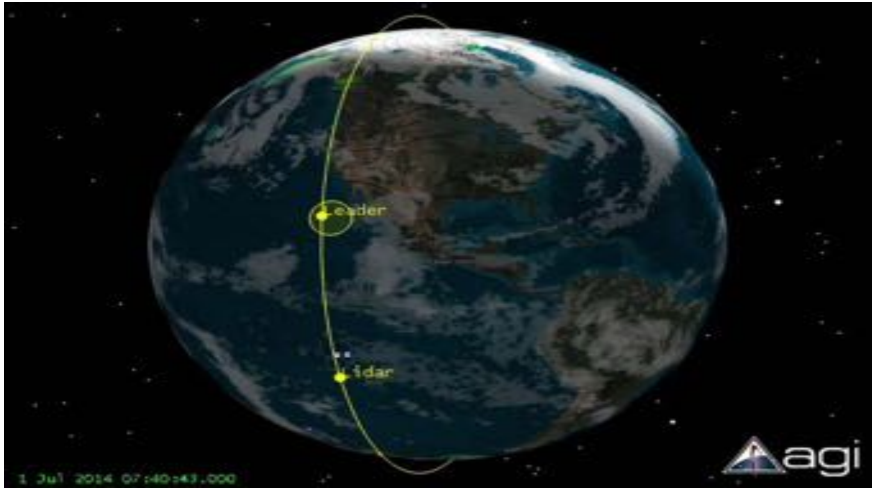


Potential Concept of Operations: Adaptive targeting with a wind lidar

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

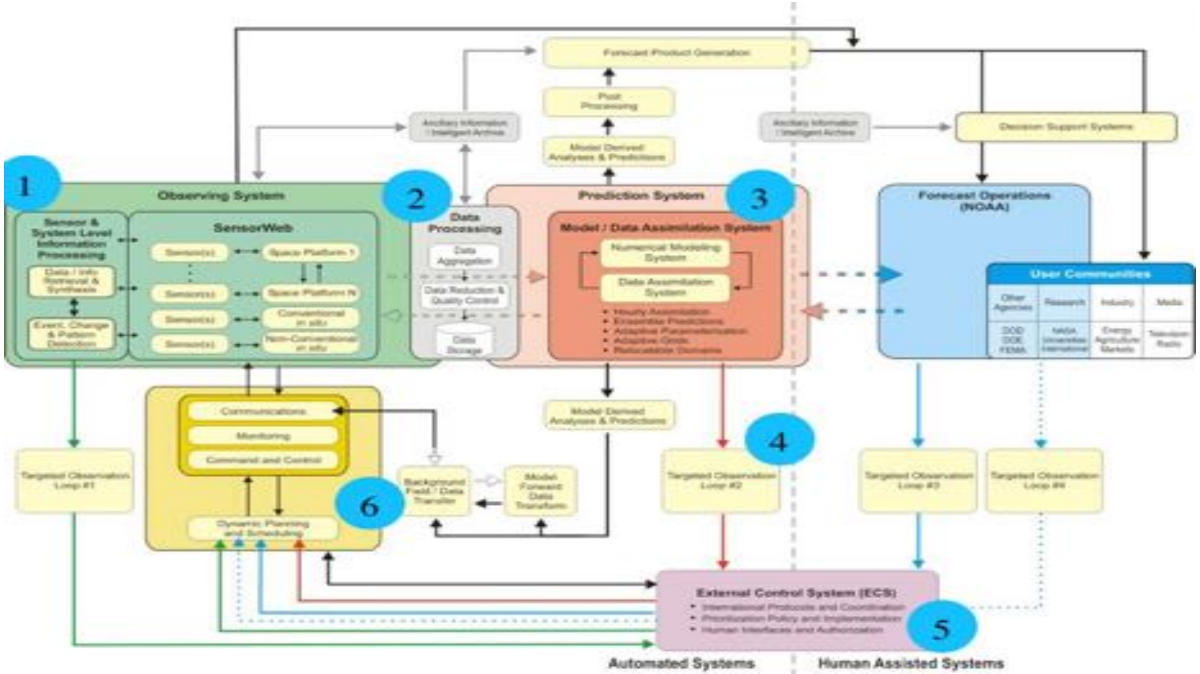
- Cryosphere
  - Glaciers changes
  - Sea Ice changes

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

Proposed Sensor Web Architecture



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)

Higgins, G., Kalb, M., Lutz, R., Mahoney, R., Mauck, R., Seablom, M., Talabac, S., 2003: "Advanced Weather Prediction technologies: Two-Way Interactive Sensor Web & Modeling System"

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

Analytic Center Framework FY18

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

Project Work Environment -JPL-

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/

Climate Impact

- Project Definition
- Plan for Investigation

Data

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

Project Work Environment

-NCCS-

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

Storage

- Local Sparc
- SGI Data Migration Facility

Computing

- Local systems
- High End Computing (DISCOVER)
- GSFC Science Cloud

POC: laura.carriere@nasa.gov

https://www.nccs.nasa.gov/services/Analytics

NASA Earth Exchange (NEX) as Analytic Center

Land Change/Use Community

- Project Definition
- Plan for Investigation

Project Work Environment -NEX-

Data

- Built into website
- Dataset Sources
  - Landsat
  - Sentinel 1A
  - Modis
  - ASTER
  - TRMM
  - AVHRR
- Climate Datasets
- Land cover
- Digital Elevation Map
- STATSGO 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
- Workflow: Jupyter Notebook

Computational Infrastructure

Storage

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

Computing

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

POC: rama.nemani@nasa.gov

https://nex.nasa.gov/nex

AIST-14 Tasks-Machine Learning

esa

PI Name	Org	Prop #	Project Title	ESTO PM	End Date	Machine Learning Elements
Kamalika Das	UC Santa Cruz	115	Uncovering Effects of Climate Variables on Global Vegetation	Oza	5/31/17	Symbolic regression
Milton Halem	*UMBC*	*96*	Computational Technologies: Feasibility Studies of Quantum Enabled Annealing Algorithms for Estimating Terrestrial Carbon Fluxes from OCO-2 and the LIS Model	Cole	5/31/17	RBM, CNN, data assimilation
Hook Hua	JPL	109	Agile Big Data Analytics of High-Volume Geodetic Data Products for Improving Science and Hazard Response	Norton	5/31/17	Fault recognition and Science processing redirection
Thomas Huang	JPL	28	OceanXtremes: Oceanographic Data-Intensive Anomaly Detection and Analysis Portal	Norton	8/29/17	NEXUS as a data delivery tool - oceanography
Kristine Larson	U of CO	4	AMIGHO: Automated Metadata Ingest for GNSS Hydrology within OODT	Hines	5/31/17	Sensor characterization
Victor Pankratius	MIT	36	Computer-Aided Discovery of Earth Surface Deformation Phenomena	Little	5/31/17	Computer aided discovery
Chaowei Yang	GMU	82	Mining and Utilizing Dataset Relevancy from Oceanographic Dataset (MUDROD) Metadata, Usage Metrics, and User Feedback to Improve Data Discovery and Access	Cole	5/31/17	Natural language processing, CNN, SVM, deep learning
Tomasz Stepinski	Cincinnati	27	Pattern-based GIS for Understanding Content of very large Earth Science datasets	Quam	6/31/17	Classification and similarity
Jonathan Gleason	LaRC	95	Ontology-based Metadata Portal for Unified Semantics (OlyMPUS)	Oza	1/31/17	Precision ontology foundation
Constantine Lukashin	LaRC	14	NASA Information And Data System (NAIADS) for Earth Science Data Fusion and Analytics	Murray	1/31/17	Scientific enterprise service bus
Aashish Chaudhary	Kitware	65	Visualization Pipelines for big-data on the NASA Earth Exchange (NEX) Prototyping Agile Production, Analytics	Hines	3/31/17	Workflow
Martyn Clark	UCAR	88	Development of Computational Infrastructure to Support Hyper-resolution Large-ensemble Hydrology Simulations from Local-to-Continental Scales	Hines	4/30/17	Assimilation and ensembles
Kwo-Sen Kuo	Bayesics	56	DERECHOS: Data Environment for Rapid Exploration and Characterization of Organized Systems	Little	4/30/17	Foundation for data delivery
Seungwon Lee	JPL	32	Climate Model Diagnostic Analyzer	Norton	4/30/17	
Christian Mattman	JPL	34	SciSpark: Highly Interactive and Scalable Model Evaluation and Climate Metrics for Scientific Data and Analysis	Norton	4/30/17	Foundation for data delivery

AIST QRS Tasks-Machine Learning

esa

PI Name	Org	Prop #	Project Title	End Date	Machine Learning Elements
Dan Duffy	GSFC	AIST-QRS-16-0002	Demo of VR Technology with live data(Phase 2) & GIS (Phase 3)	12/31/16	VR
Sreeja Nag	ARC	AIST-QRS-16-0003	Scheduling Satellite Pointing within Constellations	12/31/16	
Ved Chirayeth	ARC	AIST-QRS-16-0004	MiDAR-fused Supervised Machine Learning (SML)	1/31/17	Use of high resolution training sets to improve global scale moderate resolution data
Chris Mattman	JPL	AIST-QRS-16-0007	Deep Web Search Analytics	10/31/16	Text analytics
John Readey	HDF Group	ACCESS15 -0031	Object Store-based Data Service for Earth System Science	5/31/17	Foundation for data storage to improve access
Yehuda Bock	Scripps	AIST-QRS-16-0010	Latency test of realtime warning systems in AMCE Cloud Computing	9/30/17	Feature Detection in near real time
Amy Braverman	JPL	AIST-QRS-16-0005	Probablistic Climate Model Evaluation	9/20/16	What do you mean by similar?
Chris Lynnes	GSFC	AIST-QRS-16-0001	Experiment with Data Containers in ESDIS Context	12/20/16	Evaluation of alternative storage models for data to enable analysis

NASA AIST-16 Tasks-Machine Learning

esa

PI Name	Org	Project Title	End Date	Machine Learning Elements
Victor Pankratius	*MIT*	Computer Aided Discovery and Algorithmic Synthesis for Spatio-Temporal Phenomena in InSAR	7/30/19	Computer Aided Discovery
Jacqueline LeMoigne-Stewart	GSFC	TAT-C ML	7/30/19	Design space iterator
Branko Kosovic	NCAR	Fuel moisture content for improved fire prediction	5/31/17	Data assimilation
Andrew Michaelis	ARC	Framework for Mining and Analysis of Petabyte-size Time-series on the NASA Earth Exchange (NEX)	7/31/17	Time series analysis
Barton Forman	UMD	A mission planning tool for next generation remote sensing of snow	8/30/19	OSSE and assimilation
Ved Chirayath	*ARC*	NeMO-Net - The Neural Multi-Modal Observation & Training Network for Global Coral Reef Assessment	8/31/19	High resolution training sets improve moderate resolution imagery
Dara Entekhabi	GSFC	Autonomous Moisture Continuum Sensing Network	8/31/19	
Milton Halem	UMBC	Hybrid Quantum Annealing Approaches for Inferring &Assimilating Satellite Surface Flux Data into Global Land Surface Models.	8/31/19	Assimilation, registration
Jeffrey Morisette	USGS	Advanced Phenological Information System	8/31/19	
Walter Jetz	Yale	Software workflows for remote sensing-based biodiversity change monitoring	9/14/19	
Jonathon Hobbs	JPL	Simulation-based Uncertainty Quantification	9/30/19	Statistics
Martyn Clark	NCAR	Climate risks in water sector: Advancing the readiness of emerging technologies in climate downscaling & hydrologic modeling	9/30/19	

Simulating Observations from Future Observing Systems

- 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.

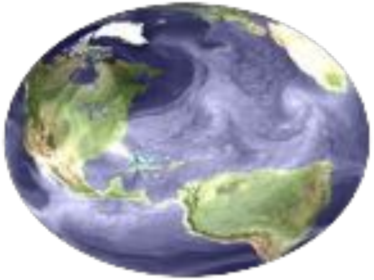
- Trade-offs in instrument or orbital configurations and methods of assimilating a new type of observing system can be determined by an OSSE and ultimately result in both time and cost savings.

- 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.



Simulated Clouds from the NASA/DAO fvGCM Nature Run

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