Towards the Development of a Global, Satellite-based, Terrestrial Snow Mission Planning Tool

Co-authors: Sujay Kumar\textsuperscript{1}, Jacqueline Le Moigne\textsuperscript{2}, and Sreeja Nag\textsuperscript{2,3}

\textsuperscript{1}=NASA GSFC - Hydrological Sciences; \textsuperscript{2}=NASA GSFC - Software Engineering; \textsuperscript{3}=Bay Area Environmental Research Institute

Bart Forman

Assistant Professor, University of Maryland

The Deborah J. Goodings Professor of Global Sustainability

Department of Civil and Environmental Engineering

December 12\textsuperscript{th}, 2017
Satellite-derived Snow “Information”
Satellite-derived Snow “Information”
Satellite-derived Snow “Information”
Satellite-derived Snow “Information”
Science and mission planning questions:

1. What **observational records** are needed (in space and time) to maximize terrestrial snow experimental utility?

2. How might observations be **coordinated** (in space and time) to maximize this utility?

3. What is the **additional utility** associated with an additional observation?

4. How can future **mission costs** be minimized while ensuring Science requirements are fulfilled?
Science and mission planning questions:

1. What observational records are needed (in space and time) to maximize terrestrial snow experimental utility?
2. How might observations be coordinated (in space and time) to maximize this utility?
3. What is the additional utility associated with an additional observation?
4. How can future mission costs be minimized while ensuring Science requirements are fulfilled?
**Science and mission planning questions:**

1. What observational records are needed (in space and time) to maximize terrestrial snow experimental utility?
2. How might observations be coordinated (in space and time) to maximize this utility?
3. What is the additional utility associated with an additional observation?
4. How can future mission costs be minimized while ensuring Science requirements are fulfilled?
Science and mission planning questions:

1. What **observational records** are needed (in space and time) to maximize terrestrial snow experimental utility?

2. How might observations be **coordinated** (in space and time) to maximize this utility?

3. What is the **additional utility** associated with an additional observation?

4. How can future **mission costs be minimized** while ensuring Science requirements are fulfilled?
Observing System Simulation Experiment

Nature Run	Snow Depth & SWE over North America

LIS + MERRA2 - model-based representation
Observing System Simulation Experiment

Nature Run | Snow Depth & SWE over North America

LIS + MERRA2 - model-based representation

TAT-C

Sub-sample in space / time

Permutation of Orbit(s) + Sensor(s)
Observing System Simulation Experiment

Observations
Objectives
OSSE
TAT-C
Hyperplanes
Eulerian Grid
Single Platform
Constellation
Trade-off Space
Machine Learning
Emulators
Variability
Experiments
Conclusions
Extra Slides
Observing System Simulation Experiment

Nature Run
LIS + MERRA2 - model-based representation

Snow Depth & SWE over North America

TAT-C
Sub-sample in space / time

Permutation of Orbit(s) + Sensor(s)

TAT-C
Mission cost estimate and risk analysis

T_{b}, \sigma_{o}, and \delta h
Operators

Machine Learning "Emulators"

Synthetic Obs.
Observing System Simulation Experiment

Nature Run
- Snow Depth & SWE over North America
  - LIS + MERRA2 - model-based representation

TAT-C
- Sub-sample in space / time
  - Permutation of Orbit(s) + Sensor(s)

LIS Open Loop
- LIS + GLDAS
  - apply representative B.C. error
  - no assimilation (a.k.a., Open Loop)
  - with assimilation (merge with observations from suite of sensors)

TAT-C
- Mission cost estimate and risk analysis
  - \( T_b, \sigma_0, \text{ and } \delta h \) Operators
  - Machine Learning "Emulators"
  - Synthetic Obs.
Observing System Simulation Experiment

- Nature Run: Snow Depth & SWE over North America
  - LIS + MERRA2 - model-based representation

- TAT-C: Permutation of Orbit(s) + Sensor(s)
  - Sub-sample in space / time
  - TB, CO, and δh Operators

- LIS Open Loop: LIS + GLDAS
  - apply representative B.C. error
  - no assimilation (a.k.a., Open Loop)
  - with assimilation (merge with observations from suite of sensors)

- LIS Assimilation: Open Loop (i.e., no assimilation)

- Mission cost estimate and risk analysis

- OSSE
  - Hyperplanes
  - Eulerian Grid
  - Single Platform
  - Constellation
  - Trade-off Space

- Machine Learning
  - Emulators
  - Variability
  - Experiments

- Conclusions
- Extra Slides
Observing System Simulation Experiment

- Nature Run: Snow Depth & SWE over North America
  - LIS + MERRA2 - model-based representation

- TAT-C
  - Sub-sample in space / time
  - Permutation of Orbit(s) + Sensor(s)

- LIS Open Loop
  - LIS + GLDAS - apply representative B.C. error
  - no assimilation (a.k.a., Open Loop)
  - with assimilation (merge with observations from suite of sensors)

- LIS Assimilation
  - Open Loop (i.e., no assimilation)
  - Data Assimilation (Bayesian merger w/ synthetic obs.)

- Mission cost estimate and risk analysis
  - T_{B}, \sigma_{0}, and \delta h Operators
  - Machine Learning "Emulators"
  - Synthetic Obs.
Observing System Simulation Experiment

Nature Run
- Snow Depth & SWE over North America
  - LIS + MERRA2 model-based representation

TAT-C
- Sub-sample in space / time
  - Permutation of Orbit(s) + Sensor(s)

LIS Open Loop
- LIS + GLDAS
  - apply representative B.C. error
  - no assimilation (a.k.a., Open Loop)
  - with assimilation (merge with observations from suite of sensors)

TAT-C
- Mission cost estimate and risk analysis
  - $T_B$, $c_0$, and $\delta h$
  - Operators

LIS Assimilation
- Open Loop (i.e., no assimilation)
- Data Assimilation
  (Bayesian merger w/ synthetic obs.)

Machine Learning
“Emulators”

Synthetic Obs.

Benchmark evaluation against “Nature Run”

Land Validation Toolkit (LVT)
TAT-C Orbital Simulator
TAT-C Orbital Simulator
“Comb” Viewing $\rightarrow$ Single Platform
“Comb” Viewing $\mapsto$ Constellation
• Explore trade-off between engineering and science
  ▶ Field-of-View (FOV)?
  ▶ Platform altitude?
  ▶ Repeat cycle?
  ▶ Single platform vs. constellation?
  ▶ Orbital configuration(s)?

• How do we get the most scientific bang for our buck?
Machine Learning “Emulators”

Observations
Objectives
OSSE
TAT-C
Hyperplanes
Eulerian Grid
Single Platform
Constellation
Trade-off Space
Machine Learning
Emulators
Variability
Experiments
Conclusions
Extra Slides

Physically-based Land Surface Model(s)

Observation Operator
(Forman et al., 2013;
Forman and Reichle, 2014;
Forman and Xue, 2016)

Multi-frequency, Multi-polarization Training Targets

Machine Learning “Emulators”

Observation Operator
(Forman et al., 2013; Forman and Reichle, 2014; Forman and Xue, 2016)

Multi-frequency, Multi-polarization Training Targets

Physically-based Land Surface Model(s)

Machine Learning “Emulators”

Physically-based Land Surface Model(s)

Observation Operator
(Forman et al., 2013; Forman and Reichle, 2014; Forman and Xue, 2016)

Multi-frequency, Multi-polarization Training Targets

Xue and Forman, 2015
Remote Sensing of Environ.
Spatiotemporal Variability

- Observations
- Objectives
- OSSE
- TAT-C
- Hyperplanes
- Eulerian Grid
- Single Platform
- Constellation
- Trade-off Space
- Machine Learning
- Emulators
- Variability
- Experiments
- Conclusions
- Extra Slides
Relevancy Scenarios

- **Scenario 1**: Benchmark Analysis
  - Passive MW Assimilation only

- **Scenario 2**: Comparative Analysis
  - Passive MW vs. Active MW vs. LIDAR

- **Scenario 3**: Multi-sensor Analysis
  - single-sensor platform
  - multi-sensor platform
  - constellation of sensors
• Global snow mission will require **evidence of achievable science** via OSSE . . . or some other means

• NASA LIS provides “nature run” plus assimilation framework

• TAT-C provides **spatiotemporal sub-sampling** of observations, including cost estimates and risk assessments

• **Machine learning** maps model state(s) into observation space (i.e., $T_b$ and $\sigma_0$)
  ▶ Enables integration of $T_b$, $\sigma_0$, and $\delta h$ in geophysical realm (i.e., SWE and snow depth)
  ▶ Multiple frequencies/polarizations/observations allow for flexibility and modularity in DA framework

• **Snow OSSE is on-going** → open to ideas + suggestions!
Global snow mission will require evidence of achievable science via OSSE . . . or some other means.

NASA LIS provides “nature run” plus assimilation framework.

TAT-C provides spatiotemporal sub-sampling of observations, including cost estimates and risk assessments.

Machine learning maps model state(s) into observation space (i.e., $T_b$ and $\sigma_0$)

- Enables integration of $T_b$, $\sigma_0$, and $\delta h$ in geophysical realm (i.e., SWE and snow depth).
- Multiple frequencies/polarizations/observations allow for flexibility and modularity in DA framework.

Snow OSSE is on-going → open to ideas + suggestions!
Global snow mission will require evidence of achievable science via OSSE . . . or some other means

NASA LIS provides “nature run” plus assimilation framework

TAT-C provides spatiotemporal sub-sampling of observations, including cost estimates and risk assessments

Machine learning maps model state(s) into observation space (i.e., $T_b$ and $\sigma_0$)

- Enables integration of $T_b$, $\sigma_0$, and $\delta h$ in geophysical realm (i.e., SWE and snow depth)
- Multiple frequencies/polarizations/observations allow for flexibility and modularity in DA framework

Snow OSSE is on-going → open to ideas + suggestions!
Global snow mission will require evidence of achievable science via OSSE . . . or some other means.

NASA LIS provides “nature run” plus assimilation framework.

TAT-C provides spatiotemporal sub-sampling of observations, including cost estimates and risk assessments.

Machine learning maps model state(s) into observation space (i.e., $T_b$ and $\sigma_0$).
  
  ▶ Enables integration of $T_b$, $\sigma_0$, and $\delta h$ in geophysical realm (i.e., SWE and snow depth).
  
  ▶ Multiple frequencies/polarizations/observations allow for flexibility and modularity in DA framework.

Snow OSSE is on-going → open to ideas + suggestions!
Global snow mission will require evidence of achievable science via OSSE . . . or some other means

NASA LIS provides “nature run” plus assimilation framework

TAT-C provides spatiotemporal sub-sampling of observations, including cost estimates and risk assessments

Machine learning maps model state(s) into observation space (i.e., $T_b$ and $\sigma_0$)

- Enables integration of $T_b$, $\sigma_0$, and $\delta h$ in geophysical realm (i.e., SWE and snow depth)
- Multiple frequencies/polarizations/observations allow for flexibility and modularity in DA framework

Snow OSSE is on-going → open to ideas + suggestions!
Thank You.

Questions and/or Comments?

Financial support provided by:

NASA New Investigator Program (NNX14AI49G)
NASA GRACE-FO Science Team (NNX16AF17G)
NASA High Mountain Asia Science Team (NNX17AC15G)

High-performance computing support provided by

UMD’s Division of Information Technology
For parameters $C > 0$ and $\varepsilon > 0$, the **standard (primal)** form is:

$$
\begin{align*}
\text{minimize} & \quad \frac{1}{2} \langle w \cdot w \rangle + C \sum_{i=1}^{m} (\xi_i + \xi_i^*) \\
\text{subject to} & \quad \langle w \cdot \phi(x_i) \rangle + \delta - z_i \leq \varepsilon + \xi_i \\
& \quad z_i - \langle w \cdot \phi(x_i) \rangle - \delta \leq \varepsilon + \xi_i^* \\
& \quad \xi_i, \xi_i^* \geq 0, i = 1, 2, \ldots, m.
\end{align*}
$$

where $m$ is the available number of $T_b$ measurements in time (for a given location in space), $z_i$ is a $T_b$ measurement at time $i$, and $\xi$ and $\xi^*$ are slack variables.
Primal optimization is commonly solved in **dual form** as:

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} \sum_{i,j=1}^{m} (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle \phi(x_i) \cdot \phi(x_j) \rangle \\
& \quad + \varepsilon \sum_{i=1}^{m} (\alpha_i + \alpha_i^*) - \sum_{i=1}^{m} z_i (\alpha_i - \alpha_i^*) \\
\text{subject to} & \quad \sum_{i=1}^{m} (\alpha_i - \alpha_i^*) = 0, \\
& \quad \alpha_i, \alpha_i^* \in [0, C], \ i = 1, 2, \ldots, m
\end{align*}
\]

where \(\alpha_i\) and \(\alpha_i^*\) are Lagrangian multipliers, \(\langle \phi(x_i) \cdot \phi(x_j) \rangle\) is the inner dot product of \(\phi(x_i)\) and \(\phi(x_j)\), \(\varepsilon\) is the specified error tolerance, and \(C\) is a positive constant that dictates a penalized loss during training.