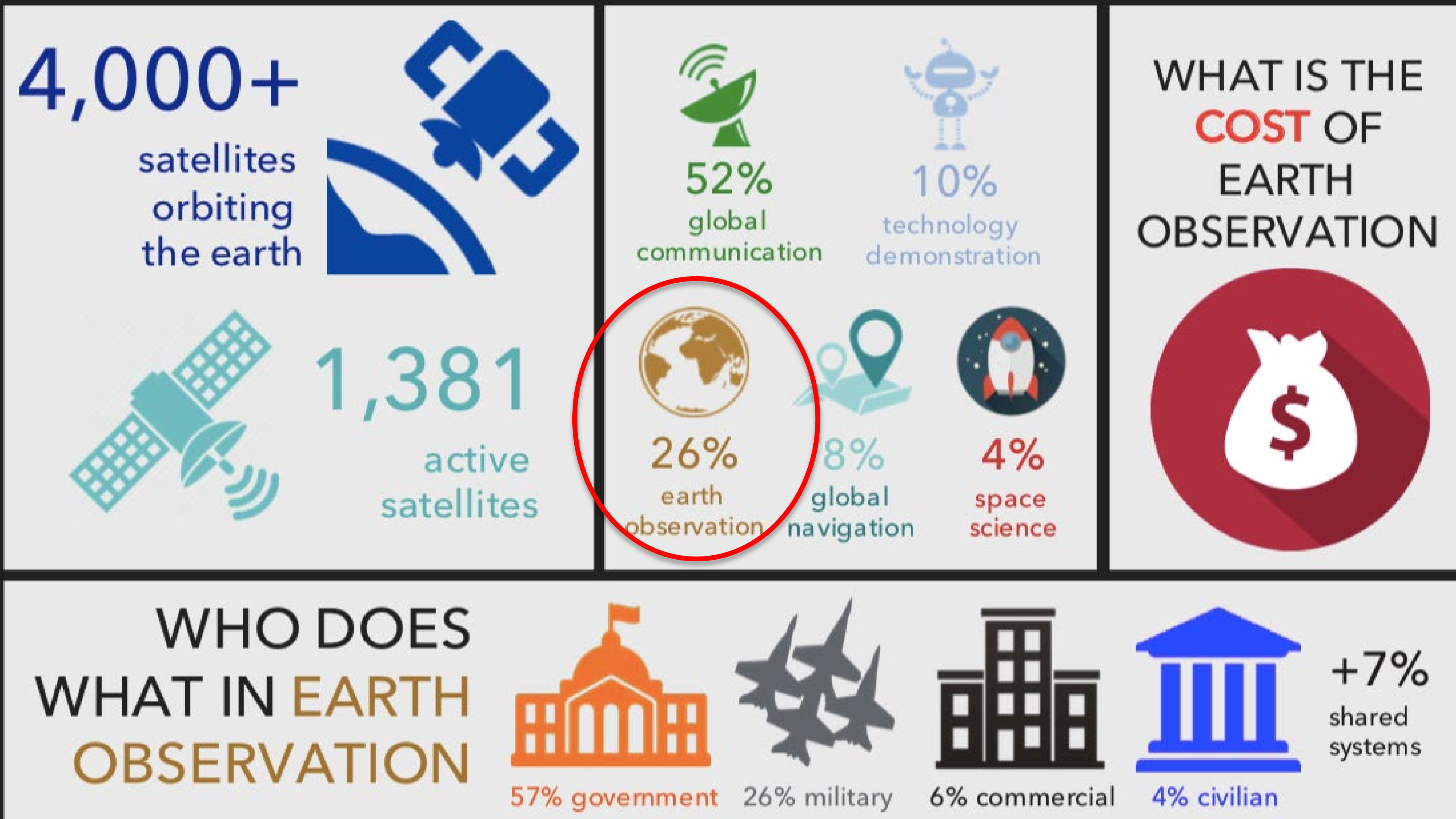
INTELLIGENCE 2025
to improve life on Earth

Efi Foufoula-Georgiou
University of California Irvine (UCI)

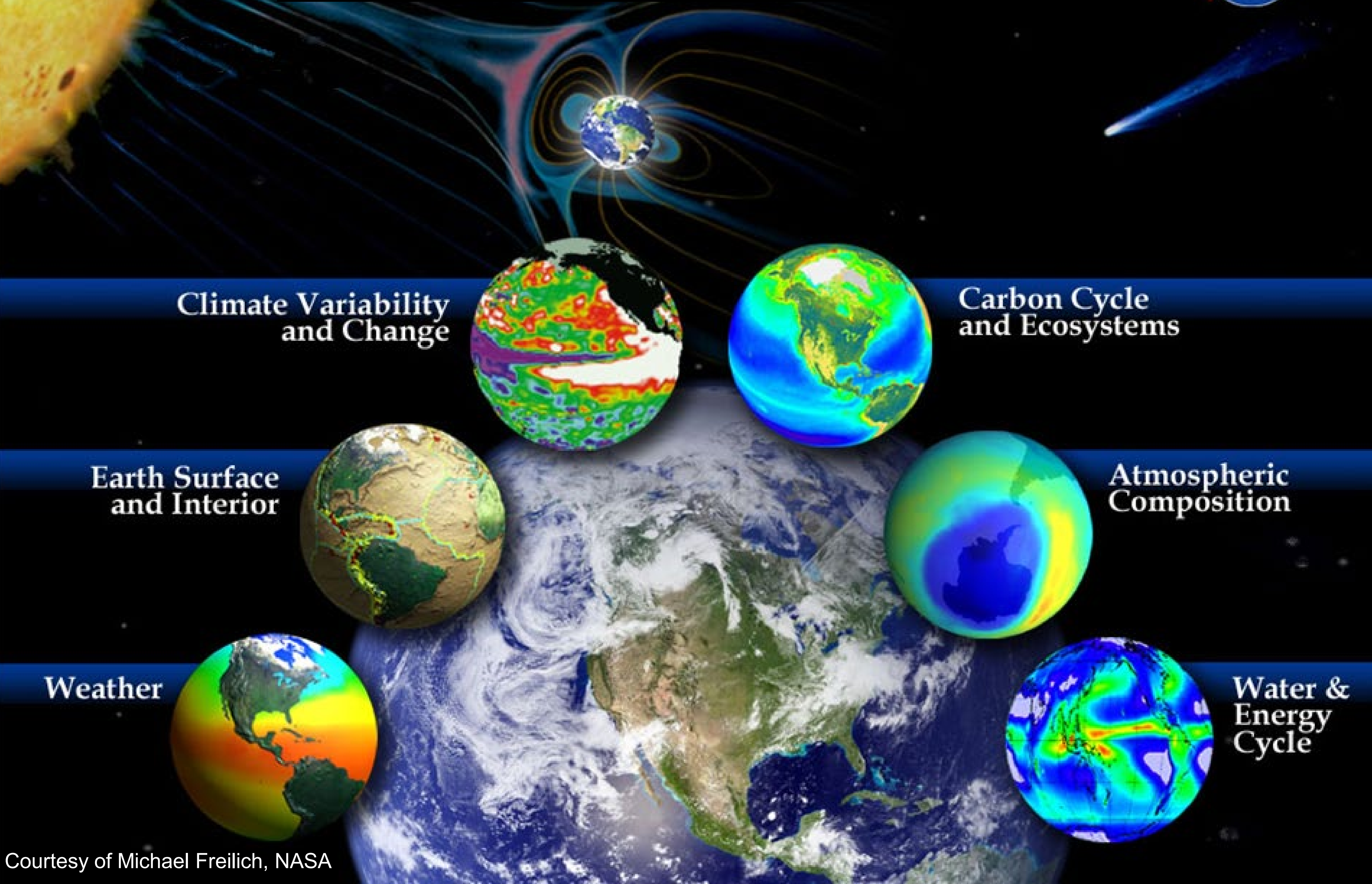
TAU-UCI
Nov 13, 2018

THE STATE OF PLAY IN SPACE TODAY



Details from http://www.ucsusa.org/nuclear-weapons/space-weapons/satellite-database#.Vzo1eBV96t8 for 1/1/2016

Understanding Earth from Space



Courtesy of Michael Freilich, NASA

Earth Science Missions
FY17 Program of Record

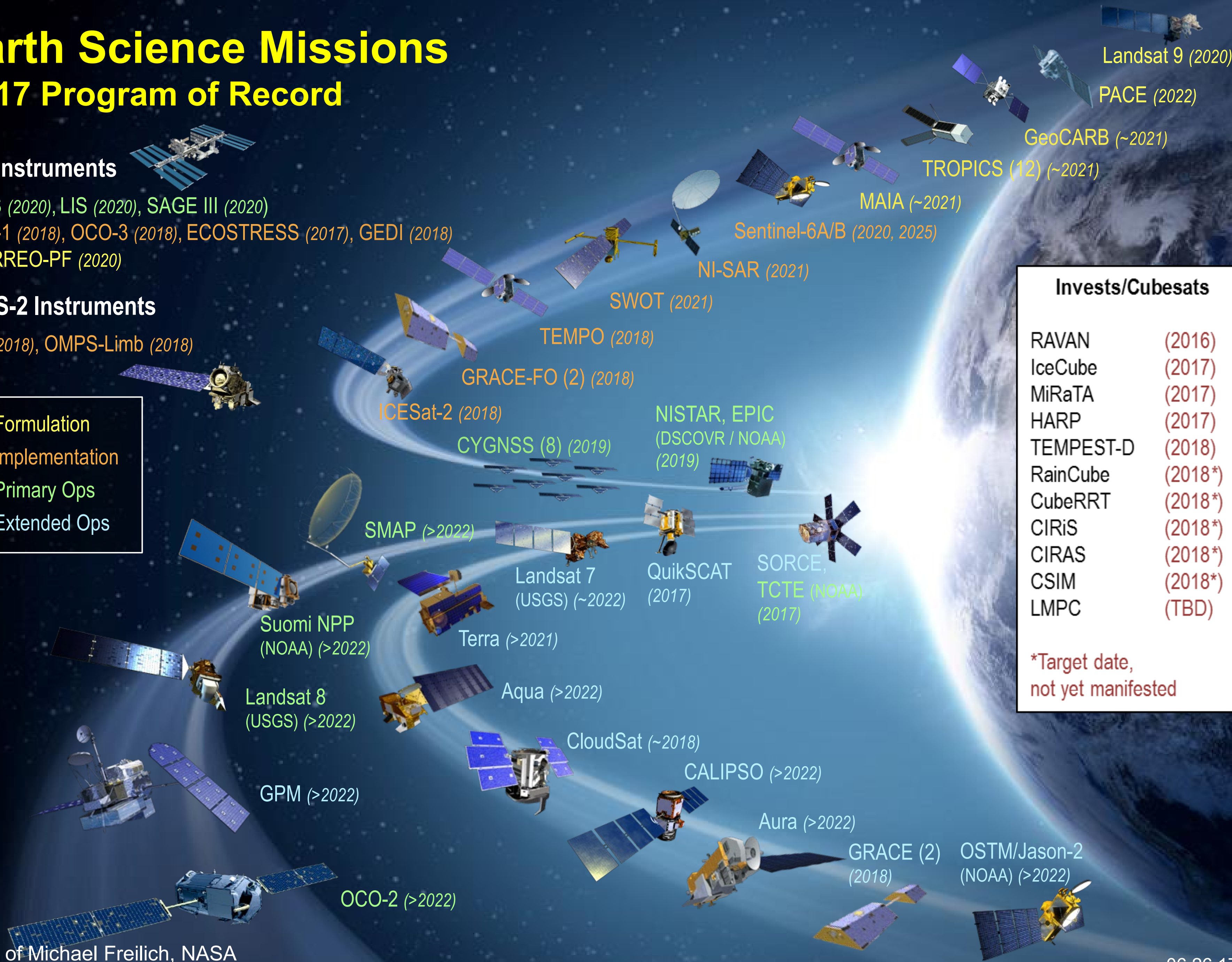
ISS Instruments

CATS *(2020)*, LIS *(2020)*, SAGE III *(2020)*
TSIS-1 *(2018)*, OCO-3 *(2018)*, ECOSTRESS *(2017)*, GEDI *(2018)*
CLARREO-PF *(2020)*

JPSS-2 Instruments

RBI *(2018)*, OMPS-Limb *(2018)*

- Formulation
- Implementation
- Primary Ops
- Extended Ops



Invests/Cubesats

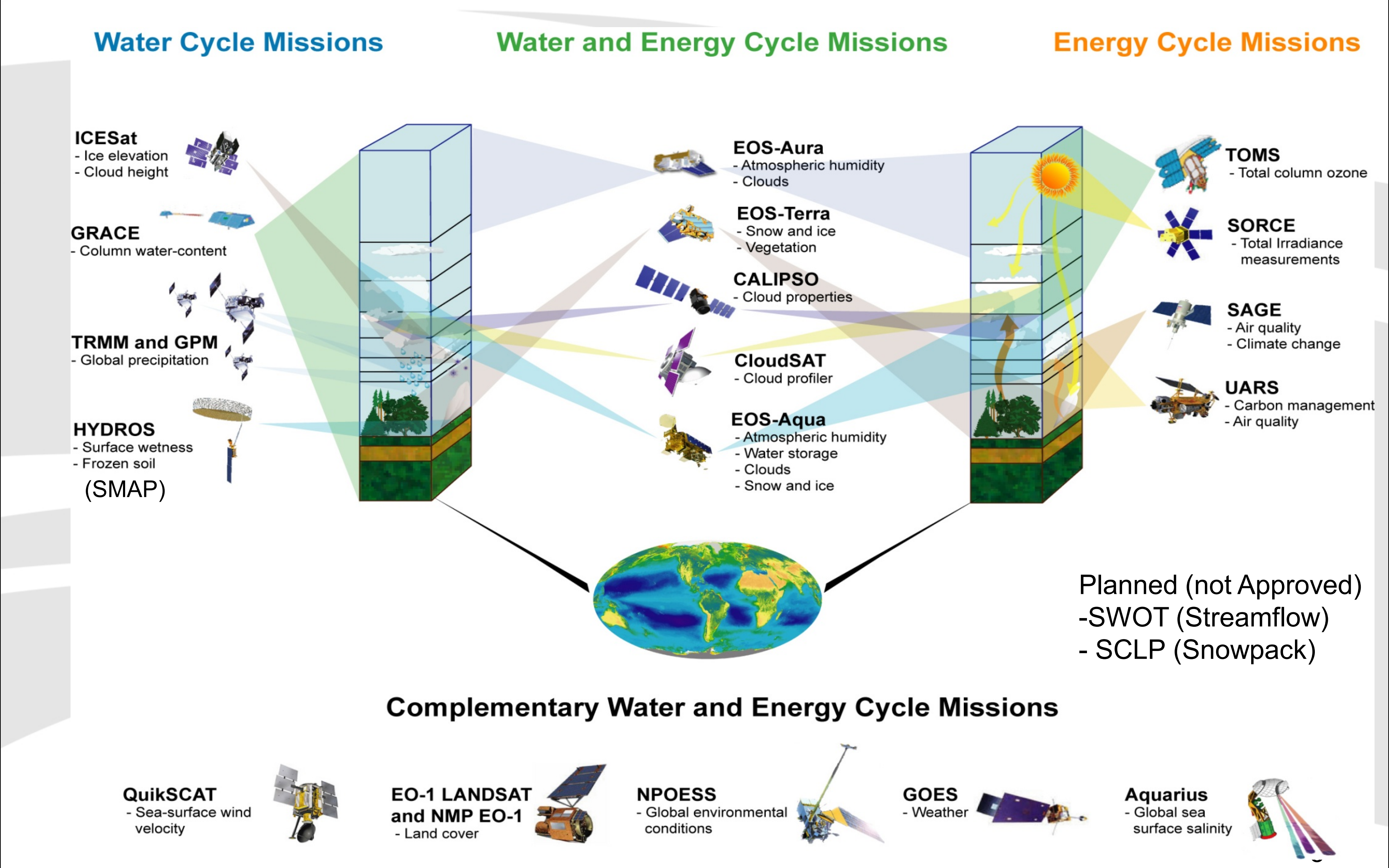
RAVAN	(2016)
IceCube	(2017)
MiRaTA	(2017)
HARP	(2017)
TEMPEST-D	(2018)
RainCube	(2018*)
CubeRRT	(2018*)
CIRiS	(2018*)
CIRAS	(2018*)
CSIM	(2018*)
LMPC	(TBD)

*Target date, not yet manifested

Courtesy of Michael Freilich, NASA

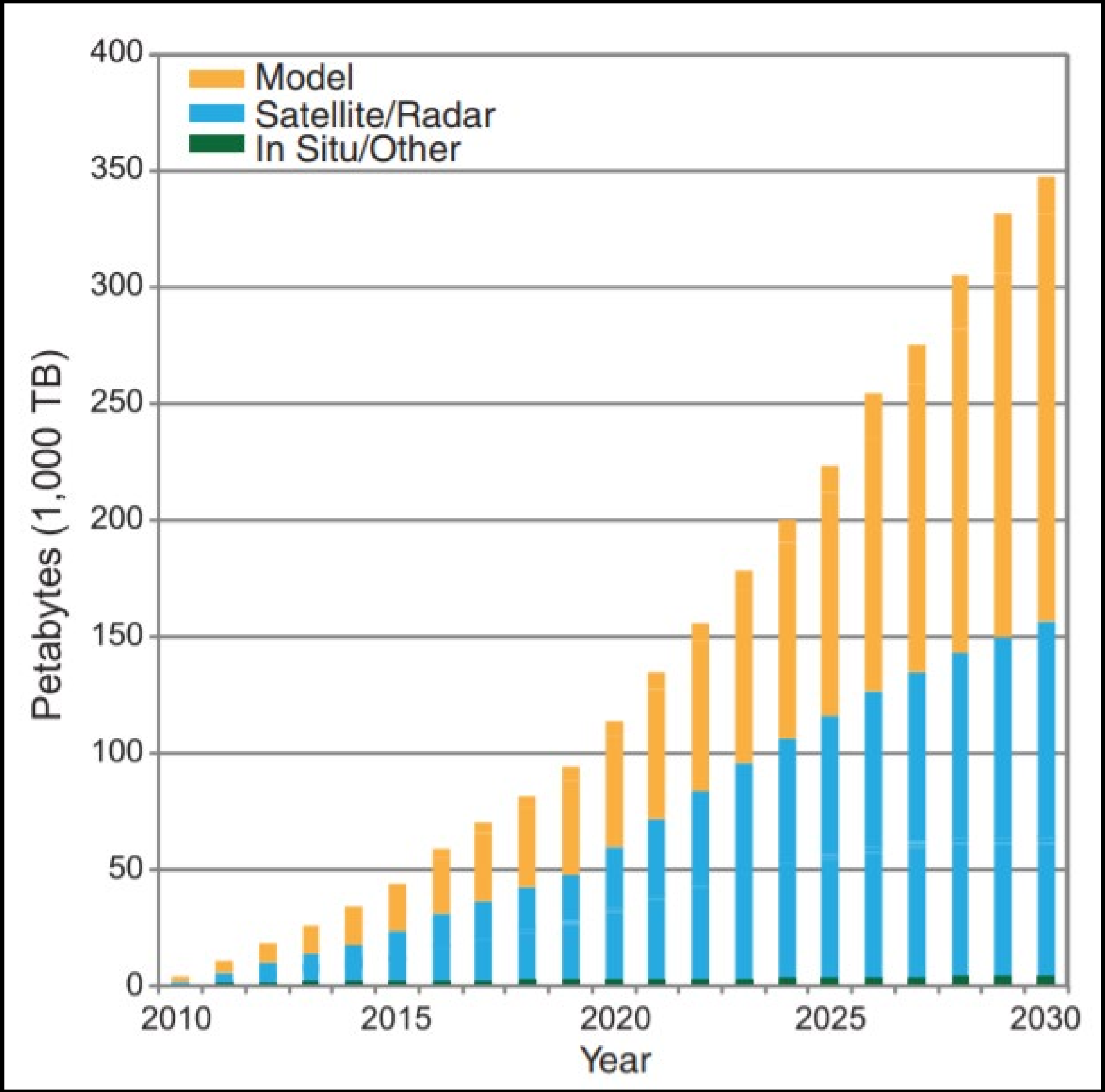
06.26.17

NASA's Water and Energy Cycle Missions



Courtesy of Michael Freilich, NASA

Exploding Volume of Climate Data from Space

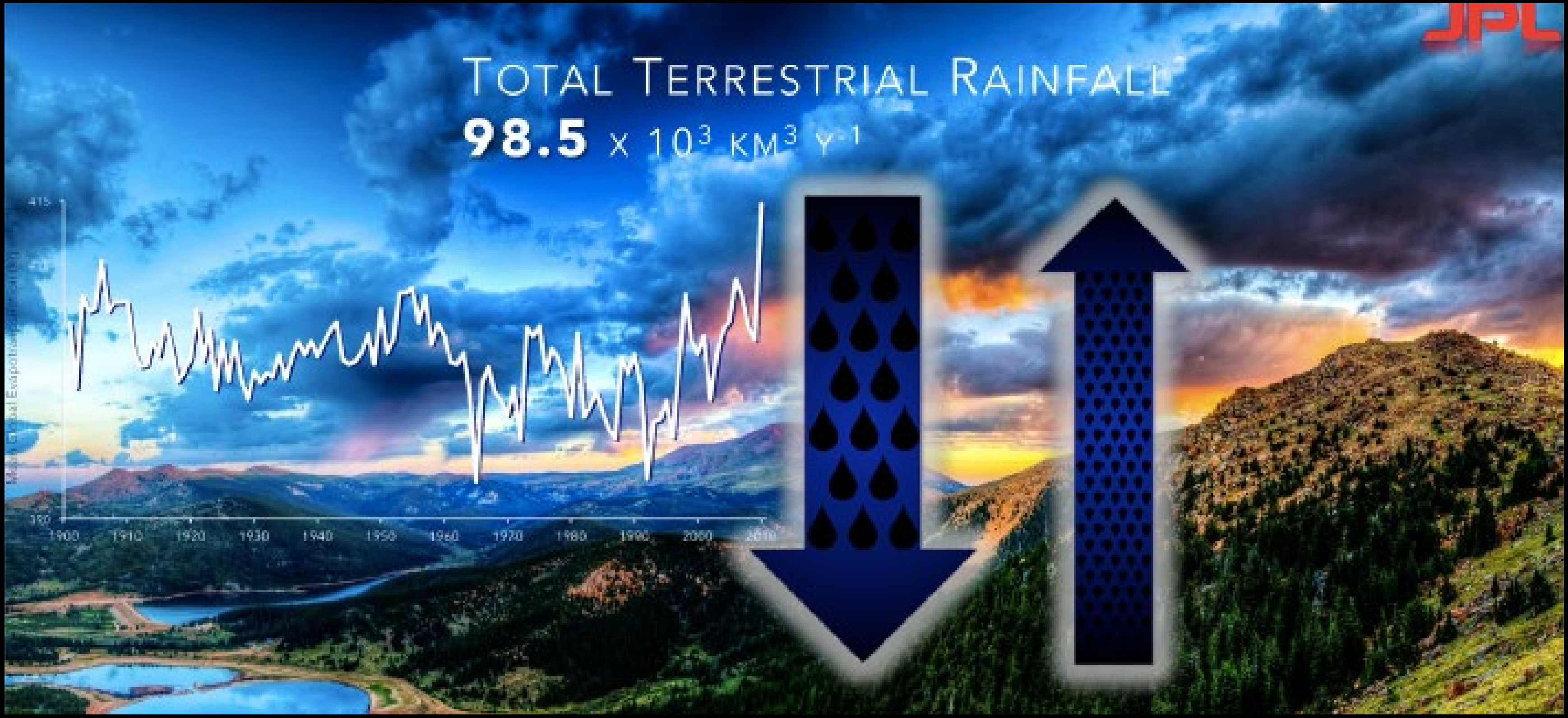


Overpeck et al,
Science, 2011

3 Themes @ UCI

1. Precipitation estimation from space
2. Climate dynamics for prediction
3. Landforms supporting life (rivers, deltas)

1. PRECIPITATION



Water cycle dynamics at global to regional scales
Monitoring extremes (hurricanes, tropical storms)
Improving weather and climate models

From TRMM to GPM

Covering 35S to 35N
Microwave Imager (TMI)
-- 9 channels
-- frequencies 10.7-to-85.5 GHz
-- swath width 878 km
Precipitation radar (PR)
-- single-frequency Ku band (13 GHz)
-- swath width 247 km

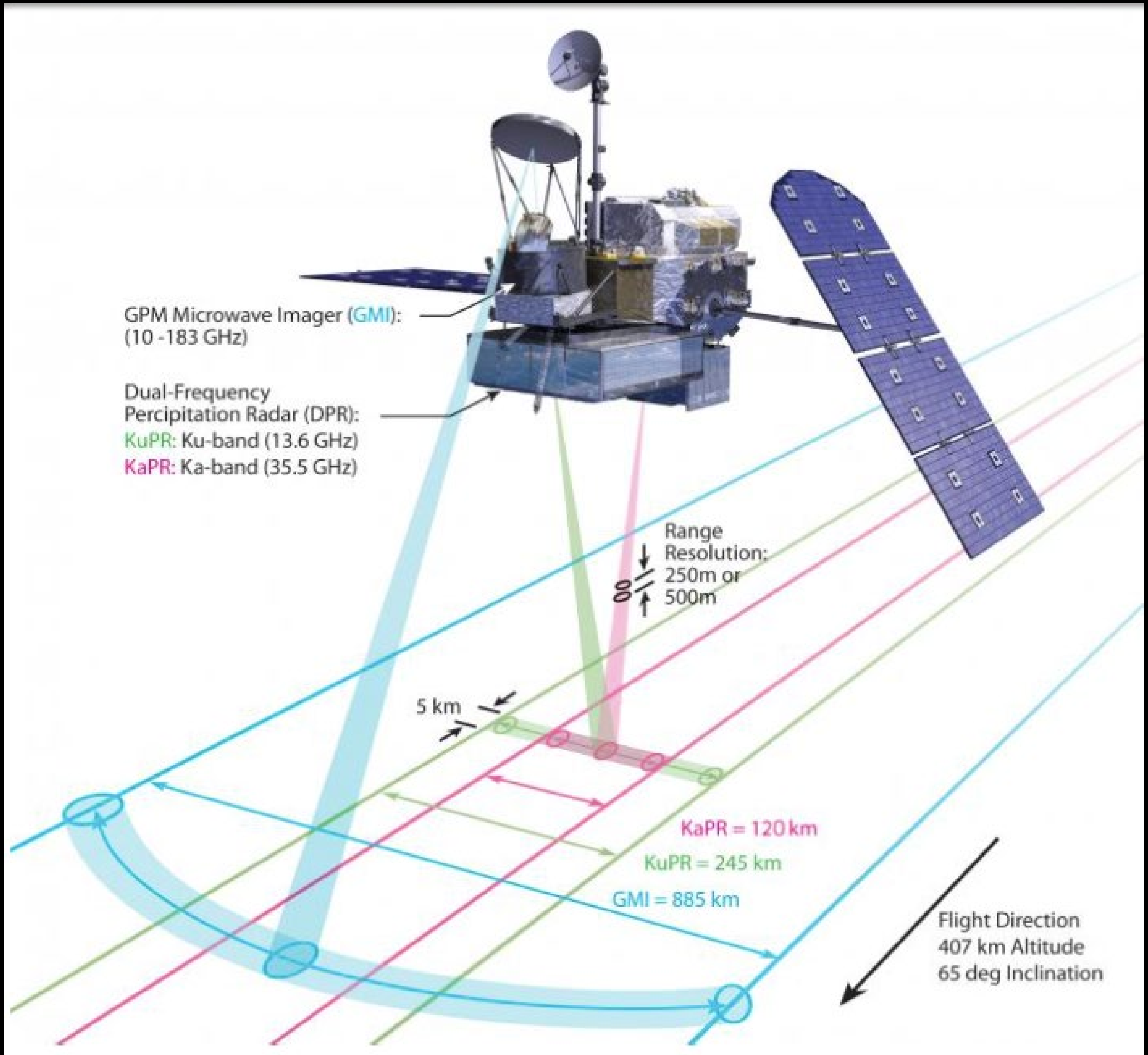
Covering 68S to 68N
GPM Microwave Imager (GMI)
-- 13 dual-polarized channels
-- frequencies 10.65-183.3 GHz
-- swath width 885 km
Dual Polarization radar (DPR)
-- dual-frequency Ku & Ka (13 and 35 GHz)
-- swath width 120, 245 km

From TRMM to GPM

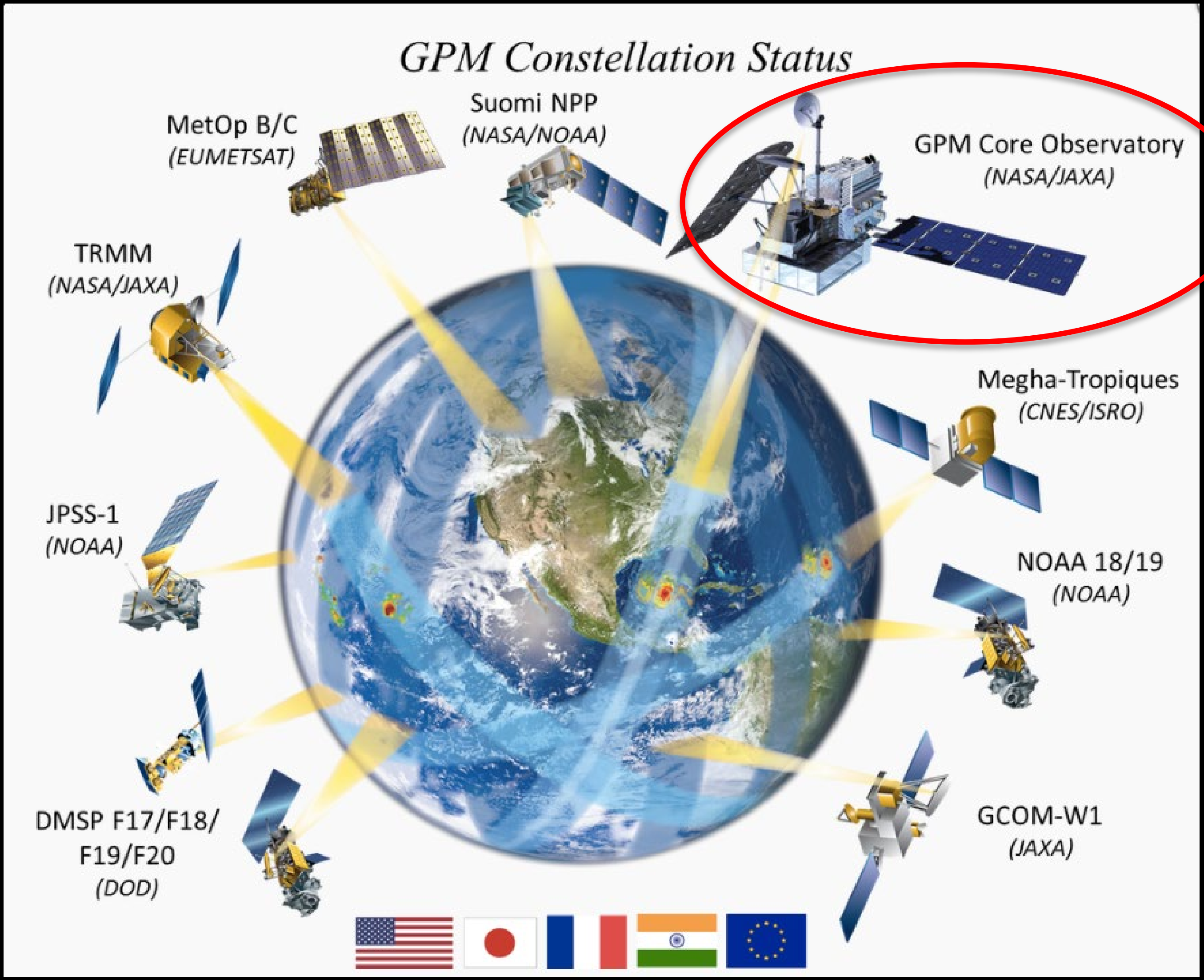
Covering 35S to 35N
Microwave Imager (TMI)
-- 9 channels
-- frequencies 10.7-to-85.5 GHz
-- swath width 878 km
Precipitation radar (PR)
-- single-frequency Ku band (13 GHz)
-- swath width 247 km

Covering 68S to 68N
GPM Microwave Imager (GMI)
-- 13 dual-polarized channels
-- frequencies 10.65-183.3 GHz
-- swath width 885 km
Dual Polarization radar (DPR)
-- dual-frequency Ku & Ka (13 and 35 GHz)
-- swath width 120, 245 km

GPM Core Satellite



GPM Constellation



Passive Microwave Retrieval: An underdetermined Inverse problem

The direct problem:

$$TB = f(R_C, E_S)$$

Surface emissivity

Hydrometeors profile

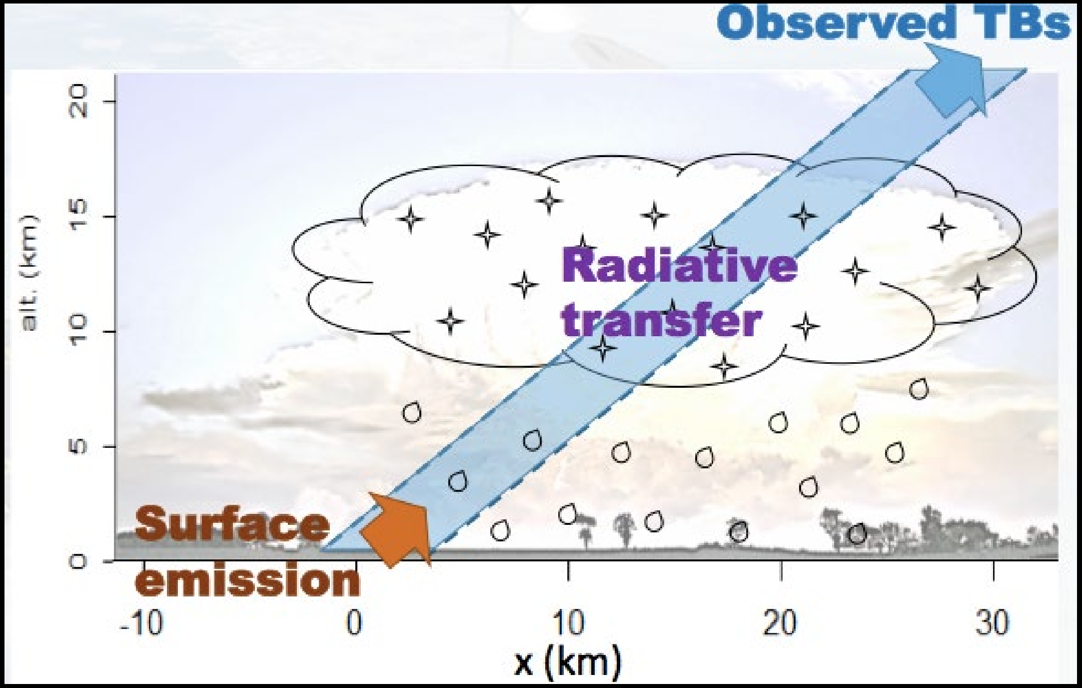
Accurately resolved by numerical (physical) radiative transfer models

The inverse problem:

$$R_C = g(TB, E_S)$$

?

Underdetermined



Passive Microwave Retrieval: an Inverse Problem

Learn patterns from data for retrieval

Spectral BT

13-dim space
(each point is a
BT vector)
Manifold of BT

"Manifold learning"

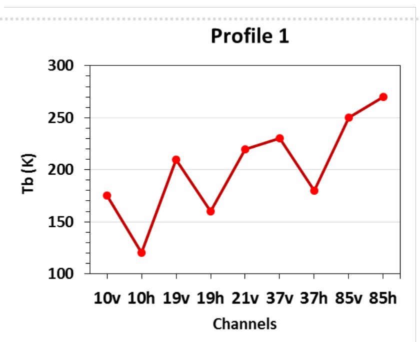
Rainfall Profiles

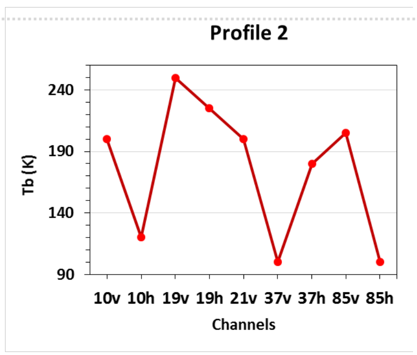
n-dim space
(each point is a Z,
surf R vector)
Manifold of R

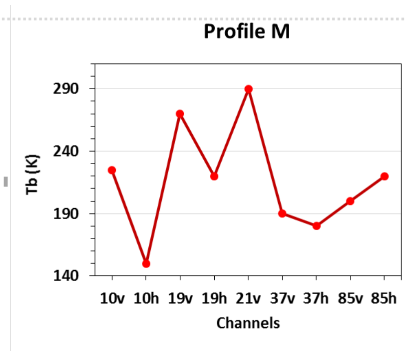
Machine Learning and Regularized Estimation in High Dimensional Spaces

Database

Spectral BT

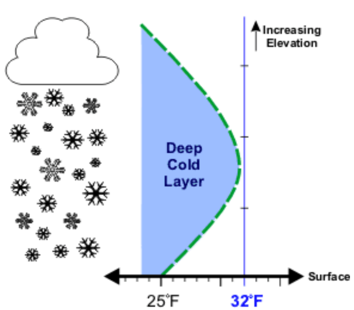




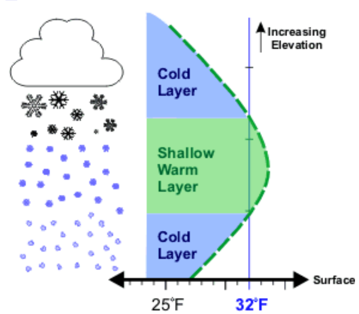


Rainfall Profiles

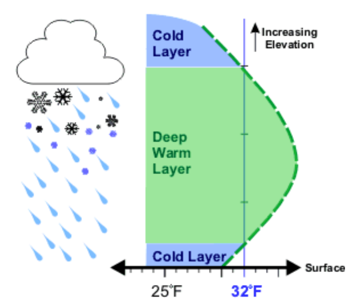
$\mathbf{x}_1$



$\mathbf{x}_2$

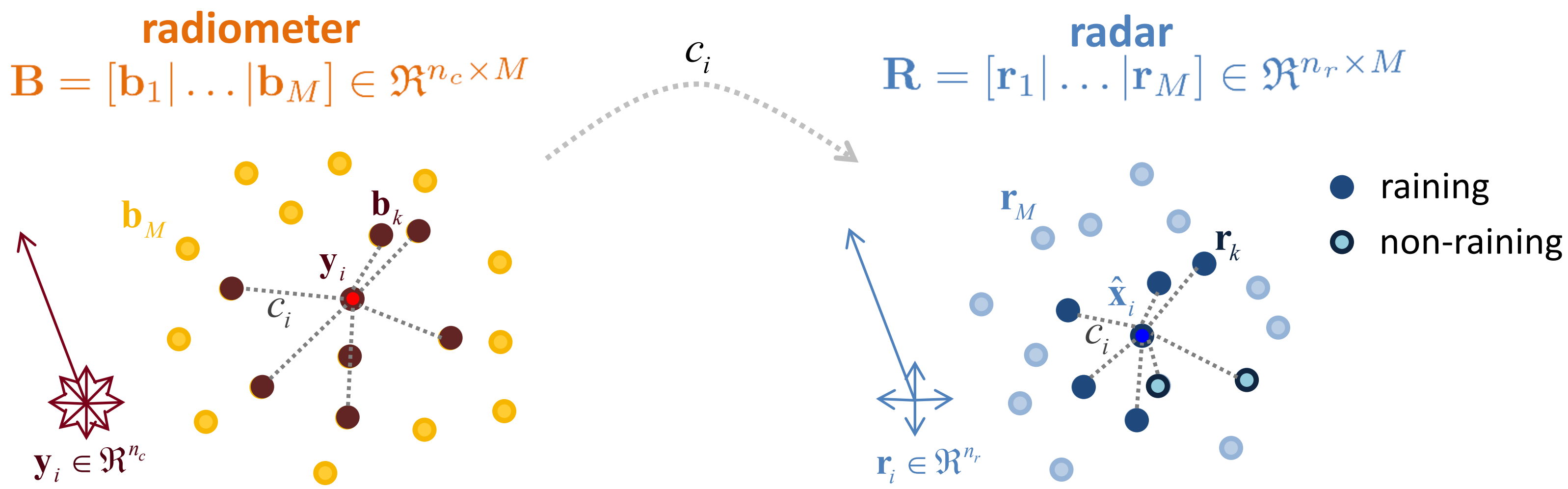


$\mathbf{x}_M$



ShARP: Locally linear embedding for rainfall retrieval

- **Inversion Algorithm based on Regularization:**
  - Concept of the locally linear embedding (supervised manifold learning):



  - Search for the **K-nearest neighbors** to detect raining signatures

$$\mathbf{B}_{\mathcal{S}} = [\mathbf{b}_1 | \dots | \mathbf{b}_K] \in \Re^{n_c \times K} \qquad \mathbf{R}_{\mathcal{S}} = [\mathbf{r}_1 | \dots | \mathbf{r}_K] \in \Re^{n_r \times K}$$

  - Estimate the **representation coefficients** and thus the rainfall profile

$$\mathbf{y}_i = \Sigma_{k=1}^{K} c_k \mathbf{b}_k + \mathbf{v}_k \longrightarrow \hat{\mathbf{x}}_i = \Sigma_{k=1}^{K} c_k \mathbf{r}_k$$

Foufoula-Georgiou et al., *Survey in Geophysics*, 2015; Ebtehaj, Foufoula-Georgiou, Lerman, Bras, *GRL*, 2015
Ebtehaj, Bras, Foufoula-Georgiou, *IEES*, 2015; & *J. Hydrometeorology*, 2016; Takbiri, Ebtehaj, Foufoula-Georgiou, *HESS*, 2017

ShARP: Locally linear embedding for rainfall retrieval

– **Detection step:**

- K-nearest neighborhood search + a probabilistic voting rule for rain/no-rain

– **Estimation Step:**

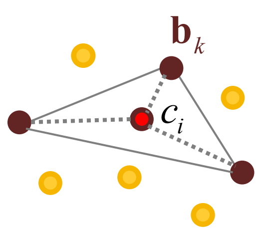
- Estimation of the representation coefficients

$$\underset{\mathbf{c}}{\text{minimize}} \quad \left\| \mathbf{W}^{1/2} \left(\mathbf{y} - \mathbf{B}_{\mathcal{S}} \mathbf{c} \right) \right\|_2^2 + \lambda_1 \left\| \mathbf{c} \right\|_1 + \lambda_2 \left\| \mathbf{c} \right\|_2^2$$

$$\text{subject to} \quad \mathbf{c} \succeq 0, \;\; \mathbf{1}^{\mathrm{T}} \mathbf{c} = 1,$$

ℓ_p -norm: $\|\mathbf{c}\|_p^p = \Sigma_i |c_i|^p$

$\lambda_1, \lambda_2 > 0$



$$\mathbf{B}_{\mathcal{S}} = [\mathbf{b}_1 | \ldots \boxed{|\mathbf{b}_{i-1}|\mathbf{b}_i|} \ldots \boxed{|\mathbf{b}_{j-1}|\mathbf{b}_j|} \ldots | \mathbf{b}_K] \in \Re^{n_c \times K}$$

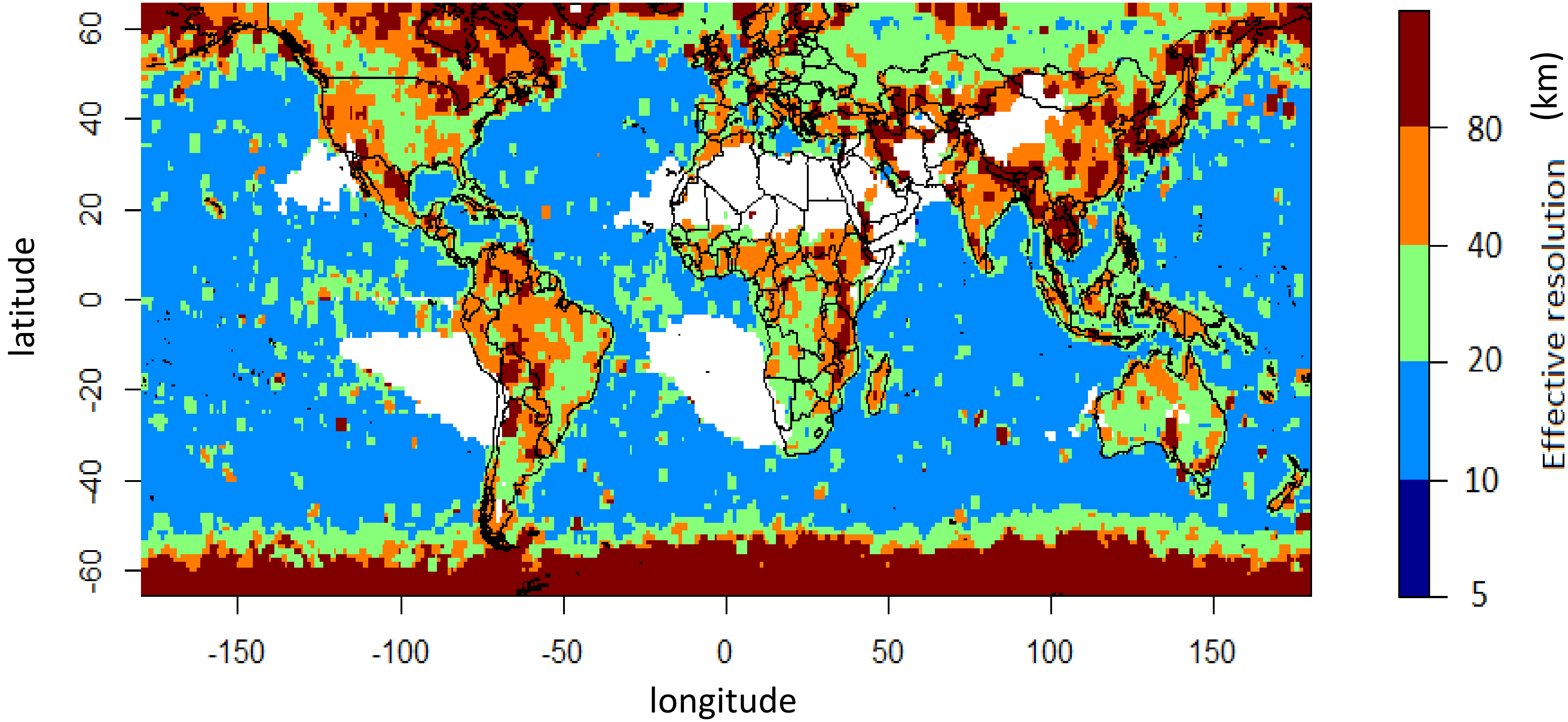
L1-L2 regularization for stability and reduced estimation error

- Rainfall estimates

$$\hat{\mathbf{x}} = \mathbf{R}_{\mathcal{S}} \hat{\mathbf{c}}$$

Yet, lacking performance in several places of the world

Effective Resolution (ER) of NASA's GPROF v7 (GMI vs KuPR)

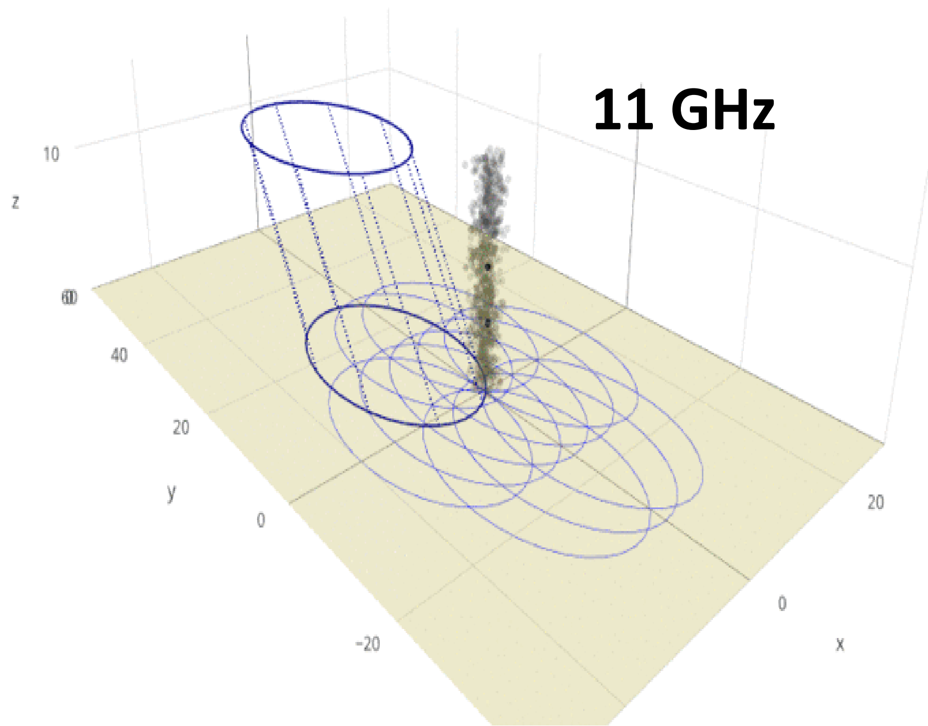


- Local values computed from all observations in 3°×3° boxes.
- March 2014 to February 2017: 16,500 GPM orbits

Guilloteau, Foufoula-Georgiou, Kummerow, *J. Hydrometeorology*, 2017.

New Direction: Retrieve patterns (not a pixel at a time)

Learn from the spatial structure in the TB space



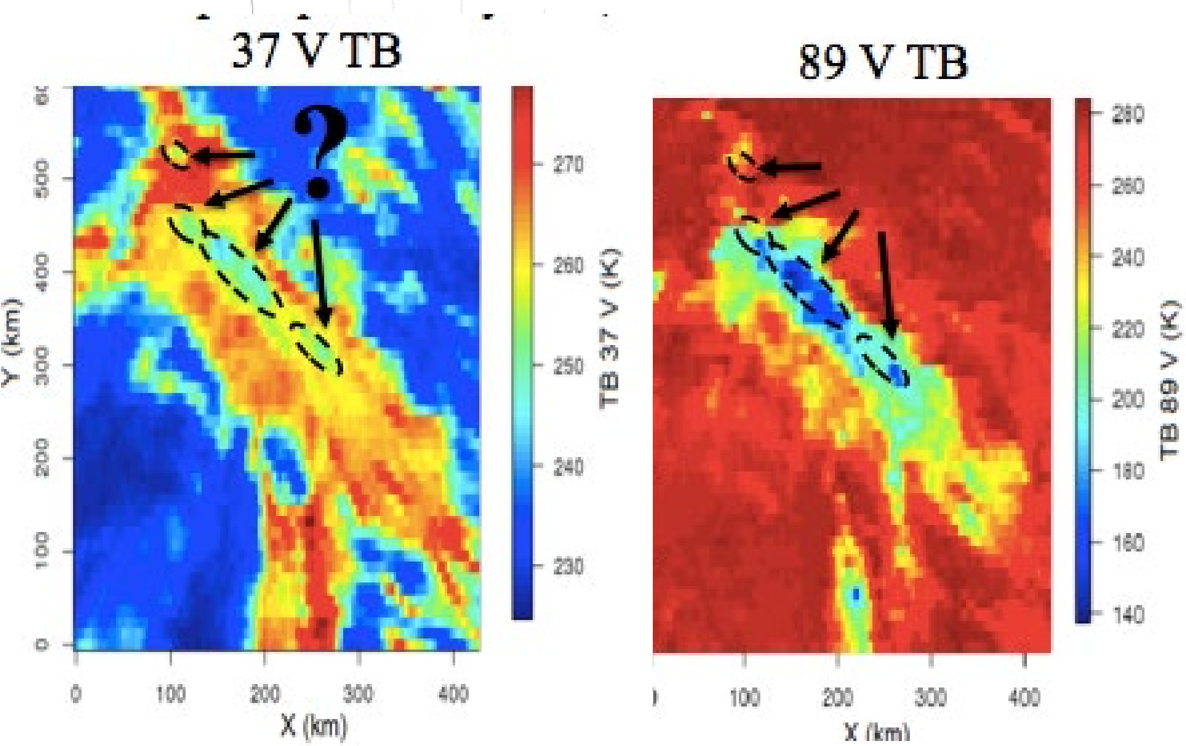
1) Sensor Geometry: with different channels responding to different altitude levels, the multi-spectral signature characterizing a given vertical column may be split across several pixels.

2) Specific spatial patterns of TBs are the signatures of specific atmospheric features

QU: HOW TO IDENTIFY THE NEIGHBOROOD AND HOW TO LEARN FROM IT?
It becomes a very high dimensional problem! Need to learn features!

New Direction: Retrieve patterns not a pixel at a time

Learn from the spatial structure in the TB space



Lower 37V =>
Lower emission signal =>
Lower precipitation?

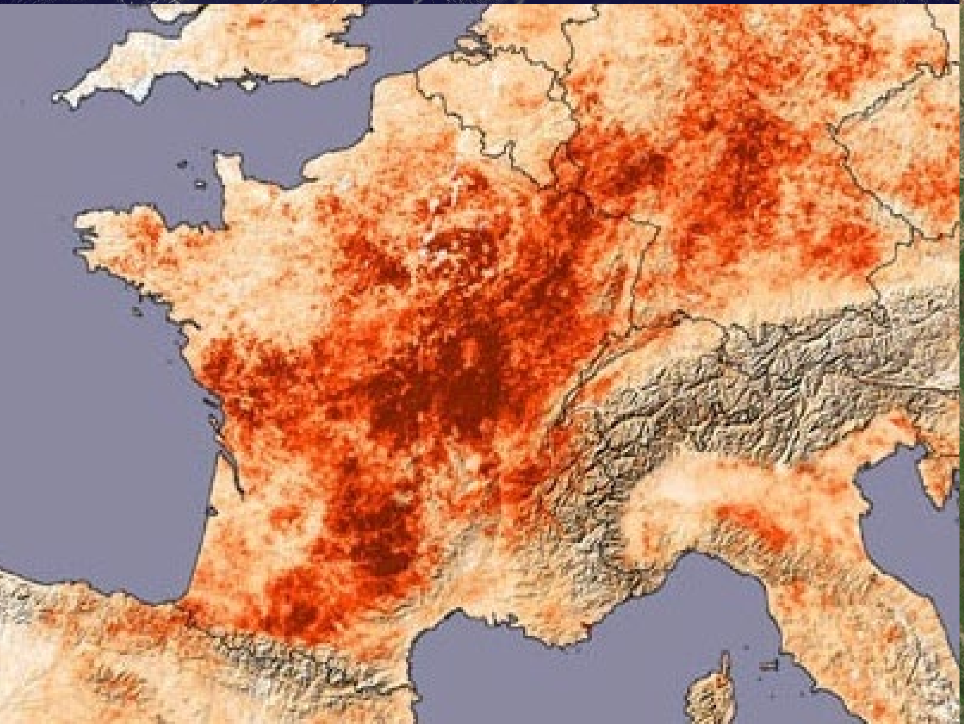
NO! It is an ice
Scattering signal=>
Very active convective
Cell

1) Sensor Geometry: with different channels responding to different altitude levels, the multi-spectral signature characterizing a given vertical column may be split across several pixels.

2) Specific spatial patterns of TBs are the signatures of specific atmospheric features

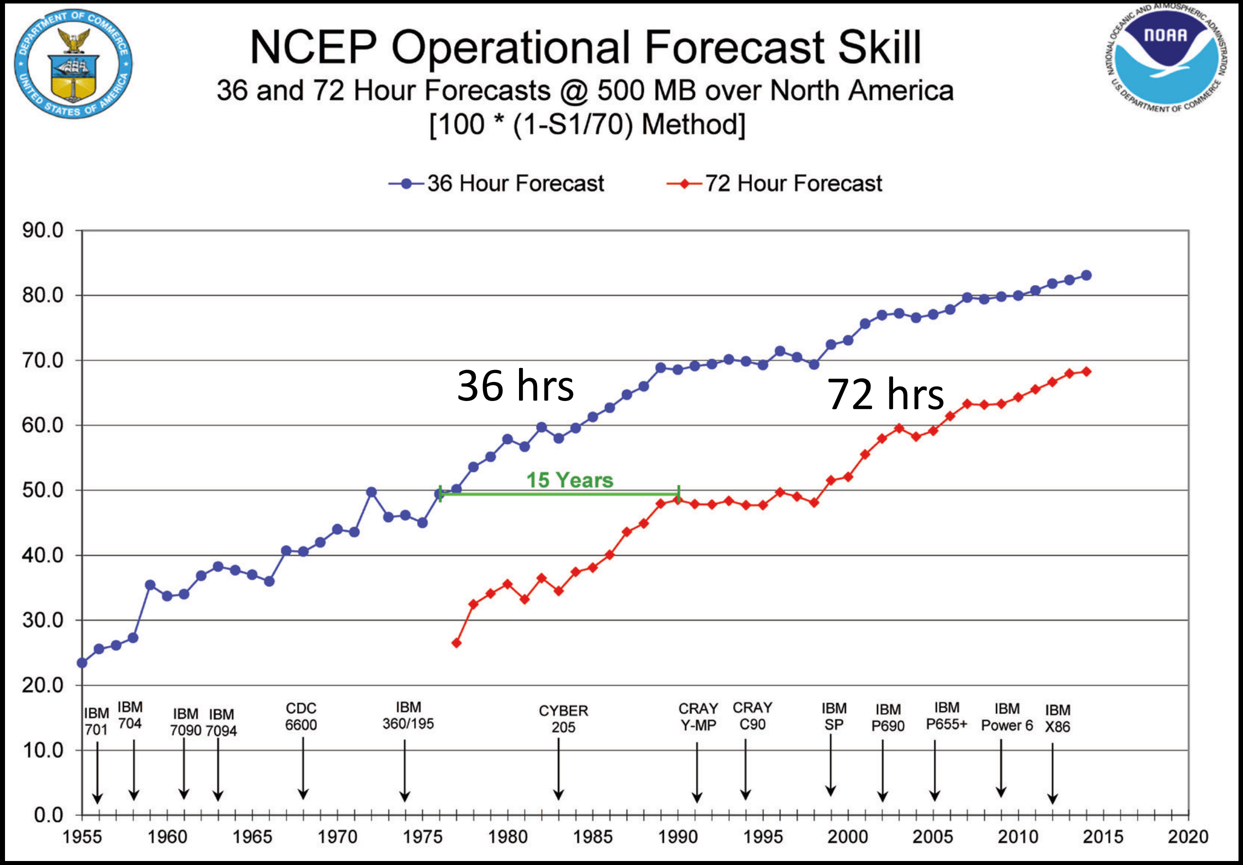
2. SEASONAL PREDICTION

How to best combine climate observations and models to improve predictions?



Bridging the gap: weather and climate

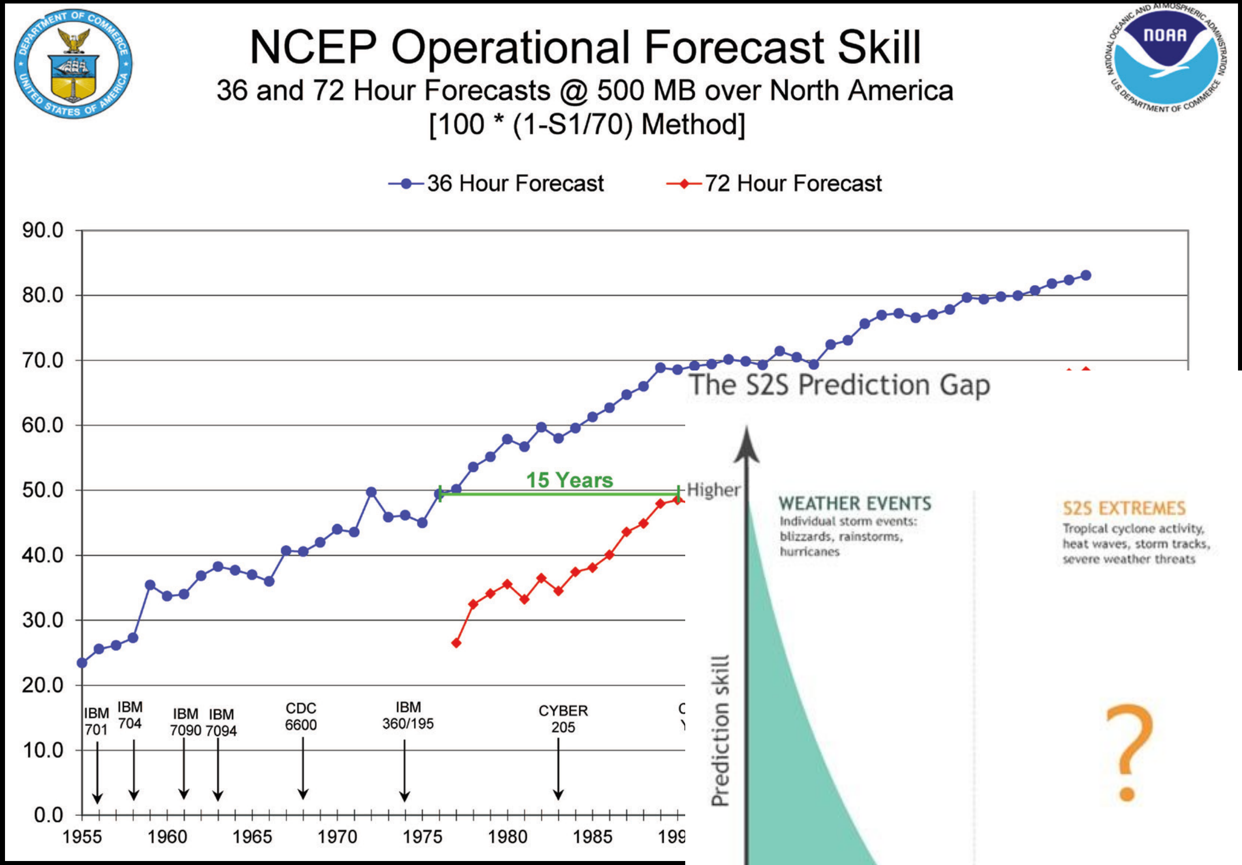
NAS (2016) *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts*



36-hr and 72-hr ahead weather forecasts are getting better and better…

Bridging the gap: weather and climate

NAS (2016) *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts*



36-hr and 72-hr ahead weather forecasts are getting better and better…

Mariotti et al (2018) Climate and Atmospheric Science, doi:10.1038/s41612-018-0014-z

Our prediction skill on subseasonal (week timescales) to seasonal (S2S) timescales is still very limited

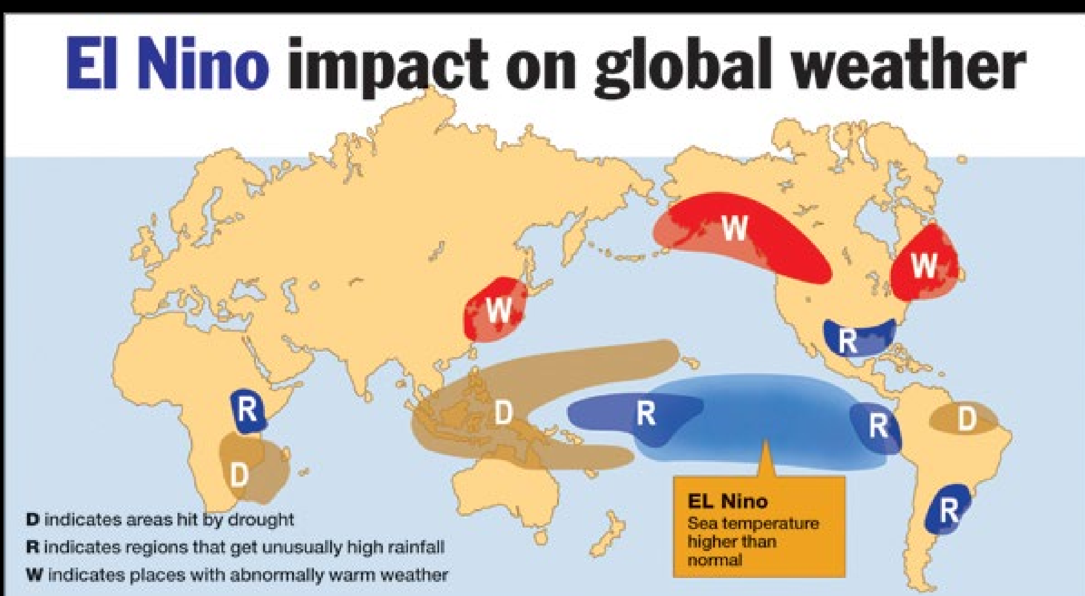
Combining physics and statistics

- Climate models show limited skill in predicting seasonal precipitation months ahead
- Best approach to predict is combining our physical understanding with statistical tools:

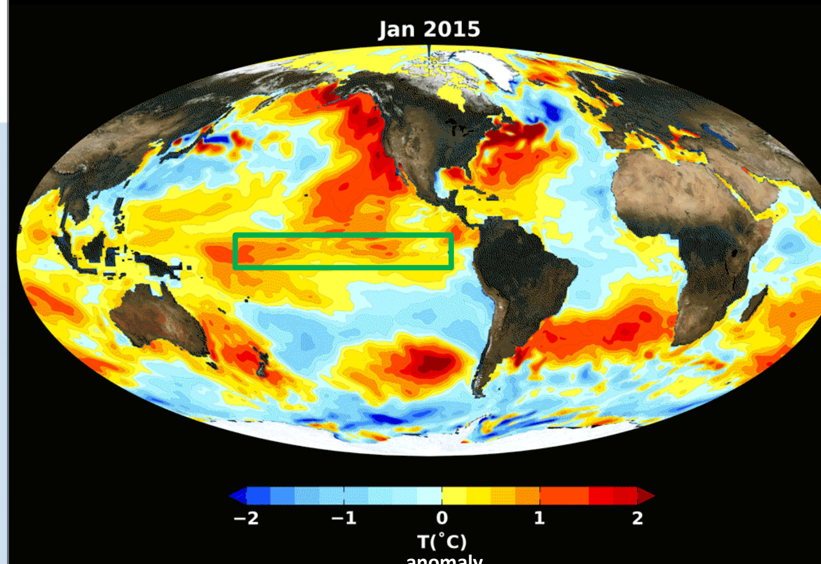
Regional hydroclimate (e.g. precipitation, temperature in California)

$$Y = f(\boldsymbol{X}) + e$$

Large-scale climate modes (e.g. ENSO see below)

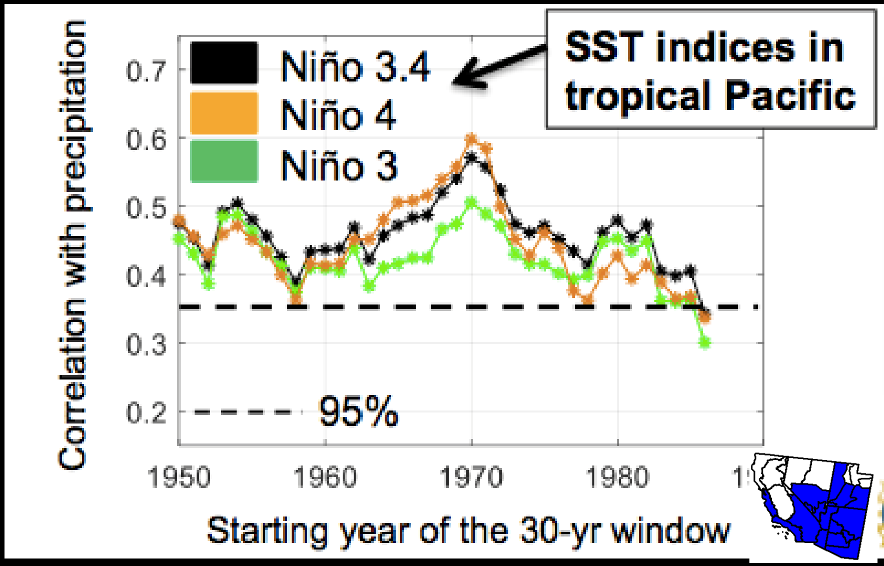


The Great El Niño in 2016



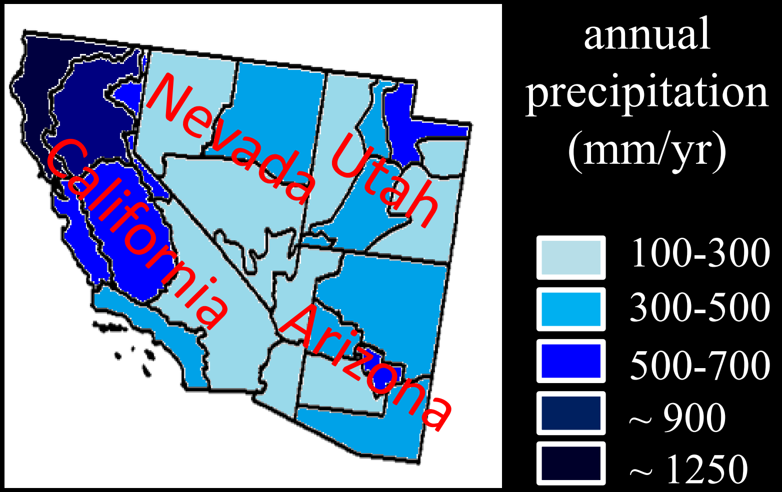
Learning from Big Data

- What are the best sources of predictability for a specific region? Are they changing?
- ML and Network analysis can extract much relevant information from the data to improve prediction

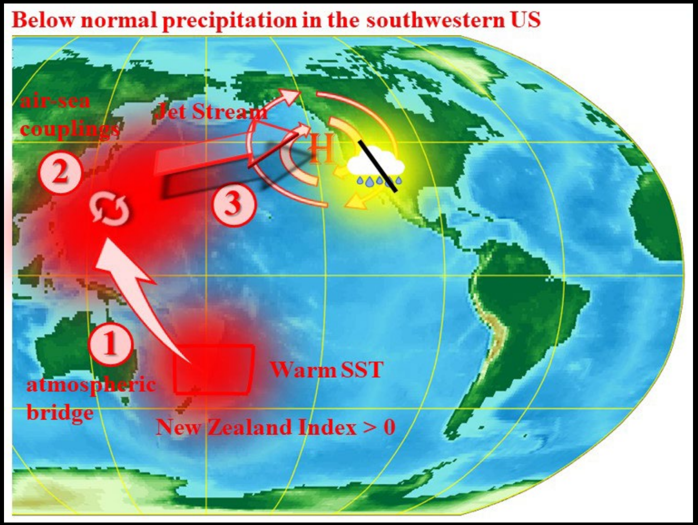


<u>Example: Precipitation in southwestern US</u>

- Dry and variable hydroclimate



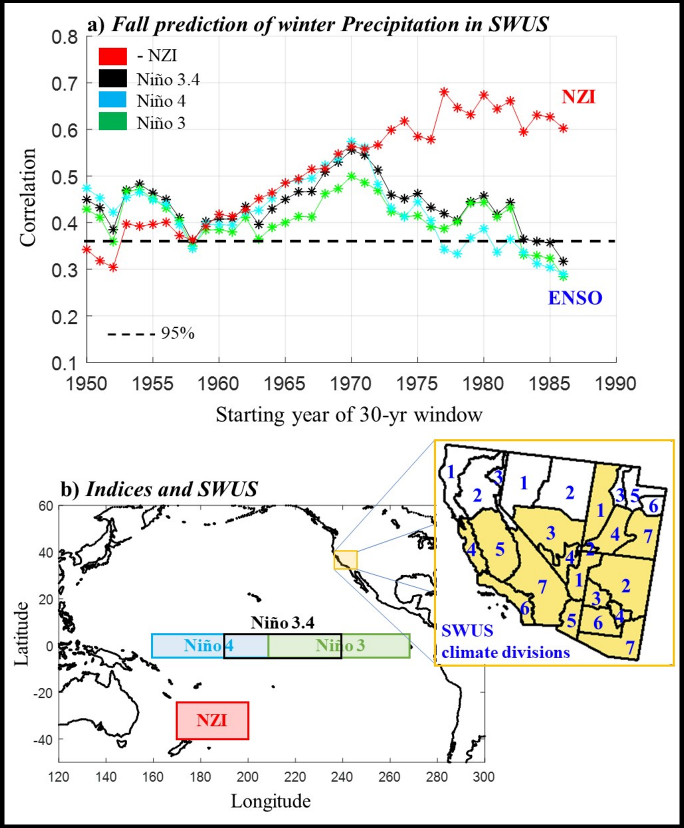
- New climate mode discovered, different than ENSO



(Mamalakis et al., 2018, *Nature Communications*)

Learning from Big Data

(Mamalakis et al., 2018, *Nature Communications*)



- This new mechanism has been more dominant in modulating SWUS precipitation in the last 3 to 4 decades:
  - ✓ Vector X may not include important/new modes if based only on our prior knowledge
  - ✓ Function f is not constant through time
- In our new project (funded by NSF), we use machine learning to address this problem:

$$\mathrm{E}[y_t] = \langle x_t, \beta \rangle$$

linear model

- Where the relative contribution of each feature in X is represented by β and is calculated by minimizing:

$$\hat{\beta} = arg\min_{\beta} \sum_{t=1}^{N} (y_t - \langle x_t, \beta \rangle)^2 + \lambda_{TV} \sum_{j,k=1}^{p} \Sigma_{j,k}^{1/2} |\beta_j - \beta_k| + \lambda_1 ||\beta||_1$$

Fit observations *Spatial dependence* *Sparsity*

TRIPODS+CLIMATE program (NSF grant DMS-1839336)

Learning from Big Data

$$\mathrm{E}[y_t] = \langle x_t, \beta \rangle$$

linear model

Spatial dependence

$$\hat{\beta} = arg\,\min_{\beta} \sum_{t=1}^{N} (y_t - \langle x_t, \beta \rangle)^2 + \lambda_{TV} \sum_{j,k=1}^{p} \Sigma_{j,k}^{1/2} |\beta_j - \beta_k| + \lambda_1 ||\beta||_1$$

Sparsity

Fit observations

- ✓ Preliminary results are promising. We can explain more that 40% of precipitation variability in the out-of-sample period.
- ✓ The patterns of β can be used to verify and reveal new mechanisms in the large-scale climate system
- ✓ Such approaches can also be used for climate model diagnostics. Do climate models capture the observed interrelations and how are these projected to change under climate change?

TRIPODS+CLIMATE program (NSF grant DMS-1839336)

3. LANDSCAPES



What can they tell us about process and how are they changing?



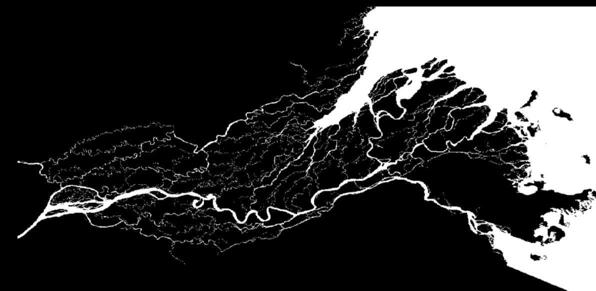


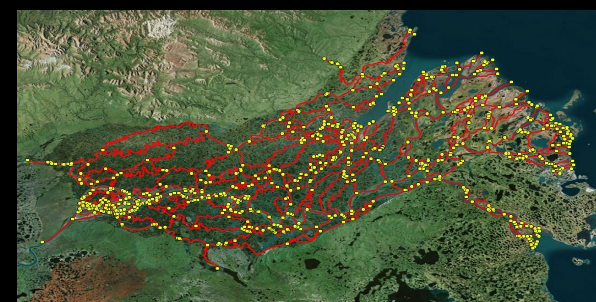

Remotely-sensed global imagery is paving the way for a Global Geomorphology

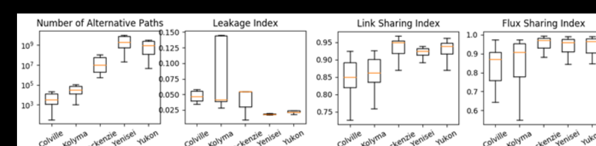
The two most critical problems:

Automatic extraction of dynamic objects

Automatic extraction of critical information

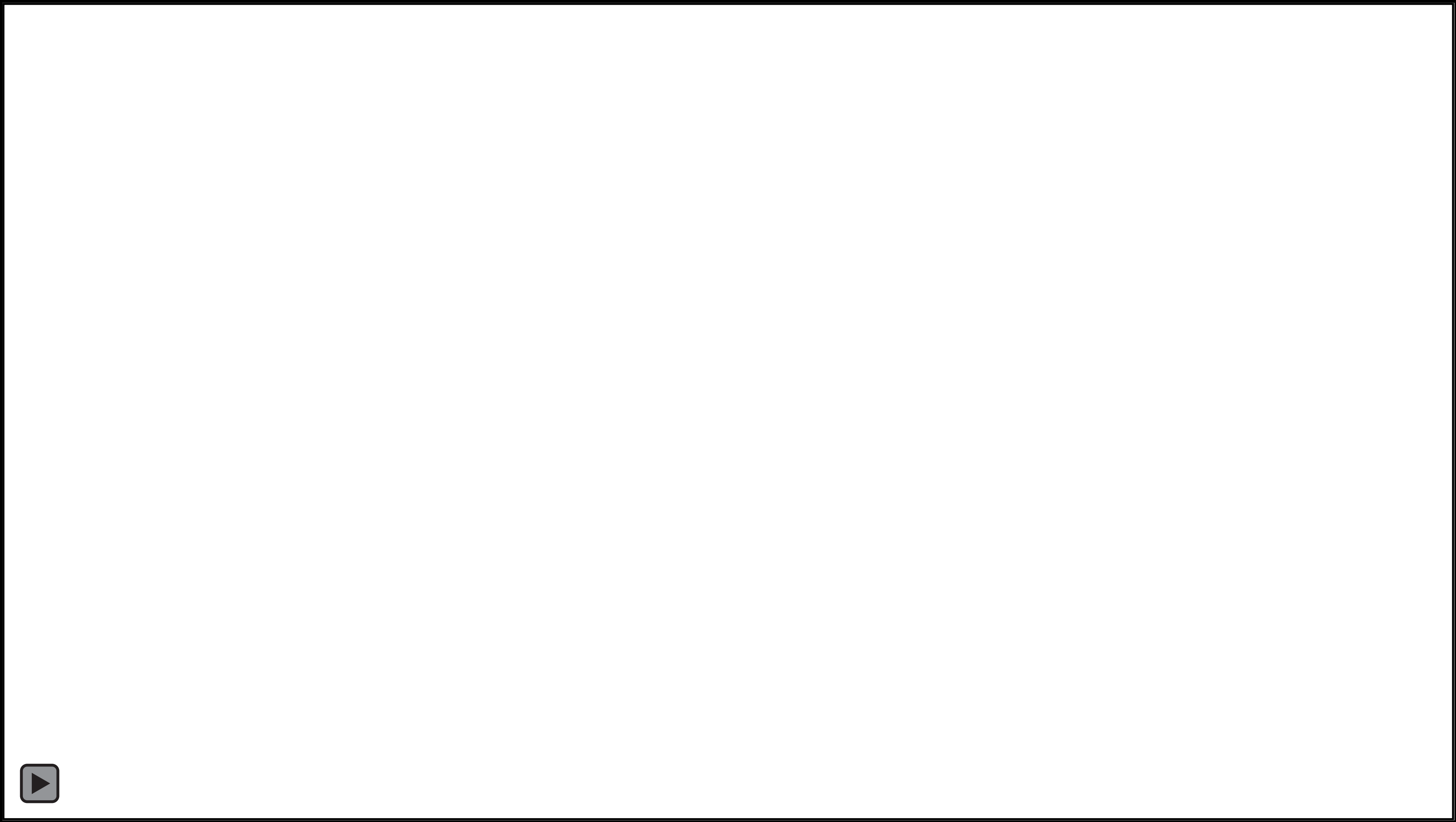






Robust extraction of rivers from multispectral imagery is a very nuanced problem that must consider: water level at time of image, exposed point bars, mixed pixels at boundaries, and clouds, shadows, snow cover, etc.

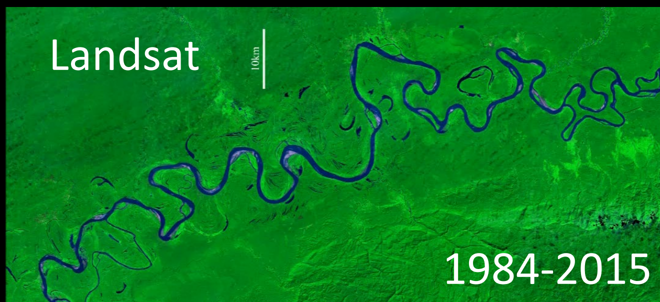
Following object identification (mask generation), robust algorithms must be capable of objectively distilling relevant metrics and insights without excessive manual intervention.

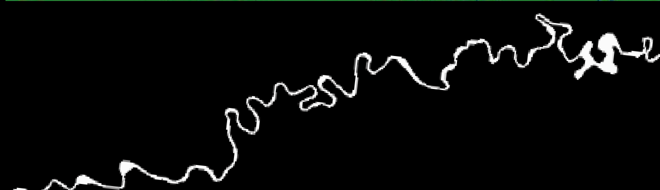


Ucayali River, Peru

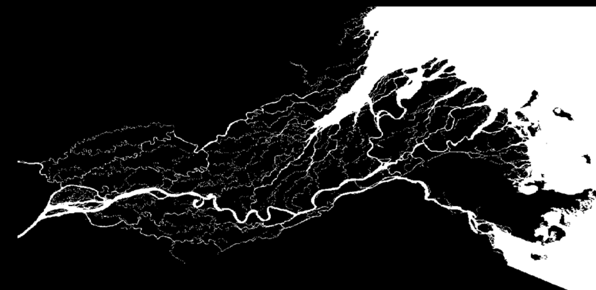
Two critical problems

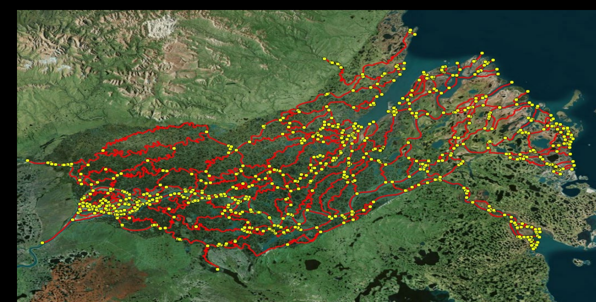
Automatic extraction of dynamic objects

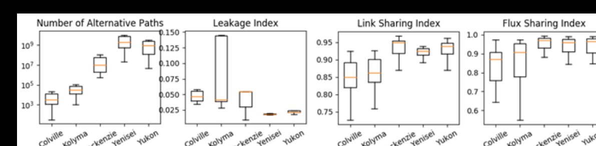




Automatic extraction of critical information



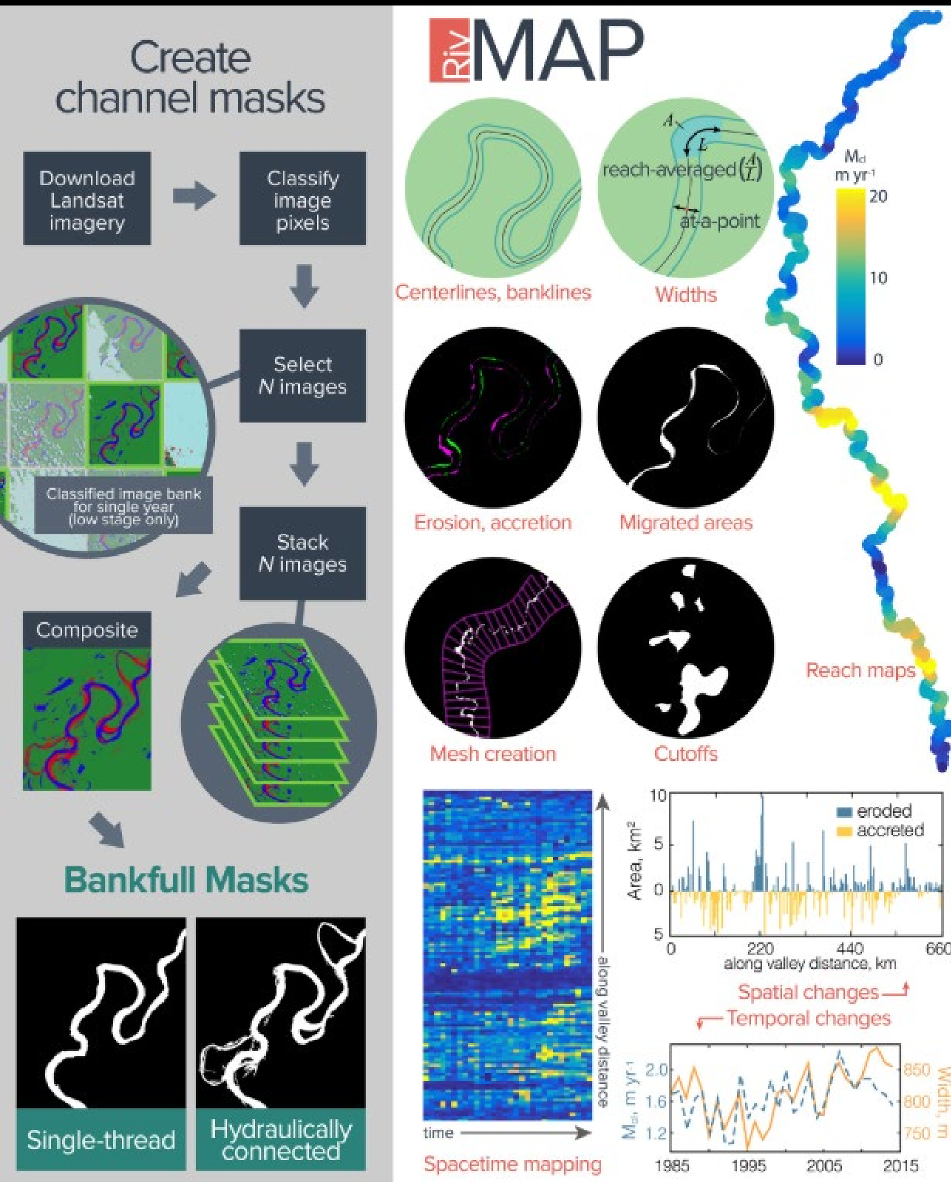


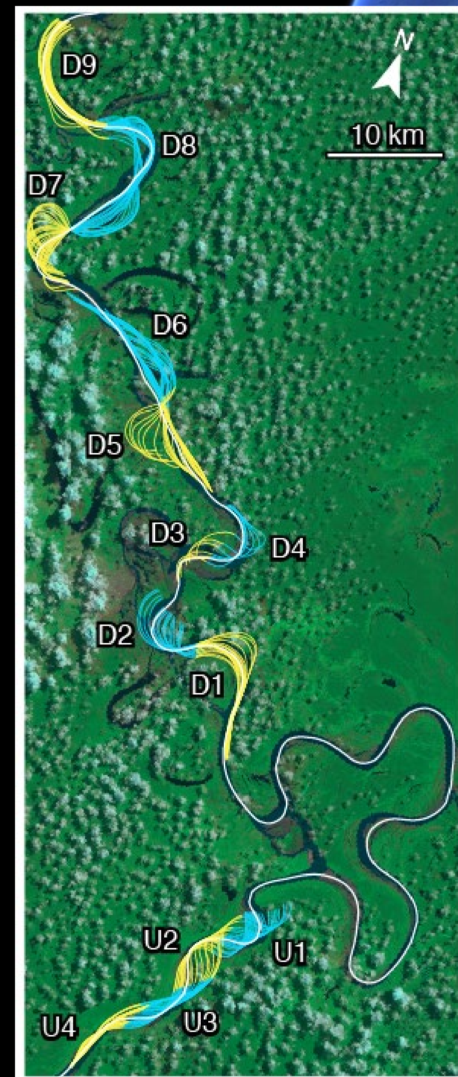


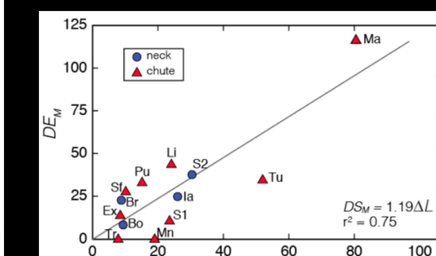
Robust extraction of rivers from multispectral must consider: water level at time of image, exposed point bars, mixed pixels at boundaries, and clouds, shadows, snow cover, etc.

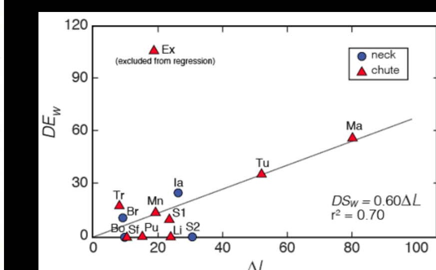
Following object identification (mask generation), robust algorithms must be capable of objectively distilling relevant metrics and insights without excessive manual intervention.

Tools for large-scale mask analysis (single rivers)









Can predict cutoffs!

J Schwenk, A Khandelwal, M Fratkin, V Kumar, E Foufoula-Georgiou. (2017) *Earth and Space Science*
J Schwenk, E Foufoula-Georgiou. (2017) *Geophysical Research Letters*

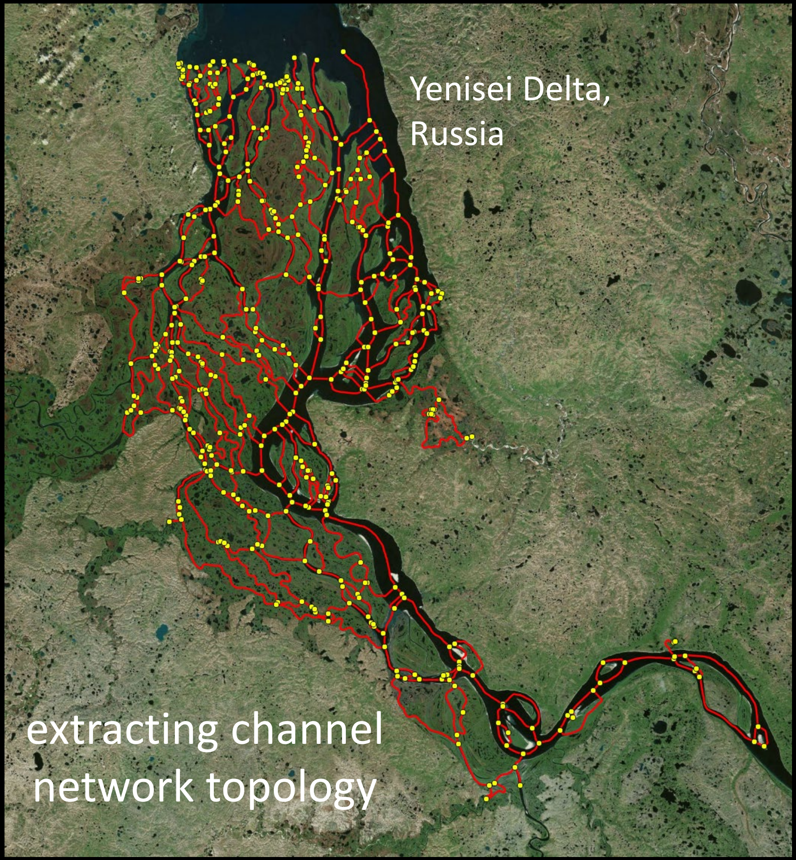
Tools for large-scale mask analysis (river networks)

RivGraph:

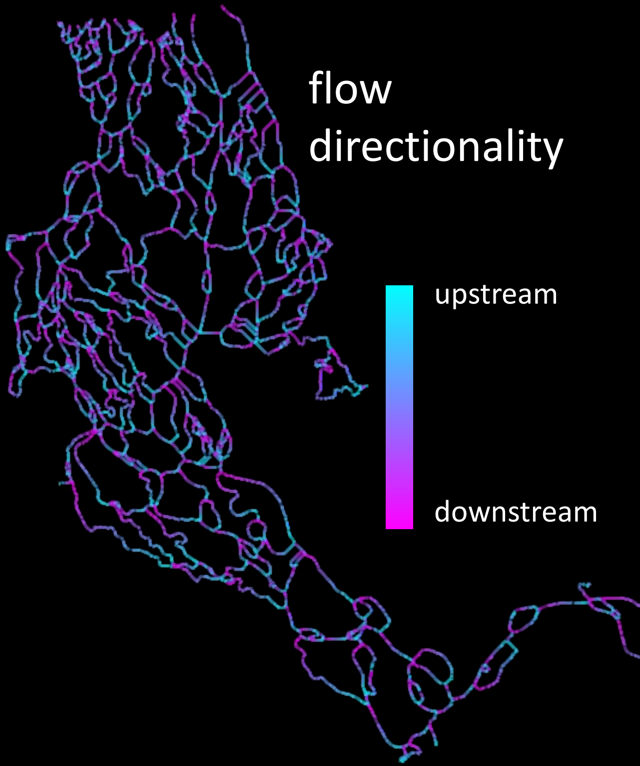
a Python toolbox for analysis of deltaic and braided river channel networks

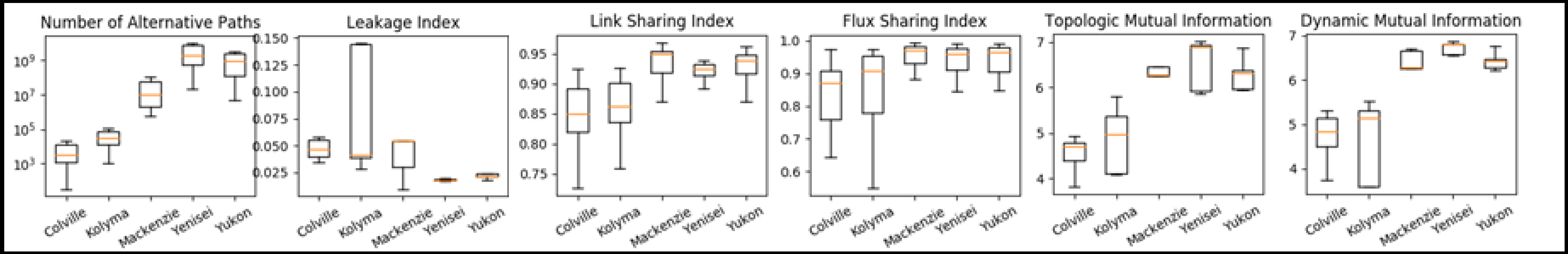
Coming soon to a Conda Repository near you…

J Schwenk, A Tejedor, A Piliourais, J Rowland, E Foufoula-Georgiou. (2018) *In preparation.*



flow directionality





morphological properties and topologic metrics

Arctic Deltas (ADs)

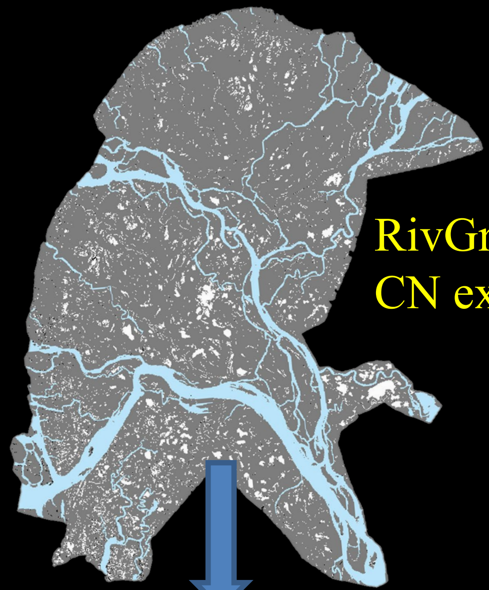
The Arctic: North of 66° 33’N

Mackenzie Delta, Source: Sam B Cornish

- Climate change affects poles with greater intensity i.e. Polar Amplification (Serreze et al. 2009)
- ADs have on the order of 91 39 Pg-Carbon (Schuur et al., 2015)
- Lakes and ponds are significant sources of methane (i.e. further warming) (Wik, 2016)
- ADs are uniquely characterized by strong spring flooding, permafrost presence, and lake abundance (Walker, 1999)

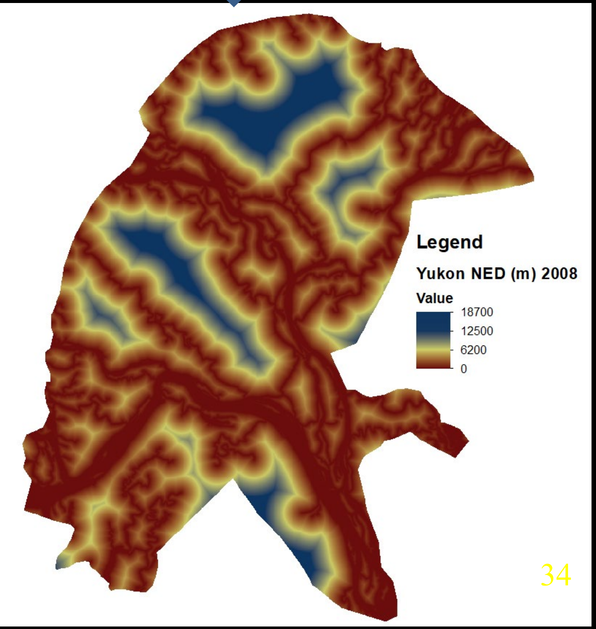
Changes in arctic deltas under climate change

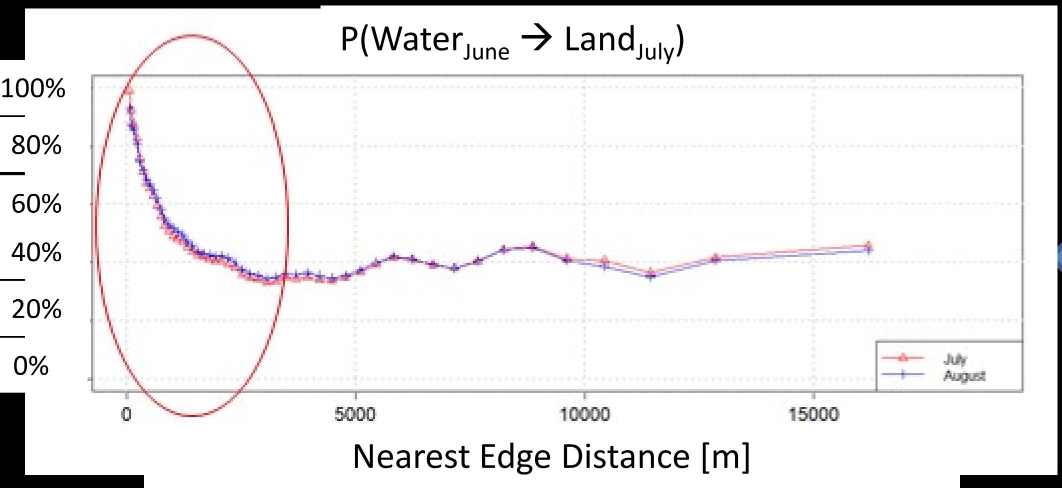
Can we infer subsurface hydrologic connectivity from the observed (surface) topology and connectivity of lakes and channels in ADs?



RivGraph for CN extraction

Approach:
We interrogate lake shrinkage rates and show that distance from the delta channel network controls lake shrinkage and thus subsurface connectivity.





INTELLIGENCE 2025
to improve life on Earth

FINDING THE SIGNAL IN THE NOISE

Workshop on
Data Analytics for Climate and Earth (DANCE):
Causality, patterns and prediction

March 27-29, 2019
Arrowhead, CA (USA)

efi@uci.edu