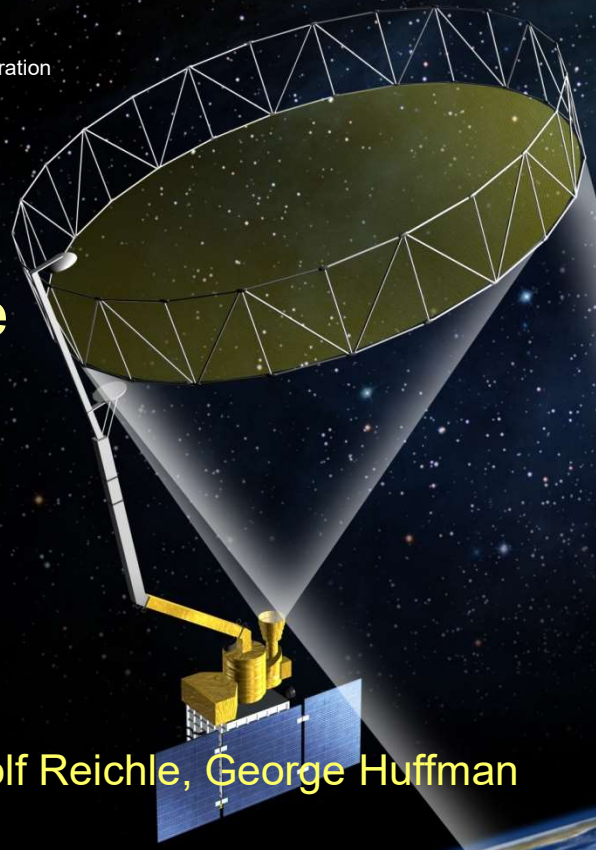
National Aeronautics and Space Administration



Soil Moisture Active Passive Mission

SMAP

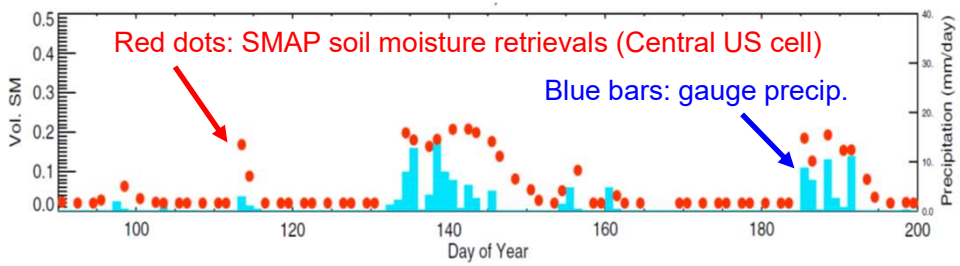
Application of SMAP-Based Precipitation Estimates to Hydrological Modeling

Randal Koster (PI), Rolf Reichle, George Huffman
(NASA/GSFC)

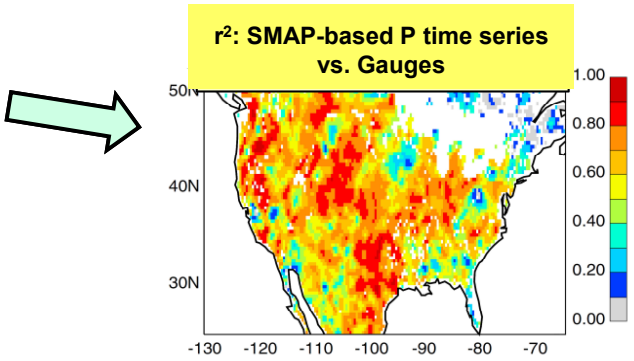
Update: June 16, 2021

Quick Refresher: Project Goal

As you would expect, SMAP soil moisture retrievals and independent precipitation measurements show significant consistency.



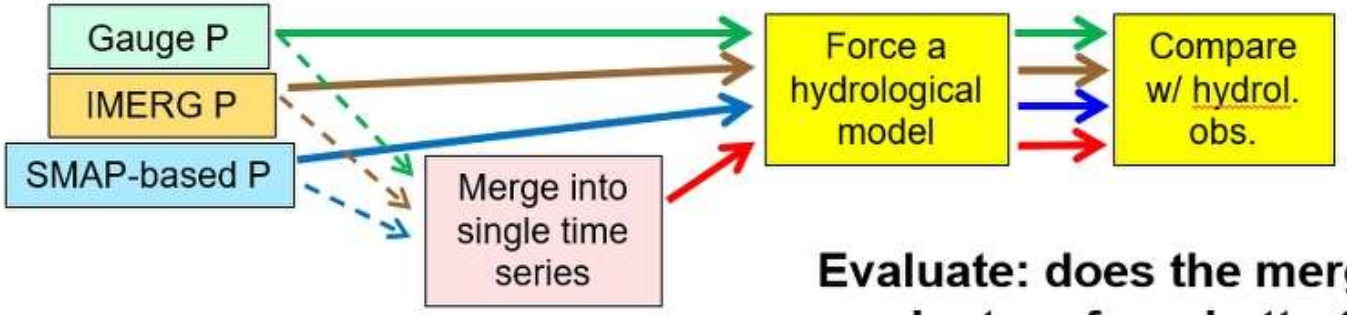
Approaches have been developed (e.g., Brocca et al. 2013) to translate soil moisture variations into estimates of precipitation time series. Over the last few years, we have applied these approaches extensively to SMAP Level-2 retrievals. The resulting SMAP-based precipitation time series appear realistic.



Project's hypothesis: *By merging different (and independent) precipitation measurement time series, we can produce a superior time series for global hydrological simulation. (Builds on idea described by Dong et al. 2020)*

Precip. data source	Advantages	Disadvantages
Rain gauges (CPCU)	High accuracy	Incomplete coverage, spatial repres. error
Satellite retrievals (GPM/IMERG)	High temporal, spatial coverage	Retrieval error, temporal repres. error
SMAP-based estimates	High (though coarse) spatial coverage, time-integral of P effects	Low temporal coverage misses high P events

Approach:



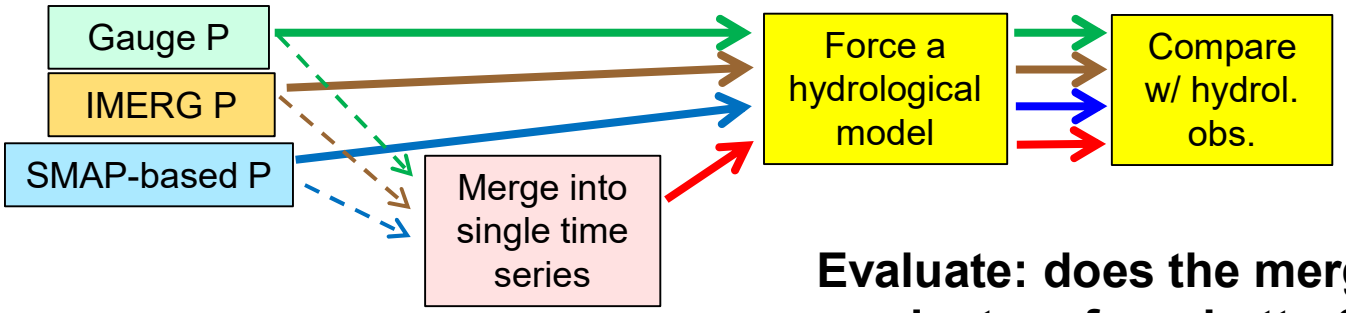
Evaluate: does the merged product perform better?

Can SMAP data contribute to an improved precipitation dataset?

Project's hypothesis: *By merging different (and independent) precipitation measurement time series, we can produce a superior time series for global hydrological simulation. (Builds on ideas described by Dong et al. 2020)*

Precip. data source	Advantages	Disadvantages
Rain gauges (CPCU)	High accuracy	Incomplete coverage, spatial repres. error
Satellite retrievals (GPM/IMERG)	High temporal, spatial coverage	Retrieval error, temporal repres. error
SMAP-based estimates	High (though coarse) spatial coverage, time-integral of P effects	Low temporal coverage, misses high P events

Approach:



Evaluate: does the merged product perform better?

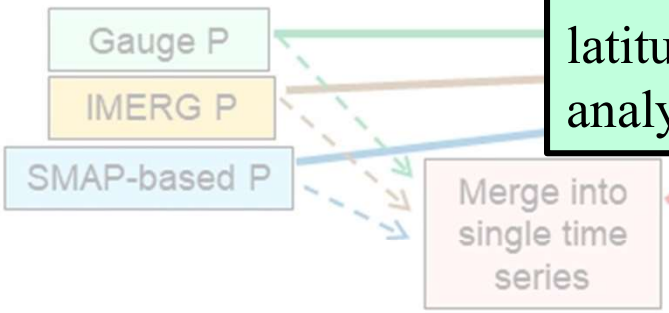
Can SMAP data contribute to an improved precipitation dataset?

Project's hypothesis: *By merging different (and independent) precipitation measurement time series, we can produce a superior time series for global hydrological simulation. (Builds on ideas described by Dong et al. 2020)*

Precip. data source	Advantages	Disadvantages
Rain gauges (CPCU)	High accuracy	Incomplete coverage, spatial repres. error
Satellite retrievals (GPM/IMERG)	mporal, spatial coverage	Retrieval error, temporal
SMAP-based estimates	High (thou coverage,	

Quick note regarding Africa and high latitudes: I'll be referring to "gauge data" in this presentation, though in fact the gauge data in Africa and high latitudes are replaced by model analysis data.

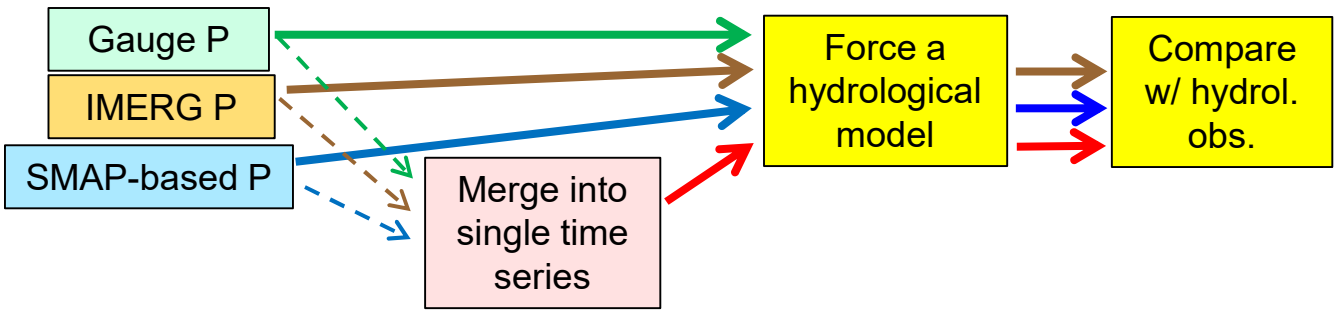
Approach:



Evaluate: does the merged product perform better?

Can SMAP data contribute to an improved precipitation dataset?

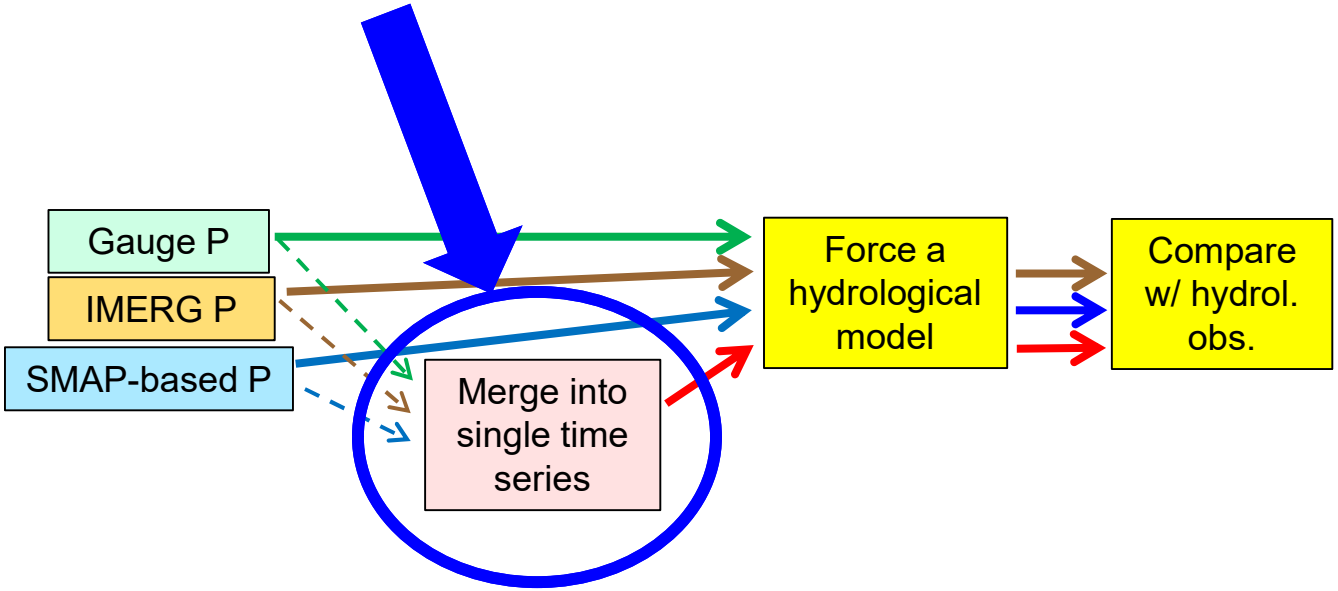
Research reported last October: Some preliminary results obtained with our first (preliminary!) version of merged precipitation data.

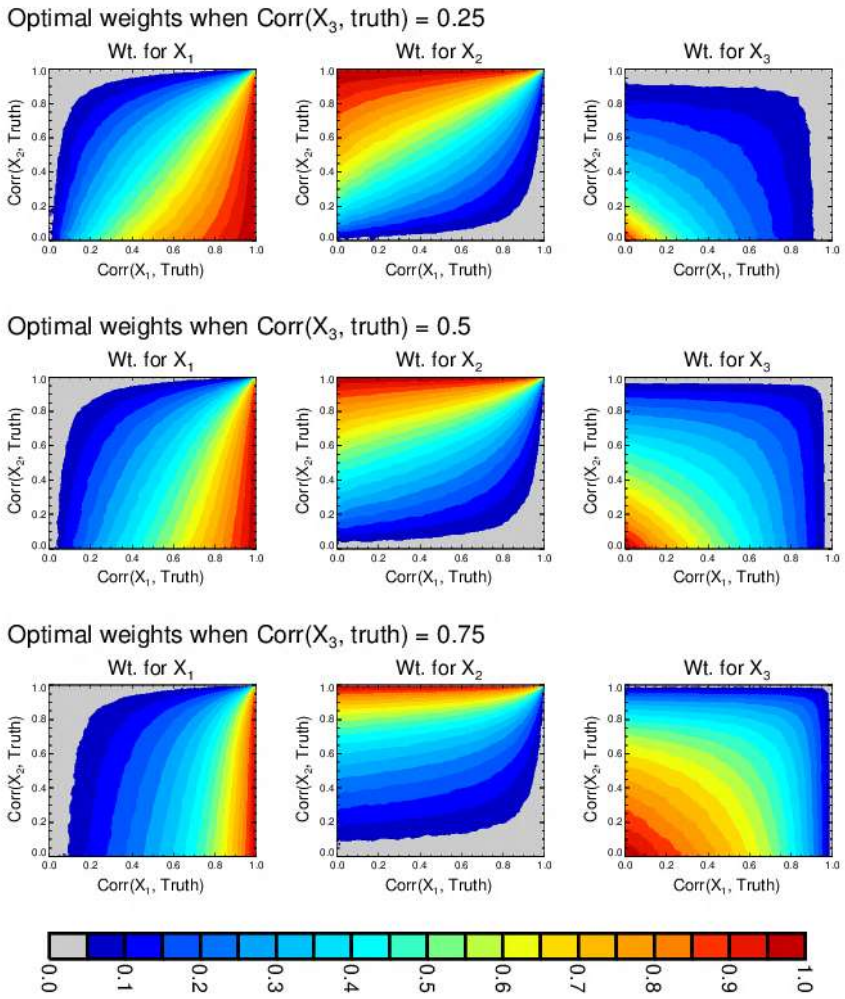


Research reported last October: Some preliminary results obtained with our first (preliminary!) version of merged precipitation data.

Since last presentation:

1. Optimized merging procedure
(devised new approach)





New merging procedure:

- Apply triple collocation procedure to logarithms of pentad data
- Given "correlations vs. unknown truth", determine optimal weights for use in combining datasets – weights that maximize $Corr(P_{merged}, P_{truth})$.

Contours showing optimized weights to apply to three independent datasets as a function of underlying accuracy.

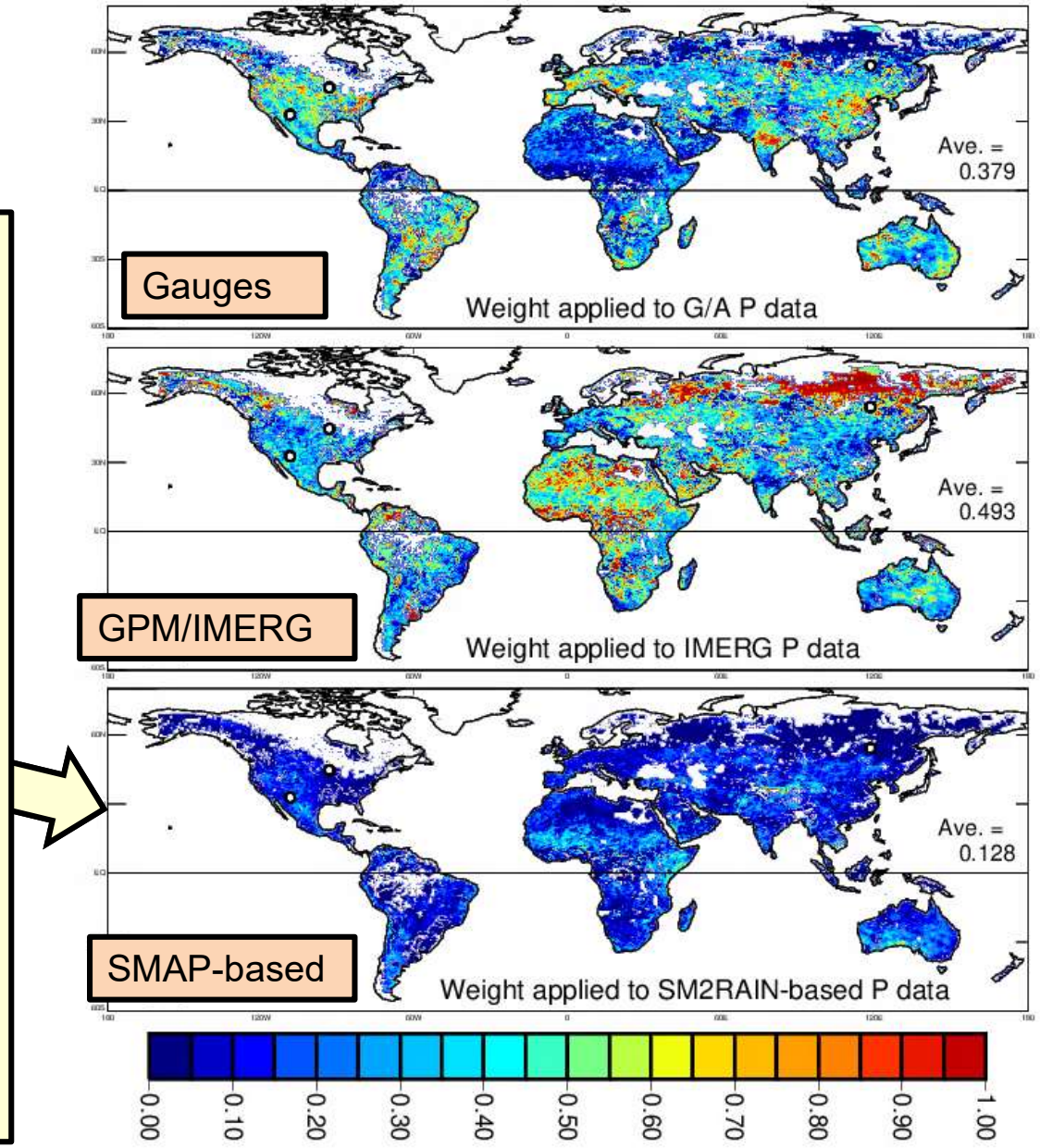
Derived weights for merging

SMAP weights are small, but remember:

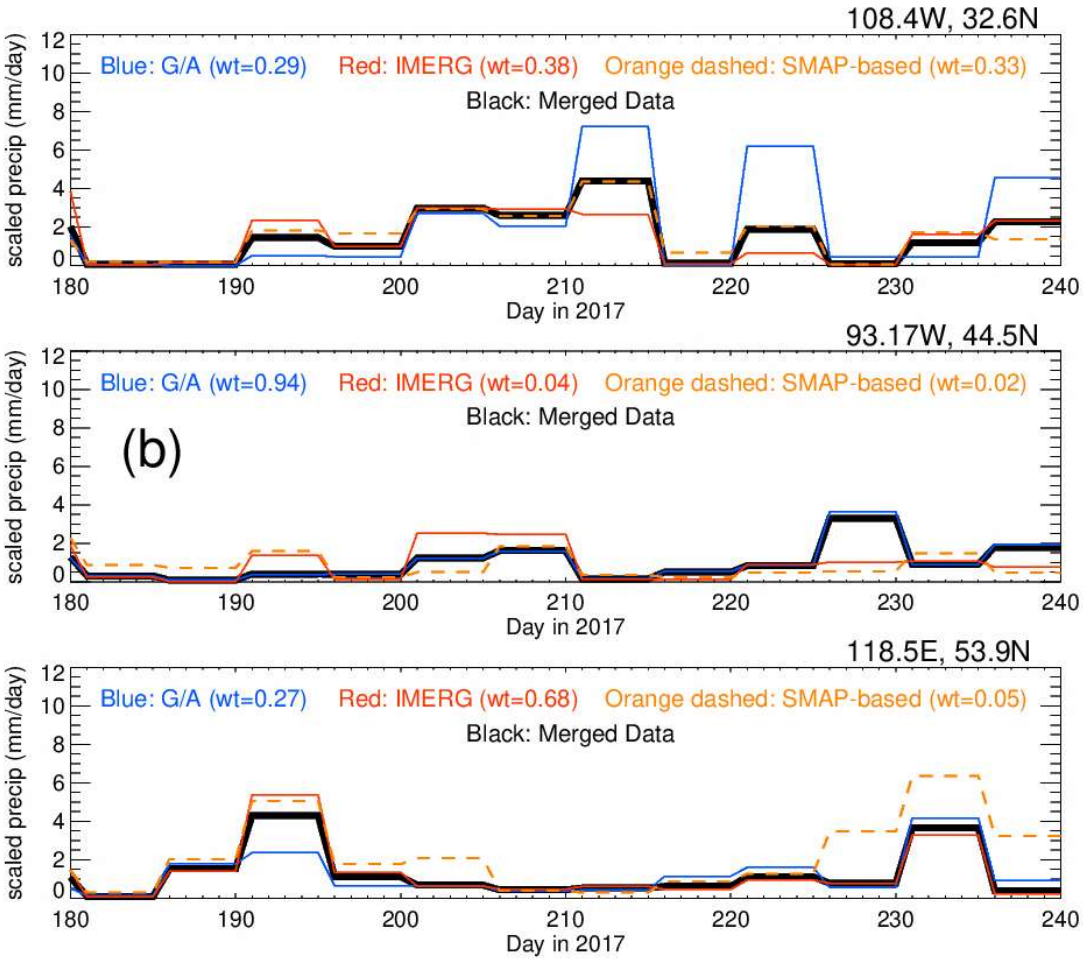
(1) Unlike the gauges and the IMERG data, SMAP was not designed to measure precipitation.

(2) If it weren't for the SMAP-based rainfall data, we wouldn't have information (through triple collocation) on the accuracy of the Gauge and IMERG data.

(3) The SMAP weights are nonzero ⇨ SMAP does contribute!

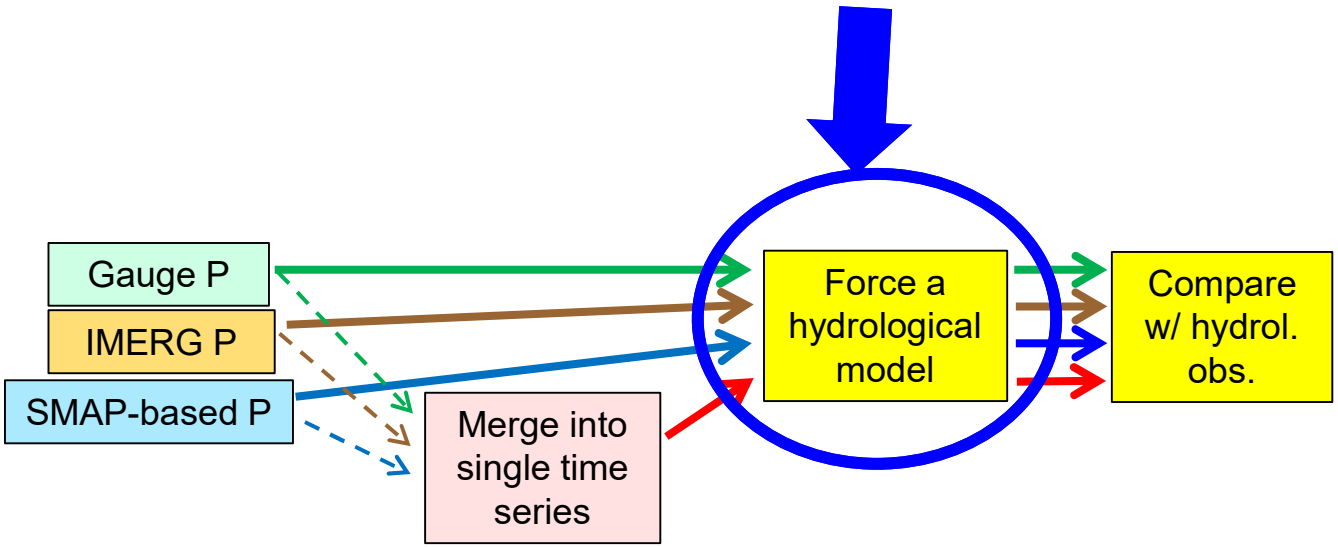


Sample time series

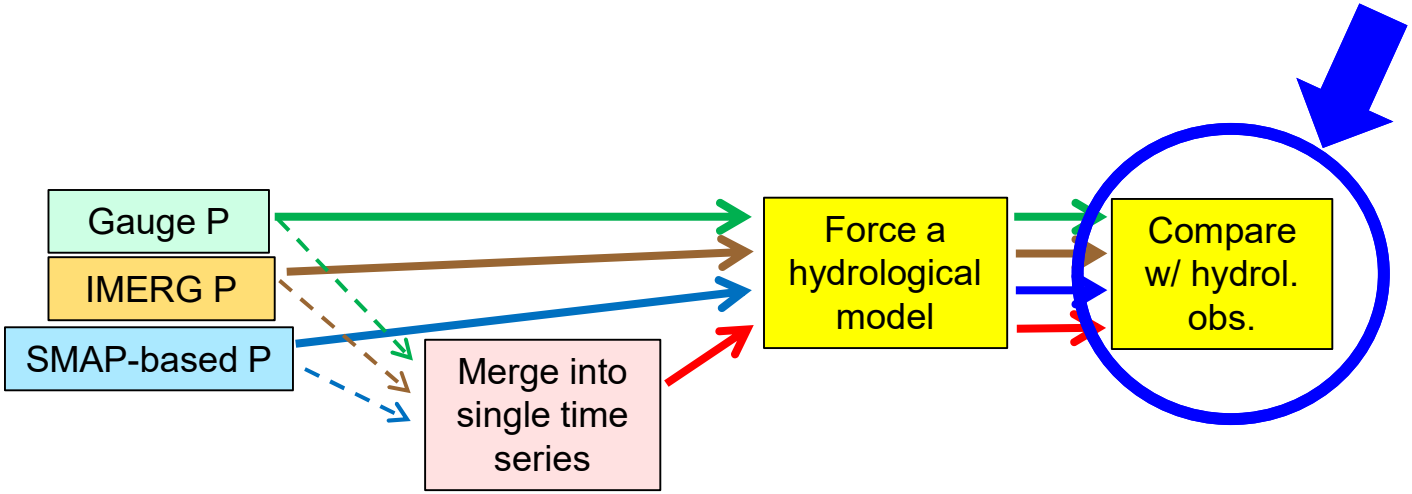


2. Expanded simulation period

(May-Sept. of 2015-2018; 2019-2020 used for SM2RAIN calibration)



3. Revamped evaluation procedure

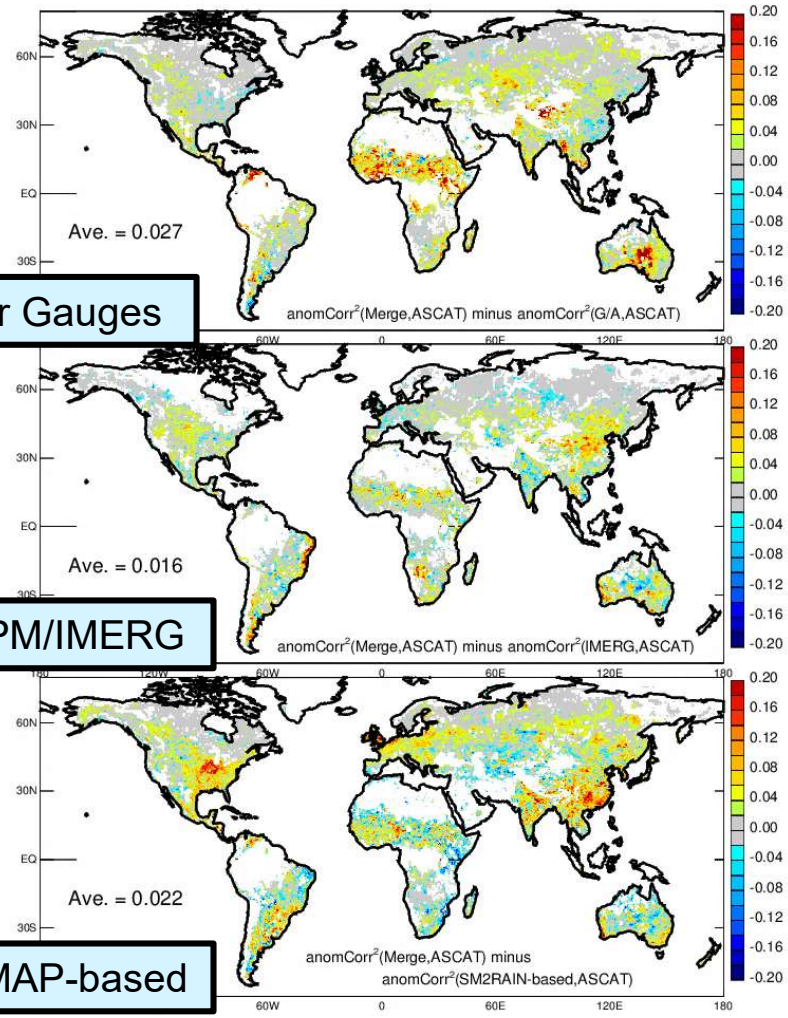


Hydrological observations examined here: ASCAT soil moisture

Although ASCAT soil moisture is far from perfect, it is completely independent of the three datasets (gauges, IMERG, SMAP-based) contributing to the merged precipitation product.

⇨ If the merged product agrees better with the ASCAT data than do each of the three contributors, the merged product can be deemed more accurate.

r^2 skill differences:
r^2 for merged product minus r^2 for individual precip. datasets



Improvement over Gauges

Improvement over GPM/IMERG

Improvement over SMAP-based

ASCAT comparisons

Mostly yellow and orange ⇨ Merged product does perform better!

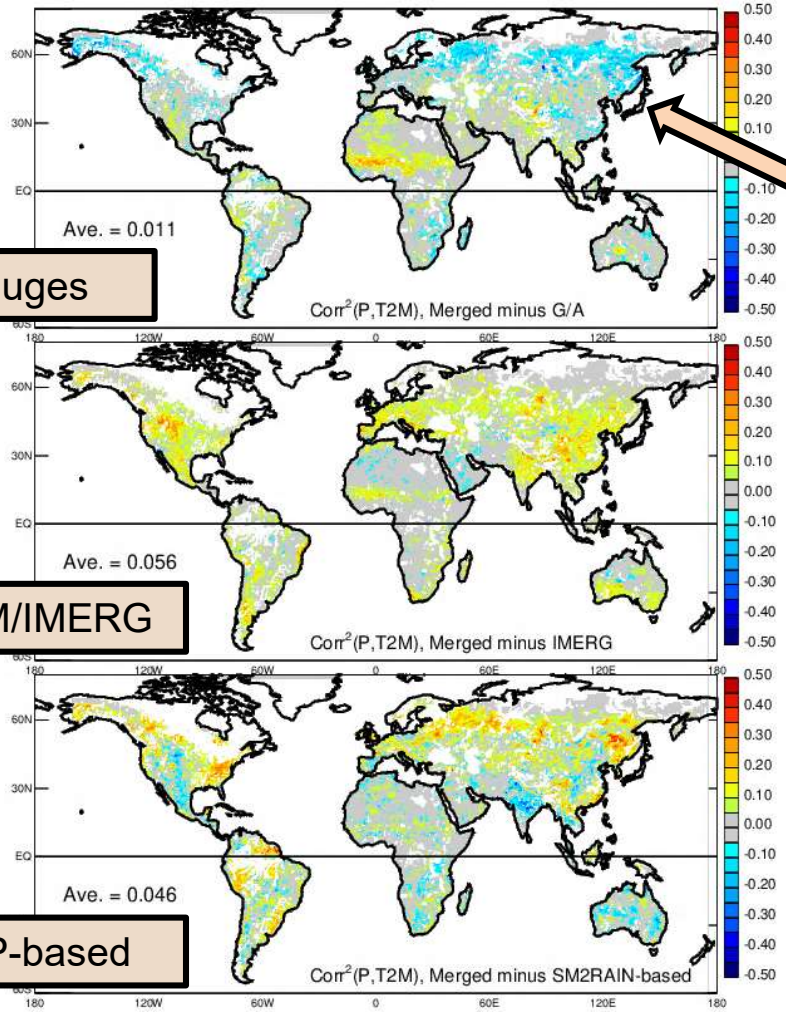
We also tested for consistency against air temperature (T2M) station observations. In theory, precipitation and air temperature should be anti-correlated for two reasons:

i) More rainfall ⇨ more ET ⇨ more evaporative cooling

ii) More rainfall ⇨ more cloudiness ⇨ less solar forcing

This is ***not*** as clean and straightforward a validation as we obtained with the ASCAT data. Still, the results are interesting.

r^2 skill differences:
r^2 for merged product minus r^2 for individual precip. datasets



Improvement over Gauges

Improvement over GPM/IMERG

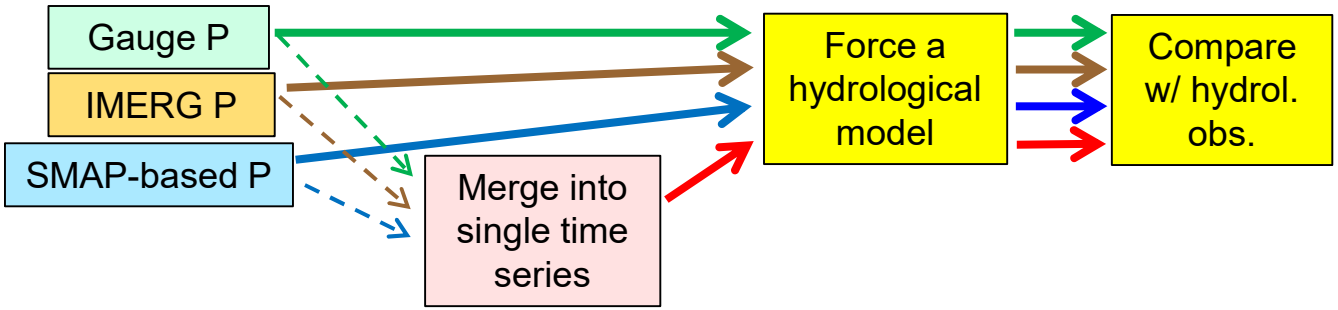
Improvement over SMAP-based

T2M comparisons

Technical reasons for some of these blue colors.

Mostly yellow and orange ⇨ Merged product does perform better!

4. Wrote paper (under review at WRR; preprint available on request)



Summary (transcribed from the “three key points” in our WRR paper):

- A novel strategy for combining three fully independent precipitation datasets into a single merged precipitation dataset is presented.
- A merged dataset is built from: (i) gauge-based, (ii) satellite-based and (iii) soil moisture retrieval-based precipitation estimates.
- The merged dataset validates better against independent data than does each contributor, benefitting from their complementary strengths.

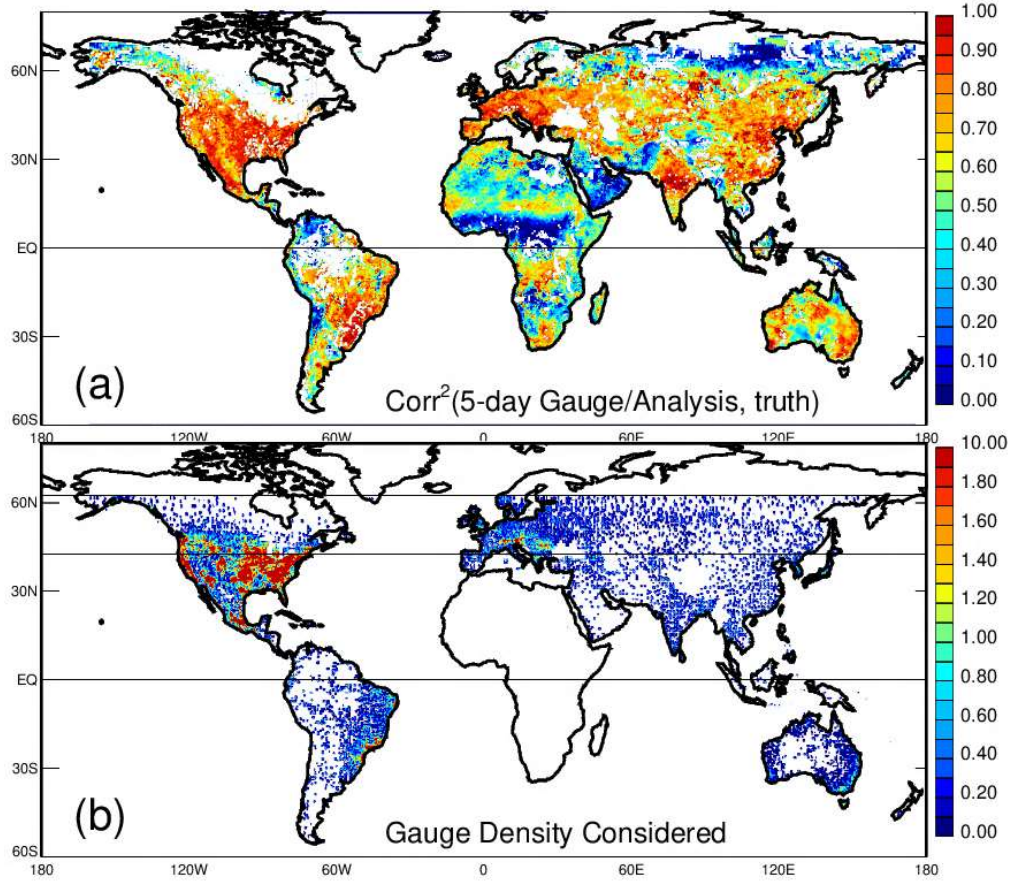
Extra Slides

Technical project advancements:

1. Revised focus on pentad precipitation
2. Use of log(P) in triple collocation analysis
3. Development of optimized approach for defining weights applied to contributing datasets
4. [Miscellaneous bug corrections, etc.]
5. Expansion of evaluation period (SMAP-based precipitation calibrated on 2019-2020; evaluation period expanded to May-Sept. of 2015-2018)
6. Revised global-scale evaluation approach:
    a. ASCAT data
    b. Air temperature data

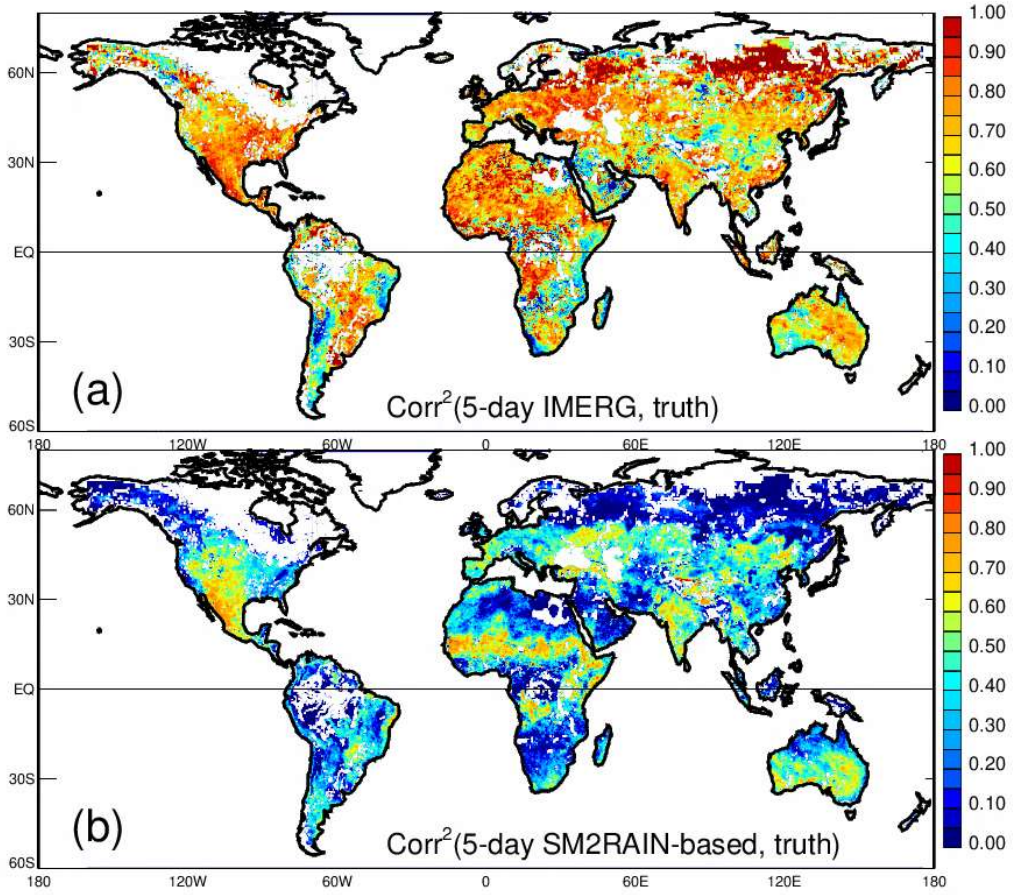
Our estimates of where the gauges are accurate relative to the unknown truth…

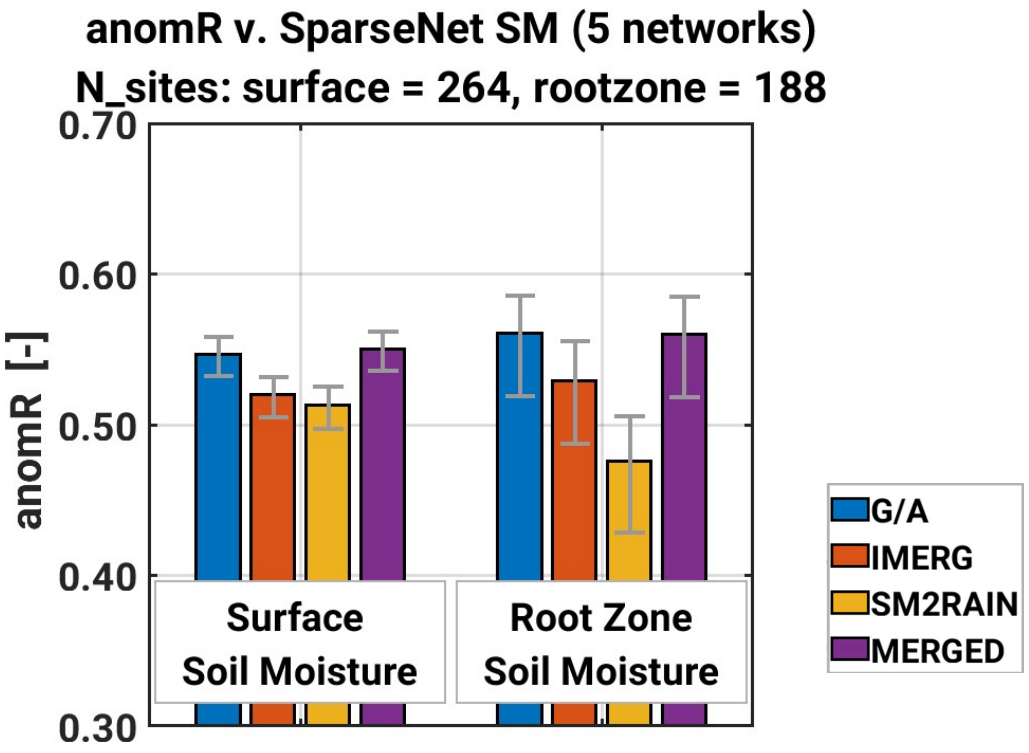
… are quite consistent with maps of rain gauge density!

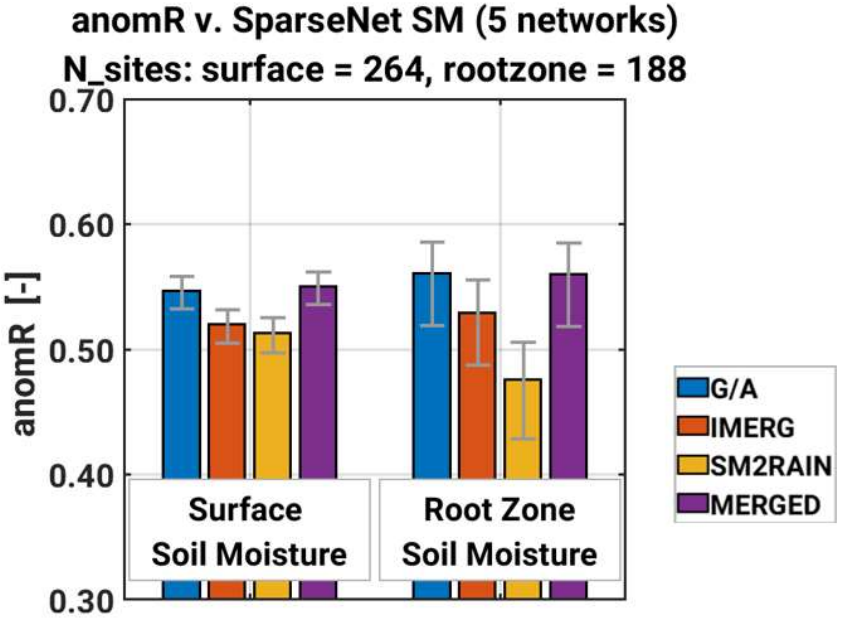


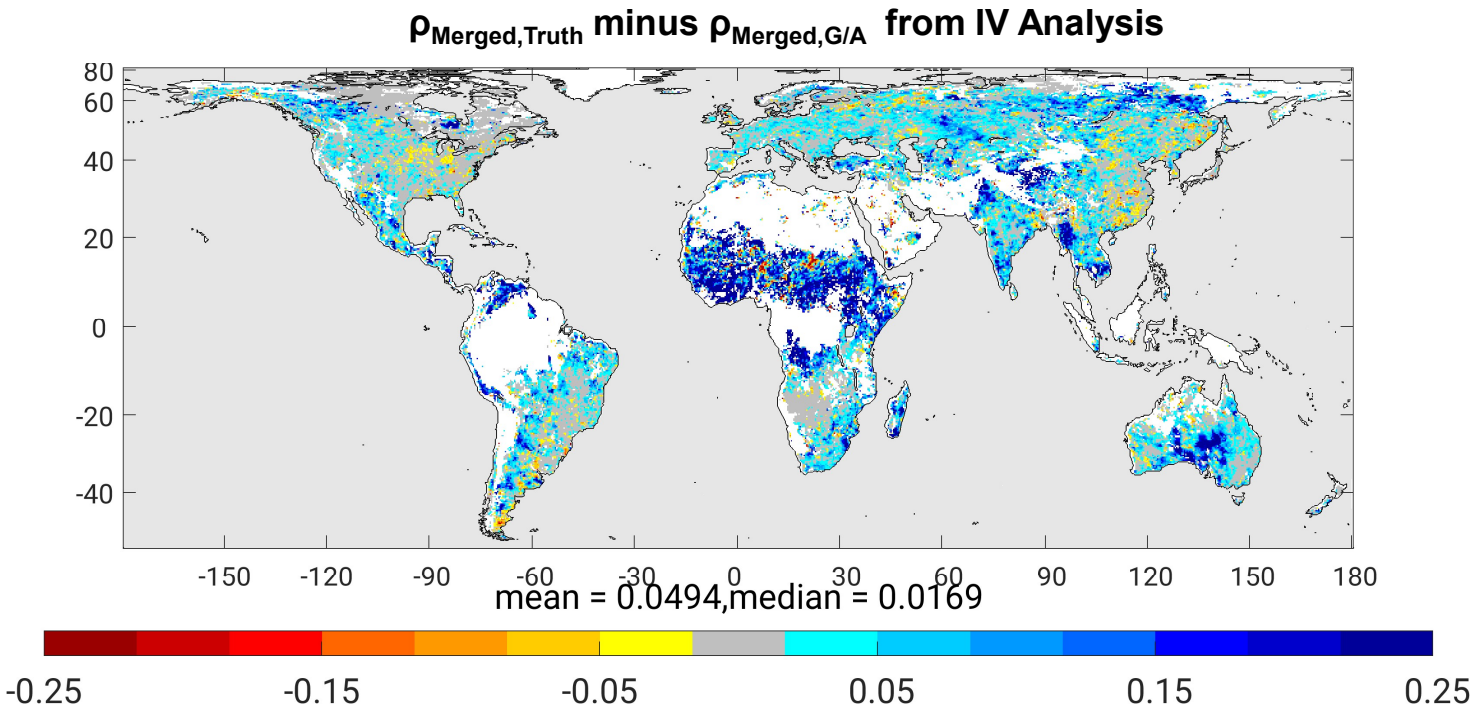
Estimates of the accuracy of the GPM/IMERG product against the unknown truth.

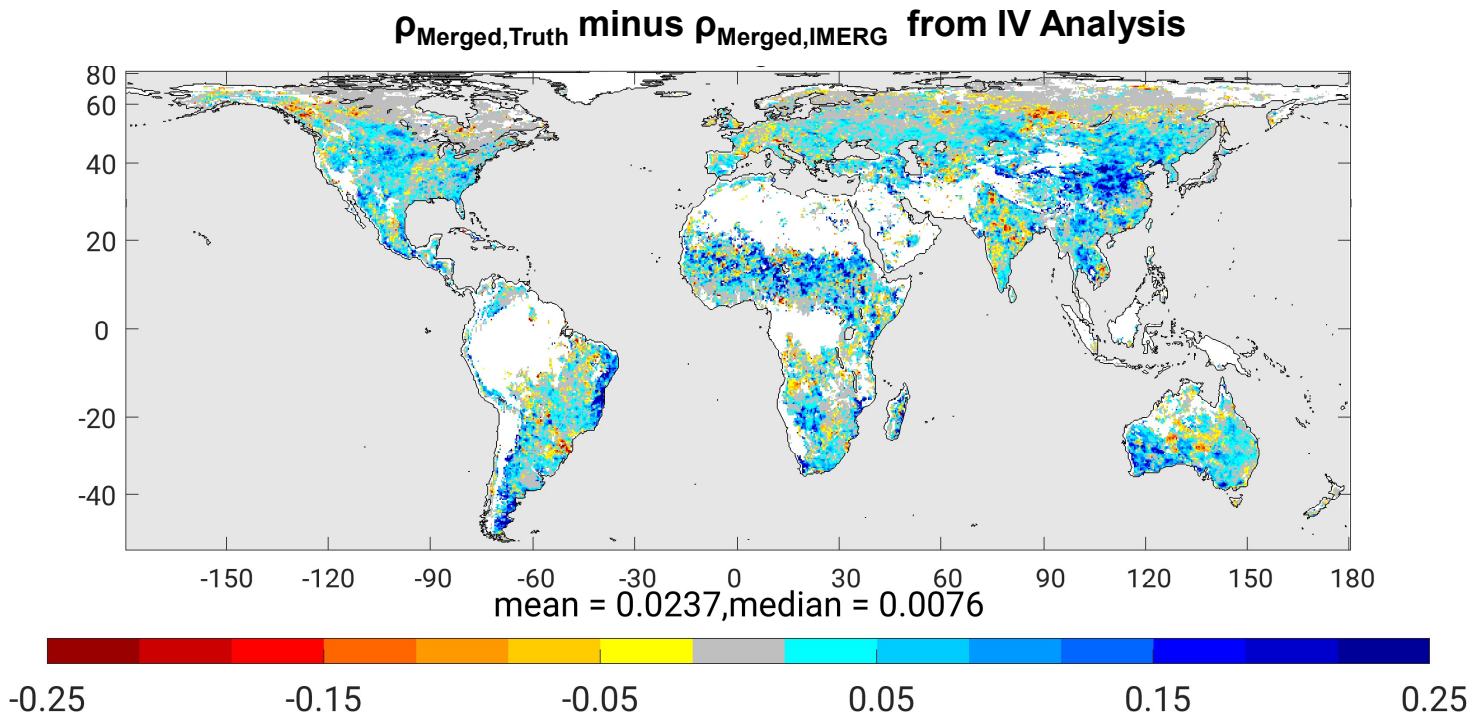
Estimates of the accuracy of the SMAP-based rainfall product against the unknown truth.

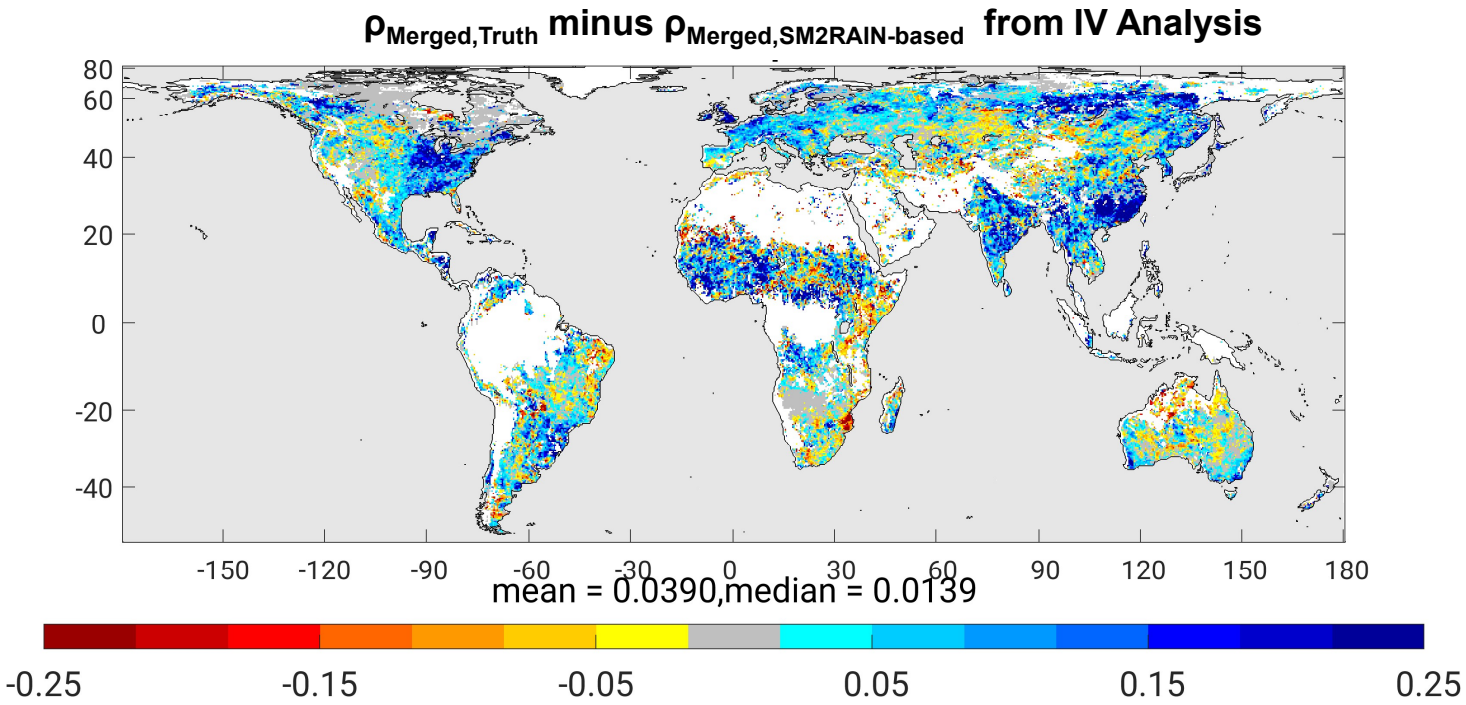


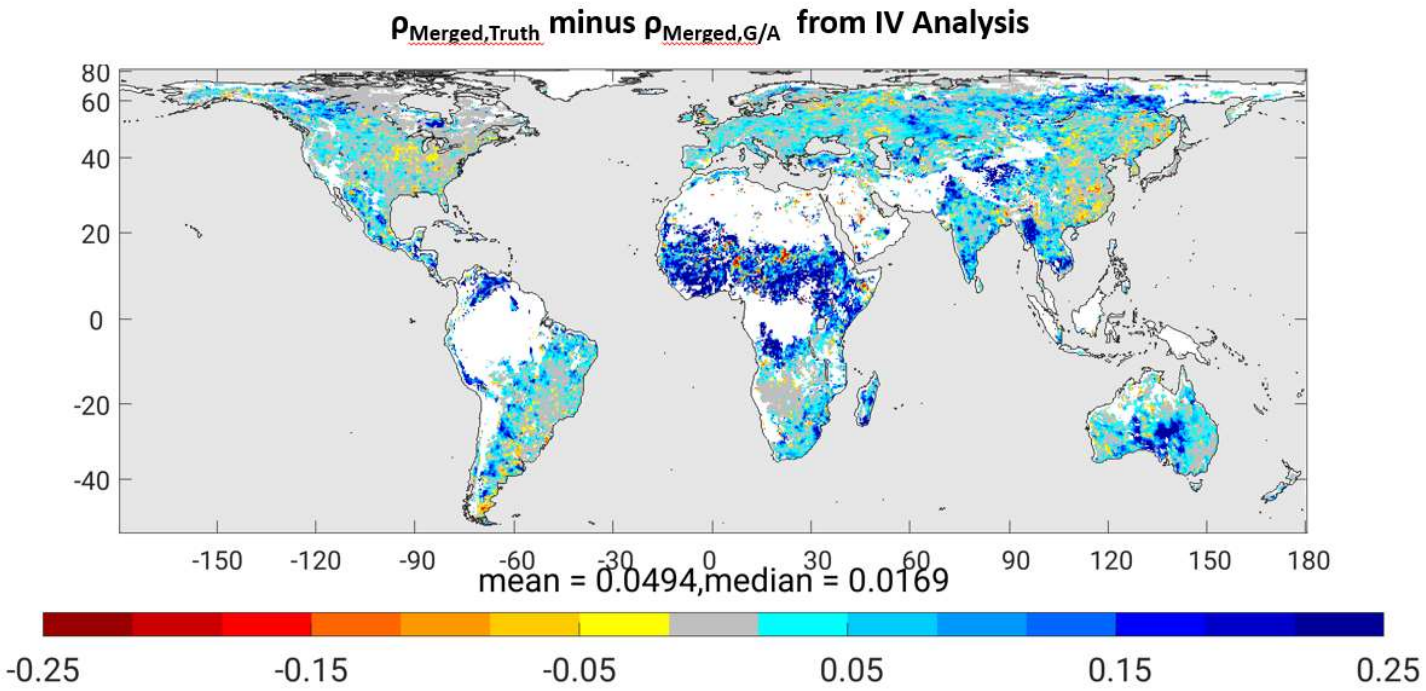

anomR v. SparseNet SM (5 networks)
N_sites: surface = 264, rootzone = 188
anomR [-]
0.70
0.60
0.50
0.40
0.30
Surface Soil Moisture
Root Zone Soil Moisture
G/A
IMERG
SM2RAIN
MERGED

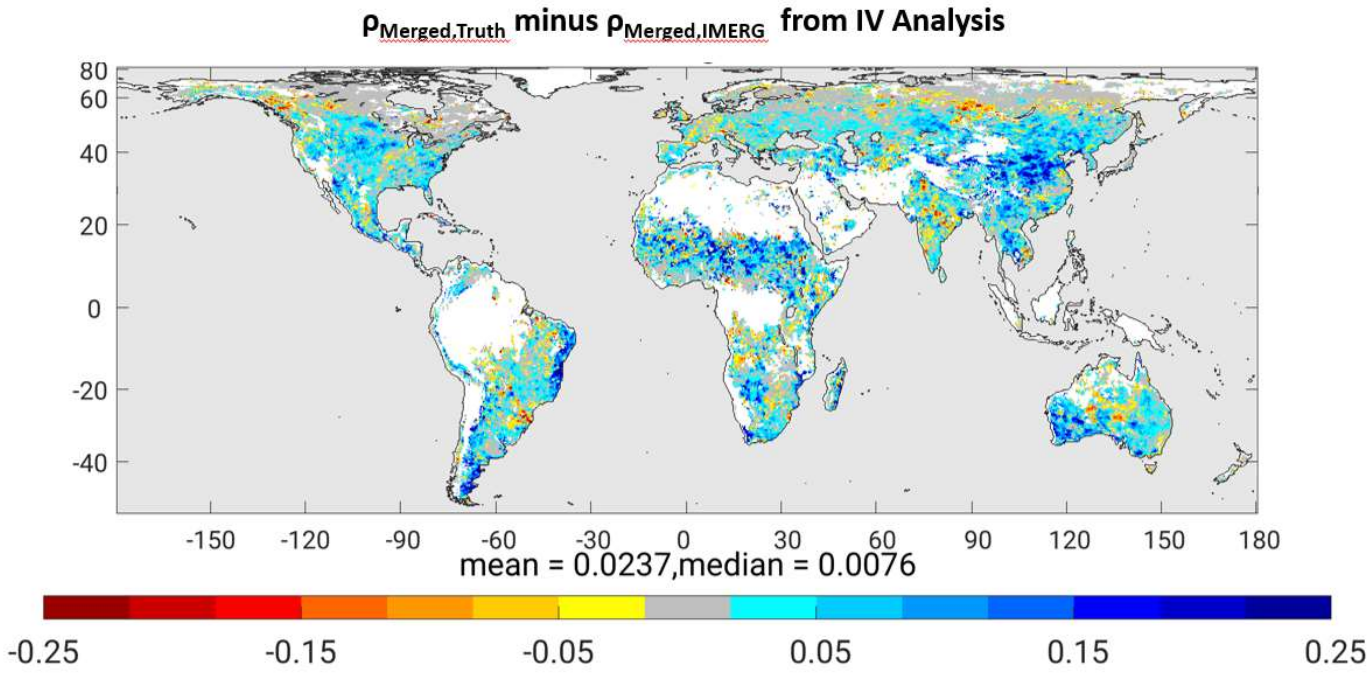

anomR v. SparseNet SM (5 networks)
N_sites: surface = 264, rootzone = 188
anomR [-]
0.70
0.60
0.50
0.40
0.30
Surface Soil Moisture
Root Zone Soil Moisture
G/A
IMERG
SM2RAIN
MERGED

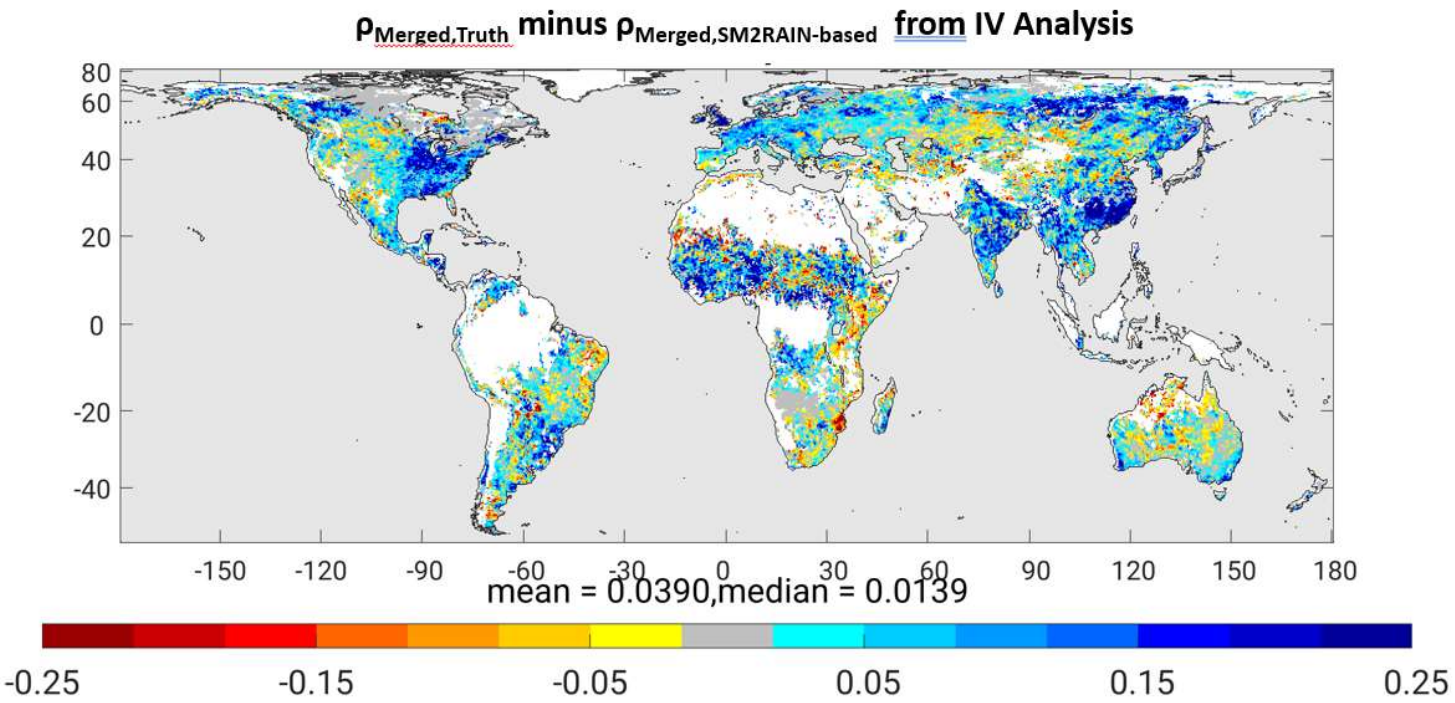

ρMerged,Truth minus ρMerged,G/A from IV Analysis
80
60
40
20
0
-20
-40
-150
-120
-90
-60
-30
0
30
60
90
120
150
180
mean = 0.0494,median = 0.0169
-0.25
-0.15
-0.05
0.05
0.15
0.25


ρMerged,Truth minus ρMerged,IMERG from IV Analysis
80
60
40
20
0
-20
-40
-150
-120
-90
-60
-30
0
30
60
90
120
150
180
mean = 0.0237,median = 0.0076
-0.25
-0.15
-0.05
0.05
0.15
0.25


ρMerged,Truth minus ρMerged,SM2RAIN-based from IV Analysis
80
60
40
20
0
-20
-40
-150
-120
-90
-60
-30
0
30
60
90
120
150
180
mean = 0.0390,median = 0.0139
-0.25
-0.15
-0.05
0.05
0.15
0.25


ρMerged,Truth minus ρMerged,G/A from IV Analysis
80
60
40
20
0
-20
-40
-150
-120
-90
-60
-30
0
30
60
90
120
150
180
mean = 0.0494,median = 0.0169
-0.25
-0.15
-0.05
0.05
0.15
0.25


ρMerged,Truth minus ρMerged,IMERG from IV Analysis
80
60
40
20
0
-20
-40
-150
-120
-90
-60
-30
0
30
60
90
120
150
180
mean = 0.0237,median = 0.0076
-0.25
-0.15
-0.05
0.05
0.15
0.25


ρMerged,Truth minus ρMerged,SM2RAIN-based from IV Analysis
80
60
40
20
0
-20
-40
-150
-120
-90
-60
-30
0
30
60
90
120
150
180
mean = 0.0390,median = 0.0139
-0.25
-0.15
-0.05
0.05
0.15
0.25

Precipitation Datasets Considered

We will merge the following datasets:

"Gauges": CPCU-based rain gauge data as used in SMAP Level 4 Nature Run. (Note: In Africa and north of 60N, CPCU data are poor and are replaced in the Nature Run and here by analysis data)

"IMERG": Satellite-derived rain rates produced for NASA's GPM project; no rain gauge corrections

"SMAP-based": Derived using the approach of Koster et al. (2018)

Because the three contributing datasets have errors that are completely independent, we can use triple collocation (e.g., Dong et al. 2020) to estimate the accuracy (r^2, versus the unknown truth) of each dataset after each is aggregated to a common grid.

We generate a merged precipitation time series using accuracy-dependent weights for each contributor:

$$P_{merged}(t) = W_{gauge}\, P_{gauge}(t) + W_{IMERG}\, P_{IMERG}(t) + W_{SMAP}\, P_{SMAP}(t)$$

gauge-based time series

GPM-based time series

SMAP-based time series

The weights are computed as

$$W_{gauge} = \rho_{gauge}^2 / (\rho_{gauge}^2 + \rho_{IMERG}^2 + \rho_{SMAP}^2)$$
$$W_{IMERG} = \rho_{IMERG}^2 / (\rho_{gauge}^2 + \rho_{IMERG}^2 + \rho_{SMAP}^2)$$
$$W_{SMAP} = \rho_{SMAP}^2 / (\rho_{gauge}^2 + \rho_{IMERG}^2 + \rho_{SMAP}^2)$$

where ρ_X is the estimated correlation against the unknown truth (from triple collocation) of time series X.

Yet to Come:

- Optimize merging approach with help from GPM folks (actually, a big job)
- Examine potential for merging near-real-time (NRT) IMERG and rain gauge data to force hydrological model, producing NRT soil moistures ⇨ how well do we do?
- Examine potential for both: (i) merging NRT IMERG and rain gauge data to force hydrological model and (ii) assimilating NRT SMAP Tb into the model to produce NRT soil moistures ⇨ how well do we do?

Can we improve L4 SM estimates?