Land Observations for Model Calibration and Data Assimilation

J. Kolassa[1,2]

R. Reichle[1], R. Koster[1] , F. Zeng, Q. Liu[1,2] and S. Mahanama[3,4]

ECMWF Annual Seminar 2021

Session 4: Assessing the impact of observation

September 16, 2021

(1) Global Modeling and Assimilation Office, NASA Goddard Space Flight Center
(2) Science Systems and Applications Inc.
(3) Science Applications International Corp.
(4) Hydrological Sciences Laboratory, NASA Goddard Space Flight Center

Introduction

- Land data assimilation (DA) in numerical weather prediction (NWP) is growing
- Emergence of new high-quality land observations
  - SMOS and SMAP for soil moisture (SM)
  - Ecostress, SBG and others for vegetation
- Land DA techniques developed in Earth System Modeling community transition into NWP systems

Image Credit: NASA

Outline

1. Merging of multi-sensor information
2. Increasing efficiency of data assimilation
3. Model parameter calibration vs. structural changes
4. Impact of land observations in NWP systems

Optimizing multi-sensor data assimilation

Multi-sensor satellite observations

Image Credit: NASA

- Combining observations from multiple satellite sensors can be very advantageous
  - For soil moisture combining active and passive microwave observations (e.g., ESA-CCI SM, SMAP)

Q: Is there an optimal method for combining observations from multiple sensors?

Multi-sensor satellite observations: Retrievals

Q: Is there an optimal method for combining observations from multiple sensors?

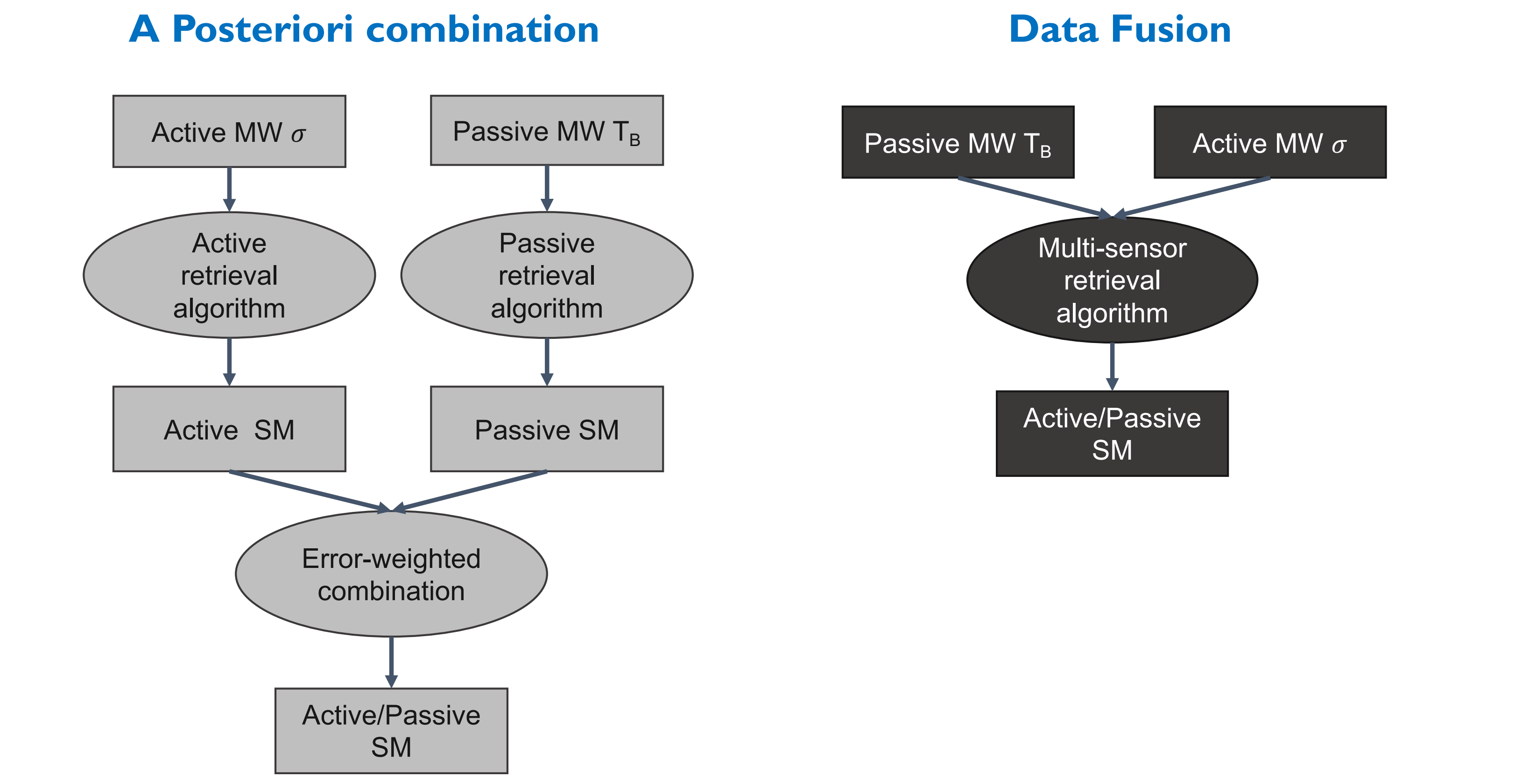
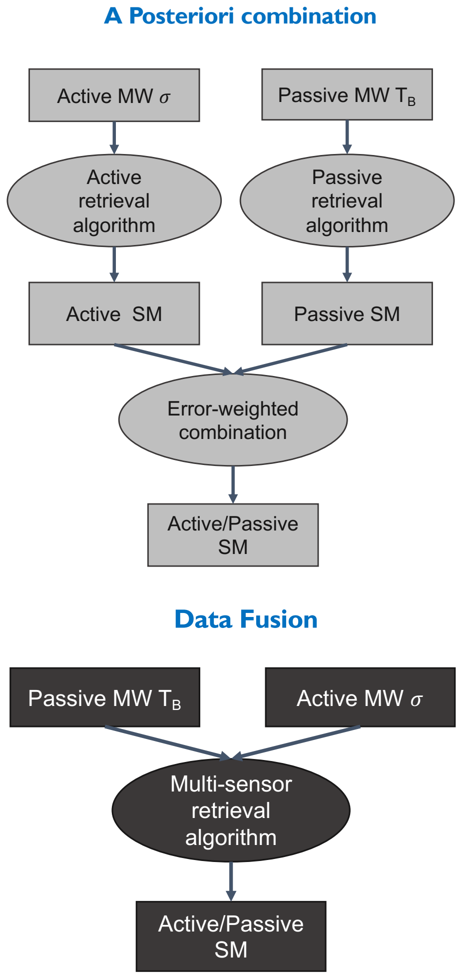


Fig 1: Comparison of methods for combining active and passive microwave data

Kolassa et al., 2013

Multi-sensor satellite observations: Retrievals

Q: Is there an optimal method for combining observations from multiple sensors?



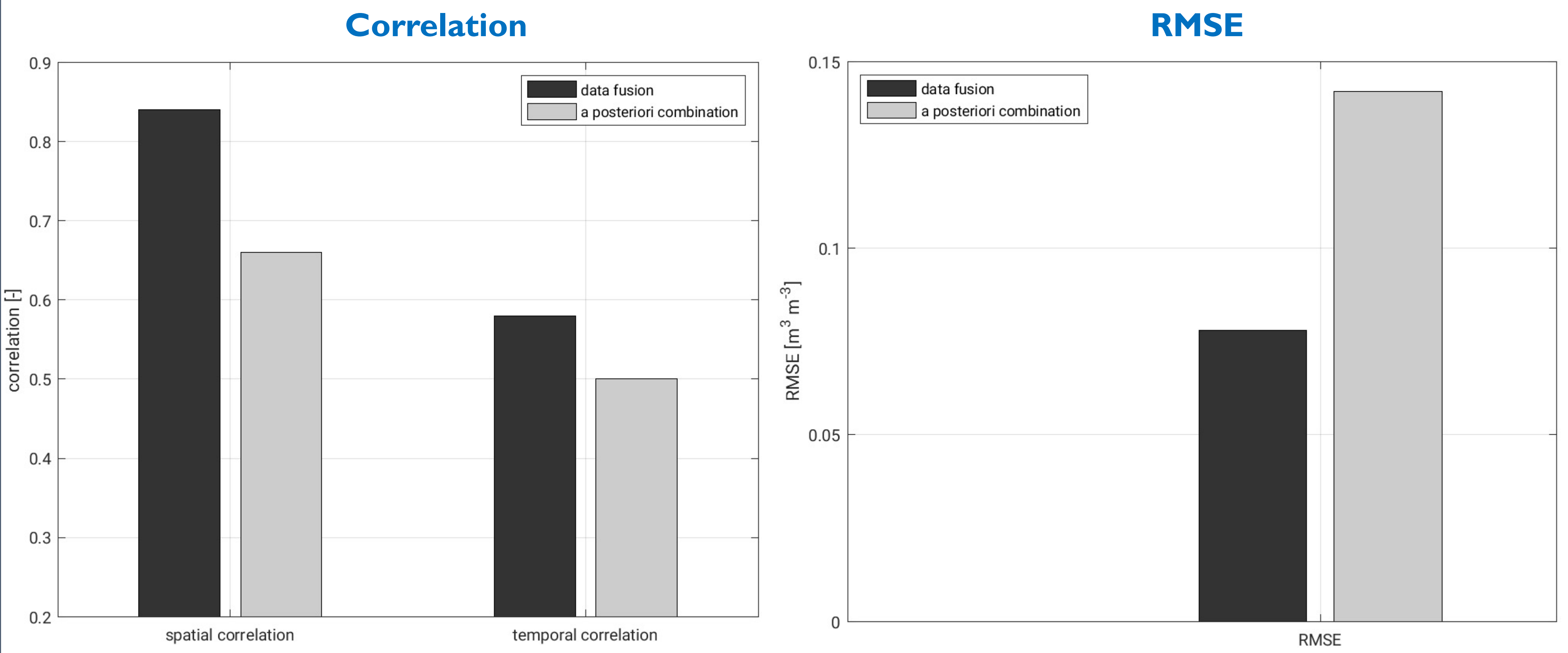


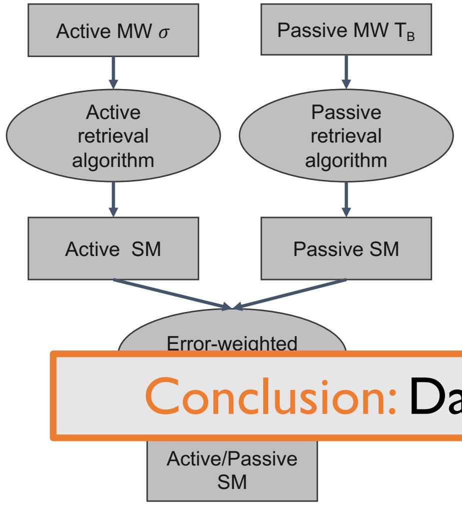
Fig 2: Soil moisture retrieval skill evaluated against *in situ* observations

Kolassa et al., 2013

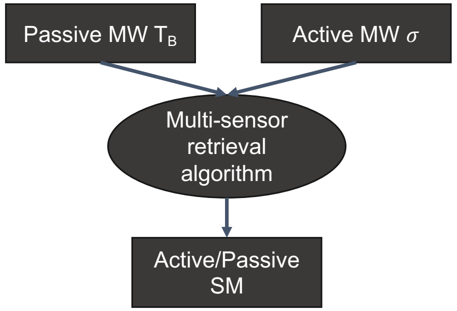
Multi-sensor satellite observations: Retrievals

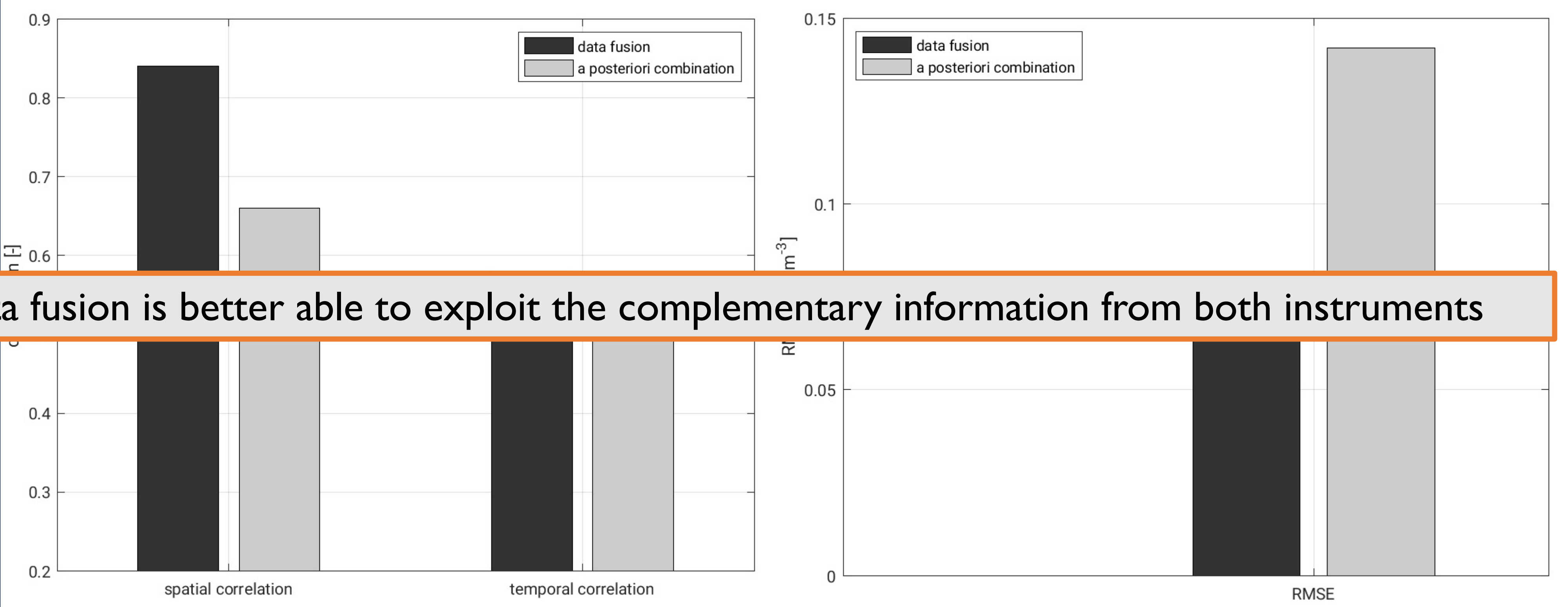
Q: Is there an optimal method for combining observations from multiple sensors?

A Posteriori combination



Data Fusion





Conclusion: Data fusion is better able to exploit the complementary information from both instruments

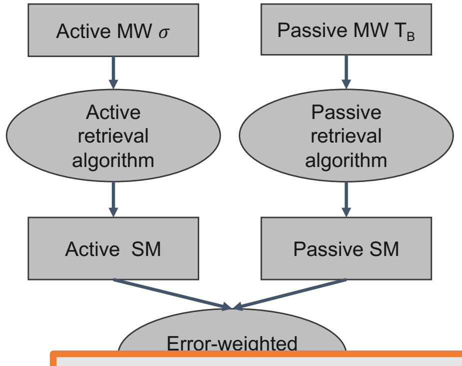
Fig 2: Soil moisture retrieval skill evaluated against *in situ* observations

Kolassa et al., 2013

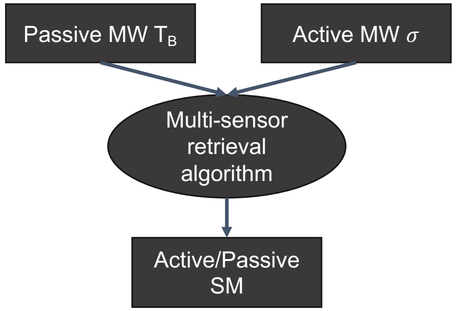
Multi-sensor satellite observations: Retrievals

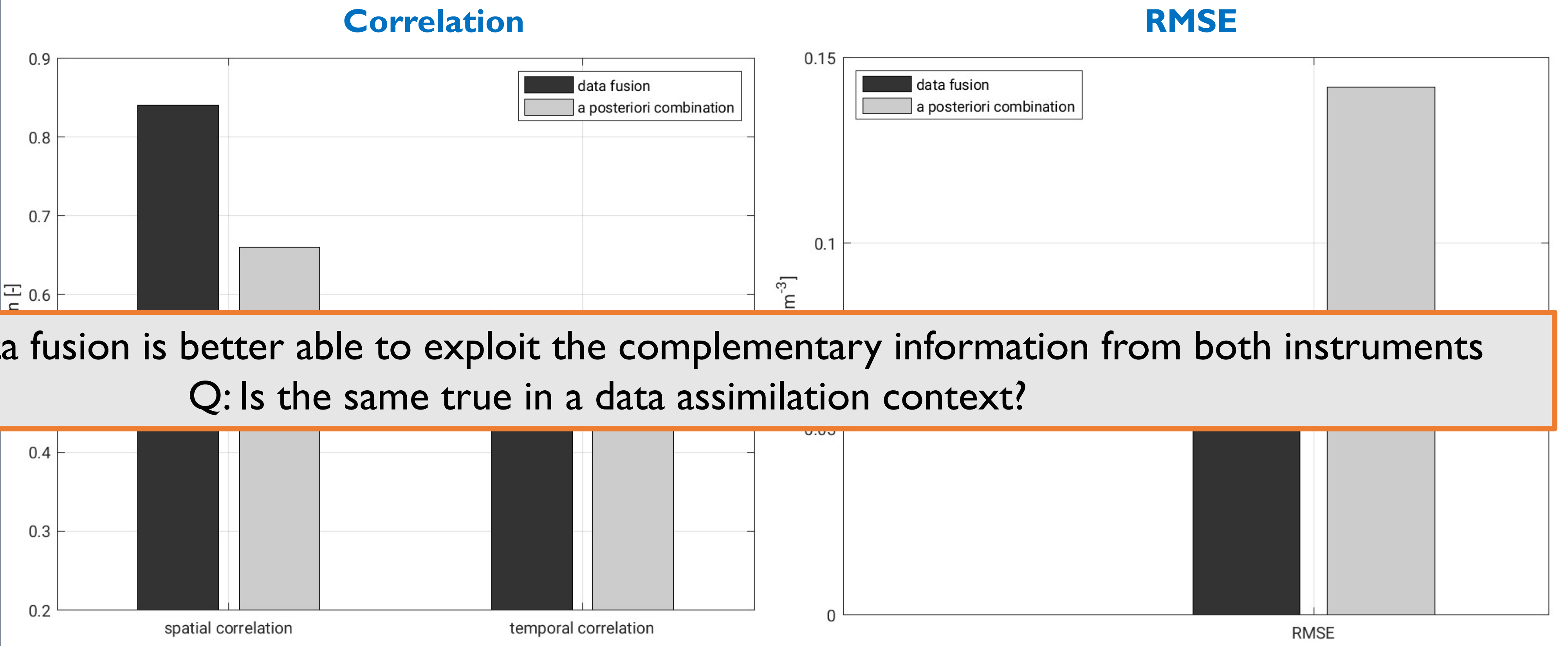
Q: Is there an optimal method for combining observations from multiple sensors?

A Posteriori combination



Data Fusion





Conclusion: Data fusion is better able to exploit the complementary information from both instruments

Q: Is the same true in a data assimilation context?

Fig 2: Soil moisture retrieval skill evaluated against *in situ* observations

Kolassa et al., 2013

Multi-sensor satellite observations: Data Assimilation

- Simultaneously assimilating active and passive microwave SM improves model estimate more than assimilating single sensor retrieval (e.g., Draper et al., 2012)

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

Exp 1: Separate Retrieval DA

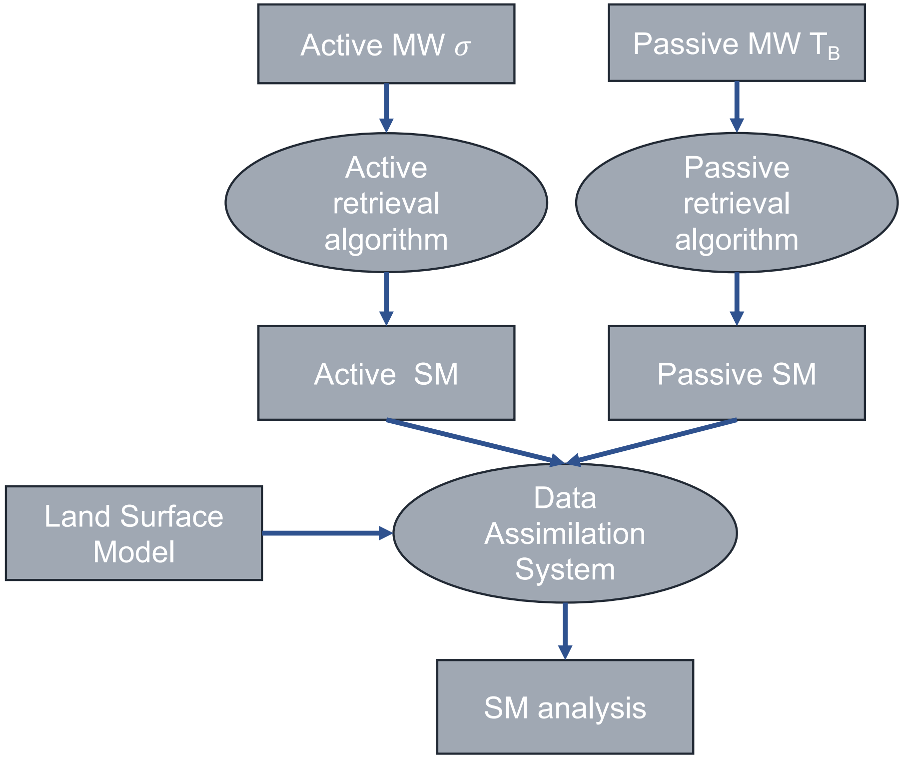


Fig 3: Comparison of methods for combining active and passive microwave data in a data assimilation context

Kolassa et al., 2017

Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

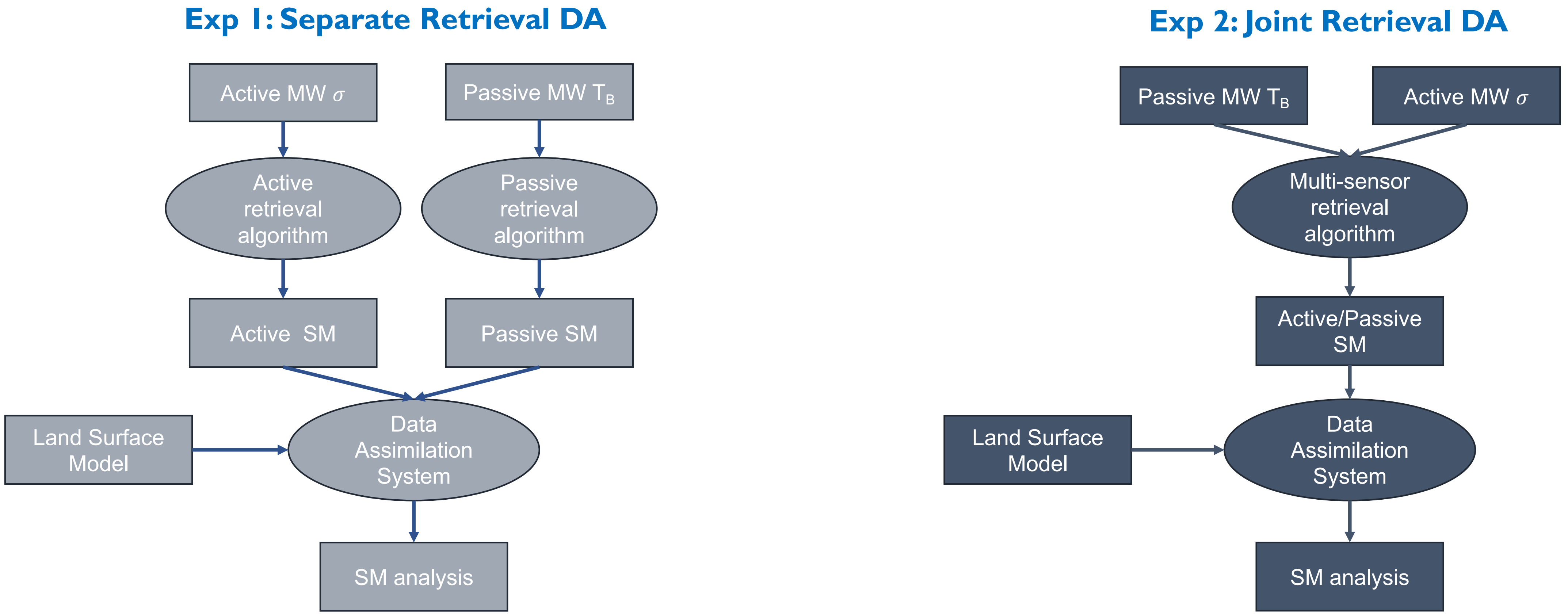


Fig 3: Comparison of methods for combining active and passive microwave data in a data assimilation context

Kolassa et al., 2017

Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

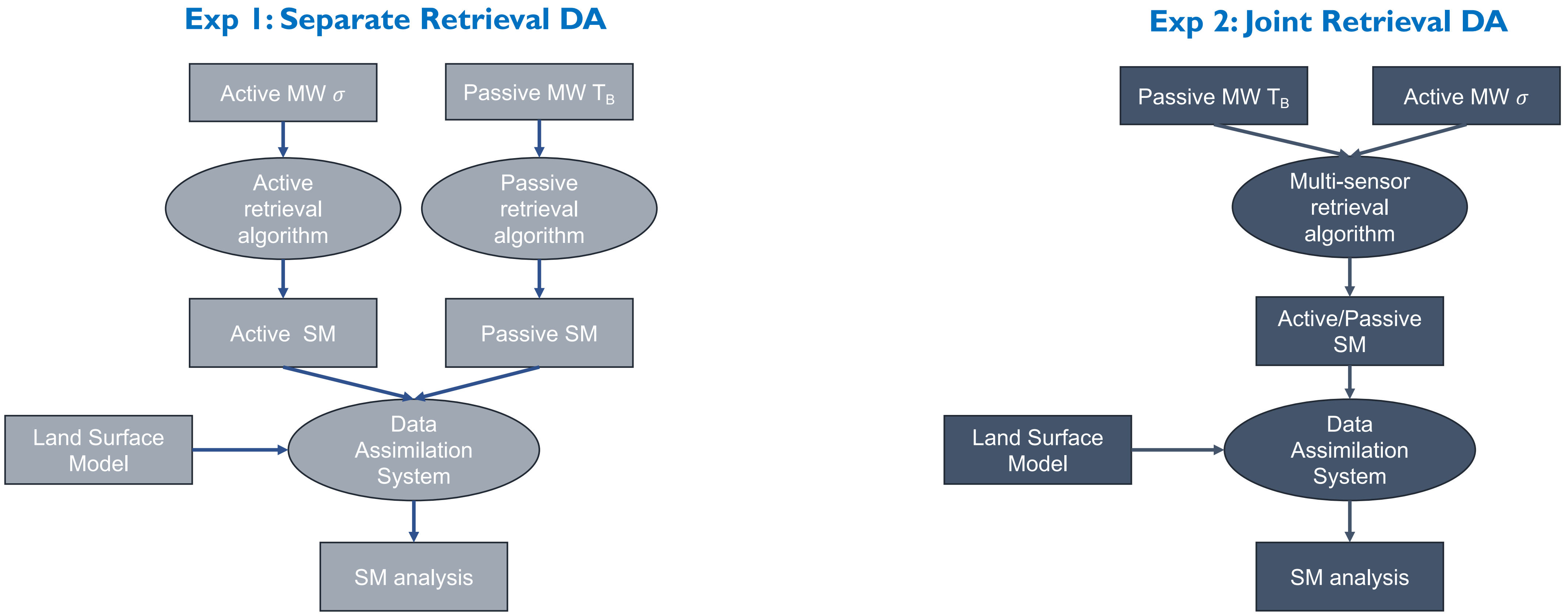


Fig 3: Comparison of methods for combining active and passive microwave data in a data assimilation context

Exp 3: Open loop (no DA)

Kolassa et al., 2017

Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

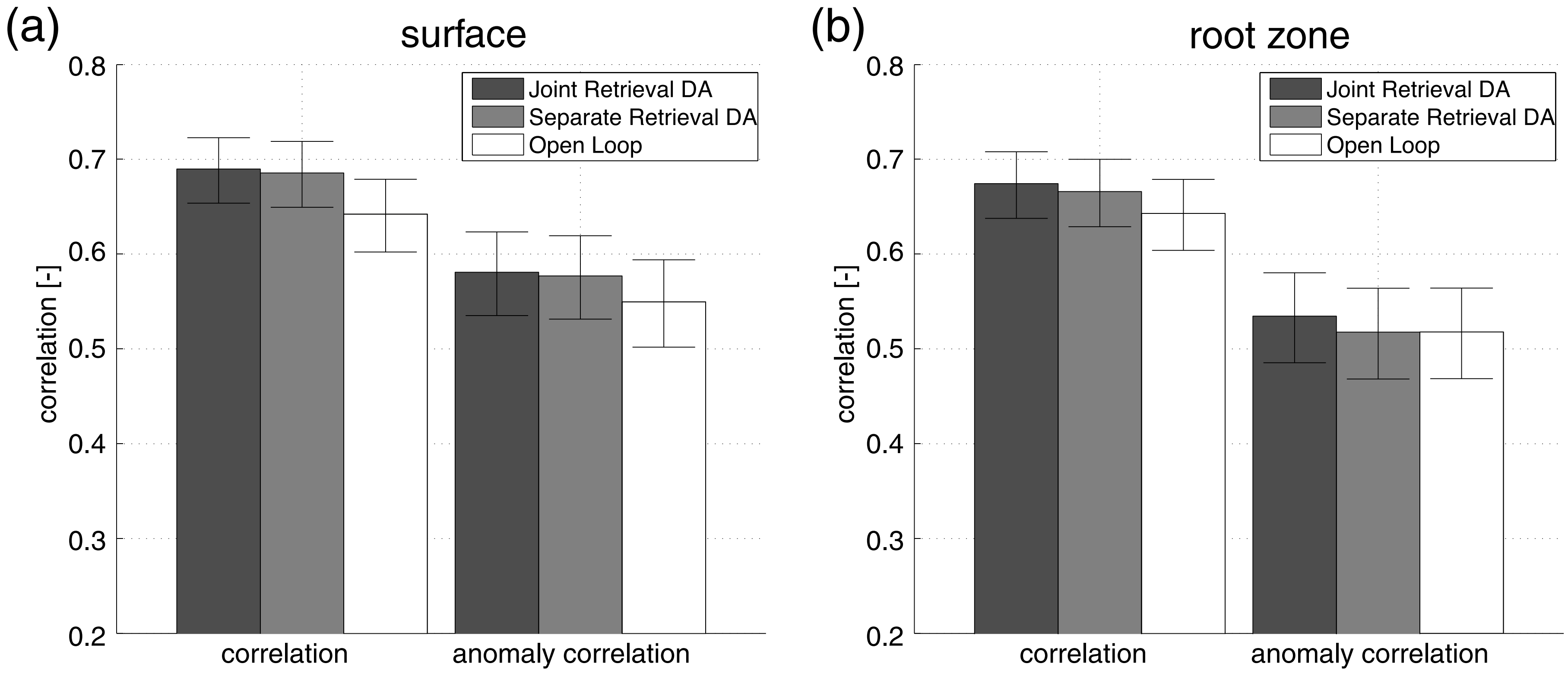
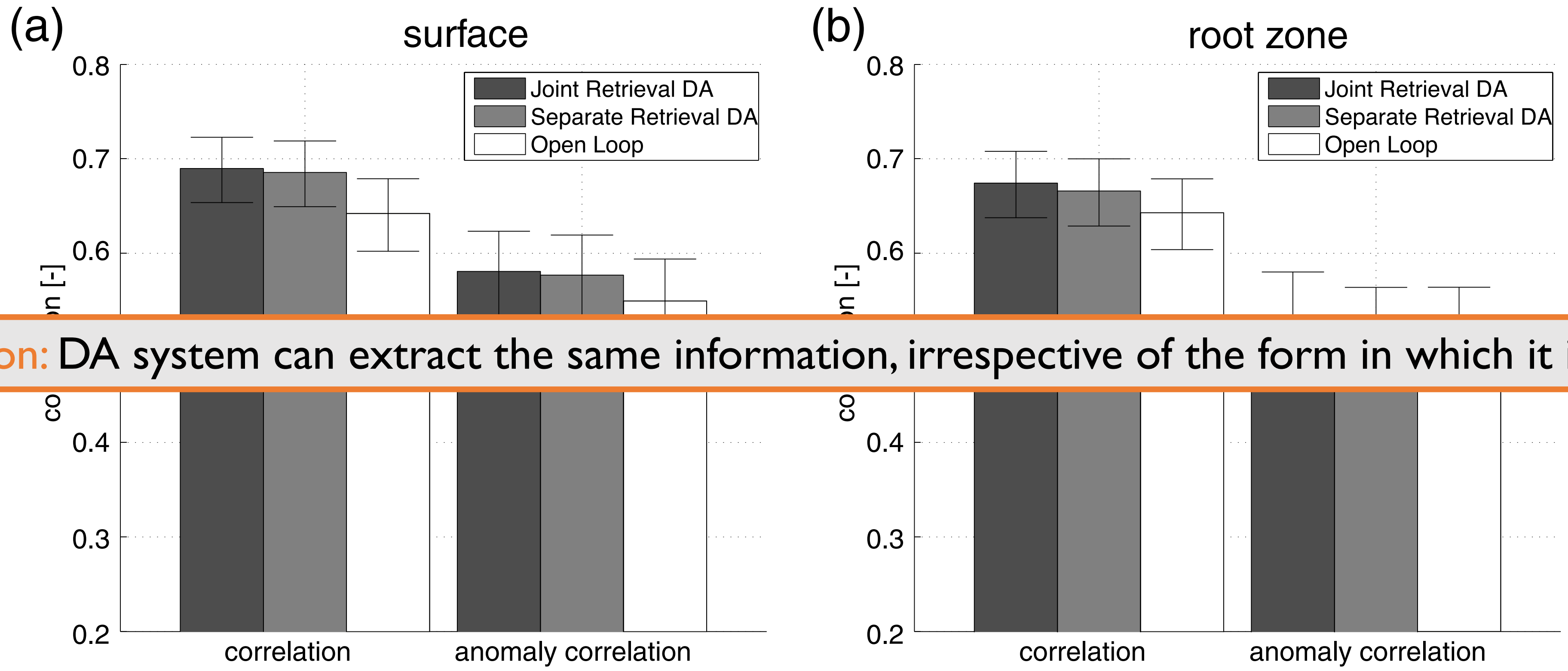


Fig 4: Soil moisture analysis skill against in situ observations

Kolassa et al., 2017

Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?



Conclusion: DA system can extract the same information, irrespective of the form in which it is provided

Fig 4: Soil moisture analysis skill against in situ observations

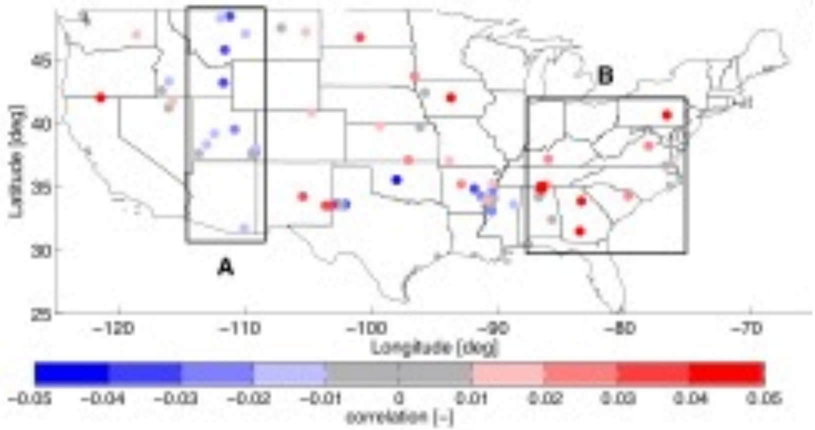
Kolassa et al., 2017

Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

Kolassa et al., 2017

Joint vs Separate DA skill



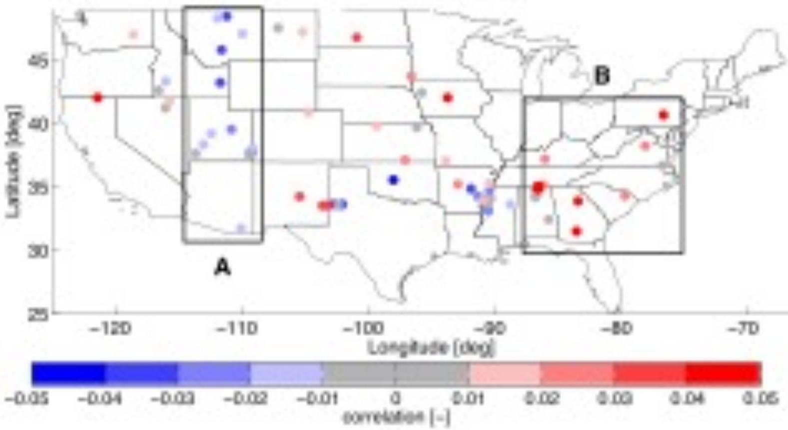
Separate retrieval DA is better ← → Joint retrieval DA is better

Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

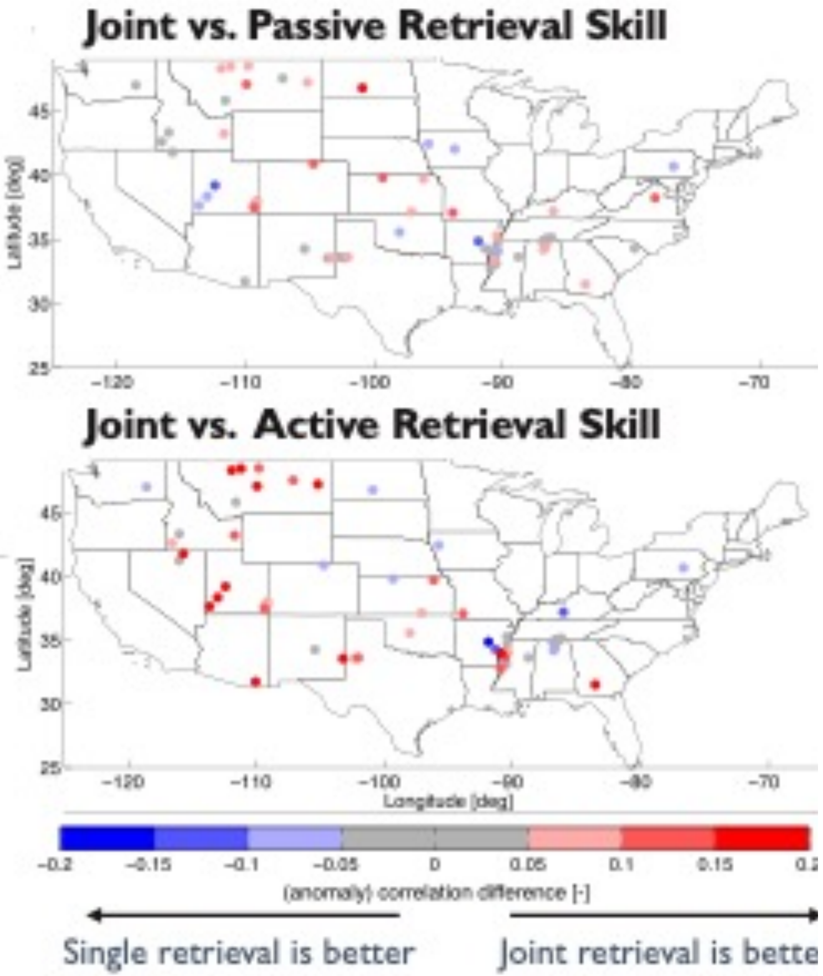
Kolassa et al., 2017

Joint vs Separate DA skill



Separate retrieval DA is better

Joint retrieval DA is better

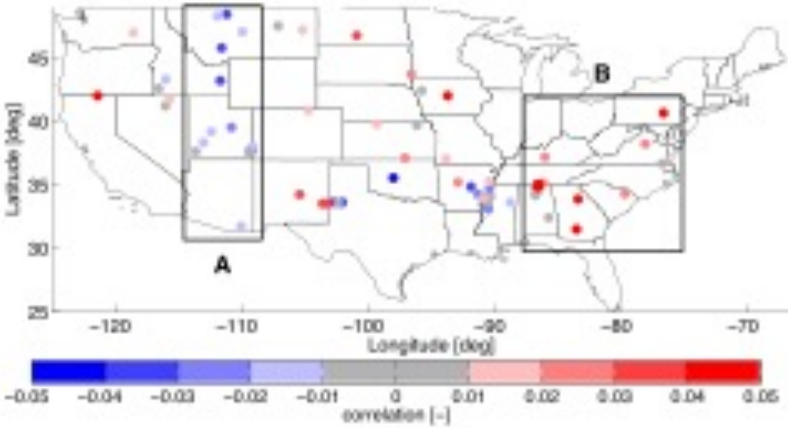


Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

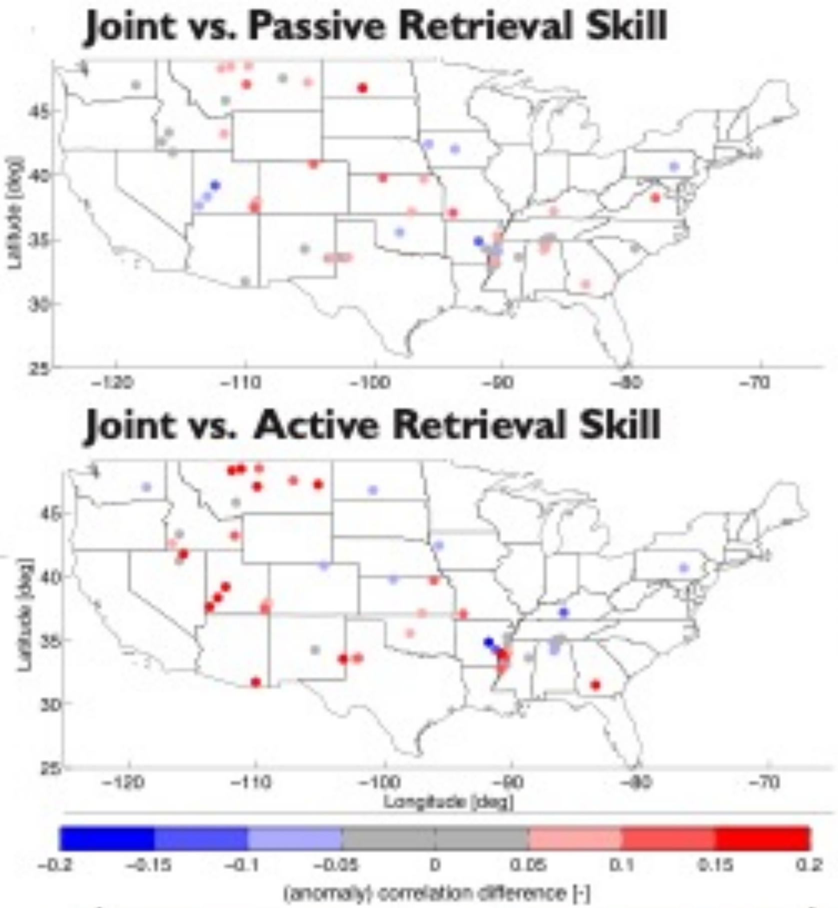
Kolassa et al., 2017

Joint vs Separate DA skill



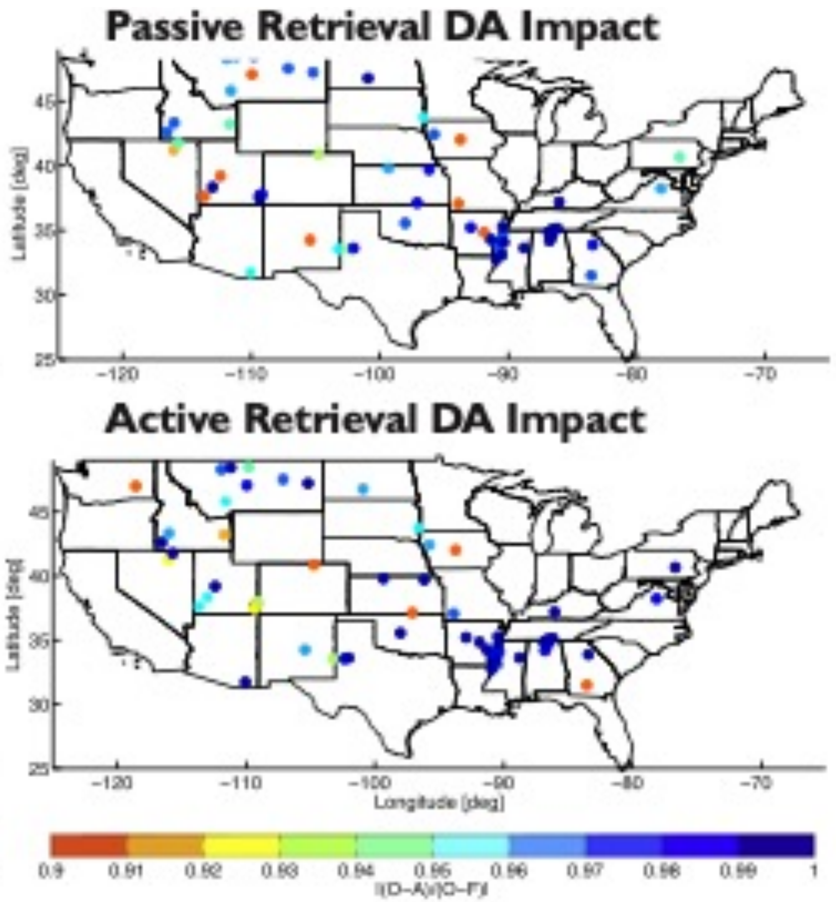
← Separate retrieval DA is better

Joint retrieval DA is better →



← Single retrieval is better

Joint retrieval is better →



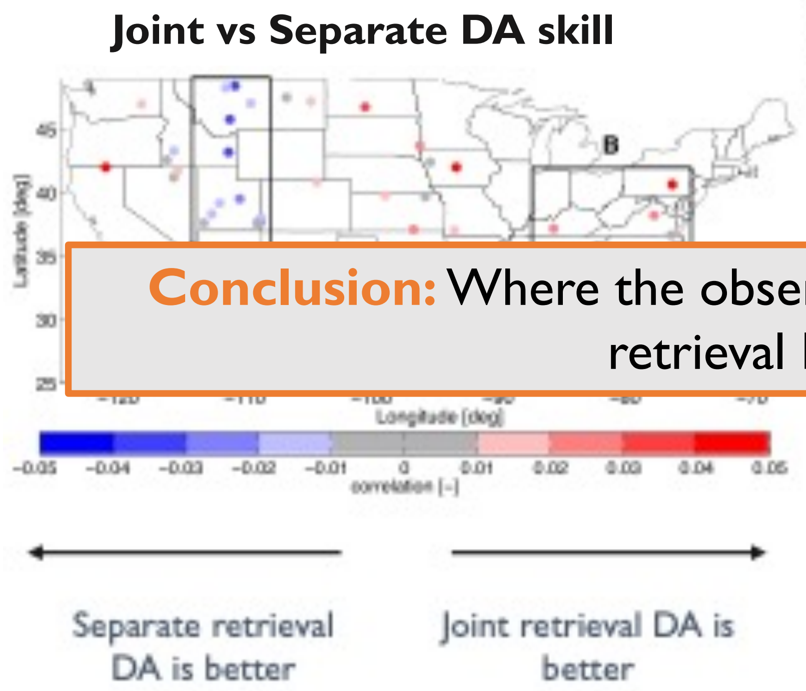
← Higher obs impact

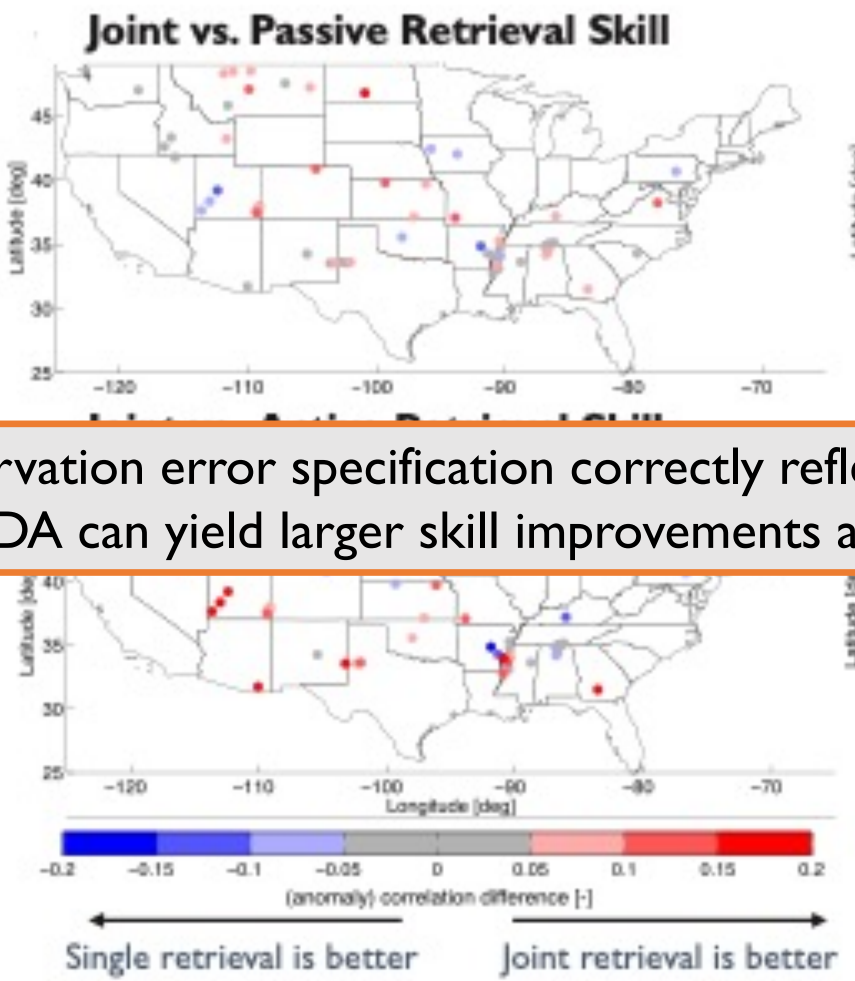
Lower obs impact →

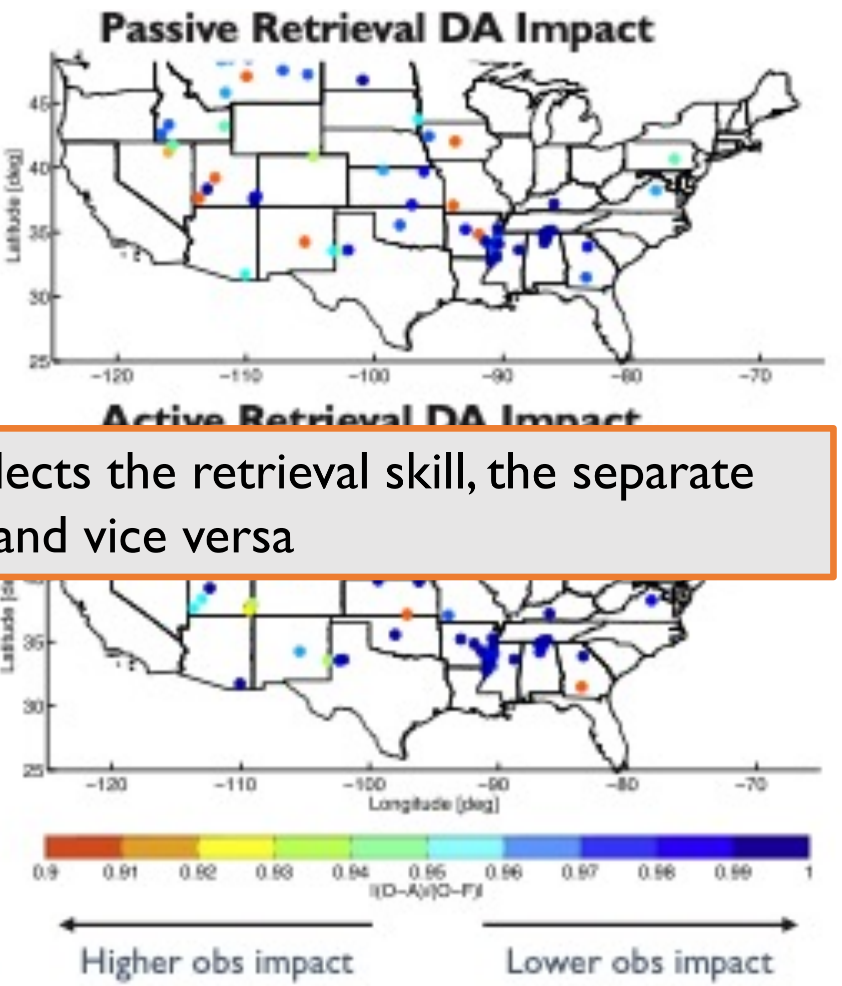
Multi-sensor satellite observations: Data Assimilation

Q: For multi-sensor data assimilation does it matter how the joint active/passive information is provided?

Kolassa et al., 2017







Conclusion: Where the observation error specification correctly reflects the retrieval skill, the separate retrieval DA can yield larger skill improvements and vice versa

Increasing data assimilation efficiency

Increasing data assimilation efficiency

- Land data assimilation requires bias-correction (rescaling) of observations prior to assimilation
    1. Assumption of DA system
    2. Different nature of observed and modeled variables (e.g., *Koster et al., 2009* for SM)

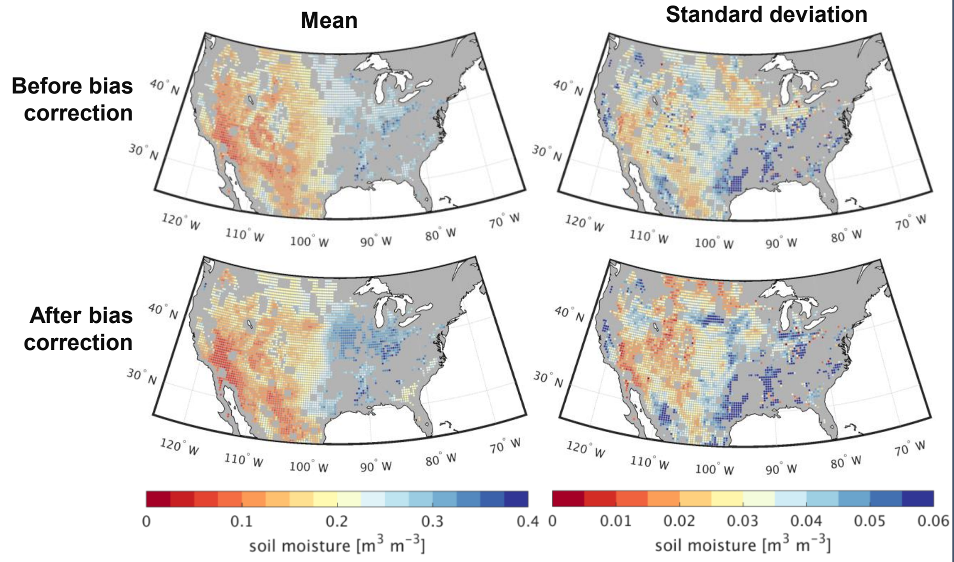


Fig 5: Effect of CDF-matching

Increasing data assimilation efficiency

- Land data assimilation requires bias-correction (rescaling) of observations prior to assimilation
  1. Assumption of DA system
  2. Different nature of observed and modeled variables (e.g., *Koster et al., 2009* for SM)

- Rescaling limits us to correcting random errors
  → Small fraction of independent satellite information is used
  → Limits efficiency of data assimilation (*Nearing et al, 2018, Kumar et al. 2012*)

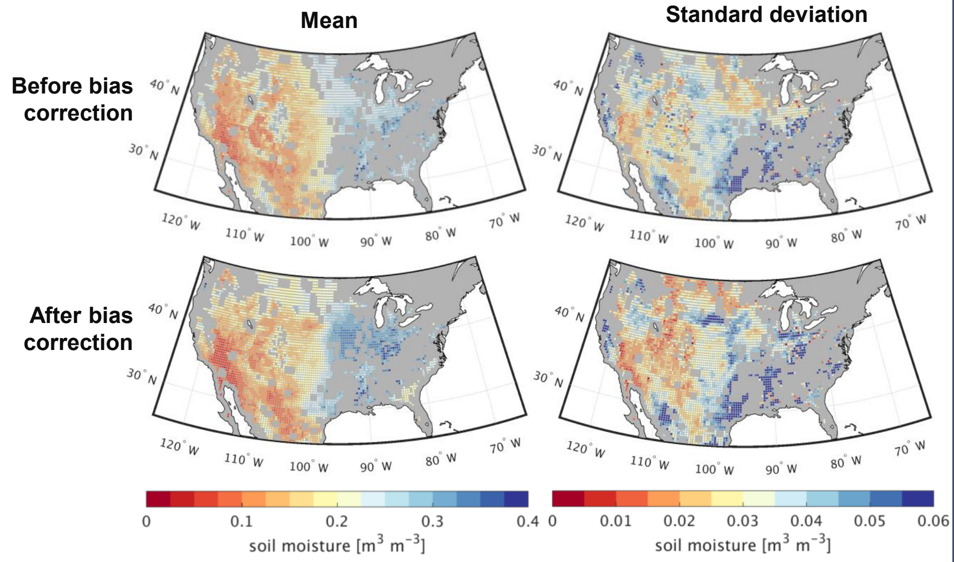


Fig 5: Effect of CDF-matching

Increasing data assimilation efficiency

- Land data assimilation requires bias-correction (rescaling) of observations prior to assimilation
  1. Assumption of DA system
  2. Different nature of observed and modeled variables (e.g., *Koster et al., 2009* for SM)

- Rescaling limits us to correcting random errors
  → Small fraction of independent satellite information is used
  → Limits efficiency of data assimilation (*Nearing et al, 2018, Kumar et al. 2012*)

- From *Nearing et al., 2018*

Total information from model and retrievals[a]	$\frac{I(Z;X,Y)}{H(Z)}$	0.18
Fraction of retrieval information lost via CDF-matching	$1-\frac{I\left(Z;Y^{\mathrm{CDF}}\mid X\right)}{I(Z;Y\mid X)}$	0.11

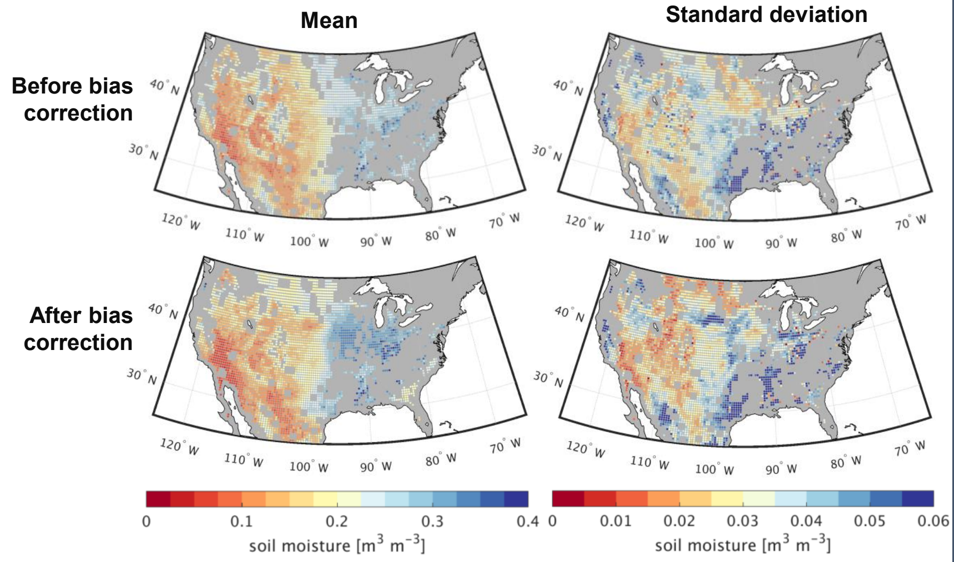


Fig 5: Effect of CDF-matching

Increasing data assimilation efficiency

- Land data assimilation requires bias-correction (rescaling) of observations prior to assimilation
  1. Assumption of DA system
  2. Different nature of observed and modeled variables (e.g., *Koster et al., 2009* for SM)

- Rescaling limits us to correcting random errors
  → Small fraction of independent satellite information is used
  → Limits efficiency of data assimilation (*Nearing et al, 2018, Kumar et al. 2012*)

- From *Nearing et al., 2018*

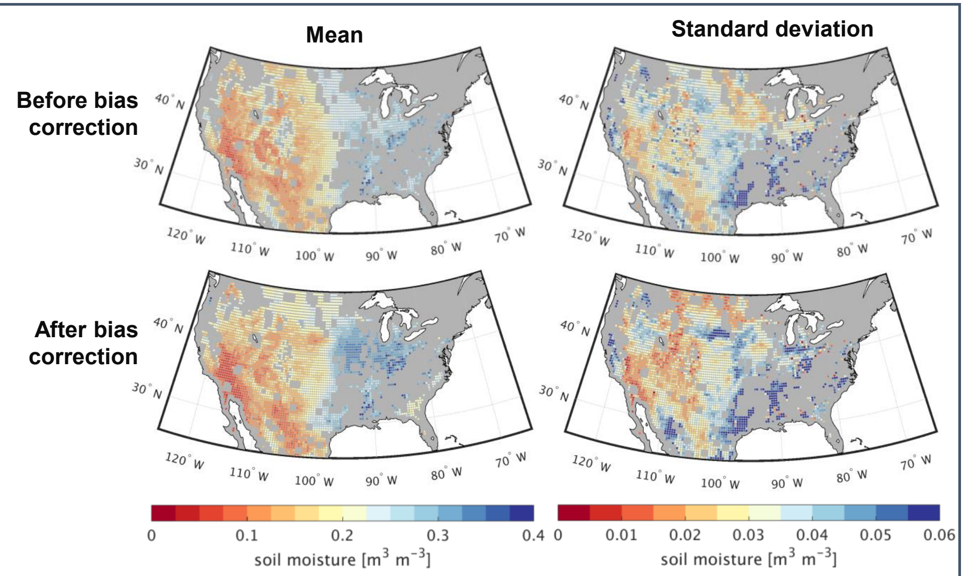


Fig 4: Effect of CDF-matching

Total information from model and retrievals[a]	$\frac{I(Z;X,Y)}{H(Z)}$	0.18
Fraction of retrieval information lost via CDF-matching	$1-\frac{I\left(Z;Y^{CDF}\middle\|X\right)}{I(Z;Y\|X)}$	0.11

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

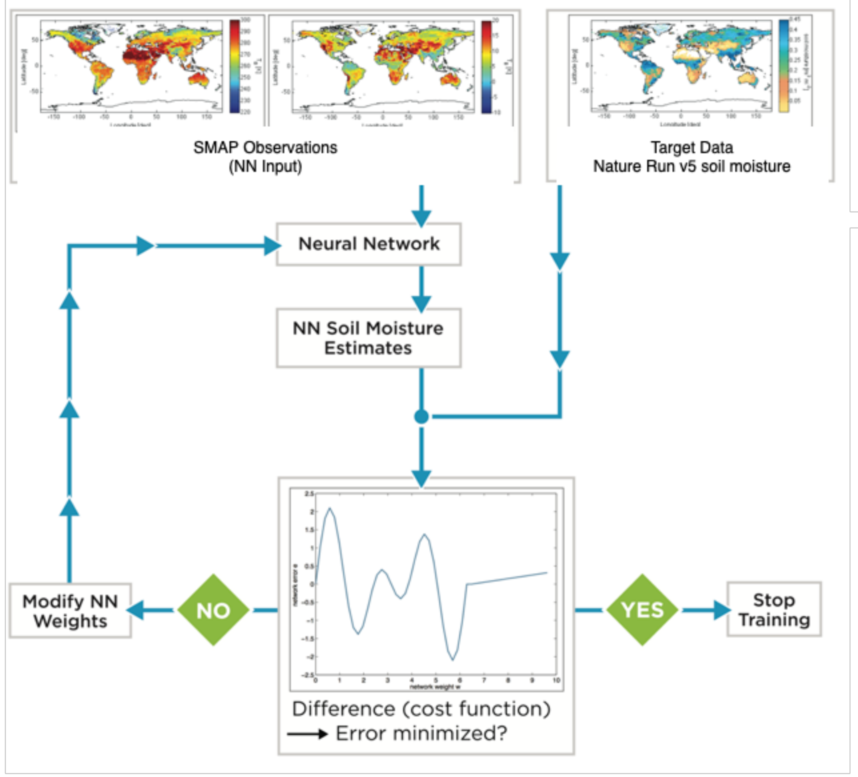


Fig 5: Schematic of NN training process
Kolassa et al., 2018

- Neural Network (NN) trained with SMAP observation inputs and modeled SM outputs

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

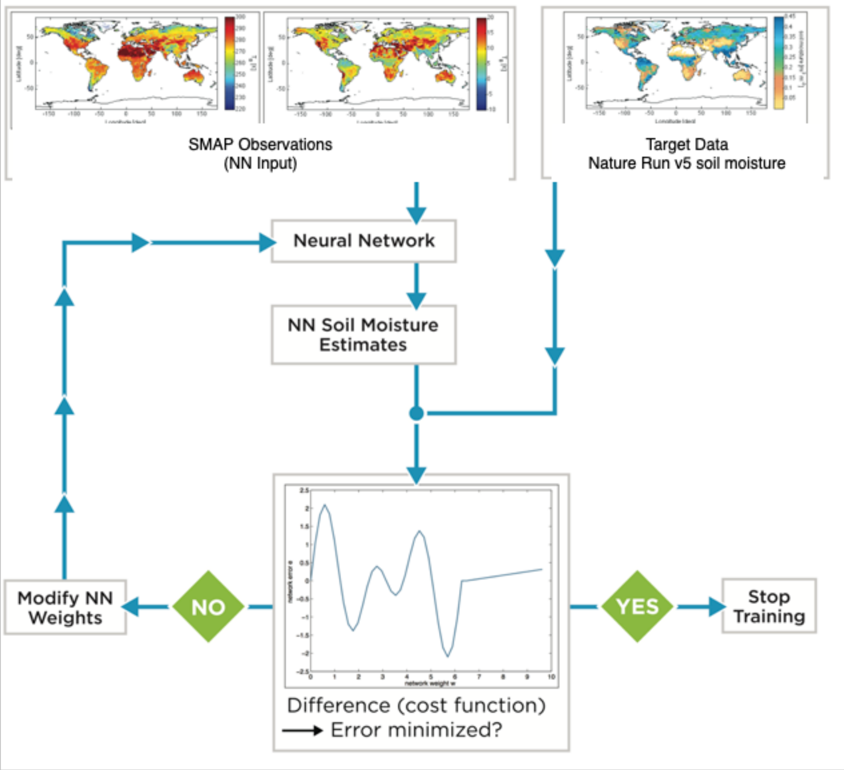


Fig 5: Schematic of NN training process
Kolassa et al., 2018

- Neural Network (NN) trained with SMAP observation inputs and modeled SM outputs

- NN output is expressed in target data climatology → globally unbiased, same dynamic range
- Spatial and temporal patterns of output are driven by satellite observations

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

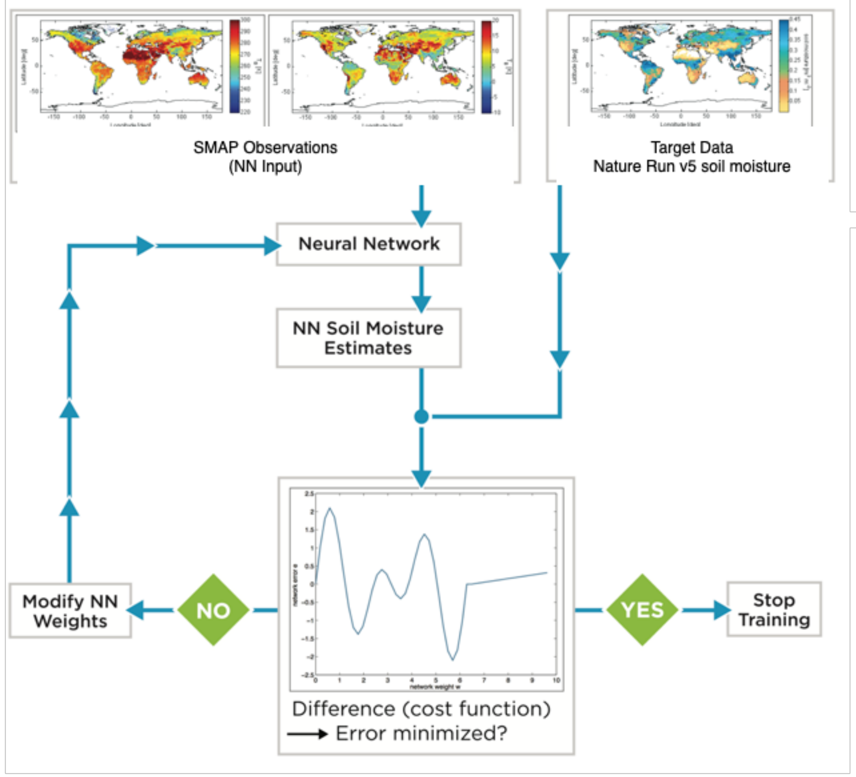


Fig 5: Schematic of NN training process
Kolassa et al., 2018

- Neural Network (NN) trained with SMAP observation inputs and modeled SM outputs

- NN output is expressed in target data climatology → globally unbiased, same dynamic range
- Spatial and temporal patterns of output are driven by satellite observations

SMAP NN DA: NN retrieval assimilation **without further bias correction**

SMAP NN DA + CDF: NN retrieval assimilation **with localized CDF-matching**

OL: Open loop; model run without assimilation

Kolassa et al., 2017b

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

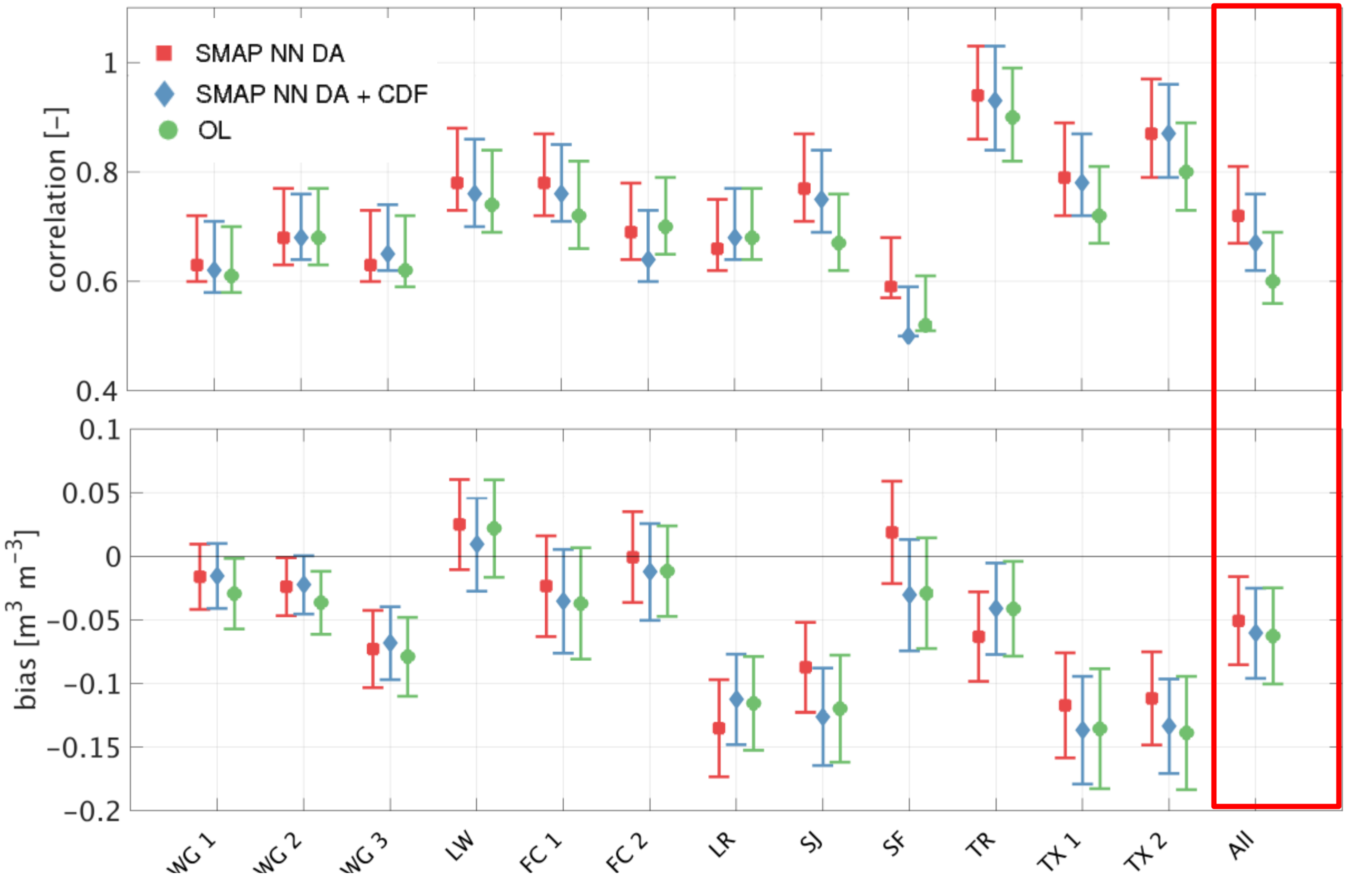


Fig 6: Skill improvements at SMAP core validation sites

Conclusion: On average, global bias correction yields slightly higher analysis skill than local bias correction

Kolassa et al., 2017b

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

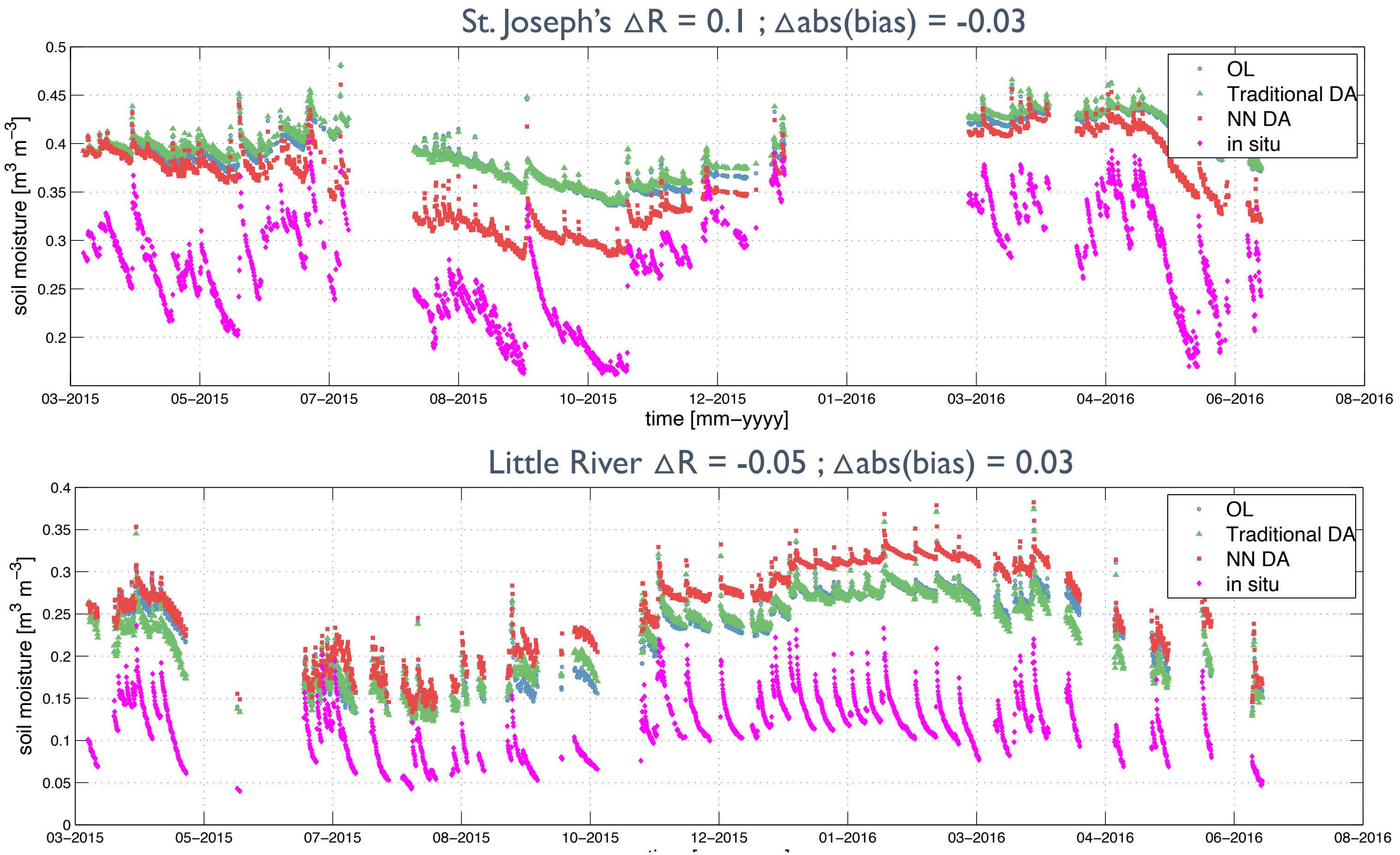


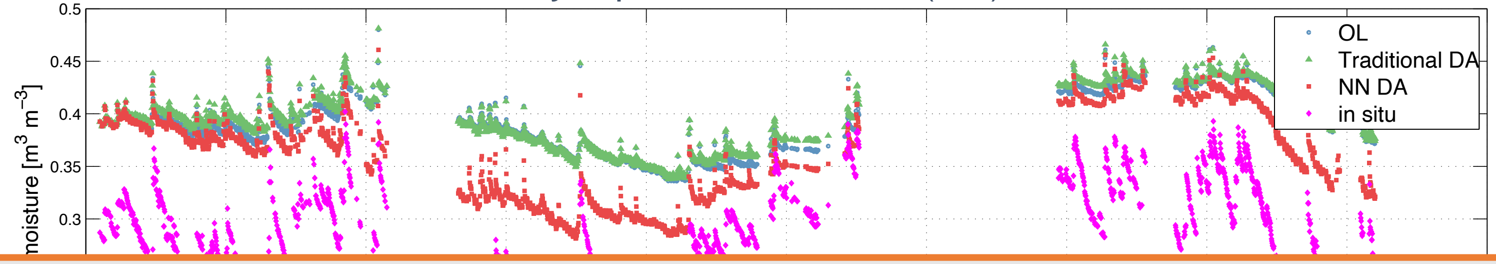
Fig 7: Experiment time series at St. Joseph and Little River stations

Kolassa et al., 2017b

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

St. Joseph's $\Delta R = 0.1$; Δabs(bias) = -0.03



Conclusions:

- NN DA is better able to capitalize on independent information provided by SMAP
- NN DA is also more susceptible to retrieval errors

→ Accurate observation error characterization is critical

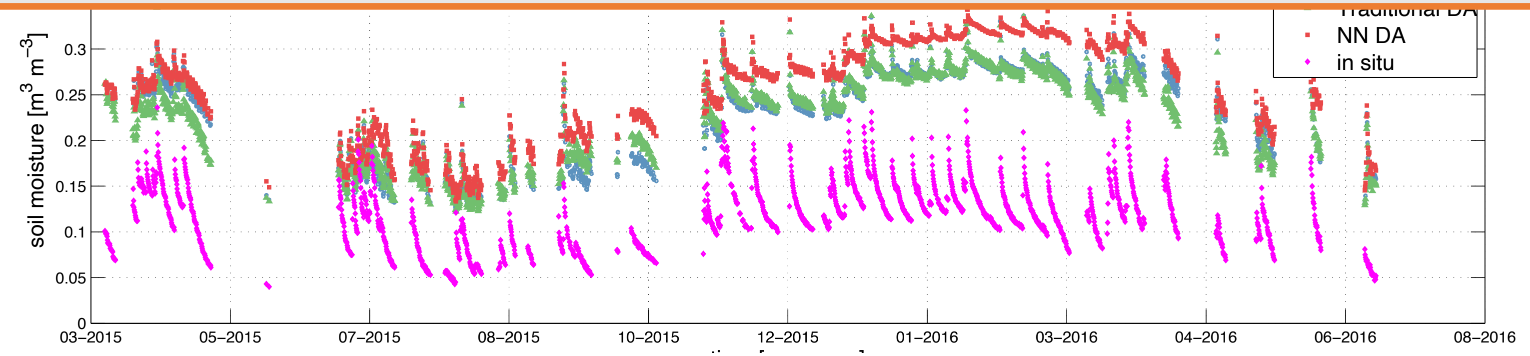


Fig 7: Experiment time series at St. Joseph and Little River stations

Kolassa et al., 2017b

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

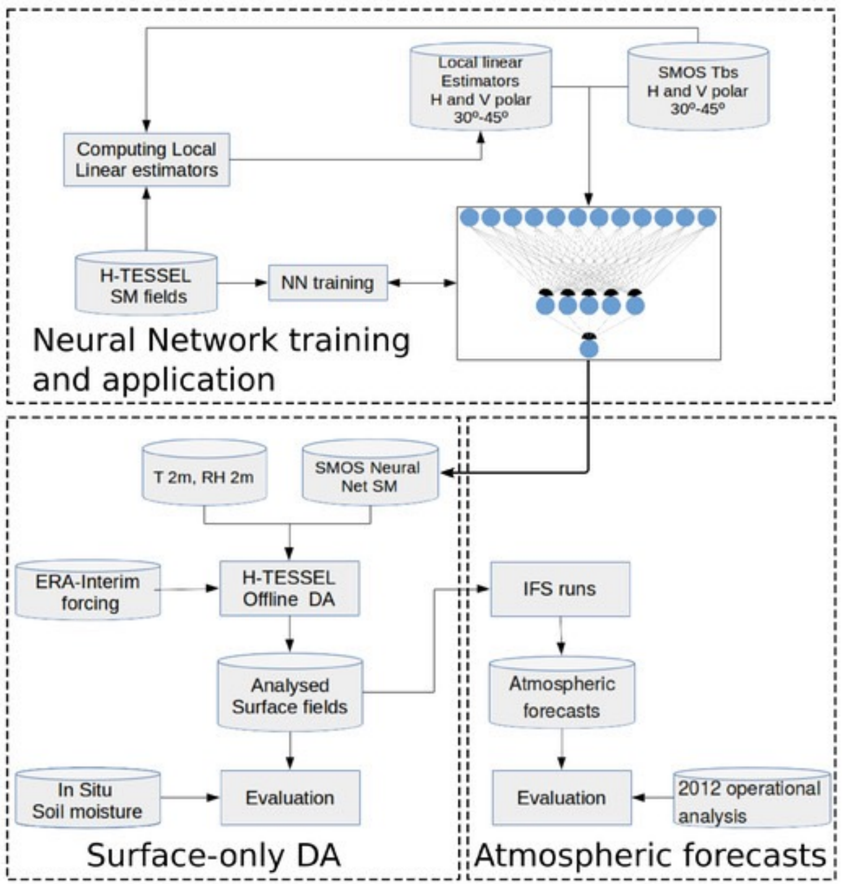


Fig 8: SMOS NN assimilation procedure

- *Rodriguez-Fernandez et al., 2019*: NN SM assimilation with spatially varying errors
    - Neutral to positive improvements in SM
    - Improvements of atmospheric forecasts in Southern Hemisphere

Increasing data assimilation efficiency

Q: How can we use more of the independent information provided by satellite observations while respecting the need for (some) bias correction?

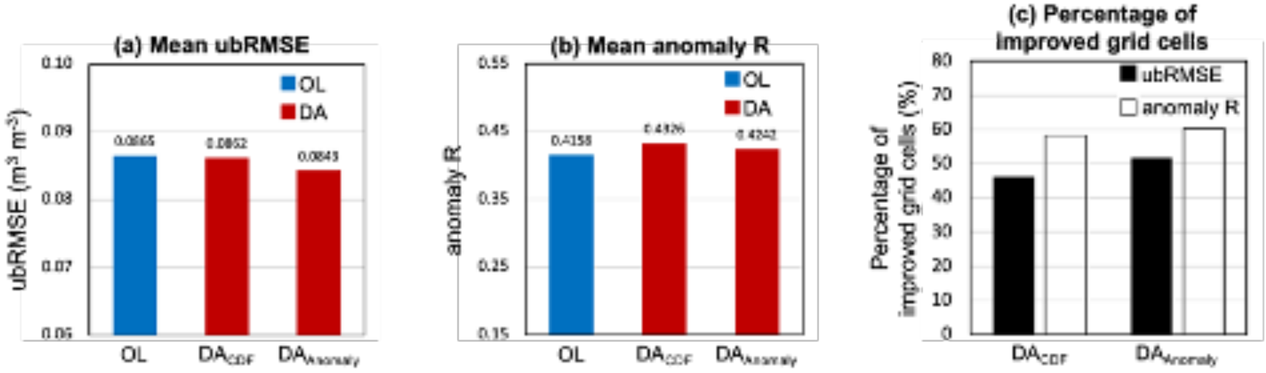


Fig 9: Average skill of two DA approaches at in situ stations

- *Kwon, Kumar et al., in review*: assimilate SMAP anomalies directly
- The anomaly DA provides comparable performance to CDF-matching and is more effective in incorporating unmodeled features such as irrigation

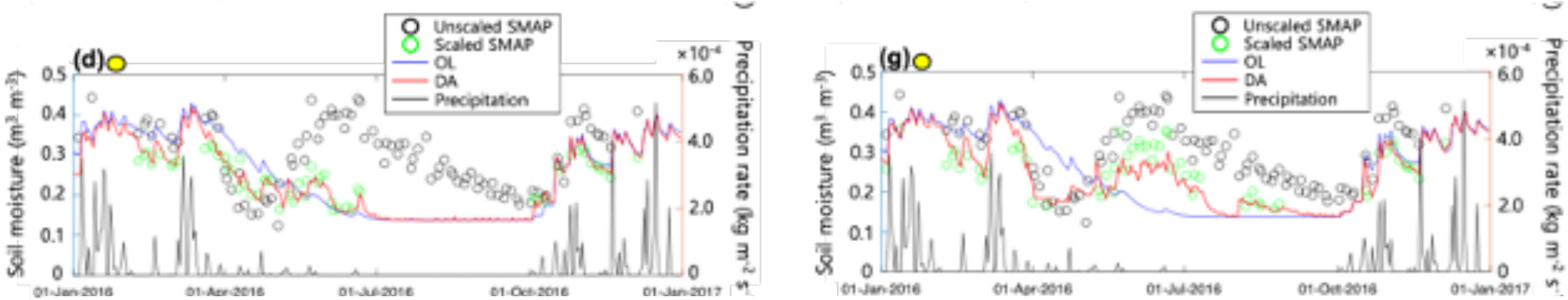


Fig 10: Comparison of two DA approaches at location with irrigation

Land observations for model calibration

Model calibration

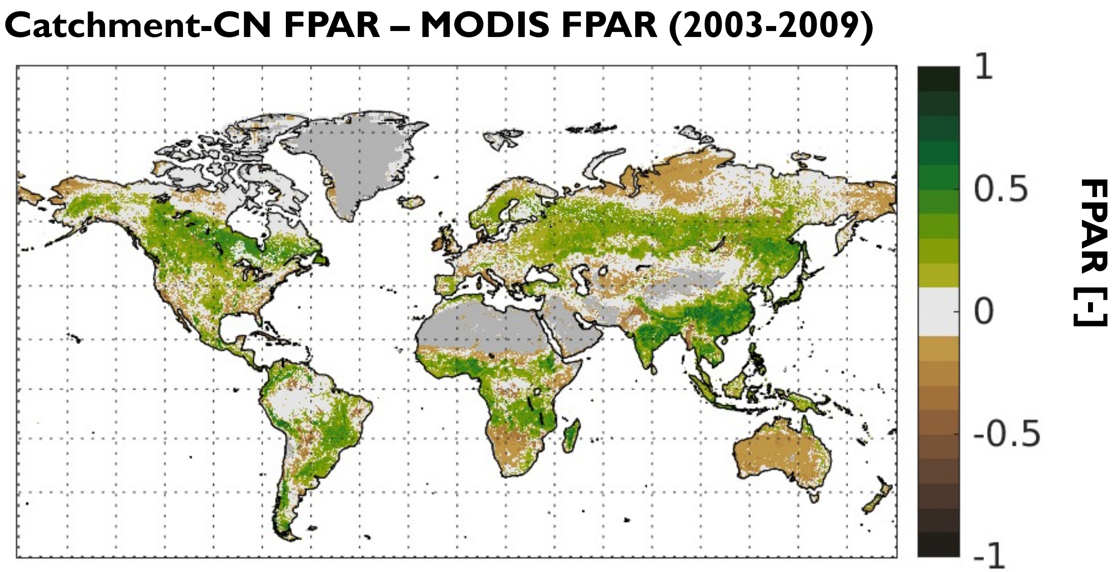


Fig 11: Catchment-CN FPAR bias

Model calibration

- Calibration parameters:
  - Timing of phenological cycle (seasonal variability)
  - Photosynthetic efficiency (bias)
  - Carbon storage/allocation (interannual variability)

Catchment-CN FPAR – MODIS FPAR (2003-2009)

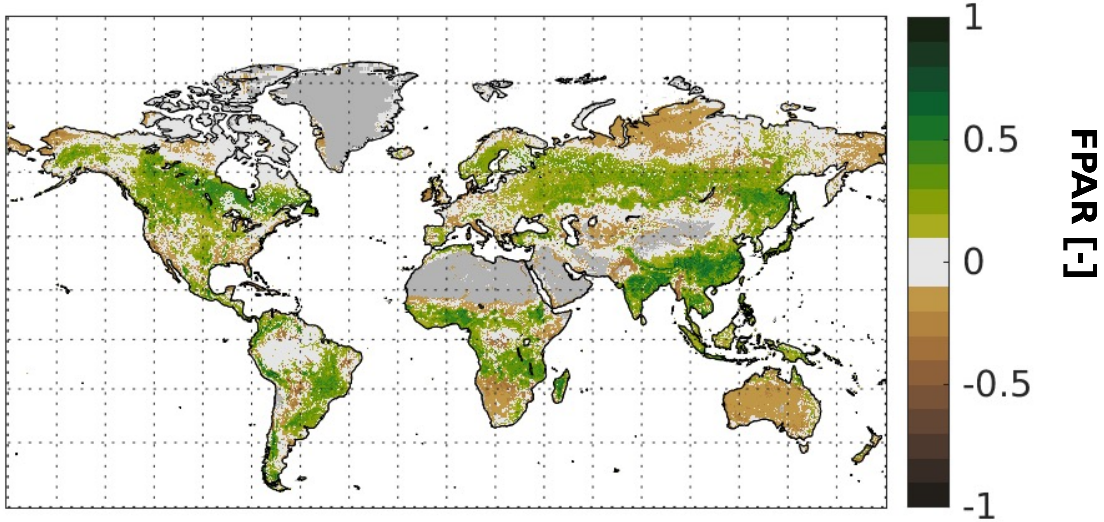


Fig 11: Catchment-CN FPAR bias

- Calibration approach:
  - Calibration period: 2003 – 2010
  - Cost function: FPAR RMSE.
  - Particle swarm (ensemble-based) optimization ***at select calibration locations***
  - Separate parameters for each Plant Functional Type (PFT)

Model calibration: Impact on modeled FPAR

- Global model simulation with new vegetation parameters evaluated against MODIS FPAR

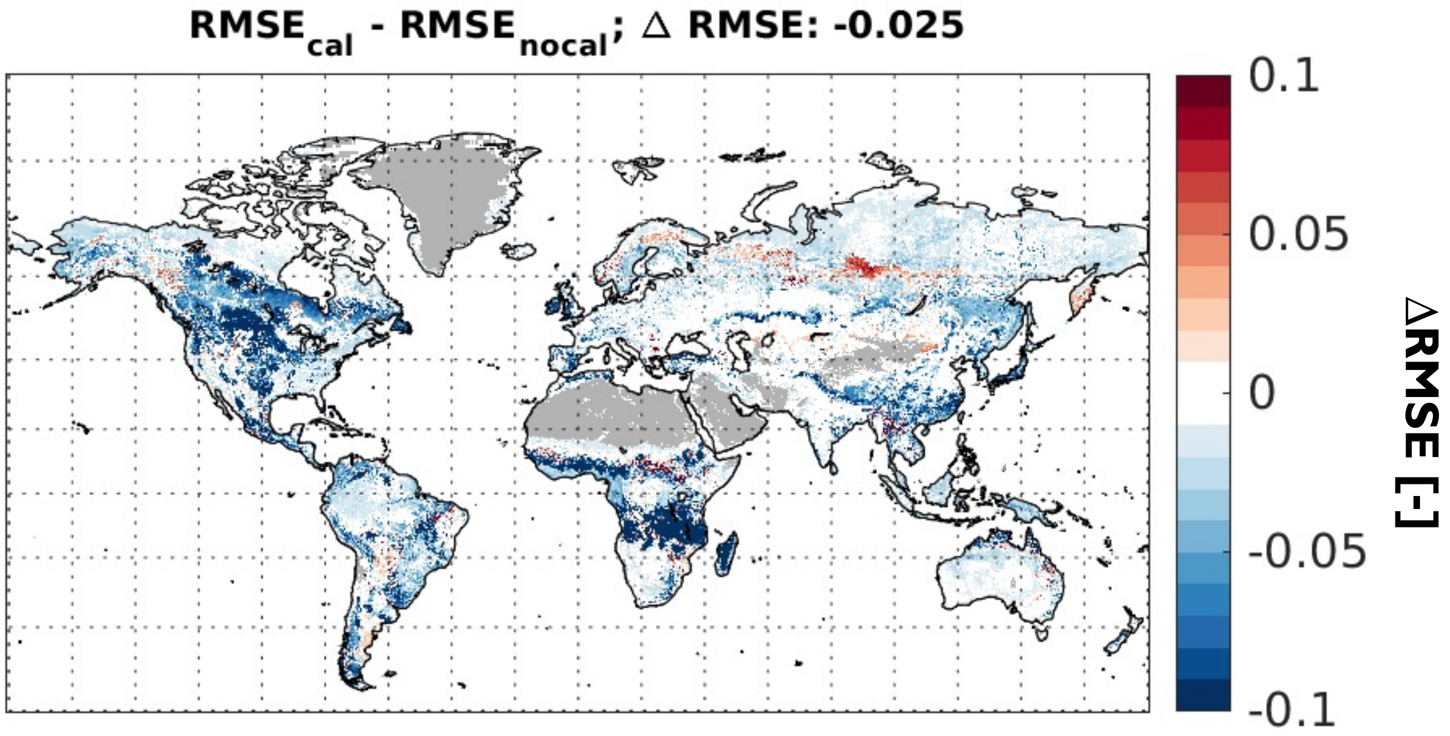


Fig 12: FPAR RMSE change resulting from calibration

Conclusion : Parameter estimation consistently reduces model RMSE with respect to MODIS FPAR

Kolassa et al., 2020

Model calibration: Impact of error distribution

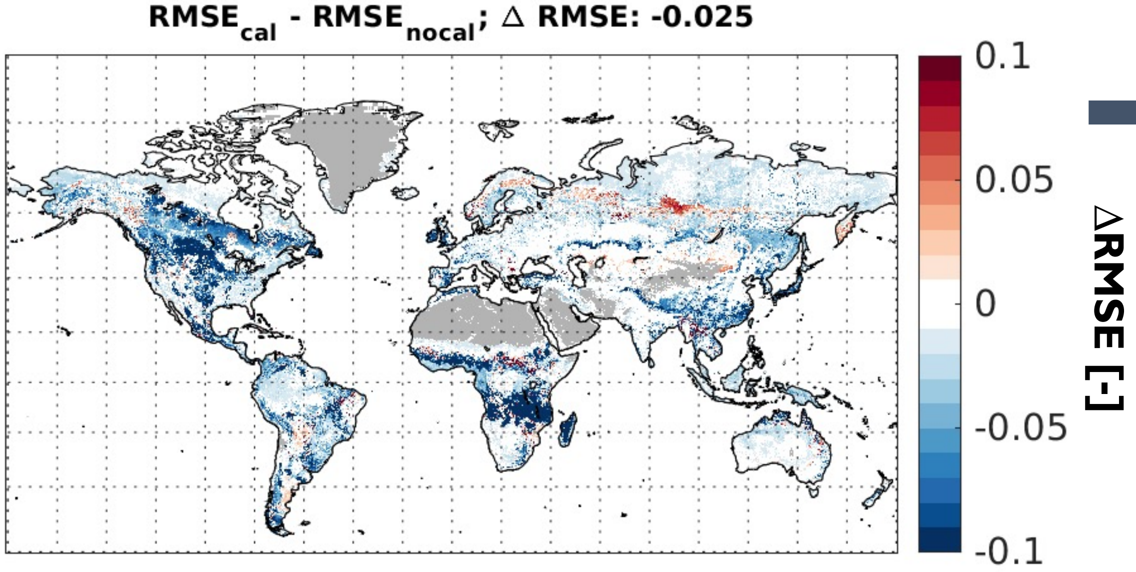
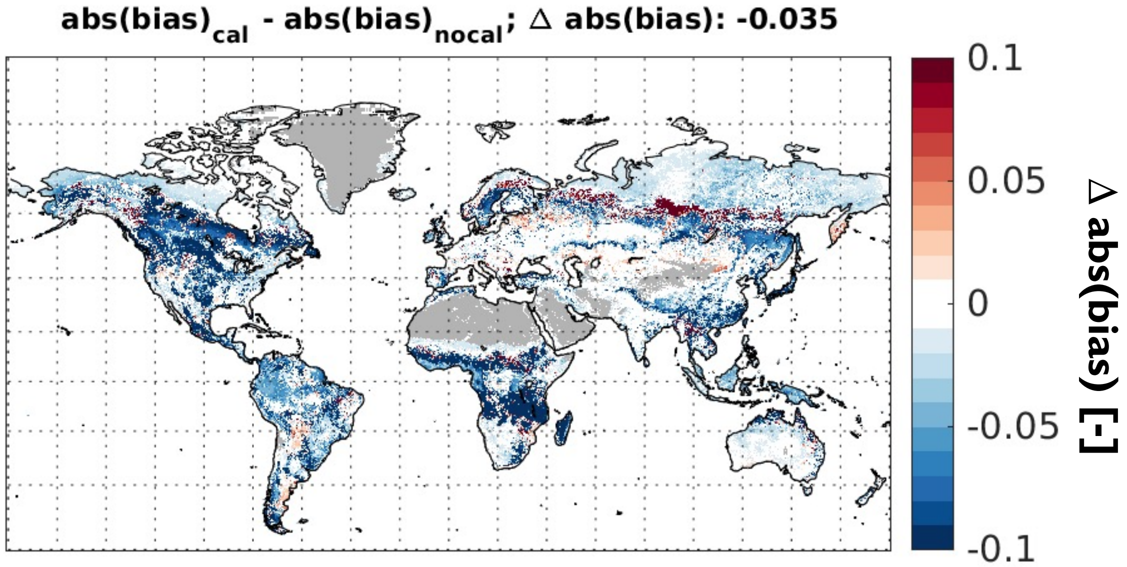
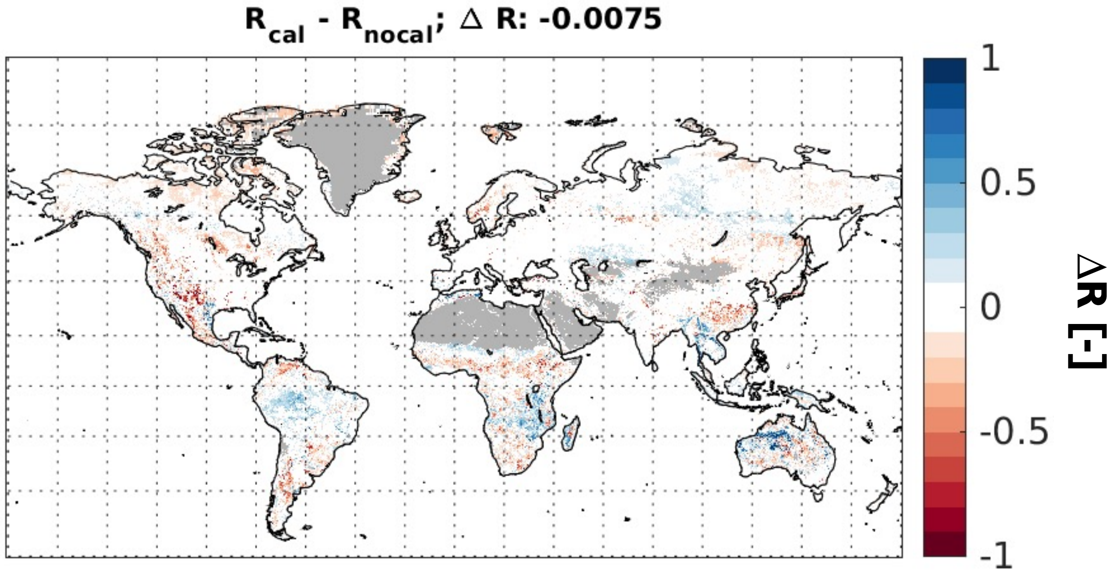


Fig 13: RMSE change breakdown



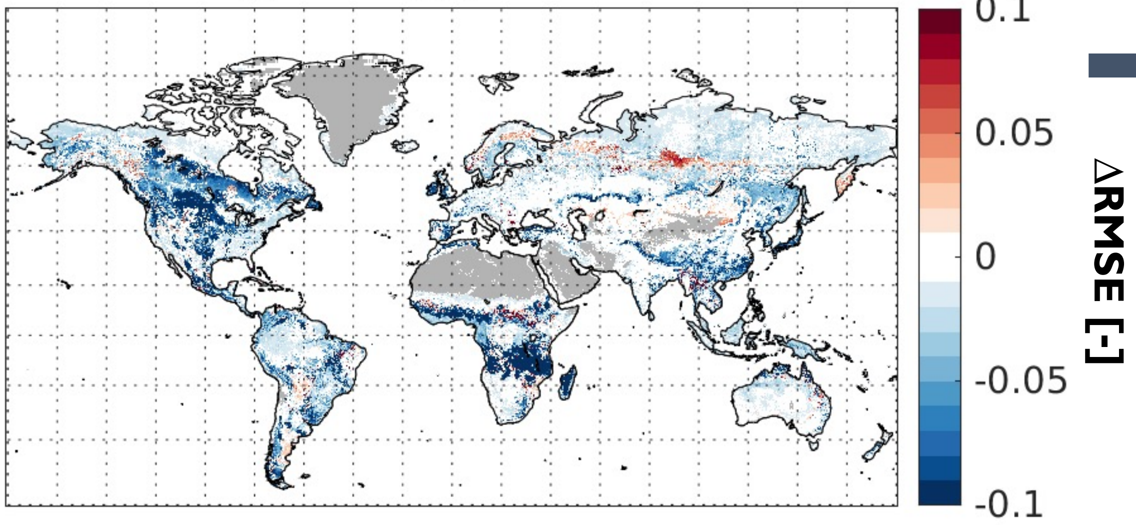


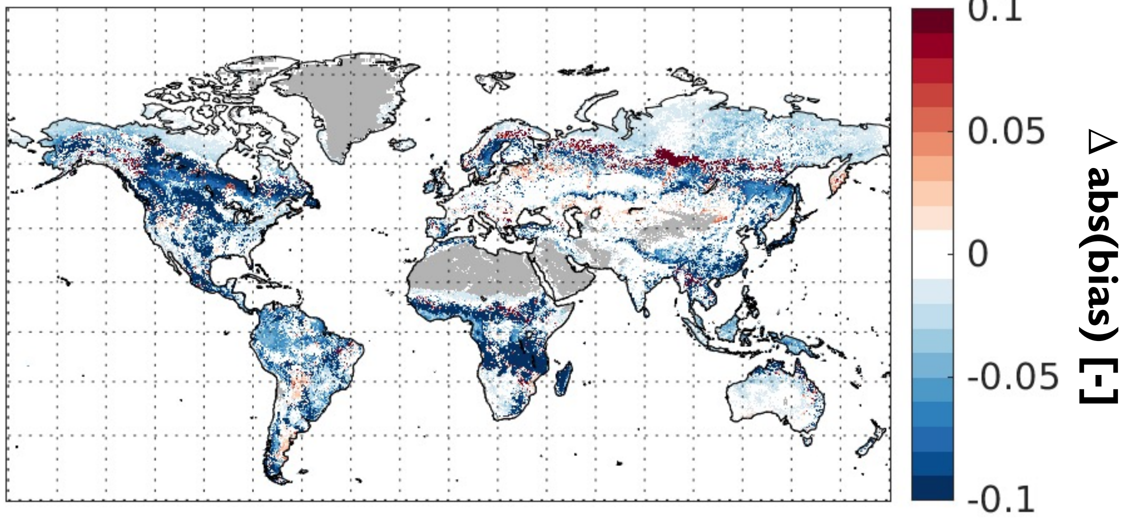
- Reduction in RMSE is driven by bias reduction
- Dominance of bias in model error skews calibration towards efficiency parameters

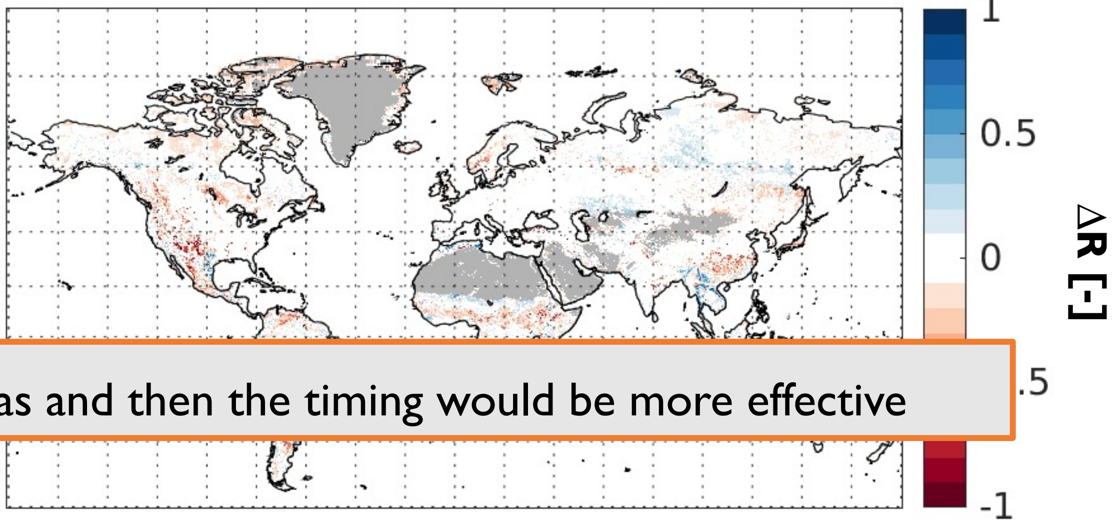
Kolassa et al., 2020

Model calibration: Impact of error distribution

Fig 13: RMSE change breakdown







- Reduction in RMSE is driven by bias reduction
- Dominance of bias in model error skews calibration towards efficiency parameters

Conclusion : Two-stage calibration to address first the bias and then the timing would be more effective

Kolassa et al., 2020

Model Calibration: Impact of structural errors

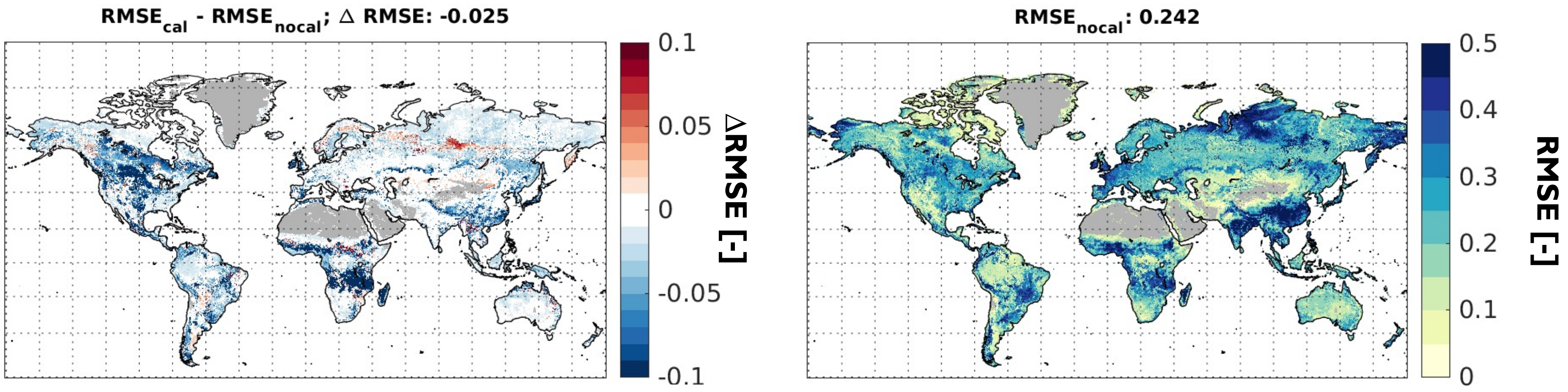


Fig 15: RMSE change relative to total RMSE

Model Calibration: Impact of structural errors

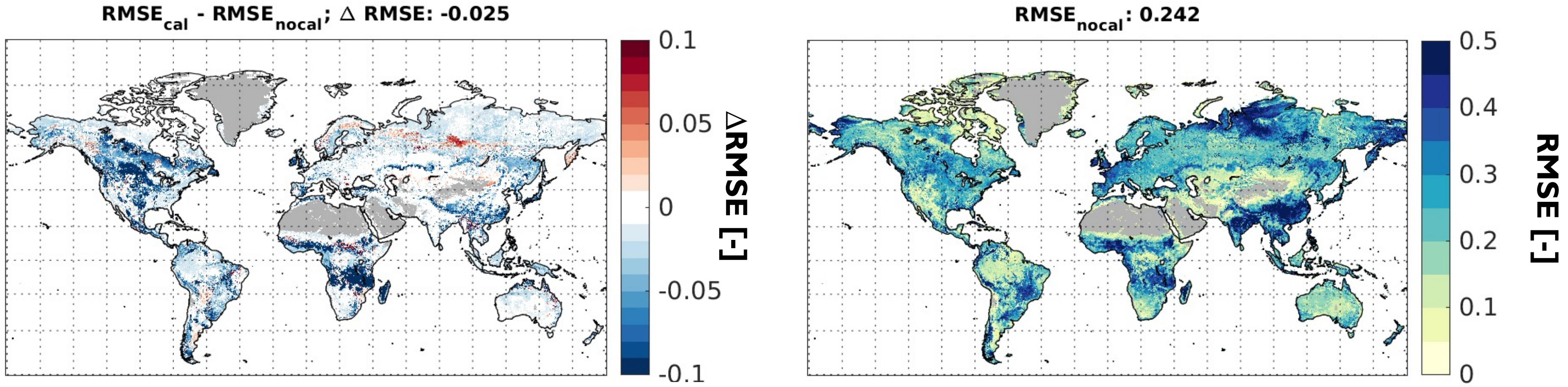
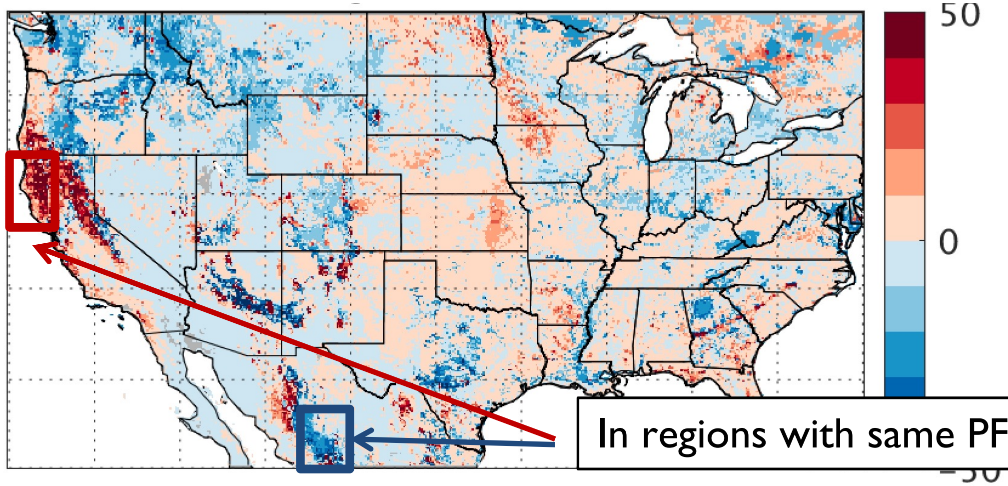


Fig 15: RMSE change relative to total RMSE

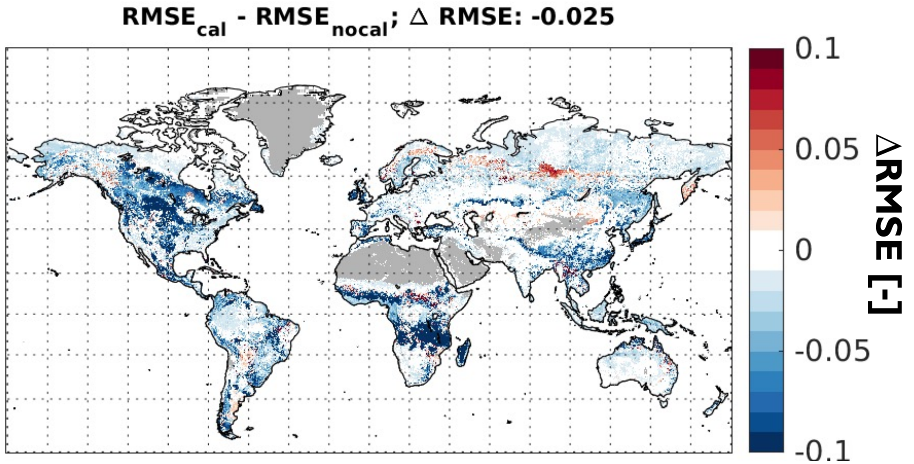
Fig 16: Initial ΔRMSE (calibrated – uncalibrated)

avg. error reduction ~5%



In regions with same PFT, calibrated parameters have very different impact.

Model Calibration: Impact of structural errors



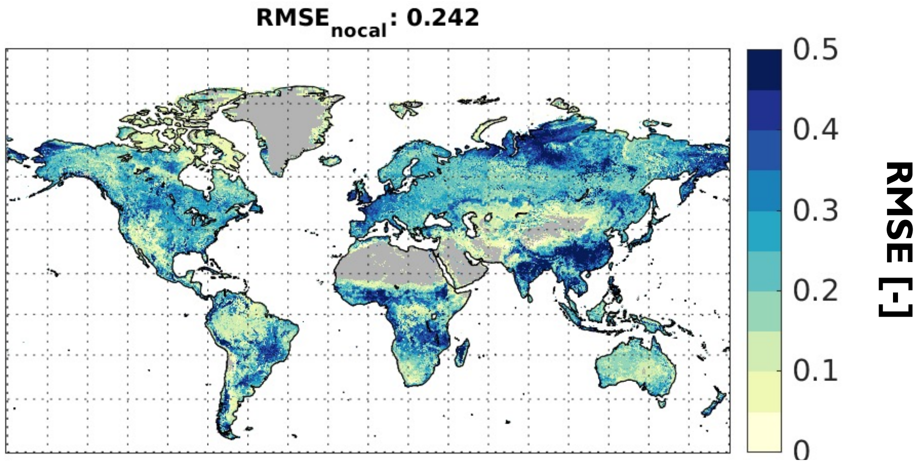
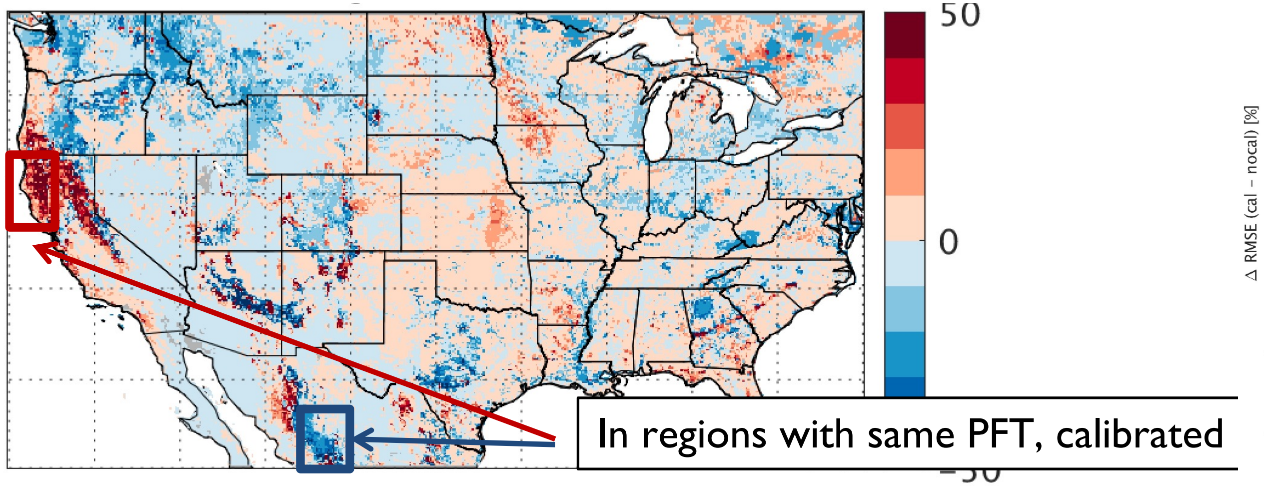


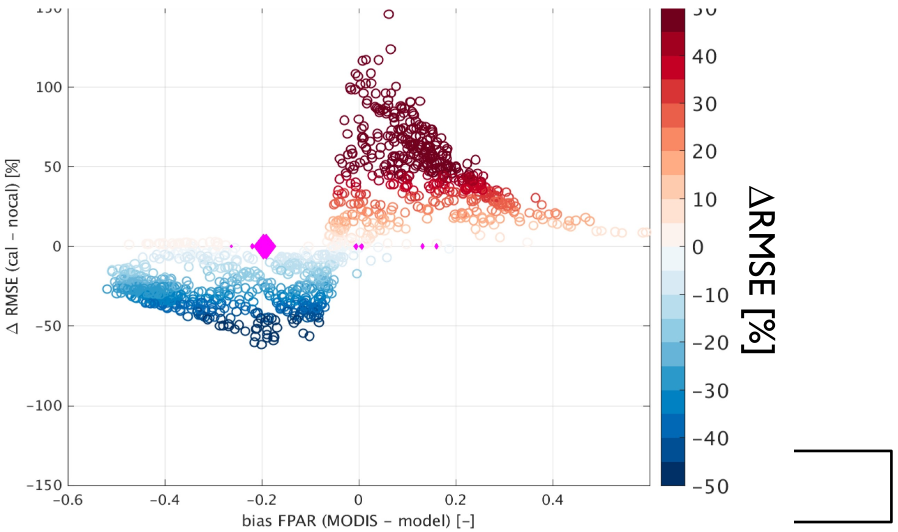
Fig 15: RMSE change relative to total

Fig 17: original bias vs. impact of updated parameters

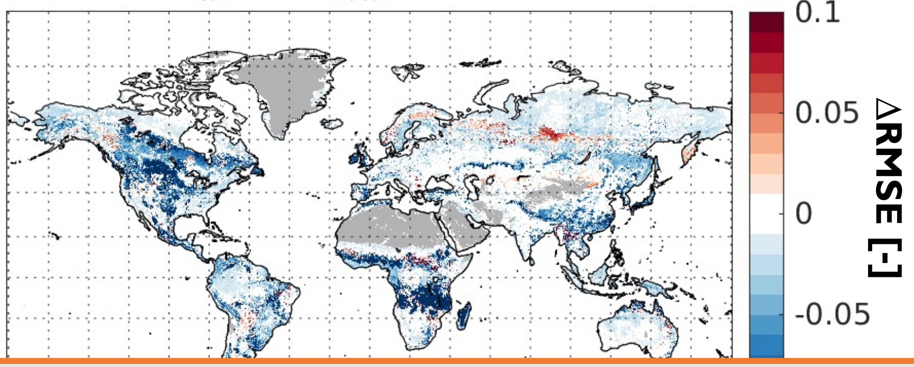
Fig 16: Initial ΔRMSE (calibrated – uncalibrated)

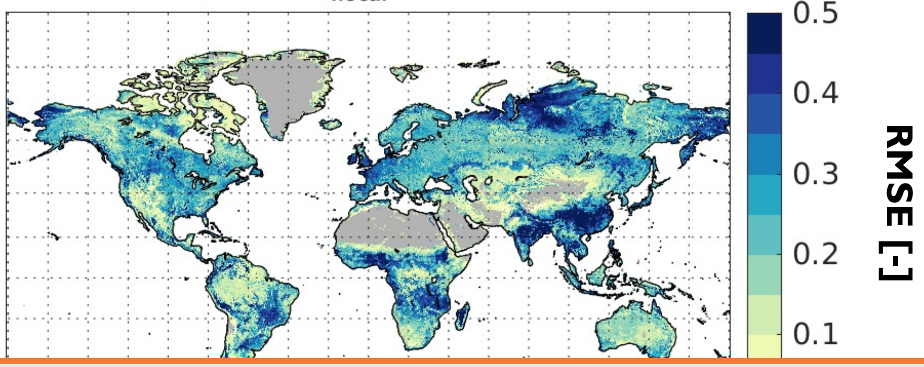
avg. error reduction ~5%





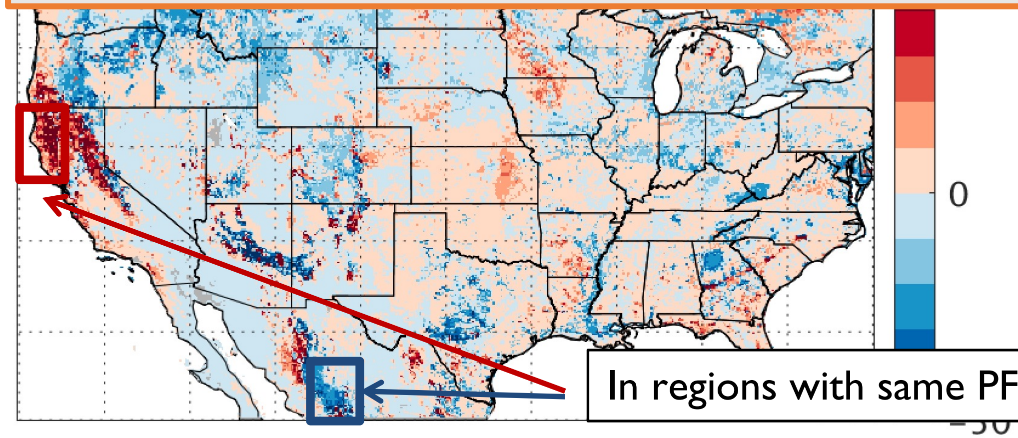
Model Calibration: Impact of structural errors



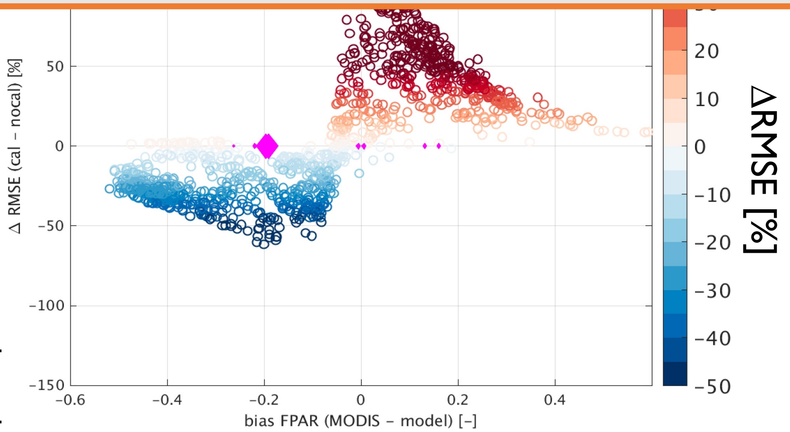


Conclusion :

- Model structural errors are 'aliased' onto parameters during calibration
- Remaining errors are due to errors in model structure



In regions with same PFT, calibrated



Model Calibration: Impact of structural errors

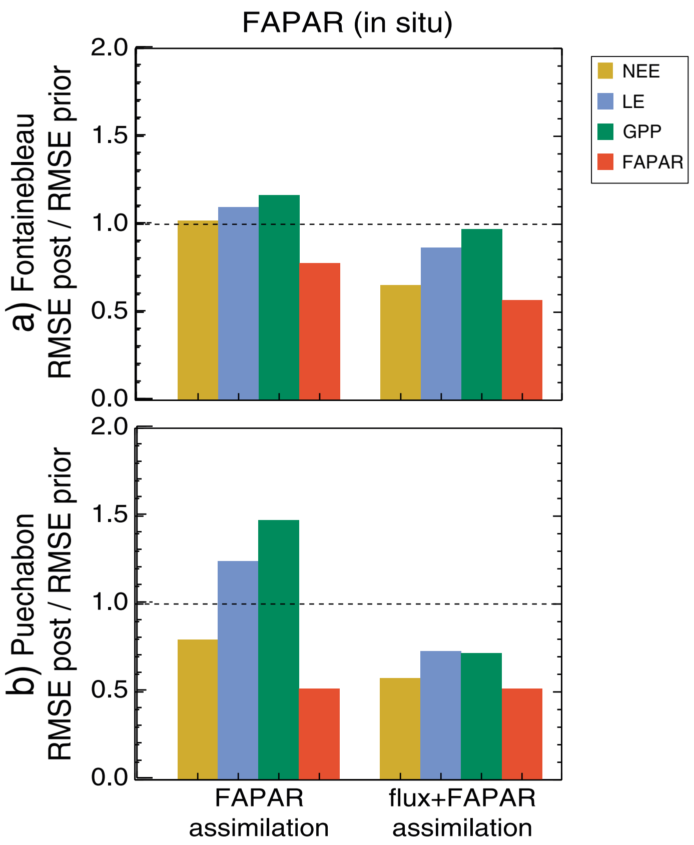


Fig 18: FPAR Calibration Impact on vegetation fluxes

Bacour et al., 2015: parameter estimation with FPAR alone improved FPAR skill, degraded skill of other vegetation variables

Model Calibration: Impact of structural errors

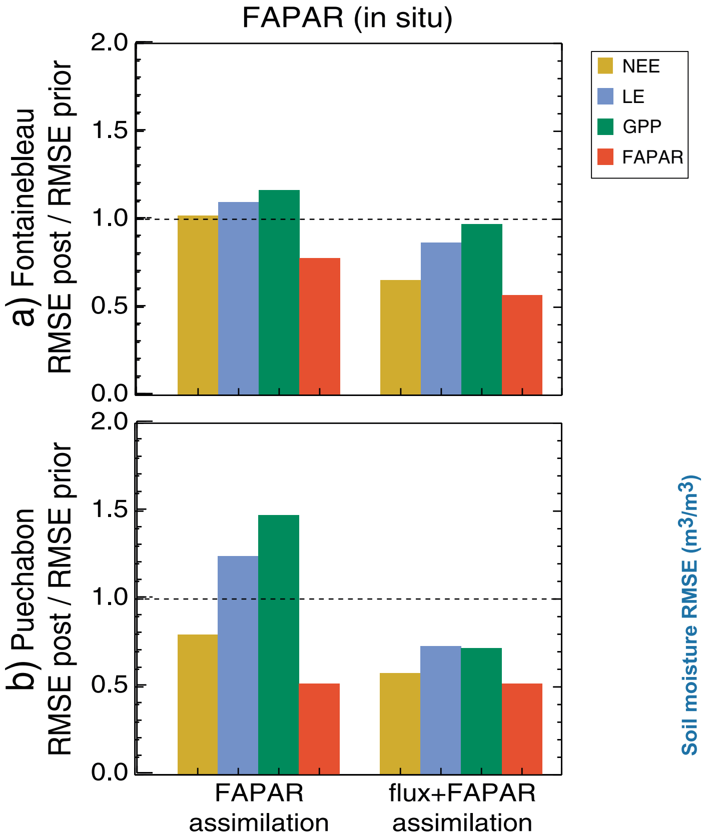


Fig 18: FPAR Calibration Impact on vegetation fluxes

Bacour et al., 2015: parameter estimation with FPAR alone improved FPAR skill, degraded skill of other vegetation variables

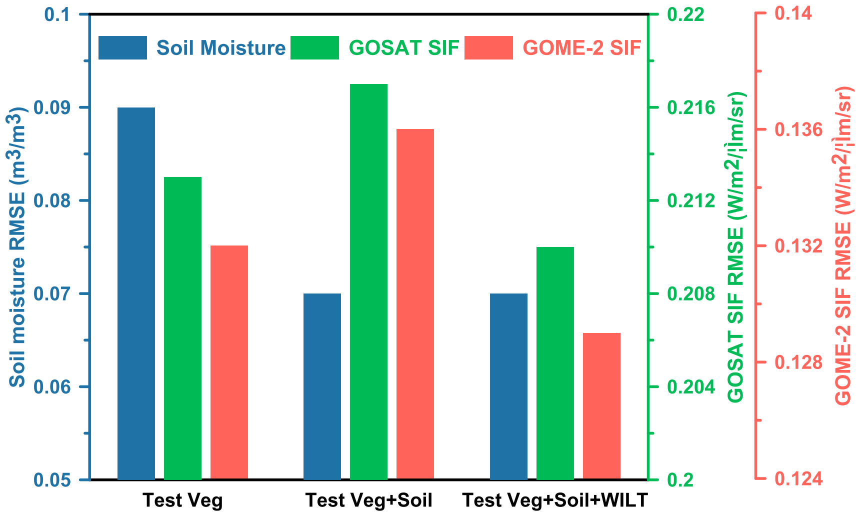


Fig 19: Parameter calibration impact on hydrology and vegetation

Qiu et al., 2017: hydraulic parameter calibration alone improved soil moisture skill, but degrades modeled SIF

Impact of land observations in NWP systems

Impact of land observations in NWP systems

- Impact of assimilating land observations on forecast skill in NWP systems is often positive, but very small (e.g., Draper et al., 2012, Rodriguez-Fernandez et al. 2019, Carrera et al., Reichle et al. in review)

  1. Land influence on atmosphere is not universal
  2. Impact of single new data type in system that assimilates ~10^6 observations

Impact of land observations in NWP systems

- Impact of assimilating land observations on forecast skill in NWP systems is often positive, but very small (e.g., Draper et al., 2012, Rodriguez-Fernandez et al., Carrera et al. 2019, Reichle et al. in review)

    1. Land influence on atmosphere is not universal
    2. Impact of single new data type in system that assimilates ~10^6 observations

- In certain conditions land influence on atmosphere is significantly more important than usual : extreme conditions
    1. Droughts and heatwaves (e.g., Seneviratne et al., 2015)
    2. Tropical Cyclones

Impact of land observations in NWP systems

- Impact of assimilating land observations on forecast skill in NWP systems is often positive, but very small (e.g., Draper et al., 2012, Rodriguez-Fernandez et al., Carrera et al. 2019, Reichle et al. in review)

  1. Land influence on atmosphere is not universal
  2. Impact of single new data type in system that assimilates ~10^6 observations

- In certain conditions land influence on atmosphere is significantly more important than usual : extreme conditions
  1. Droughts and heatwaves (e.g., Seneviratne et al., 2015)
  2. Tropical Cyclones

- Very wet land surface → help to sustain or re-intensify TC ("Brown Ocean Effect")
- Dry land surface → faster TC dissipation
- Soil moisture gradients → different TC over-land track

- SMAP data assimilation → better land surface initial conditions → better TC forecasts → societal benefit

Q: Can the assimilation of SMAP observations into a global numerical weather prediction (NWP) model improve the prediction of tropical cyclone (TC) evolution prior to and after landfall?

Impact of land observations in NWP systems

Observing System Experiment to determine the potential of SMAP data assimilation to improve forecasts of tropical cyclone structure and precipitation surrounding landfall.

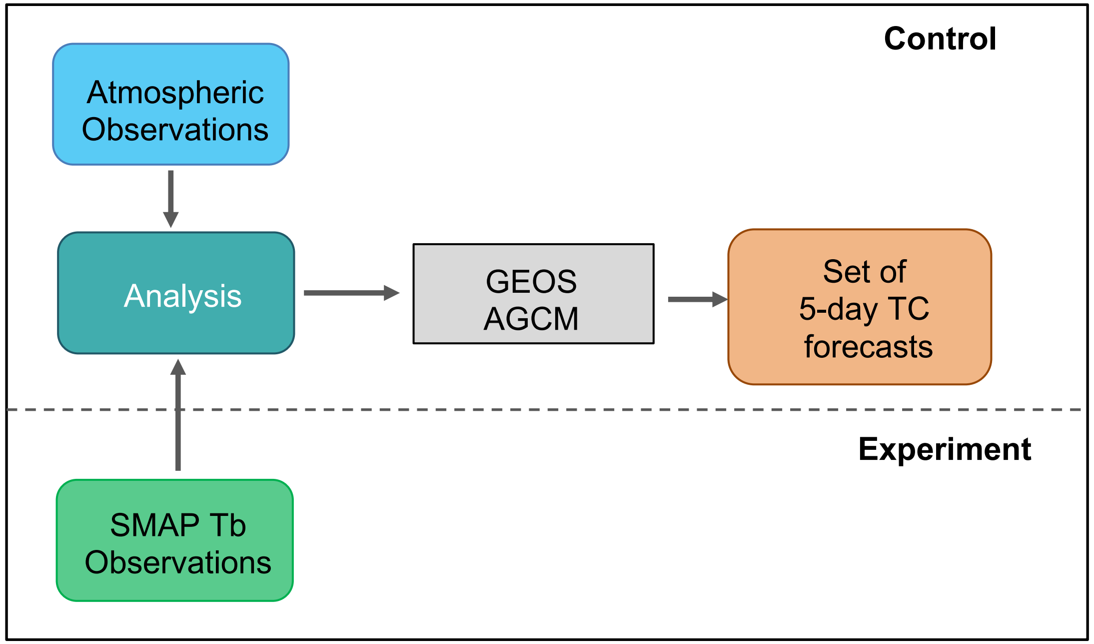


Fig 21: Experiment design schematic

Control:

- Forecasts of TC from analysis constrained by standard suite of atmospheric observations

Experiment:

- Additional constraint through SMAP Tb observations

Evaluation:

- Combination of global skill metrics, regional tailored metrics and phenomenological approaches to evaluate impact on TC forecast skill

Impact of land observations in NWP systems

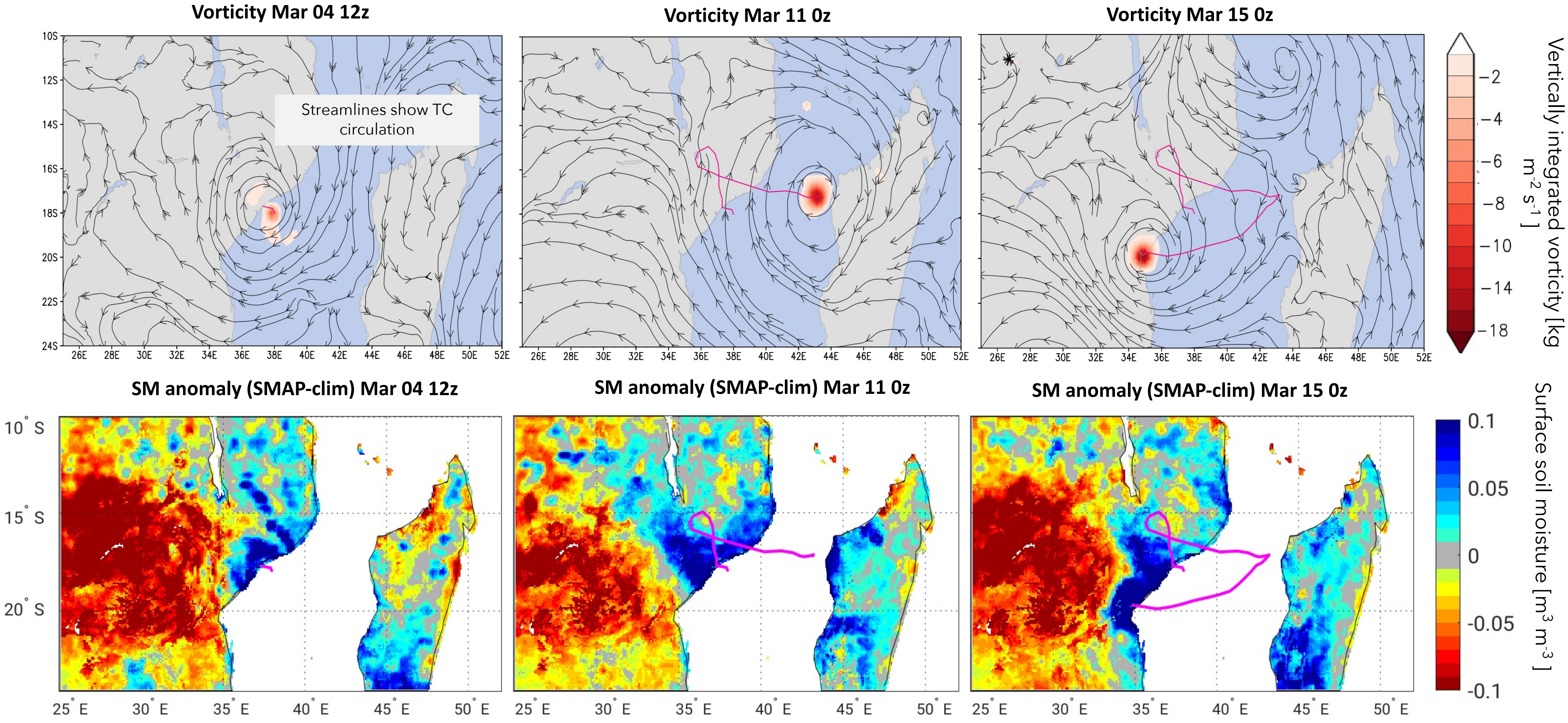


Fig 22: TC Idai evolution of vorticity, circulation and soil moisture conditions

Impact of land observations in NWP systems

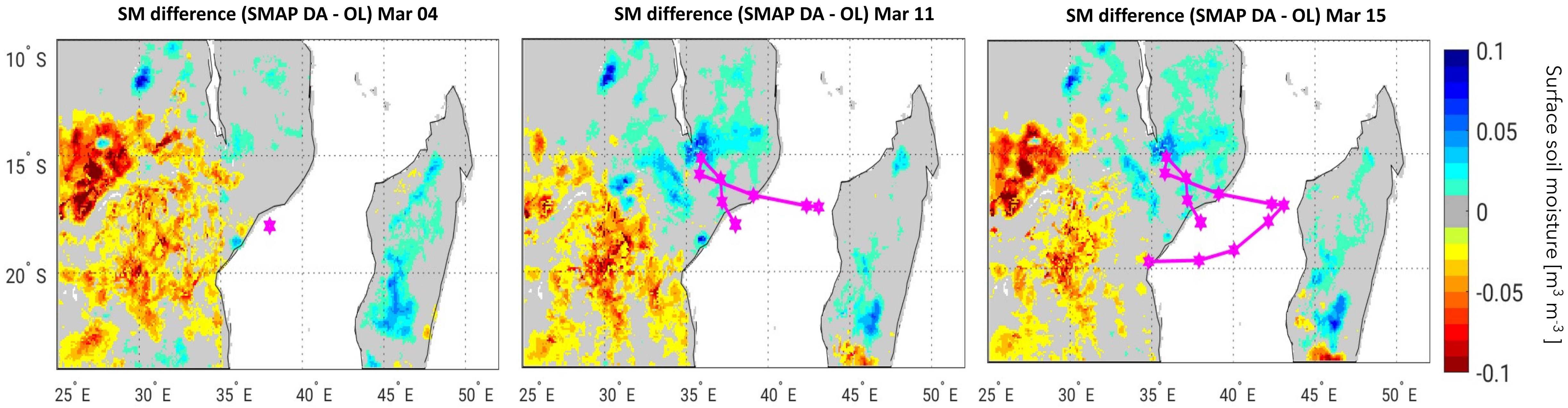


Fig 23: SMAP L4 difference with respect to open loop for Idai period

- The SMAP DA run captures the wetter than normal conditions better than a model run that is not constrained by SMAP

→ By assimilating SMAP we could improve the forecast of Idai's behavior

Evaluating data assimilation impact

- For soil moisture, evaluation is frequently against in situ observations

1. At locations with in situ stations, models are already very skillful

2. Satellite observations (and models) contain very little information about point scale soil moisture variability

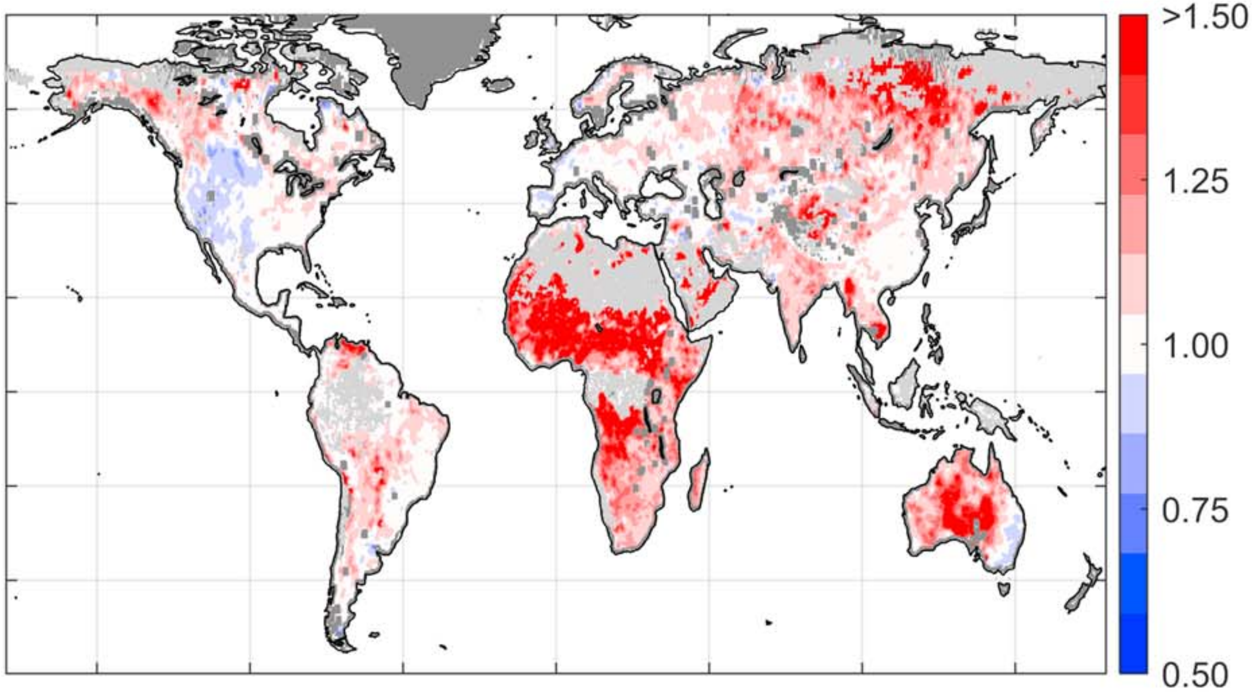
Table 3

Breakdown of the Information-Use Efficiency Metrics From an EnKF Assimilation of LPRM Retrievals Into the Noah-MP Land Surface Model as Evaluated Against SCAN Data

Measurement	Metric	Value
Information in model simulations[a]	$\frac{I(Z;X)}{H(Z)}$	0.13
Information in retrievals[a]	$\frac{I(Z;Y)}{H(Z)}$	0.08
Total information from model and retrievals[a]	$\frac{I(Z;X,Y)}{H(Z)}$	0.18

Evaluating data assimilation impact

- *Dong et al., 2019*: Relative skill evaluation using a third (independent) product



SMAP L4 product is better

OL is better

Fig 24 : Relative skill of SMAP Level 4 product and Open Loop

- "ground validation conducted in data-rich areas does not adequately sample the added value of SMAP data assimilation … and substantially underestimates the added skill provided by the SMAP level 4 system."

Conclusions and Outlook

Conclusions:

- Accurate observation error characterization is critical
- Need to improve efficiency of DA systems → better DA approaches, better observations, better correspondence between model and observations
- Parameter calibration should be performed in conjunction with improvements to model structure
- Alternative ways to assess the impact of land observations in NWP systems
- Reconsider how we evaluate land analyses

Outlook:

- Coupling in NWP systems and ESMs
- Increasing correspondence between modeled and observed variables
- Use of machine learning or hybrid techniques when process knowledge is uncertain
- Land DA community to improve internal collaboration as well as coordination with other communities

  → AIMES Land Data Assimilation Working Group https://aimesproject.org/ldawg/

References:

[1] Bacour, C., Peylin, P., MacBean, N., Rayner, P.J., Delage, F., Chevallier, F., Weiss, M., Demarty, J., Santaren, D., Baret, F. and Berveiller, D., 2015. Joint assimilation of eddy covariance flux measurements and FAPAR products over temperate forests within a process-oriented biosphere model. *Journal of Geophysical Research: Biogeosciences*, 120(9), pp.1839-1857.

[2] Carrera, M.L., Bilodeau, B., Bélair, S., Abrahamowicz, M., Russell, A. and Wang, X., 2019. Assimilation of passive L-band microwave brightness temperatures in the Canadian Land Data Assimilation System: Impacts on short-range warm season numerical weather prediction. *Journal of Hydrometeorology,* 20(6), pp.1053-1079.

[3] Dong, J., Crow, W., Reichle, R., Liu, Q., Lei, F. and Cosh, M.H., 2019. A global assessment of added value in the SMAP Level 4 soil moisture product relative to its baseline land surface model. *Geophysical Research Letters*, 46(12), pp.6604-6613.

[4] Draper, C.S., Reichle, R.H., De Lannoy, G.J.M. and Liu, Q., 2012. Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters,* 39(4).

[5] Kolassa, J., Aires, F., Polcher, J., Prigent, C., Jimenez, C. and Pereira, J.M., 2013. Soil moisture retrieval from multi-instrument observations: Information content analysis and retrieval methodology. *Journal of Geophysical Research: Atmospheres*, 118(10), pp.4847-4859.

[6] Kolassa, J., Reichle, R.H. and Draper, C.S., 2017. Merging active and passive microwave observations in soil moisture data assimilation. *Remote Sensing of Environment,* 191, pp.117-130.

[7] Kolassa, J., Reichle, R.H., Liu, Q., Cosh, M., Bosch, D.D., Caldwell, T.G., Colliander, A., Holifield Collins, C., Jackson, T.J., Livingston, S.J. and Moghaddam, M., 2017b. Data assimilation to extract soil moisture information from SMAP observations. *Remote sensing*, 9(11), p.1179.

[8] Kolassa, J., Reichle, R.H., Liu, Q., Alemohammad, S.H., Gentine, P., Aida, K., Asanuma, J., Bircher, S., Caldwell, T., Colliander, A. and Cosh, M., 2018. Estimating surface soil moisture from SMAP observations using a Neural Network technique. *Remote sensing of environment*, 204, pp.43-59.

[9] Koster, R.D., Guo, Z., Yang, R., Dirmeyer, P.A., Mitchell, K. and Puma, M.J., 2009. On the nature of soil moisture in land surface models. *Journal of Climate*, 22(16), pp.4322-4335.

[10] Kumar, S.V., Reichle, R.H., Harrison, K.W., Peters-Lidard, C.D., Yatheendradas, S. and Santanello, J.A., 2012. A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resources Research*, 48(3).

[11] Kwon, Y., S.V. Kumar, et al., Direct use of SMAP soil moisture anomalies improves irrigation characterization through data assimilation, in review

[12] Nearing, G., Yatheendradas, S., Crow, W., Zhan, X., Liu, J. and Chen, F., 2018. The efficiency of data assimilation. *Water resources research*, 54(9), pp.6374-6392.

Qiu, B., Xue, Y., Fisher, J.B., Guo, W., Berry, J.A. and Zhang, Y., 2018. Satellite chlorophyll fluorescence and soil moisture observations lead to advances in the predictive understanding of [13] global terrestrial coupled carbon-water cycles. *Global Biogeochemical Cycles,* 32(3), pp.360-375.

[14] Reichle, R., Zhang, Q.S., et al. (2021) Assimilation of SMAP Brightness Temperature Observations in the GEOS Land-Atmosphere Data Assimilation System, *JSTARS SMAP Special Issue* (in review)

[15] Rodríguez-Fernández, N., de Rosnay, P., Albergel, C., Richaume, P., Aires, F., Prigent, C. and Kerr, Y., 2019. SMOS neural network soil moisture data assimilation in a land surface model and atmospheric impact. *Remote Sensing*, 11(11), p.1334.

Thank you!