Session C24 -T203B

Joint Uncertainty Assessment and Statistical Merging of Multiple Model and Remote Sensing Data Records for Biogeochemical Modeling

Lucas A. Jones[1,2], John S. Kimball[1,2], E. F. Wood[3], R. Reichle[4]

[1] *University of Montana Flathead Lake Biological Station*, Polson, MT ;[2]*Numerical Terradynamic Simulation Group,* University of Montana, Missoula, MT ; [3]*Princeton University,* Princeton, NJ; [4]*NASA GSFC/GMAO,* Greenbelt, MD;

Email: lucas@ntsg.umt.edu Websites: www.umt.edu/flbs & www.ntsg.umt.edu

Abstract

Ecosystem carbon and hydrological process models require gridded input surface meteorology, usually provided by model reanalysis and remote sensing observations. However, reanalysis accuracy can be rather poor for some surface data fields, such as soil moisture, and reanalyses typically do not take full advantage of available remotely sensed information, such as soil moisture, or precipitation, or ET over land. A major factor limiting use of remotely-sensed surface information are unknown error fields associated with these observations. Nevertheless, remotely sensed observations contain significant independent information that can potentially improve the accuracy of reanalysis fields provided their error structure is known. Here we investigate a statistical technique to jointly estimate errors and combine ≥3 independent datasets to improve regional soil moisture state estimates over the continental USA. We include model reanalysis (MERRA), satellite microwave remote sensing retrievals of surface soil moisture (AMSR-E) and precipitation (TRMM 3B42), and remote-sensing based ET (MODIS) datasets. The results are comparable to a more sophisticated assimilation of similar datasets within the MERRA land model [1,2]. This indicates that knowing the observation error structure and combining observations within a simple state-space framework benefits soil moisture state estimates as much as data assimilation using a detailed land model. The results provide important feedback for diagnosing physical land model and remote sensing algorithm inaccuracies.

Objectives and Hypotheses

1) Develop a relatively simple, computationally efficient method for jointly estimating surface soil moisture relative error & 'true' states using several independent datasets (≥3) that are sensitive to the same underlying signal but have different error structures.

2) If 1) is achieved, then the skill & accuracy of the estimated 'true' soil moisture state will be greater than the most skillful estimate from the component datasets.

3) Determine to what extent 1) can be successfully applied to longer time-scale components (seasonality).

Methods

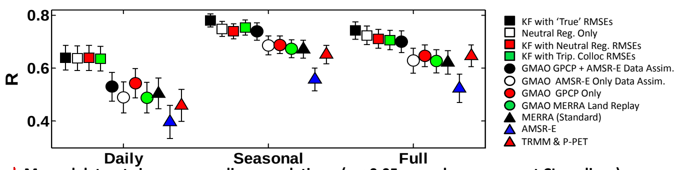
- **State-Space Representation:** The three datasets represent each element of the observation vector. The statistical system state equation represents uncertain system dynamics with a simple AR(1) time-series model, in contrast to the physics-based differential equations typically used in data assimilation. The Kalman filter (KF) provides an algorithm to optimally estimate the mean of the state vector when the system and observations are linear equations and the uncertainty is from additive Gaussian noise. The 'optimal' weighting depends on the relative uncertainty between the system and observation equations.
- **Timescale Separation:** Soil moisture timeseries contain characteristic timescales of variability, while cycles and auto-correlated errors can pose problems for meeting independence assumptions for error estimation. Timeseries are rescaled and decomposed into components, which are then separately assimilated. After filtering, the components are reassembled (summed) to provide the full state estimate.
- **Soil Water Index:** A simple antecedent time series model [3] is used to integrate precipitation and ET data into an indicator of soil moisture storage prior to filtering.
- **Determining dataset errors and weighting: Triple Collocation** (TC) is one method used to determine relative RMSE's of three datasets [4]. **Neutral Regression** (NR), an alternative to TC, uses an eigenvalue-eigenvector decomposition to find the relative value weights for each dataset, which can either be used directly to calculated a merged state or used to calculate error variances as input to a Kalman filter [5]. Eigenvector and Kalman gain weights are similar but not equivalent. Neither TC, nor NR techniques are theoretically optimal estimates under conditions present for most real applications.

Datasets

- **MERRA GEOS-5 Reanalysis:** The hourly surface soil moisture state variable [% Saturation] is averaged to daily values.
- **AMSR-E VU 10.7 GHz Soil Moisture:** Daily morning overpass soil moisture provided by the Vrjie University Amsterdam dual-polarization algorithm [6].
- **TRMM 3B42 Precipitation:** A merged and bias-corrected multi-sensor TRMM and IR precipitation product [7]. The 3-hourly precipitation rates [mm hr-1] are averaged to daily average rates and then used to calculate a soil water index. This dataset is merged with satellite-based P-ET prior to merging with MERRA and AMSR-E.
- **Satellite-based P-ET:** This dataset uses GPCP precipitation and a satellite-derived daily ET dataset to calculate monthly P-ET [8, 9]. The monthly values are used to calculate a monthly soil water index. Monthly values are then interpolated to smooth daily seasonal values using cubic splines. This dataset is merged with TRMM soil water index prior to merging with MERRA and AMSR-E.
- **All datasets are co-registered to a 25-km global EASE grid.**

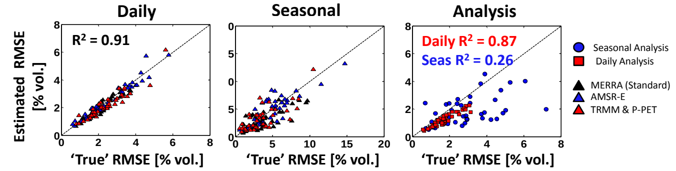
A. Soil Moisture Validation at 42 FLUXNET Stations

Surface Soil Moisture (≤5cm) Timeseries Correlations



- Merged datasets improve median correlations ($p < 0.05$; error bars represent CI medians) relative to GMAO MERRA land model data assimilation [1,2]. Correlation improvement is most substantial for the shorter timescale (daily) component .

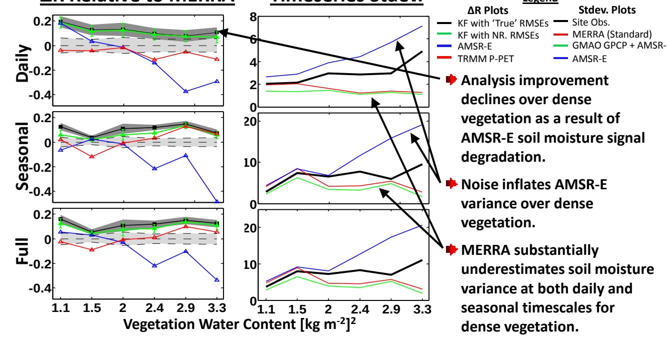
Soil Moisture Error Performance[1]



- The success of dataset merging depends on how well the RMSEs of each dataset are estimated. Neutral regression shows favorable performance in estimating RMSEs at daily timescales, while seasonal performance is lower.

[1]Data re-scaled to site soil moisture and corrected for point-to-pixel scaling error [10].

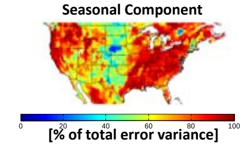
ΔR Relative to MERRA / Timeseries Stdev.



- Analysis improvement declines over dense vegetation as a result of AMSR-E soil moisture signal degradation.
- Noise inflates AMSR-E variance over dense vegetation.
- MERRA substantially underestimates soil moisture variance at both daily and seasonal timescales for dense vegetation.
- Despite success for dataset merging at daily timescales, the seasonal timescales contain the bulk of overall soil moisture variance. This ultimately limits potential improvement for methods that focus only on short timescales.

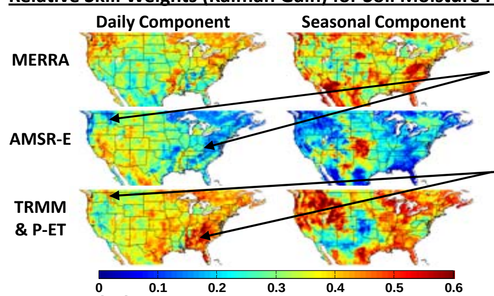
[2]Estimated from AMSR-E 10.7 GHz vegetation optical depth retrievals [11].

Error Variance (%) Attributed to the Seasonal Component

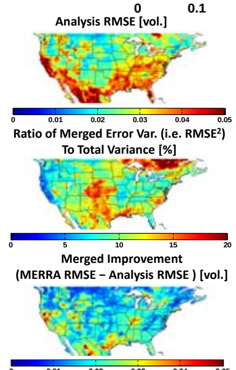


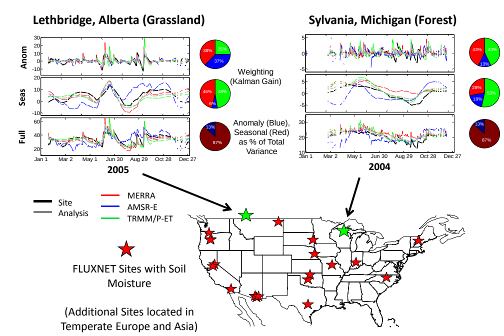
B. Application to the Continental USA Domain (2003-2006)

Relative Skill Weights (Kalman Gain) for Soil Moisture Prediction



- Lower (blue) weighting patterns reflect expected AMSR-E soil moisture signal degradation over dense vegetation.
- Loss of AMSR-E soil moisture signal is offset by higher (red) weighting skill for satellite-based P-ET information.





Conclusions

- Merging independent soil moisture information from MERRA, AMSR-E & TRMM/P-ET datasets using the proposed method significantly (median p<0.05) improves surface soil moisture correlations relative to site observations.
- Overall improvement in soil moisture estimation accuracy compares with more sophisticated data assimilation techniques.
- Errors are favorably estimated & improvement is greatest for short (daily) timescale components, but with reduced seasonal performance; however, much of the overall variance is associated with the seasonal component.
- MERRA generally underestimates the surface soil moisture variance of daily & seasonal components relative to site observations for densely vegetated areas.
- The information content & relative value of remotely sensed soil moisture & precipitation are complementary & offset each other depending on vegetation biomass levels.
- The presented error estimation techniques are not theoretically optimal, unbiased estimators for several reasons, but clearly provide useful predictive power.
- The results indicate that the planned NASA SMAP & GPM missions will provide complementary information for improved regional soil moisture predictions.
- The methods are flexible & can be applied to other land parameters in addition to soil moisture.

References:

[1] Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. Toure (2011), Assessment and enhancement of MERRA land surface hydrology estimates, *J. Climate*, in press. doi:10.1175/JCLI-D-10-05033.1. **[2]** Liu, Q., R. H. Reichle, R. Bindlish, M. H. Cosh, W. T. Crow, R. de Jeu, G. J. M. De Lannoy, G. J. Huffman, and T. J. Jackson (2011), The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system, Journal of Hydrometeorology, 12, 750-765, doi:10.1175/JHM-D-10-05000.1. **[3]** Wagner, W., G. Lemione, H. Rott. 1999. A method for estimating soil moisture from ERS Scatterometer and soil data. *Rem. Sens. Environ.* 70: 191-207. **[4]** Janssen, P. et al. 2007. Error estimation of buoy, statellite, and model wave height data. *J. Atmos. Oceanic Technol.* 24, 1665-77. **[5]** Mardsen, R. F. 1999. A proposal for neutral regression. *J. Atmos. Oceanic Technol.* 16, 876-83. **[6]** Owe, M., R. de Jeu, J. Walker. 2001. A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. *IEEE Trans. Geosci. Rem. Sens.* 39(8): 1643-54. **[7]** Huffman, G.J., et al., 2007: The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. J. Hydrometeor., 8(1), 38-55. **[8]** Huffman, G.J., et al., 1997. The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.* 78: 5-20. **[9]** Zhang, K. et al. (2010). A continuous satellite-derived global record of land surface evapotranspiration from 1983-2006. *Water Res. Research*. 46, W09522, doi:10.1029/2009WR008800. **[10]** Miralles, D. G., W. T. Crow, M. H. Cosh. (in review). Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations. *J. Hydromet.* 11, 1423-9, DOI: 10.1175/2010JHM1285.1 **[11]** Jones, L.A., and J.S. Kimball, 2010. Daily Global Land Surface Parameters Derived from AMSR-E. Boulder Colorado USA: National Snow and Ice Data Center. Digital media (http://nsidc.org/data/nsidc-0451.html).

Acknowledgements:

This work is supported by a NASA Earth System Science Fellowship to L. A. J. The authors are grateful to flux tower site investigators for the use of FLUXNET/EUROFLUX/AsiaFLUX soil moisture data available through ww.fluxnet.org. The number of PI's are too numerous to list in the limited space available here for the 43 sites that were used. Discussions with W. T. Crow significantly contributed to this work.