**Project Title:** Improving land evaporative processes and land-atmosphere interactions in the NCEP Global Forecast System (GFS) and Climate Forecast System (CFS)
**PI/Co-PI:** Eric F. Wood, Michael Ek, and Justin Sheffield
**Report Year:** Final project report
**Grant #:** NA11OAR4310175 (GC11-590a)

## Project Overview

Land surface evapotranspiration plays a central role in the water, energy, and carbon cycles. It provides the link between the energy and water budgets at the land surface and the link between the water and carbon cycle through vegetation transpiration. Accurately modeling terrestrial evapotranspiration processes is fundamental to numerical weather prediction, seasonal forecasting, and global circulation models. The goal of this project is to analyze, evaluate, and improve land evaporative processes in the NOAH land surface component of the NCEP Global Forecast System (GFS) and Climate Forecast System (CFS). The focus is on warm season terrestrial evaporative processes including: free evaporation from water bodies and canopy intercepted precipitation, evaporation of soil water, and transpiration by vegetation. This has been achieved in this project through the following tasks:

*Task 1: Data set selection and compilation*
*Task 2. Generation of off-line model runs*
*Task 3: Diagnostic analyses for assessing process deficiencies*
*Task 4: Developing and testing new ET parameterizations*

It was also recognized during the course of the research that parameters that impact the turbulent latent heat fluxes are also relevant to the turbulent sensible heat fluxes, and accuracy in parameterizing evaporative (latent heat) processes are critical to the sensible heat parameterizations. This is especially true for the aerodynamic resistance parameterization, which is discussed under Tasks 3 and 4. Thus, a Task 5 was added, namely: *Developing and Testing new Sensible Heat parameterizations*.

## Results and Accomplishments

*Task 1: Data set selection and compilation*

To determine the weaknesses in Noah's evapotranspiration module, a special emphasis is placed on minimizing the uncertainty in the model input meteorological data and observed latent heat fluxes. The FLUXNET global network of eddy covariance towers provides a wealth of high quality information that helps accomplish this goal. FLUXNET is an active global network of meteorological sites comprised of over 650 sites in 30 regional networks covering 5 continents. These sites measure the exchange of water vapor, carbon dioxide, and energy between terrestrial ecosystems using the eddy covariance method [*Baldocchi*, 2008]. In an effort to provide a dataset that can be used by the global earth science community, the FLUXNET community has harmonized, standardized, and gap-filled these sites to create the La Thuile dataset. This summary database contains 253 eddy covariance stations with a total of 960 site-years of data at a 30-minute time resolution [*ORNL DAAC*,

2013]. This data is invaluable for land surface modeling as it provides high quality input data including incoming shortwave and longwave radiation, air temperature, wind speed, specific humidity, and precipitation. But more importantly, it provides data to diagnose the land surface model output including latent heat, sensible heat, net radiation, ground heat flux, surface friction velocity, and outgoing longwave radiation, among others. Figure 1 shows the spatial coverage of the La Thuile dataset over the globe and the number of years of data available per site; the average temporal coverage per site is ~4 years. To minimize misinterpretations of the model results, the model input meteorology and observations of latent heat are thoroughly quality controlled.

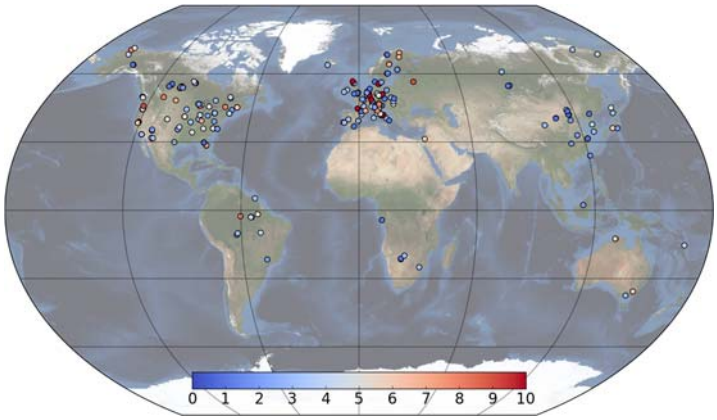


*Figure 1. Map of the 253 eddy covariance sites in the FLUXNET La Thuile dataset. Each sites's color represents the number of years of data available.*

*Task 2: Generation of off-line model runs*

The high quality FLUXNET dataset is then used to run the Noah land surface model run at each eddy covariance site using the local meteorological data. For these simulations the model is run using the default Noah evapotranspiration module and the default model parameters. The 30-minute latent heat simulations at each eddy covariance station are then compared to the quality-controlled observations. To adequately determine model performance, the biases in the temporal mean, temporal standard deviation, and linear correlation are assessed. These different metrics are then combined via the KGE metric to provide a summary of the Noah land surface model's performance. The results as shown in Figure 2 illustrate how the model captures well the linear correlation with the default parameters. This result is due to the model's ability to capture the seasonal and diurnal cycles. There are substantial biases in the temporal mean and standard deviation for many of the stations in FLUXNET. This illustrates uncertainties in either the model parameterizations or the model parameters.

*Task 3: Model experiments for assessing process deficiencies*

***Model Parameter Sensitivity –*** The uncertainty in Noah's latent heat estimates is driven by the large number of uncertain model parameters. However, each model parameter will have a different impact and contribute in different aspects to the model biases. A potential path forward is to tune the model parameters to improve model performance.

However, given the limited number of eddy covariance sites, tuning all the model parameters at each site appears to be ill advised. To address this concern, we first seek to reduce the number of tunable model parameters by identifying the most sensitive parameters. The model parameters that we determine to be insensitive can be ignored since they don't contribute heavily to the model biases. To determine the sensitivity of the model parameters, we use the Sobol sensitivity analysis. Prior studies have successfully used this technique to discern the role of the Noah model parameters at a limited number of sites [*Rosero et al.*, 2010; *Hou et al.*, 2015].

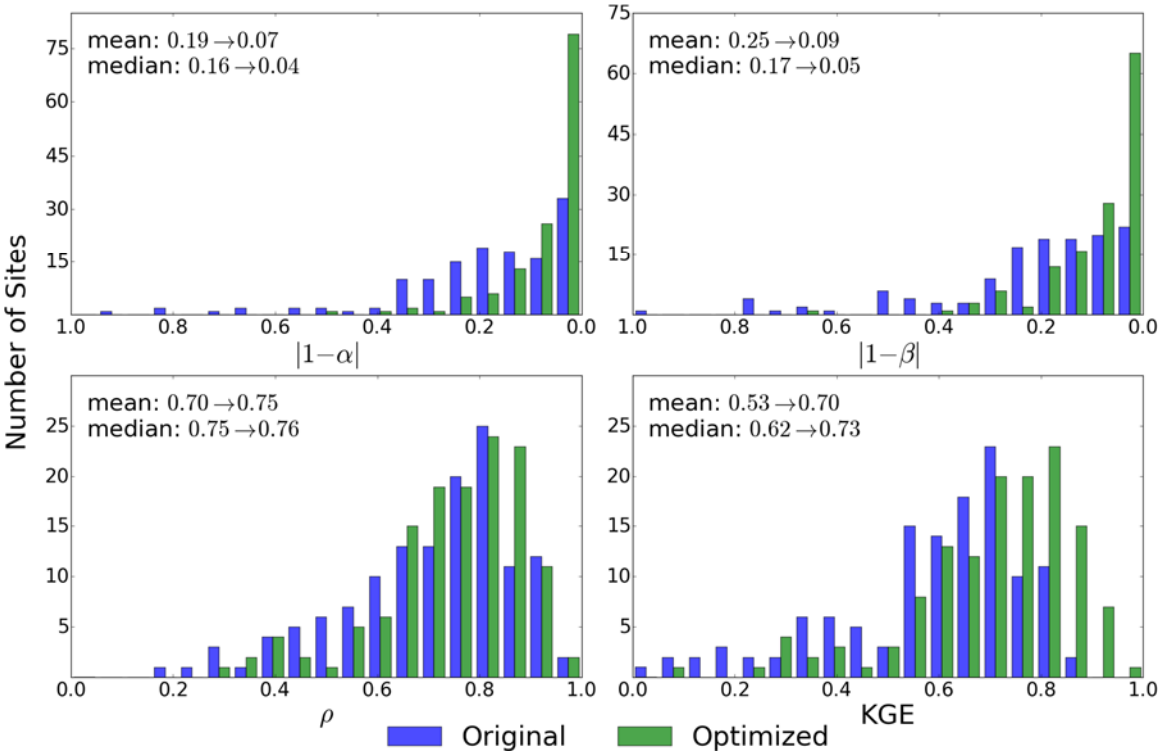


*Figure 2. The Noah land surface model is run at 130 eddy covariance sites in the FLUXNET network using the default and optimized parameters sets. The four panels show - via histograms - the difference in model performance before and after parameter optimization. The upper left panel shows the bias in standard deviation, the upper right panel shows the bias in the mean, the lower left panel shows the linear correlation, and the lower right panel shows the KGE metric derived from the other three components.*

The Sobol sensitivity analysis [Sobol 1993, 2001] is a global method that decomposes the variance of the model output $Y$ into contributions from each parameter $X_i$ and its interactions with other parameters. The first-order index $S_i$ represents the expected reduction in variance if parameter $X_i$ were fixed, not accounting for interactions with other parameters. The total-order index $S_{Ti}$ represents the reduction in variance ($V_{\sim i}$) that would occur if all parameters except $X_i$ were fixed, accounting for all interactions with other parameters. Ensemble members are constructed by sampling two ensembles of size $n$ from the Sobol quasi-random sequence [*Sobol*, 1993; *van Werkhoven et al.*, 2008], then "cross-sampling" by holding one parameter fixed at a time for a total of $n(k+2)$ parameter sets, where $k$ is the number of parameters. The Noah land surface model is run

independently for each parameter set, and a performance metric is computed for each ensemble member. In this study, the chosen performance metric to compare the observed and simulated latent heat fluxes is the Kling-Gupta efficiency (KGE) metric [*Gupta et al.*, 2009]. This multi-objective metric combines three terms: linear correlation, temporal variability, and mean bias.

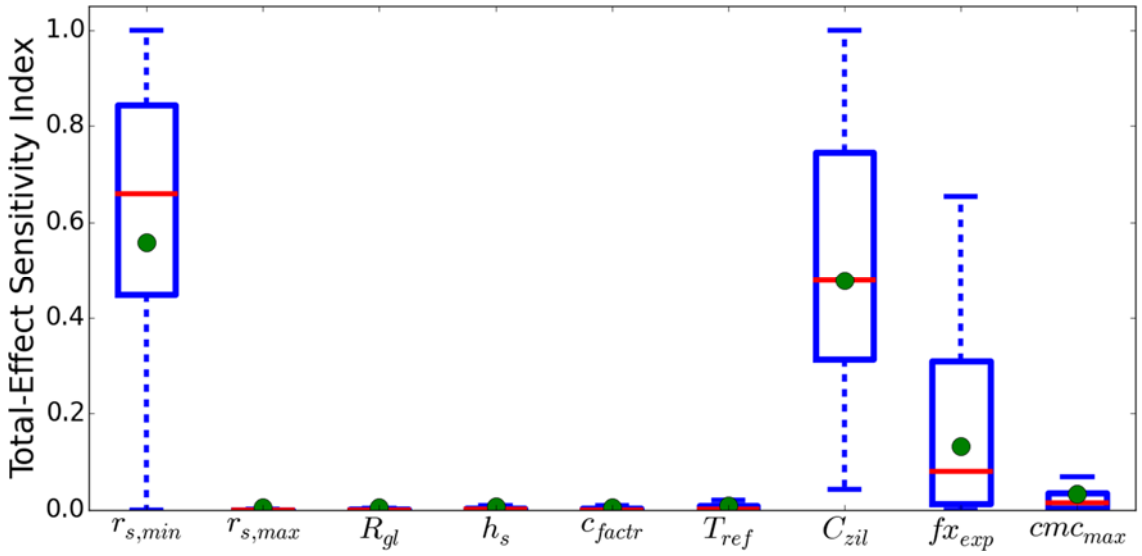


*Figure 3. The summary of the total-effect sensitivity index across all stations is shown via a boxplot per model parameter. The red line indicates the median value and the green dot indicates the mean.*

$$KGE = 1 - \sqrt{\left(r_{pearson} - 1\right)^2 + \left(\frac{\sigma_{model}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{model}}{\mu_{obs}} - 1\right)^2}$$

$$correlation = r_{pearson};\ variability = \frac{\sigma_{model}}{\sigma_{obs}};\ bias = \frac{\mu_{model}}{\mu_{obs}}$$

After calculating the first-order and total-effect sensitivity indices for each site, the results are summarized across all eddy covariance stations to assess the parameter sensitivity across the FLUXNET network; the results are shown in Figure 3. The total-effect index is computed at each 130 site for each of the chosen 9 model parameters. The boxplots in Figure 3 are the summary of all total-effect index values in the network per parameter. Noah's ability to accurately simulate latent heat flux is highly sensitive to both $r_{s,min}$ and $C_{zil}$. The same is true for $fx_{exp}$ when we focus on sites that have a strong seasonality in vegetation coverage or have minimal vegetation coverage year round. The model performance is strongly insensitive to the rest of the parameters. This provides a strong argument that model calibration should focus on the $r_{s,min}$, $C_{zil}$, and $fx_{exp}$ parameters, fixing the remaining parameters with values from the model look-up tables.

*Task 4: Developing and testing new ET parameterizations*

***Model Parameter Optimization*** - Reducing the number of parameters using the Sobol

sensitivity analysis simplifies the calibration exercise at each eddy covariance site. The next step is to determine the optimal parameter values that minimize the biases in the simulated latent heat. To obtain approximately optimal model performance while assessing the role of model parameter equifinality, the Latin Hypercube Sampling technique (LHS) [*McKay et al.*, 1979] is used to assess model performance across the reduced model parameter space. We assess the model performance for each parameter set by computing the KGE between the observed and simulated latent heat fluxes.

Figure 2 shows how the biases and linear correlation change when selecting the best model parameter set from the Latin Hypercube Sample compared to the default parameters. The results suggest that optimization leads to a large reduction in the bias in the mean and the standard deviation at each site. The network average of the mean bias shifts from 25% to 9%; the network average of the standard deviation bias shifts from 19% to 7%. Unfortunately, the improvement in the linear correlation is not as dramatic. The network average linear correlation shifts from 0.70 to 0.75. Upon combining these components to create the KGE metric we notice a shift in the network average KGE from 0.53 to 0.70. The network median of the KGE shifts from 0.62 to 0.73. The difference between the mean and the median suggests that there are a number of sites that perform very poorly using the look-up table parameter values that improve significantly after optimization. The inability to improve many stations that have a KGE under 0.5 after optimization suggests room for further improvement. This could be indicative of biases and uncertainties in the input data or structural deficiencies in Noah's evapotranspiration module and its parameterizations.

***Model Parameter Regionalization*** **-** Although parameter calibration leads to a large reduction in bias in Noah's estimates of latent heat flux, it does not ensure that the optimized parameters ($r_{s,min}$, $C_{zil}$, and $fx_{exp}$) can be transferred to other sites using climate, vegetation, and soil characteristics. To test whether this is possible, an Extra-Trees model (Extra-Trees; [*Geurts et al.*, 2006]) with 13,000 extremely randomized trees is fit using the local environmental characteristics to the optimal parameter sets. The model is validated using a leave-one-out cross validation for the 130 eddy covariance sites. The leave-one-out cross-validation results for the optimal parameter dataset are shown in Figure 4. When the 13,000 trees are used to estimate the Noah model parameters, the coefficient of determination is above 0.96 for all three parameters. When only the trees that were not trained on a given site (100 trees) are used to estimate the parameters at that site, the skill decreases. In this case, the coefficient of determination of $r_{s,min}$ is 0.27, $C_{zil}$ is 0.17, and $fx_{exp}$ is 0.32. This provides a robust evaluation of the model's ability to estimate the optimal parameters at sites not included in the training process. Although there is clearly room for improvement, the cross-validation results suggest that indeed a skillful functional relationship between the optimized model parameters and the local environmental characteristics does exist.

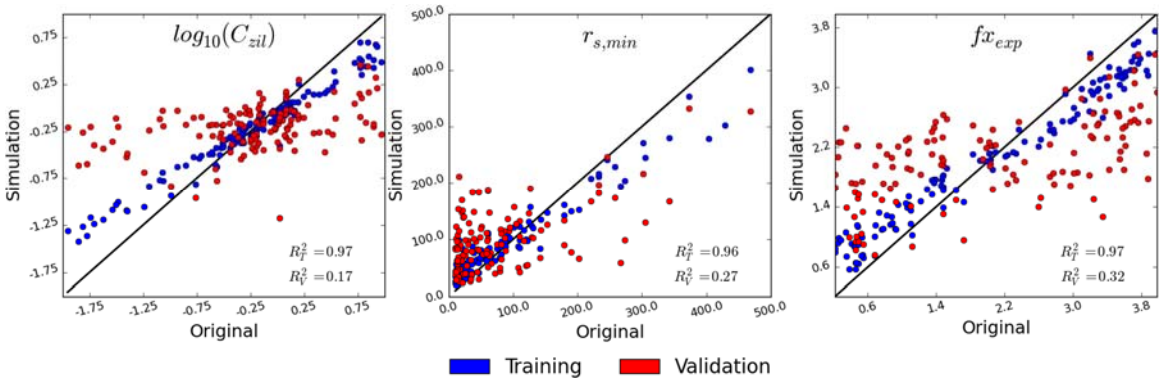


*Figure 4. An Extra-Trees model with 13,000 trees is fit to the 130 optimal parameter sets obtained via the Latin Hypercube Sample. A leave-one-out cross-validation is used to provide insight into the model's skill to estimate parameters not used during training. These three panels illustrate the model's skill to estimate the training and validation parameter sets.*

The use of global datasets of environmental characteristics allows for an estimation of these parameters using the optimal Extra-Trees model over the globe. Figure 5 shows the mapped estimates for both the mean (prediction) and the standard deviation (uncertainty estimate) of $r_{s,min}$, $C_{zil}$, and $fx_{exp}$ at a 1 km spatial resolution over the globe. The mean and standard deviation are calculated from the predictions at each grid cell of the 13,000 decision trees in the fitted Extra-trees model. The predominance of high $C_{zil}$ values over dry climates and short vegetation is encouraging – this agrees with the physical understanding of the $C_{zil}$ parameter [*Chen and Zhang*, 2009]. However, the physical consistency of the spatial properties of $r_{s,min}$ are not as apparent. We would expect the highest values to be in the water and energy limited regions and the lowest in the areas that are not water or energy limited – this does not seem to be always the case. Although further investigation is required, this suggests that the role that optimized $r_{s,min}$ parameter values play in the model might not be related to its physical meaning but simply as a bias correction term that absorbs other sources of uncertainty in Noah's estimates of evapotranspiration - including the errors in the resistance functions in the model's Jarvis type formulation of canopy resistance. Finally, the $fx_{exp}$ parameter shows distinct spatial patterns. The values are highest in regions where we don't expect a large role of bare soil in the latent heat flux. This suggests that the parameter optimization attempts to shut off the signal of bare soil evaporation in the model. Over drier regions, there is a tendency towards lower $fx_{exp}$ parameter values; it appears that the model attempts to increase the contribution of bare soil evaporation by allowing it to extract more water from the soil's top layer. This outcome for both the high and low $fx_{exp}$ values could be indicative of model weaknesses in the reliance on the green vegetation fraction to define the bare soil contribution to evaporation.

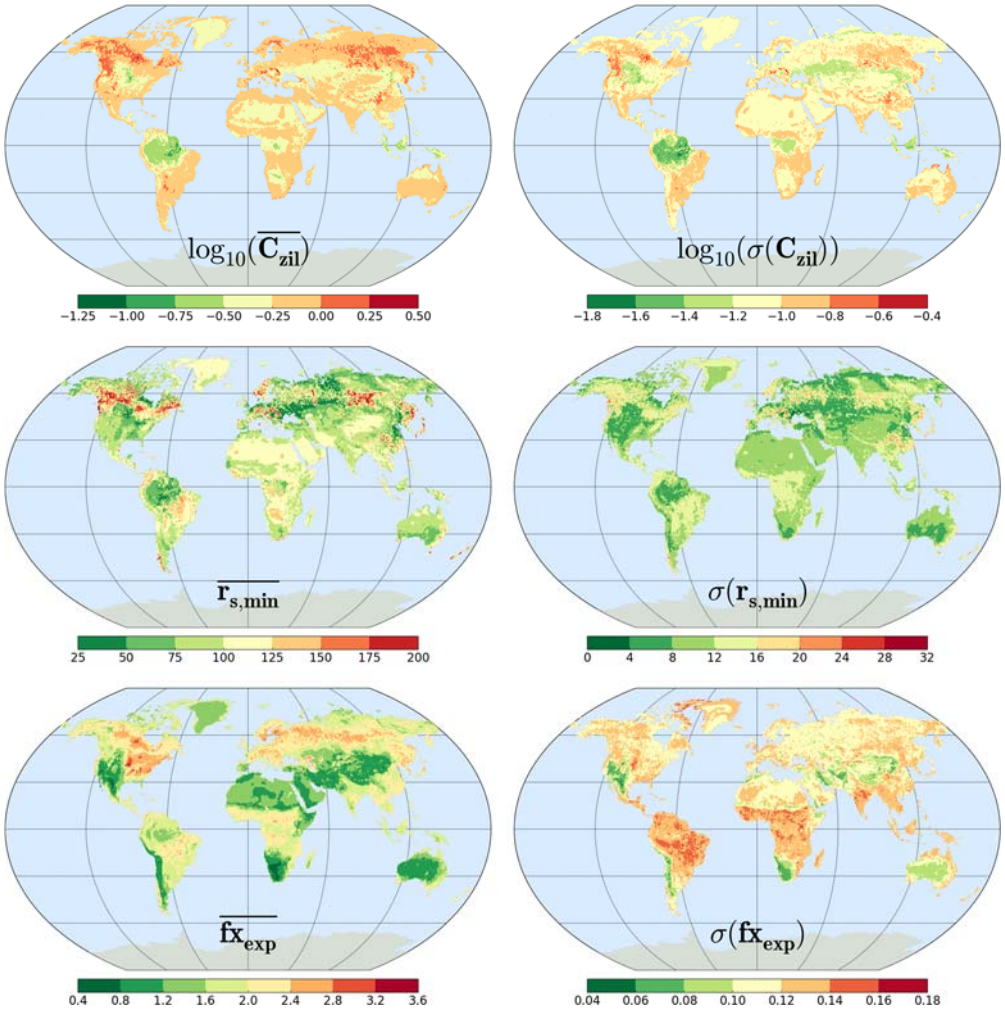


*Figure 5. The fitted Extra-Trees model is used to map the predicted values (mean) and uncertainty estimates (standard deviation) of $r_{s,min}$, $fx_{exp}$, $C_{zil}$ globally at a 1 km spatial resolution.*

*Task 5: Developing and Testing new Sensible Heat parameterizations*

The general parameterization for sensible heat flux, *H*, is: $H=\rho\, Cp\, (Ts-Ta)/r_a$ , where $\rho$ is air density [kg $m^{-3}$], *Cp* is specific heat [J $kg^{-1}$ $K^{-1}$], *Ts* and *Ta* are surface and 2m air temperatures [K] respectfully and $r_a$ is the aerodynamic resistance [s $m^{-1}$]. The parameterization of the aerodynamic resistance follows that formulated in Noah-MP. In particular, $r_a = (C_h\, u_z)^{-1}$ where $u_z$ is the wind speed at height *z*, and $C_h$ is a bulk heat transfer coefficient whose parameterization is based on the Monin and Obukhov stability criteria (Monin and Obukhov, 1954), and can be written as (Chen et al., 1997):

$$C_h = \frac{\frac{k^2}{R}}{\left[\ln\left(\frac{z}{z_{0m}}\right) - \Psi_m\left(\frac{z}{L}\right) + \Psi_m\left(\frac{z_{0m}}{L}\right)\right]\left[\ln\left(\frac{z}{z_{0h}}\right) - \Psi_h\left(\frac{z}{L}\right) + \Psi_h\left(\frac{z_{0h}}{L}\right)\right]}$$

Where $k$ is the von Karman constant that is ≈0.40, $z$ is height above the surface, $z_{0m}$ and $z_{0h}$ are roughness height for momentum and heat respectively, $L$ is the Obukhov length, often represented by

$$L = -\frac{u_*^3 \bar{\theta}_v}{kg\overline{w'\theta_v'}}$$

where $u*$ is the friction velocity, $\overline{\theta_v}$ is the mean virtual temperature, ($\overline{w'\theta v'}$) is the surface virtual potential temperature flux, and ψ(.) are the stability functions for momentum and heat. The ratio of the roughness heights for momentum and heat have been parameterized as (Zilitinkevich, 1995)

$$\frac{z_{0m}}{z_{0h}} = \exp\left(kC\sqrt{Re^*}\right)$$

$$Re^* = \frac{u_* z_{0m}}{\nu}$$

where $C$ is referred to as the Zilitinkevich constant that depends on the vegetation characteristics. In Task 4 this parameter (referred to there as *Czil*) was optimized based on FluxNet tower data. Here we used an iterative scheme to correct $z_{0h}$, $u*$ and $L$ to solve for $C_h$.

For the data set developed under this task, we desired to maintain input and satellite consistency across data sets with other GEWEX products. Thus, the sensible heat data product uses a land surface temperature (LST) data set consistent with retrievals from the High Resolution Infrared Radiation sounder (HIRS). Cloud-contamination and limits of swatch coverage cause sparse coverage in the retrievals, so the HIRS retrievals are merged with the NCEP Climate Forecast System Reanalysis (CFSR) LST estimates to form a global, hourly, 0.5° resolution LST data product (Coccia et al. 2015). These HIRS-consistent estimates are validated against the Baseline Surface Radiation Network (BSRN)-based LST (Siemann et al. 2016). Multiple reanalysis-based air temperature estimates are used to form the surface temperature gradient for the sensible heat flux estimates.

The aerodynamic resistance is based on the iterative solver used in the Noah land surface model (LSM) and optimized values at 70 global FluxNet towers (as earlier described under Task 4), which are extended into a global gridded-dataset using an objective analysis with climate and land cover covariates. A climatology of roughness length is taken from the UMD land cover data set for each vegetation type, and sensible heat estimates consist of a weighted average based on fractional coverage of sensible heat fluxes for each vegetation type which is then weighted based on green-vegetation fraction and bare soil for the final weighted average for each grid. The final product is a global, hourly, terrestrial, 0.5° resolution sensible heat flux data set for each of six different air temperatures (see Figure 6).

Global annual averages of four of the six sensible heat flux products fall within the range of values from the recent literature (i.e. Wild et al. 2015; L'Ecuyer et al. 2015; Jung et al. 2011), including our "baseline" product calculated with the 2-meter air temperature, which is consistent with ECMWF Interim Re-Analysis (ERA-Interim) and Climate Research Unit

(CRU) Time Series 3.10 monthly data (Wang and Zeng 2013). The sensible heat flux displays reasonable spatial patterns in many regions, including negative sensible heat fluxes in some areas with seasonal snow cover, low sensible heat flux over heavily forested areas, and high sensible heat flux for regions with shorter vegetation, such as Australia, the Horn of Africa, South Africa, and the Southwest of CONUS. Other regions display over-estimation in our products relative to reanalysis estimates including the CONUS Midwest region, the Canadian Prairies, and portions of central and eastern Eurasia. These products are also compared with FluxNet sensible heat flux estimates at 34 stations, and while several monthly correlations are above 0.5 for various products at 10 stations, correlations are low and biases are large at several other stations. Mis-match in the sensible heat flux products and reanalysis products as well as FluxNet products can be attributed to the sensitivities of the sensible heat flux to several input variables as well as the uncertainties in sensible heat derived from uncertainties in these inputs, including roughness length, the Zilinkitevich empirical constant, the temperature gradient, and the wind speed.

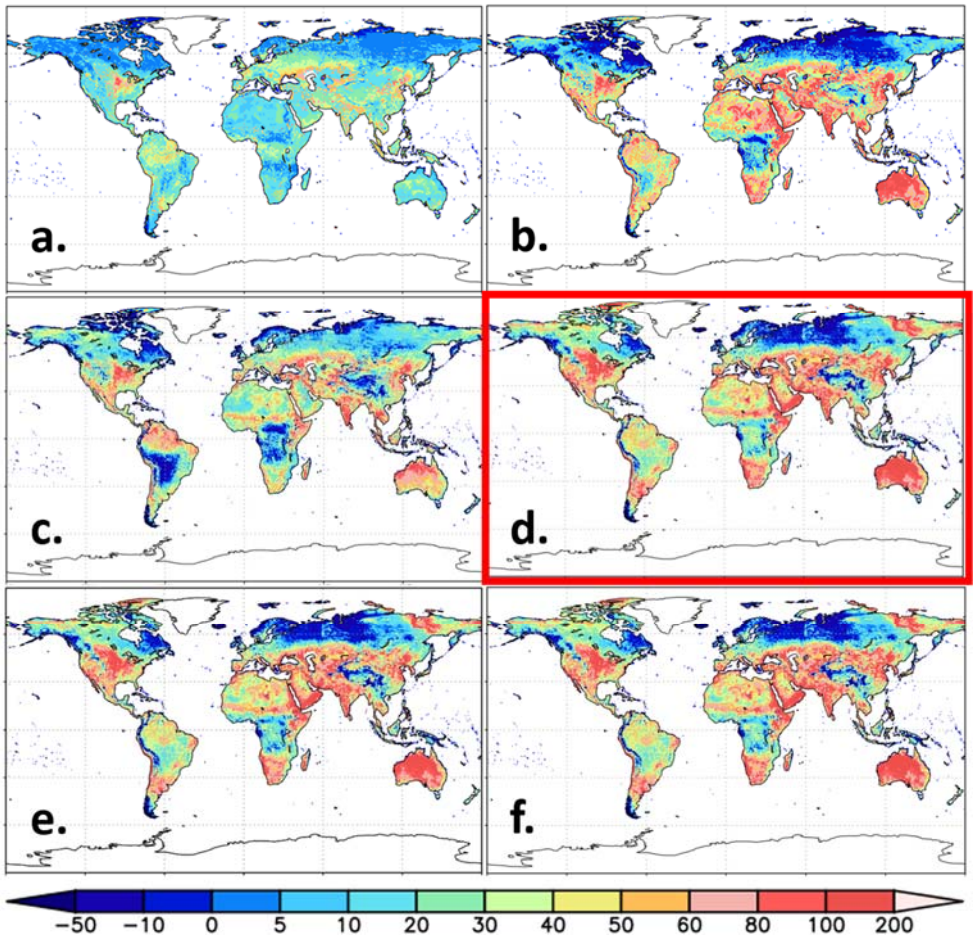


*Figure 6: Annual average sensible heat flux (1979-2009) for products using (a.) the CFSR temperature gradient, (b.) the CFSR air temperature, (c.) the National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO) MERRA air temperature, and (d.)-(f.) air temperatures consistent with CRU monthly data and based on ERA-Interim, National Centers for Environmental Prediction (NCEP) – National Center for Atmospheric Research (NCAR) reanalysis 1 (NRA), and MERRA, respectively.*

## Project Summary and Conclusions

This project has used the global network of eddy covariance sites (FLUXNET) to validate and improve the parameters of the evapotranspiration module in the Noah land surface model. A comprehensive sensitivity analysis using Sobol's method shows that the model's skill to simulate evapotranspiration is strongly tied to the $r_{s,min}$, $C_{zil}$, and $fx_{exp}$ model parameters. A 1,000 Latin Hypercube Sample is used to find the optimal values of these three parameters at each of the quality controlled 130 eddy covariance sites. Overall, calibration of the most sensitive parameters leads to a large decrease in biases in both the mean and temporal variability with a relatively small improvement in linear correlation. The Extra-Trees machine-learning algorithm is then used to relate the optimal parameter sets at each site to local environmental characteristics. Evaluation of the fitted Extra-Trees model shows that this parameter estimation technique can lead to improved simulations of the latent heat fluxes at both sites that were used to train the parameter regionalization model and those that were left out for validation. Finally, this functional relationship is used to produce 1-km maps of the $r_{s,min}$, $C_{zil}$, and $fx_{exp}$ parameters over the globe. The parameter regionalization technique developed in this study has the potential to enable the land surface modeling community to move beyond outdated parameter look-up tables. This research demonstrates the potential for integrating the extensive and growing network of FLUXNET observations to improve the parameterization and process representation of terrestrial evapotranspiration in macroscale land surface models. Further work will look into the impact of including the updated model parameters in coupled simulations to determine the impact of model parameter biases on the land-atmosphere coupling.

## Highlights of Project Accomplishments

- Gathered and curated the FLUXNET database of eddy covariance sites.
- Ran a SOBOL sensitivity analysis to determine how sensitive the model estimates of latent heat flux are to the Noah model parameters.
- Used the Latin Hypercube sampling technique to estimate the optimal parameter set per FLUXNET site and determine the role of model parameter equifinality.
- Developed, implemented, and validated an extra-trees regression model to estimate the optimal parameter sets globally at a 1 km spatial resolution.

## Conference Presentations from the Project

- N.W. Chaney, Liang. M., Wood, E. F., Validation of a suite of process-based models of evapotranspiration using FLUXNET, AGU 2012, (Contributed poster)
- N. W. Chaney and Wood, E. F., Improving Conductance Schemes in Macroscale Land Surface Models, EGU 2014, (Contributed talk)
- A. Siemann, G Coccia, N Chaney, D Miralles, C Jimenez, M F McCabe, E F Wood 2015 Analysis of a global terrestrial sensible heat flux dataset and global energy budget closure, presented at Earth Observation for Water Cycle Science 2015, ESA-ESRIN, Frascati, 20th-23rd October 2015

## Publications from the Project

- Coccia, G., A. Siemann, M. Pan, and E.F. Wood, 2015: Creating consistent datasets combining remotely-sensed data and Land Surface Model estimates through a Bayesian uncertainty post-processing: the case of Land Surface Temperature from HIRS. *Remote Sensing of the Environment*, 170, 290-305, doi:10.1016/j.rse.2015.09.010.
- Chaney, N.W., Herman, J. D., Ek, M. , and Wood, E. F., 2016. Deriving global parameter estimates for the Noah land surface model using FLUXNET and Machine Learning, *J Geophys. Res. – Atmsoph.* (in revision).
- Siemann, Amanda L., Gabriele Coccia, Ming Pan, Eric F. Wood. 2016 Development and Analysis of a Long Term, Global, Terrestrial Land Surface Temperature Dataset Based on HIRS Satellite Retrievals, *J Climate,* 29(10):3589-3606, May 2016.

## PI Contact Information

Eric F Wood
Department of Civil and Environmental Engineering
Princeton University
Princeton, NJ 08544
Email: efwood@princeton.edu
Tel: +1-609-258-4675

## References

Baldocchi, D. (2008), Breathing of the Terrestrial Biosphere: Lessons Learned from a Global Network of Carbon Dioxide Flux Measurement Systems, *Aust. J. Bot.*, *56*.

Betts, A. K., F. Chen, K. E. Mitchell, and Z. Janjic (1997), Assessment of the Land Surface and Boundary Layer Models in Two Operational Versions of the NCEP Eta Model Using FIFE Data, *Mon. Weather Rev.*, *125*.

Beven, K. (2006), A manifesto for the equifinality thesis, *J. Hydrol.*, *320*, 18–36.

Chaney, N. W., J. D. Herman, P. M. Reed, and E. F. Wood (2015), Flood and drought hydrologic monitoring: the role of model parameter uncertainty, *Hydrol. Earth Syst. Sci. Discuss.*, *12*(2), 1697–1728.

Chen, F., and Y. Zhang (2009), On the coupling strength between the land surface and the atmosphere: From viewpoint of surface exchange coefficients, *Geophys. Res. Lett.*, *36*(L10404), doi:10.1029/2009GL037980.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts (1996), Modeling of land surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res.*, *101*(D3), 7251–7268.

Chen, F., J. Janjic, and K. E. Mitchell (1997), Impact of Atmospheric Surface-Layer Parameterizations in the New Land-Surface Scheme of the NCEP Mesocale ETA model, *Boundary-Layer Meteorol.*, *85*, 391–421.

Coccia, G., A. Siemann, M. Pan, and E.F. Wood, 2015: Creating consistent datasets combining remotely-sensed data and Land Surface Model estimates through a Bayesian uncertainty post-processing: the

case of Land Surface Temperature from HIRS. *Remote Sensing of the Environment*, 170, 290-305, doi:10.1016/j.rse.2015.09.010.

Ek, M., and L. Mahrt (1991), *OSU 1-D PBL model user's guide*, Corvallis, Oregon.

Ek, M., K. E. Mitchel, Y. Lin, R. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational Eta model, *J. Geophys. Res.*, *108*(D22), 8851.

Galelli, S., and A. Castelletti (2013), Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling, *Hydrol. Earth Syst. Sci.*, *17*, 2669–2684.

Geurts, P., D. Ernst, and L. Wehenkel (2006), Extremely Randomized Trees, *Mach. Learn.*, *63*(1), 3–42.

Gupta, V. H., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, *377*(1-2), 80–91.

Hou, T., Y. Zhu, H. Lu, E. Sudicky, Z. Yu, and F. Ouyang (2015), Parameter sensitivity analysis and optimization of Noah land surface model with field measurements from Huaihe River Basin, China, *Environ. Res. Risk Assess.*, doi:10.1007/s00477-015-1033-5.

Hou, Z., M. Huang, L. R. Leung, G. Lin, and D. M. Ricciuto (2012), Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to the Community Land Model, *J. Geophys. Res.*, *117*.

Jacquemin, B., and J. Noilhan (1990), Sensitivity study and validation of a land surface parameterization using the Hapex-Mobilhy data set, *Boundary-Layer Meteorol.*, *52*, 93–134.

Jarvis, P. G. (1976), The interpretation of the Variations in Leaf Water Potential and Stomatal Conductance Found in Canopies in the Field, *Philos. Trans. B*, *370*(1665), doi:10.1098/rstb.1976.0035.

Jung M., M. Reichstein, H. A. Margolis, A. Cescatti, A. D. Richardson, M. A. Arain, A. Arneth, C. Bernhofer, D. Bonal, J. Chen, D. Gianelle, N. Gobron, G. Kiely, W. Kutsch, G. Lasslop, B. E. Law, A. Lindroth, L. Merbold, L. Montagnani, E. J. Moors, D. Papale, M. Sottocornola, F. Vaccari and C. Williams.2011. "Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological estimates." *Journal of Geophysical Research* 116. doi:10.1029/2010JG001566

Koren, V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, and J. M. Baker (1999), A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, *104*(D16), 19569–19585, doi:10.1029/1999JD900232.

L'Ecuyer T. S., H. K. Beaudoing, M. Rodell, W. Olson, B. Lin, S. Kato, C. A. Clayson, E. Wood, J. Sheffield, R. Adler, G. Huffman, M. Bosilovich, G. Gu, F. Robertson, P. Houser, D. Chambers, J. S. Famiglietti, E. Fetzer, W. T. Liu, X. Gao, C. A. Schlosser, E. Clark, D. P. Lettenmaier and K. Hilburn. 2015. "The observed state of the energy budget in the early twenty-first century." *Journal of Climate* 28: 8319-8346. doi:10.1175/JCLI-D-14-00556.1

Mahrt, L., and M. Ek (1984), The Influence of Atmospheric Stability on Potential Evaporation, *J. Clim. Appl. Meteorol.*, *23*, 222–234.

Mahrt, L., and H. Pan (1984), A Two-layer model of soil hydrology, *Boundary-Layer Meteorol.*, *29*, 1–20.

McKay, M. D., R. J. Beckman, and W. J. Conover (1979), A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, *21*(2), 239–245.

Monin, A. S., and A. M. Obukhov (1954), Basic laws of turbulent mixing in the ground layer of the atmosphere (in Russian), *Tr. Akad. Nauk SSSR Geophiz. Inst.*, *24*(151), 163–187.

ORNL DAAC (2013), FLUXNET Web Page, Available from: http://fluxnet.ornl.gov

Pan, H.-L., and L. Mahrt (1987), Interaction betwen soil hydrology and boundary-layer development, *Boundary-Layer Meteorol.*, *38*, 185–202.

Rosero, E., Z. Yang, T. Wagener, L. E. Gulden, S. Yatheendradas, and G. Niu (2010), Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, *J. Geophys. Res. Atmos.*, *115*(D3), D03106.

Saltelli, A. (2002), Making best use of model evaluations to compute sensitivity indices, *Comput. Phys. Commun.*, *145*(2), 280–297.

Saltelli, A., P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola (2010), Variance Based Sensitivity Analysis of Model Output. Design and Estimator for the Total Sensitivity Index, *Comput. Phys. Commun.*, *181*(2), 259–270.

Schaake, J. C., V. I. Koren, and Q. Y. Duan (1996), Simple water balance model for estimating runoff at different spatial and temporal scales, *J. Geophys. Res.*, *101*(D3), 7461–7475.

Siemann, A. L., G. Coccia, M. Pan, and E. F. Wood. 2016. "Development and analysis of a long-term, global, terrestrial land surface temperature dataset based on HIRS satellite retrievals." *Journal of Climate* 29: 3589-3606. doi:10.1175/JCLI-D-15-0378.1

Sobol, I. M. (1993), Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. Exp.*, *1*, 407–417.

Van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2008), Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, *44*(W01429), doi:10.1029/2007WR006271.

Wang, A., and X. Zeng. 2013. "Development of global hourly 0.5° land surface air temperature datasets." *Journal of Climate* 26: 7676-7691. doi:10.1175/JCLI-D-12-00682.1.

Wild, M., D. Folini, M. Z. Hakuba, C. Schar, S. I. Seneviratne, S. Kato, D. Rutan, C. Ammann, E. F. Wood, and G. Konig-Langlo. 2015. "The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models." *Climate Dynamics* 44: 3393-3429. doi:10.1007/s00382-014-2430-z.