# Development and evaluation of global land surface evaporative flux records

<u>M. F. McCabe</u><sup>a</sup>, A. Ershadi<sup>a</sup>, C. Jimenez<sup>b</sup>, D Michel<sup>c</sup>, D. G. Miralles<sup>d,e</sup> and E.F. Wood<sup>f</sup>

<sup>a</sup> Division of Biological and Environmental Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia <sup>b</sup> Estellus, Paris, France

<sup>c</sup> ETH Zurich, Switzerland

<sup>d</sup> Department of Earth Sciences, VU University Amsterdam, Amsterdam, The Netherlands <sup>e</sup> Laboratory of Hydrology and Water Management, University of Ghent, Ghent, Belgium

<sup>f</sup>Department of Civil and Environmental Engineering, Princeton University, Princeton, United States of

America

Email: matthew.mccabe@kaust.edu.sa

Abstract: Understanding the spatial and temporal behavior of evaporation at regional and global scales is important for improving our understanding of the water, energy and carbon cycles. In recent years, there have been a number of efforts to develop global products of land surface heat flux, spanning a range of space and time scales and utilising a variety of different forcing and underlying model structures. In the context of the GEWEX-LandFlux and WACMOS-ET projects, the efforts have been guided by a need for independent and observation driven predictions of the heat fluxes, using remotely sensed data where available and supplementing with other forcing when required. The GEWEX based project has provided long-term records of fluxes that will enable an assessment of trends and variability to be examined, while the WACMOS-ET project, operating over a shorter period, has offered insights into the inter-model variability of flux response over many regions. These data sets provide a means to better understand the cycling of water through the Earth system, examining not only the absolute range of flux observed, but also how this changes in space and time. Here we will present some of the key outcomes and findings of these efforts, together with some of the lessons learned in delivering long-term, consistent flux records. Results focusing on examining the multimodel and multi-scale assessment of flux products will be presented, together with an evaluation of product differences and areas where future work is required to improve flux characterization. Issues of product independence, model structure, forcing data and model sensitivity will also be explored.

Keywords: LandFlux, WACMOS-ET, GEWEX, latent heat flux, global evaporation

# 1. INTRODUCTION

The long-term and large-scale estimation of land surface heat fluxes is a topic of considerable interest to the Earth and environmental science community, as it is plays an important role in characterizing the water, energy and carbon cycles. Over the last few years, significant developments in the production of regional and global data sets based largely on remote sensing satellite systems has provided a wealth of information to describe a range of Earth system processes. For instance, spatially and temporally distributed maps of soil moisture (Liu et al., 2011a), vegetation (Liu et al., 2011b; Liu et al., 2013) and precipitation (Nesbitt et al., 2004) have provided enhanced descriptions of our understanding of regional and global scale water movement and distribution (Pan et al., 2008; McCabe et al., 2008). Understanding the spatial and temporal variability of such variables is particularly important in the context of advances in regional and global climate model simulations (Evans and McCabe, 2010), which provide a capacity to simulate components of the water and energy cycle in great detail. In order to gain greater confidence in the products of regional and global models, results need to be evaluated against independent observations or at least benchmarked against independent model results.

The need for assessment and evaluation of regional (El Kenawy et al., 2015) and global model (El Kenawy and McCabe, 2015; Mueller et al., 2011) simulations, as well as to provide an observationally based capacity for characterising the trends and variability in water and energy cycles, has provided the motivation for recent efforts towards developing long term global land surface heat flux estimates. The LandFlux project (Jiménez et al., 2009; Mueller et al., 2013), developed out of the Global Energy and Water Exchanges (GEWEX) Data and Assessments Panel (GDAP), together with the European Space Agencies (ESA) Water Cycle Observation Multi-mission Strategy – Evapotranspiration (WACMOS-ET) (Miralles et al., 2015), are two such efforts that have sought to provide regional to global scale land surface flux data sets. These projects have undertaken parallel but complimentary efforts towards providing observationally driven flux response across a range of spatial and temporal scales. Using largely the same process model descriptions, the project outputs also share similar forcing data sets, allowing an intercomparison of flux response to be undertaken. Importantly, the challenges and lessons learned from these exercises overlap, providing an opportunity to reflect on needed improvements, establish inherent limitations and guide future efforts to refine global products. Here we will review the findings from these recent products, focusing on a range of evaluation metrics and including assessments at the local, regional and continental scales.

# 2. MODEL DESCRIPTIONS AND FORCING DATA

There are a wide range of approaches available for monitoring surface heat fluxes from space and much research has been directed towards the development of such techniques over the last few decades (Kalma et al., 2008; Wang and Dickinson, 2012). In terms of developing globally distributed products however, the wide range of possible estimation approaches collapses quite rapidly, as forcing data requirements, model theoretical limitations and physical restrictions impose application constraints. Indeed, of those approaches that have attempted to develop global land surface evaporation products, the bulk are based on either Penman-Monteith or Priestley-Taylor type approaches. For the LandFlux and WACMOS-ET projects, four models were examined for analysis, including the Penman-Monteith based Mu model (PM-Mu) (Mu et al., 2011), the Priestley-Taylor-Jet Propulsion Laboratory (PT-JPL) model (Fisher et al., 2008), the Global Land surface Evaporation: the Amsterdam Method (GLEAM) (Miralles et al., 2011) and the Surface Energy Balance System (SEBS; Su, 2002). Further details on the models and a thorough description of their underlying physical basis and structure can be found in Ershadi et al. (2014) and Michel et al. (2015).

In terms of the data used to run the individual models, both projects aimed for a degree of consistency. For LandFlux, this was focused on external consistency, where data were selected preferentially from existing GEWEX forcing sets developed as part of the GDAP. For WACMOS-ET, a stronger preference towards internally consistent forcing was made, with data re-used as much as possible in the calculation of co-related variables. Even within the different projects, there is still a degree of overlap in the forcing data sets, given that there are relatively few available sources for large scale forcing. Table 1 outlines some of the similarities and differences between these data and their sources.

## 3. LOCAL TO GLOBAL SCALE MODEL ASSESSMENT

Understanding the behavior of model response at different scales is a key motivation behind both the LandFlux and WACMOS-ET projects. The inherent assumption that models are scale invariant, in that the big-leaf assumption can be applied equally at the field or catchment scale, is one that is routinely implemented but rarely closely examined. Of course, evaluating the impact of scale on model results,

**Table 1.** Forcing data requirements and sources for the LandFlux and WACMOS-ET products. LandFlux forcing data can be downloaded from <u>http://hydrology.princeton.edu/data/landflux/</u> and the Reference Input Data set (RIDS) of the WACMOS project can be accessed via an online request to <u>http://wacmoset.estellus.eu</u>

Variable	LandFlux Product	WACMOS-ET Product	Model
Air temperature	Princeton Univ (0.5° and 3 hourly)	ERA-Interim(75 km and 3 hourly)	All models
Humidity	Princeton Univ (0.5° and 3 hourly)	ERA-Interim (75 km and 3 hourly)	All but GLEAM
Pressure	Princeton Univ (0.5° and 3 hourly)	ERA-Interim (75 km and 3 hourly)	All models
Net radiation	GEWEX SRB v3 at 1° and 3 hourly	GEWEX SRB v3 at 1° and 3 hourly	All models
Ground heat flux	From net radiation + fractional vegetation cover data, 0.5° and 3-hourly	From net radiation + fractional vegetation cover data, 0.5° and 3-hourly	All models
Land cover	IGBP Land Cover Data set	ESA GlobCover and IGBP	All but GLEAM
NDVI	GIMMS NDVI at 0.5° and bi-monthly	From GlobAlbedo LAI (8 days and 1 km)	All but GLEAM
Leaf area index	Princeton Univ (0.5° and monthly)	From GlobAlbedo (8 days and 1km)	All but GLEAM
Fractional Vegetation	Calculated from NDVI	From GlobAlbedo FAPAR ((8 days and 1km))	All but GLEAM
Precipitation	CFSR-Land (hourly and 40 km)	CFSR-Land (hourly and 40 km)	GLEAM only
Soil moisture thresholds	IGBP-DIS data aggregated to 0.5°	ESA CCI-WACMOS (daily and 25 km)	GLEAM only
Optical depth	ESA CCI-WACMOS (daily and 25 km)	ESA CCI-WACMOS (daily and 25 km)	GLEAM only
Soil moisture	ESA CCI-WACMOS (daily and 25 km)	ESA CCI-WACMOS (daily and 25 km)	GLEAM only
Soil depth	GlobSnow (daily and 25 km)	GlobSnow (daily and 25 km)	GLEAM only
Cover fractions	MOD44B data at 0.5°	MOD44B data at 0.5°	GLEAM only

particularly at scales beyond the field, is exceptionally challenging, due to the lack of validation data at the appropriate scales. Where such data does exist, usually in the form of tower-based eddy-covariance observations, they are generally poorly spatially distributed and do not have extensive temporal extents. Here, quantitative analysis is undertaken at the tower scale, while a more qualitative assessment of model behavior is presented at the regional and global scales. Regardless of the current assessment, additional evaluation metrics are required for greater confidence to be gained in such global scale products.

# 3.1. Flux Comparison Against Tower-based Observations

Although there are well known issues in using tower-based flux observations as a validation source, they still remain the benchmark against which model approaches can be assessed. In the context of both the LandFlux and WACMOS-ET projects, eddy covariance based tower data form the backbone of model assessments. With an expanding network of flux observation available as part of the AmeriFLUX and FluxNet projects (Baldocchi et al., 2001) there is the potential for examination against a range of towers located over differing land surface and climate zones. With this in mind, a collection of tower data was selected as a validation source against which model simulations could be assessed. While there are several hundred available towers globally, the number that could be used here was limited by the fact that towers needed to provide observations required by all models, to ensure consistency in statistical evaluation. For the models used here, surface temperature required by SEBS was the leading constraint.

Figure 1 highlights a sample of the model evaluation undertaken as part of the WACMOS-ET project. Here, 24 towers were used for model assessment, collected from a range of biomes including Mixed Forests, Croplands, Deciduous Broadleaf Forests, Savannas, Evergreen Needleaf Forests, Grasslands, Wood Savannas and Shrublands. Models were run against the tower-based meteorology and then compared against independent observations using the eddy-covariance derived latent heat flux. As can be seen, there is quite a degree of variability between the different model response. SEBS shows the most variability and generally overestimates the flux response relative to the towers, while PM-Mu tends to underestimate. The PT-JPL and GLEAM models, maintain a degree of consistency in evaporative response across the tower locations.



**Figure 1.** Models output using tower-based forcing across different biomes, undertaken as part of the WACMOS-ET assessment (see Michel et al. 2015 for further information on selected towers and the period used). Plot shows the model determined evaporative fraction (EF) using tower-based meteorological observations at 24 tower locations. The grey band encompasses the range between uncorrected eddy-covariance measurements and corrected values using the energy residual approach (i.e. LE = Rn - H - G). Biomes are arranged from wet-to-dry (L-R) based on the biome averaged EF, with individual towers within each biome sorted by their tower-based EF.

While tower-based data provide the benchmark for model assessment, ultimately the global runs are performed at grid-scales, varying anywhere from several-to-tens of kilometers. Although restricted by assumptions related to observation fetch and within-pixel heterogeneity at the scale of simulation, the towers offer a useful means to assess model performance: particularly if run in parallel against tower-driven meteorological forcing. Figure 2 presents such an analysis using a Taylor diagram, with LandFlux based models driven by both tower-based forcing and LandFlux grid-based forcing (see Table 1). In this case, an expanded number of towers were collated, with 45 eddy-covariance observations spread across a globally distributed network and covering a range of biome and climate types. See Ershadi et al. (2014) for further details.

As can be seen, results improve when models are run with tower-based forcing a not entirely unexpected result, but one which can only be confirmed through such an analysis. The Penman-Monteith and Priestley-Taylor based models are **Figure 2.** Taylor diagram showing the LandFlux model runs using meteorological forcing from the towers (red) as well as global gridbased forcing (centered at the tower locations) (blue). Tower data is based on eddy-covariance observations at 45 tower locations (see Miralles et al. 2015 for further information on selected towers and period used). Figure shows the standard deviation, correlation coefficient and the root mean square difference (RMSD).

relatively tightly grouped, with PM-Mu reflecting slightly poorer statistics than the PT-JPL and GLEAM models. SEBS is an outlier with higher standard deviations in both tower and grid based evaluations. GLEAM performs slightly better at the tower and slightly worse at the grid scale compared to PT-JPL.

#### 3.2. Regional Scale Simulation over Australia's Murray Darling Basin

While local scale assessment is critically important to analyzing the different model performance over a range of land surface and climate types, regional scale assessment provides a capacity to better understand larger scale features and impacts of linked hydrological processes (McCabe et al., 2005). Here we examine the different model responses over the Murray Darling Basin (MDB) in southeast Australia, focusing on results from the WACMOS-ET simulations for the year 2007. Figure 3 presents seasonal maps of these

models over the basin, allowing for an intercomparison of model behavior. Although no in-situ data is used to evaluate the models in this assessment (the limited footprint of flux tower data limits the insight gained when comparing against 25 km or larger grid response) the spatial patterns themselves offer some guidance on model performance. For instance, the transition from a drier interior to a wetter (and hence more evaporative) eastern boundary is reflected well in both the PT-JPL and GLEAM models simulations, especially in January-March (JFM) and October to December (OND), when greater rainfall amounts are expected across the eastern seaboard (Evans and McCabe, 2013). On the other hand, PM-Mu tends to underestimate the evaporative behavior, both in spatial pattern and temporal response (Miralles et al. 2015).

Interesting spatial features are evident across the different model simulations, including not just the east-west rainfall transition, but also a north-south effect in JFM and OND particularly, perhaps as a result of mesoscale atmospheric processes during that time. The lack of spatial structure in the PM-Mu seasonal response is also instructive. Given that the forcing requirements of the PM approach are identical to the PT-JPL scheme, it is likely that this suppressed spatial pattern is a result of model structure and parameterizations. The sensitivity of PM models to parameterization of their resistance schemes has been explored recently (Ershadi et al., 2015b), highlighting the challenges of developing appropriate model structures that capture the range of biome and climate types present in global applications.

#### **3.3.** Model Intercomparison at the Global Scale

A goal of both the LandFlux and WACMOS-ET projects was the production of global evaporation data sets spanning a range of observation periods. For WACMOS-ET, the focus was on the period from 2005-2007 and covered both local, regional and global scale evaluation with multi-resolution model runs. The LandFlux data set on the other hand, has a focus on long-term simulations at a relatively coarse 1-degree spatial resolution, making it suitable for comparison against global climate model results.

From Figure 4 it can be seen that while there is broad agreement at the global scale, there are obvious



Figure 3. Seasonally averaged maps of evaporation over the Murray Darling Basin for the year 2007. Models include the PT- JPL, PM-Mu approach and the GLEAM schemes. Colour bar is in units of mm.day<sup>-1</sup>.

regional scale differences between the modeling approaches, in spite of the fact that the models have all been run using a common forcing data set. Such a result is expected, since there are model specific forcing data requirements between the different approaches (see Table 1) as well as model parameterizations that are unique to each particular scheme. Considering developer-driven model assumptions and implementations, it is apparent that even where the underlying physical basis may be similar, obtaining consistent agreement between modeling schemes is a challenging task.

One area in particular where models seem to diverge in their agreement is over Australia. While the spatial pattern, if not the absolute magnitude, of evaporation response is relatively well preserved over places such as the Amazon and South America, the spatial response varies considerably over Australian or Africa. In the case of the PM-Mu approach, there is a significant underestimation of evaporation over much of the Australian continent. Likewise, the PT-JPL approach tends to homogenize the evaporative response, especially in the coastal fringes and extending inland, where rainfall and evaporation should vary the most. Clearly further work is required to both understand and improve this variable model behavior.

### 4. CONCLUSIONS AND FUTURE WORK

Global-scale estimation of evaporative fluxes is a both a straightforward and a challenging task. It is straightforward in that the needed model forcing data from numerical weather prediction or operational

products are more widely available than ever before, in addition to a range of satellite based products that can be used to supplement model data requirements. Computational constraints are also largely removed, with parallel processing and high-performance computers allowing rapid and repeatable simulations in ways not

previously available to the research community. However, the ease with which products can be developed belies the challenge in implementing robust and coherent simulations. Assumptions on the validity of using models developed at the local scale for global simulation, uncertainties associated with using inconsistent forcing data sets, impacts of using model parameterizations not tested for different land surface and climate types and the inability to easily evaluate subsequent model output all seek to confound global flux monitoring efforts. Two difficulties in developing such products lay in 1) understanding the inherent forcing data uncertainties and 2) quantifying the subsequent impacts of these on the model response. To address these points effectively requires a coordinated approach to develop a truly consistent forcing data set, where atmospheric and surface variables such as air and land temperature are reflected in net longwave calculations, or where the albedo is derived in concert with vegetation parameters that in turn influence shortwave radiation estimates. Only then can effort be directed to disentangle the relative influences of different forcing data on individual model response: a process best undertaken within a comprehensive sensitivity analysis process. Once addressed, a clearer understanding of the retrieval accuracy of flux estimates will be gained. The challenge then will be to develop efficient and appropriate evaluation metrics that span spatial and temporal scales, as the current focus on using tower-based data is clearly inadequate.



**Figure 4.** Globally distributed estimates of evaporation produced by (from top-bottom) PM-Mu, GLEAM and PT-JPL model simulations using the WACMOS-ET reference input data set for the years 2005-2007 (see Miralles et al. 2015 for further details). Maps represent the three-year product average, with total annual evaporation in mm.year<sup>-1</sup>.

#### ACKNOWLEDGEMENTS

M.F.M and A.E. were supported by the King Abdullah University of Science and Technology (KAUST) and the European Space Agency through the WACMOS-ET project. D.G.M. acknowledges financial support from The Netherlands Organisation for Scientific Research (NWO) Veni grant 863.14.004.

#### REFERENCES

- Baldocchi, D., E. Falge, L. Gu, R. Olson, D. Hollinger, S. Running, . . . S. Wofsy (2001). "FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem–Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities." Bulletin of the American Meteorological Society 82(11): 2415-2434.
- El Kenawy, A., J. I. López-Moreno, M. F. McCabe, N. A. Brunsell and S. M. Vicente-Serrano (2015). "Daily temperature changes and variability in ENSEMBLES regional models predictions: Evaluation and intercomparison for the Ebro Valley (NE Iberia)." Atmospheric Research 155: 141-157.
- El Kenawy, A. M. and M. F. McCabe (2015). "A multi-decadal assessment of the performance of gauge- and modelbased rainfall products over Saudi Arabia: Climatology, anomalies and trends." International Journal of Climatology.
- Ershadi, A., M. F. McCabe, J. P. Evans, N. W. Chaney and E. F. Wood (2014). "Multi-site evaluation of terrestrial evaporation models using FLUXNET data." Agricultural and Forest Meteorology 187: 46-61.

- Ershadi, A., M. F. McCabe, J. P. Evans and E. F. Wood (2015). "Impact of model structure and parameterization on Penman–Monteith type evaporation models." Journal of Hydrology 525(0): 521-535.
- Evans, J. P. and M. F. McCabe (2010). "Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment." Journal of Geophysical Research: Atmospheres 115(14).
- Evans, J. P. and M. F. McCabe (2013). "Effect of model resolution on a regional climate model simulation over southeast Australia." Climate Research 56(2): 131-145.
- Jiménez, C., M. McCabe, S. I. Seneviratne, E. Wood, and W. Rossow (2012) 4th LandFlux Workshop, GEWEX News, 22(4), 17-18.
- Kalma, J. D., T. R. McVicar and M. F. McCabe (2008). "Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data." Surveys in Geophysics 29(4-5): 421-469.
- Liu, Y. Y., R. A. M. de Jeu, M. F. McCabe, J. P. Evans and A. I. J. M. van Dijk (2011a). "Global long-term passive microwave satellite-based retrievals of vegetation optical depth." Geophys. Res. Lett. 38(18): L18402.
- Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. De Jeu, W. Wagner, A. I. J. M. Van Dijk, M. F. McCabe, J. P. Evans (2011b). "Developing an improved soil moisture data set by blending passive and active microwave satellite-based retrievals." Hydrology and Earth System Sciences 15(2): 425-436.
- Liu, Y. Y., A. I. J. M. van Dijk, M. F. McCabe, J. P. Evans and R. A. M. de Jeu (2013). "Global vegetation biomass change (1988-2008) and attribution to environmental and human drivers." Global Ecology and Biogeography 22(6): 692-705.
- McCabe, M. F., E. F. Wood and H. Gao (2005). "Initial soil moisture retrievals from AMSR-E: Multiscale comparison using in situ data and rainfall patterns overs Iowa." Geophysical Research Letters 32(6): 1-4.
- McCabe, M. F., E. F. Wood, R. Wójcik, M. Pan, J. Sheffield, H. Gao and H. Su (2008). "Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies." Remote Sensing of Environment 112(2): 430-444.
- McCabe, M. F., A Ershadi, C. Jimenez, D. G. Miralles, D. Michel and E. F. Wood (2015). "The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally-gridded forcing data" Geoscientific and Model Development Discussions, 8, 6809-6866, doi:10.5194/gmdd-8-6809-2015.
- Michel, D., C. Jimenez, D. M. Miralles, M. Jung, M. Hirschi, A. Ershadi, B. Martens, M. F. McCabe, J. B. Fisher, Q. Mu, S. I. Seneviratne, E. F. Wood and D. Fernandez-Prieto (2015). "The WACMOS-ET project – Part 1: towerscale evaluation of four remote sensing-based evapotranspiration algorithms", Hydrology and Earth System Sciences Discussions (under review)
- Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters and A. J. Dolman (2011). "Global land-surface evaporation estimated from satellite-based observations." Hydrol. Earth Syst. Sci. 15(2): 453-469.
- Miralles, D. G., C. Jimenez, M. Jung, D. Michel, A. Ershadi, M. F. McCabe, M. Hirschi, A. J. Dolman, J. B. Fisher, B. Martens, Q. Mu, S. I. Seneviratne, E. F. Wood and D. Fernandez-Prieto (2015). "The WACMOS-ET project Part 2: Evaluation of global terrestrial evaporation data sets." Hydrology and Earth System Sciences Discussions (under review).
- Mu, Q., M. Zhao and S. W. Running (2011). "Improvements to a MODIS global terrestrial evapotranspiration algorithm." Remote Sensing of Environment 115(8): 1781-1800.
- Mueller, B., S. I. Seneviratne, C. Jimenez, T. Corti, M. Hirschi, G. Balsamo, P. Ciais, P. Dirmeyer, J. B. Fisher, Z. Guo, M. Jung, F. Maignan, M. F. McCabe, R. Reichle, M. Reichstein, M. Rodell, J. Sheffield, A. J. Teuling, K. Wang, E. F. Wood, and Y. Zhang (2011). "Evaluation of global observations-based evapotranspiration data sets and IPCC AR4 simulations." Geophysical Research Letters 38(6).
- Mueller, B., M. Hirschi, C. Jimenez, P. Ciais, P. A. Dirmeyer, A. J. Dolman, J. B. Fisher, M. Jung, F. Ludwig, F. Maignan, D. G. Miralles, M. F. McCabe, M. Reichstein, J. Sheffield, K. Wang, E. F. Wood, Y. Zhang, and S. I. Seneviratne (2013). "Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis." <u>Hydrology and Earth System Sciences</u> 17(10): 3707-3720.
- Nesbitt, S. W., E. J. Zipser and C. D. Kummerow (2004). "An examination of version-5 rainfall estimates from the TRMM Microwave Imager, precipitation radar, and rain gauges on global, regional, and storm scales." J. Appl. Meteorol. 43(7): 1016-1036.
- Pan, M., E. F. Wood, R. Wójcik and M. F. McCabe (2008). "Estimation of regional terrestrial water cycle using multisensor remote sensing observations and data assimilation." Remote Sensing of Environment 112(4): 1282-1294.
- Su, Z. (2002). "The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes." Hydrol. Earth Syst. Sci. 6(1): 85-100.