Validation of the MODIS LAI product in the Murrumbidgee Catchment, Australia

K.A. McColl [a], R.C. Pipunic [a], D. Ryu [a] and J.P. Walker [b]

[a] *Department of Infrastructure Engineering, The University of Melbourne, Victoria, Australia*
[b] *Department of Civil Engineering, Monash University, Victoria, Australia*
Email: robp@unimelb.edu.au

Abstract: Terrestrial vegetation plays an important role in the water, energy and carbon cycles at the land surface. Leaf Area Index (LAI), a measure of canopy structure, is an important biophysical parameter in ecosystem models which represent these cycles. However, field observations of LAI are rarely available at the resolution and coverage required for regional- or global-scale ecosystem modelling. Poor LAI estimates will introduce greater uncertainty into models, potentially compromising the accuracy of energy balance partitioning and related fluxes such as evapotranspiration. The MODerate resolution Imaging Spectroradiometer (MODIS) uses surface reflectance observations to provide 8-day composites of remotely-sensed LAI estimates at a nominal spatial resolution of 1-km. However, even though various sources of estimation error exist, such as sub-pixel heterogeneity and approximations in retrieval algorithms, error estimates are not provided with the remotely-sensed LAI products. Previous validation studies in the USA and Australia have yielded contradictory results, with one demonstrating a positive bias in the MODIS LAI product, and the other a negative bias. In this paper, the latest MODIS LAI estimates (MYD15A2) are compared to Australian field observations across the Murrumbidgee Catchment in New South Wales. The ground observations were collected during three field campaigns: the National Airborne Field Experiment in November 2006 (NAFE'06), the Australian Airborne Cal/val Experiments for SMOS (AACES) in January/February and September 2010 and the Soil Moisture Active Passive Experiments (SMAPEx) in July and December 2011. Using these data, we conduct a characterization of the error properties of the MODIS LAI product using ground observations over varying land surface conditions and seasons. The results show that an assumption of Gaussian LAI errors for this region is reasonable. We also show that MODIS overestimates LAI for lightly vegetated regions (LAI < ~0.6) but underestimates LAI for densely vegetated regions (LAI > ~0.6). This result suggests that the contradictory results from previous validation studies may be associated with the different ranges of LAI observed in the studies. Overall, the bias in MODIS LAI is approximately -0.35 and the standard deviation of the error is 0.82. Observed LAI error properties from this study are expected to improve model predictions of the land-surface energy balance in ensemble model prediction studies underway for the Murrumbidgee Catchment.

Keywords: LAI, MODIS, remote sensing, validation, Murrumbidgee

1. INTRODUCTION

Terrestrial vegetation plays an important role in the water, energy and carbon cycles at the land surface. Vegetation intercepts solar radiation and rainfall that would otherwise reach the soil; it generates turbulence by disrupting boundary-layer airflow, which promotes turbulent sensible and latent heat fluxes; and controls evapotranspiration, which links all three cycles. The complex interaction of these cycles results in feedbacks that affect the climate (Sellers et al. 1997). Understanding the role of vegetation in these cycles is crucial in predicting the effects of global change, including those due to global warming and landcover change.

To meet this challenge, increasingly sophisticated ecosystem models representing these cycles have been developed (e.g., Kowalczyk et al. 2006). However, they require large-scale, high-resolution datasets for determining key parameters (such as landcover type), and for validating their outputs. Furthermore, data assimilation studies, where observations and model estimates are optimally combined to yield better predictions, require estimates of observation uncertainty (Quaife et al. 2008). Obtaining these data from ground observations alone is infeasible. Earth Observing (EO) satellite missions, such as the MODerate resolution Imaging Spectroradiometer (MODIS), provide consistent global observations of biophysical variables relevant to ecosystem models.

One such variable is Leaf Area Index (LAI). LAI is defined as the one-sided green leaf area per unit ground surface area (-) in broadleaf canopies, and the maximum projected area in coniferous canopies (R. B. Myneni et al. 1997). Uncertainties in LAI have been shown to result in large errors in land-surface model estimates of heat fluxes and carbon fluxes. For example, Schwinger et al. (2010) found the percentage error in LAI approximately matched the maximum percentage error in latent heat flux estimates; Quaife et al. (2008) found an average increase in carbon flux standard deviation of 177 g C m^{-2} y^{-1} without assimilation of LAI. MODIS uses surface reflectance observations to provide 8-day composite remotely-sensed estimates of LAI at a nominal spatial resolution of 1-km. However, estimation errors are introduced due to sub-pixel heterogeneity, approximations in retrieval algorithms and other factors. Current understanding of the error properties of the LAI product, which is needed in ensemble prediction studies, is very limited due to the lack of validation studies for the product. Existing studies report contradictory results about the collection 5 MODIS LAI product. One study, conducted in spring (May-June) over a mixed coniferous forest in Oregon, USA, found the MODIS product underestimates true LAI by an average of 1.18 (De Kauwe et al. 2011). Another, conducted in the dry season (September) over savanna in the Northern Territory, Australia, found it overestimated LAI by an average of 0.15 (Sea et al. 2011).

This study compares MODIS LAI observations with ground observations from the Murrumbidgee Catchment, New South Wales. This comparison updates results from a preceding study, undertaken just outside the catchment (Leuning et al. 2005), in so much as the MODIS LAI retrieval algorithms have been altered significantly (from 'collection 4' to 'collection 5'), necessitating this new comparison. Only one validation study of the collection 5 product has been undertaken in Australia, at a site in the Northern Territory (Sea et al. 2011) with very little vegetation. This study assesses the collection 5 MODIS LAI product over a range of landcover types and seasons, providing a more comprehensive dataset for validation.

2. METHOD

2.1. Study Area

The LAI observations used in this study were taken across the Murrumbidgee River Catchment (33-37°S, 143-150°E), a sub-catchment of the Murray-Darling Basin in New South Wales, Australia (Figure 1). The catchment spans 84,000 km^2, and has a diverse climate, ranging from semi-arid to alpine. Annual rainfall varies between 400-600 mm in the western lowlands, and up to 1300 mm in the eastern mountainous areas. The land is predominantly used for agriculture: grazing, rice, fruit and various crops.

2.2. Data

The ground data were collected using the LiCOR LAI-2000 (LiCOR Biosciences) during three field campaigns: the National Airborne Field Experiment 2006 (NAFE'06, Merlin et al., 2008), the Australian Airborne Cal/Val Experiments for SMOS (AACES, Peischl et al., 2011) and the Soil Moisture Active Passive Experiments (SMAPEx, http://www.smapex.monash.edu.au/). AACES comprised two separate campaigns over summer (AACES-1, with LAI measurements taken over the period 21/01/2010-18/02/2010) and spring (AACES-2, 12/09/2010-19/09/2010). SMAPEx comprised two separate campaigns over winter (SMAPEx-1, 06/07/2010-10/07/2010) and summer (SMAPEx-2, 04/12/2010-08/12/2010). LAI observations were collected in NAFE'06 over the period 30/10/2006-17/11/2006. Ground sampling points for all
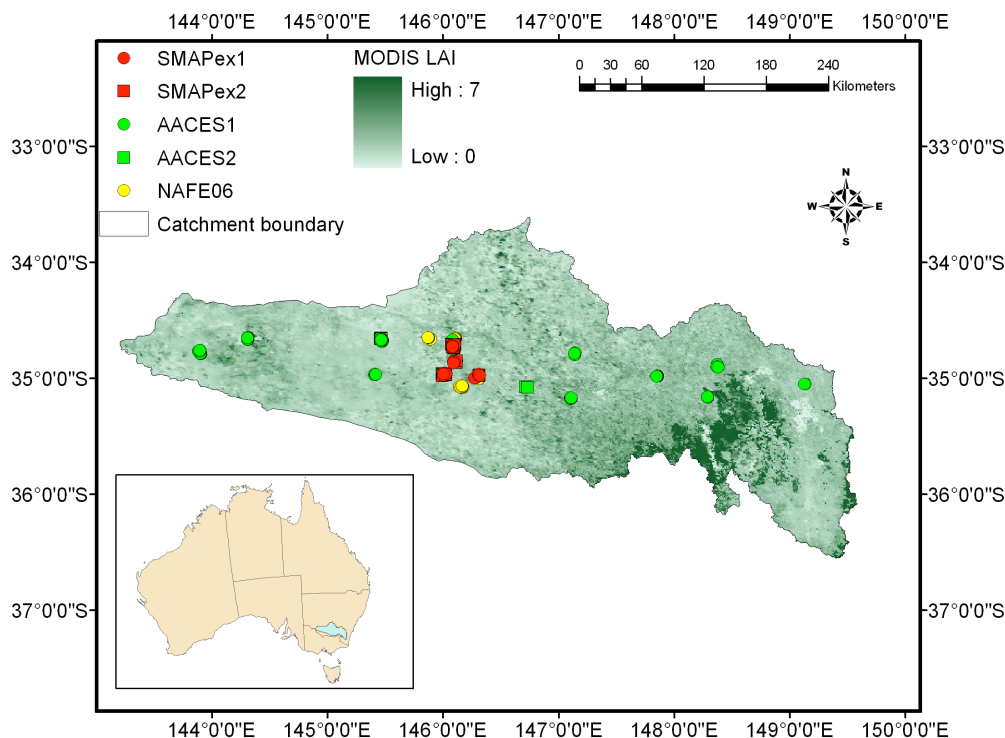
Figure 1. The study site in the Murrumbidgee Catchment, showing MODIS LAI estimates for July 5, 2010. LAI sampling sites from the SMAPEx, AACES and NAFE06 field experiments are also given.

campaigns are shown in Figure 1, with landcover types given in Table 1. Together, these data capture LAI across a range of seasons and landcover types within the catchment.

The MODIS collection 5 product MYD15A2 (Aqua) LAI/fPAR (fraction of Photosynthetically Active Radiation) is recorded at a nominal spatial resolution of ~1-km (0.0094°) as an 8-day composite. It obtains LAI using bidirectional reflectances, correcting for landcover and atmospheric effects (Knyazikhin et al. 1999). MODIS LAI product over the study area was collected over periods corresponding to the field experiments, and linearly disaggregated from 8-day to daily. In addition, 250-m MODIS NDVI (Normalized Difference Vegetation Index) observations were obtained as 16-day composites, and also linearly disaggregated from 16-day to daily.

3. RESULTS AND DISCUSSION

A simple average of LAI ground observations (LAI_g) was compared to the MODIS estimate (LAI_{MODIS}) for the 155 MODIS pixels coincident with ground data through both space and time. These results are shown in Figure 2. In general, the MODIS product overestimates LAI at lower values, and underestimates it at higher values.

Table **1**. Number of LAI samples collected for different landcover types over the study period.

Landcover type	Number of LAI samples
Woodland	28
Open shrubland	107
Pasture	97
Cropland	160
Barren/Fallow	9
Total	401

The range of vegetation conditions observed over our study site allows us to reconcile conflicting results from previous studies. Sea et al. (2011) found that the MODIS product was positively biased over lightly vegetated savanna. Yet, De Kauwe et al. (2011) found a negative bias over more densely vegetated forested regions. The reason for this discrepancy is the difference in vegetation conditions between the two studies. Sea et al. (2011) measured LAI in lightly-vegetated savannah; De Kauwe et al. (2011) measured it in much more highly-vegetated mixed coniferous forest. Results from our study site, which spans a range of seasons and agricultural landcover types, confirms the results of both studies for their respective vegetation conditions. The MODIS product is negatively-biased for most vegetated areas (with the bias increasing with increasing levels of LAI). However, for areas with LAI < ~0.6, the MODIS product is positively-biased. This bias is due in part to an apparent arbitrary minimum, non-zero value for the MODIS product. We observed 19 of the 155 points with the identical value of 0.200000002980232 in this comparison (they have been removed from Figure 3 but retained in Figure 2). Overall, there is more variation in our results compared to previous
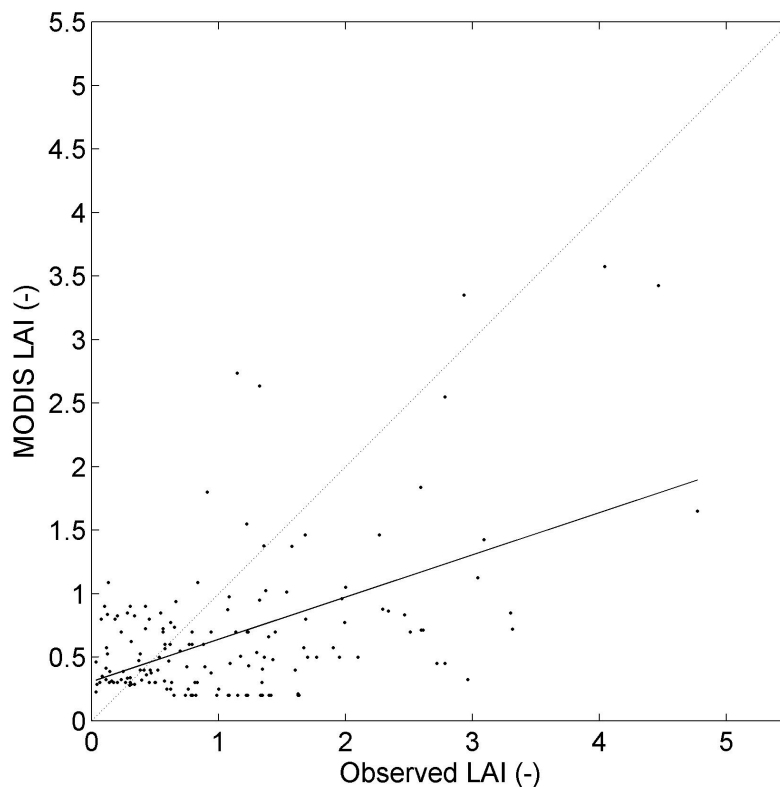

Figure 2. Ground observations of LAI (averaged within a 1-km MODIS pixel) compared to the MODIS LAI product, for 155 pixels. The dotted line is the 1:1 line; the solid line is the linear regression line ($R^2 = 0.27$; RMSE = 0.90).

studies, resulting in a poor fit of the linear regression line ($R^2 = 0.27$; RMSE = 0.90). This is associated with the fact that our data have been chosen to span a range of seasons and landcover types, and thus vegetation conditions.

Characterizing the error probability distribution is essential if MODIS LAI is to be used in ensemble ecosystem models. The error distribution was estimated across the study site (Figure 3), by calculating $LAI_{MODIS} - LAI_g$ for each 1-km pixel containing ground observations. Pixels with MODIS values at the artificial lower bound of the product were excluded to prevent artificial distortion of the calculated error distribution, resulting in $N = 136$ 1-km pixels. The sample mean error is -0.35, significantly different to zero (ascertained using a one-sample t-test with $p = 2e-6$). The sample standard deviation is 0.82. The Yule-Kendall Index, a robust and resistant estimator of skewness that uses quartiles to reduce the impact of outliers, was calculated to be -0.33. This is smaller than two standard errors of skewness (calculated as $(6/N)^{0.5}$, a typical estimate of the standard error of skewness) from zero, suggesting this value is not significant. Combined with a visual inspection of the histogram, the results indicate that we can assume normally distributed MODIS LAI errors in ensemble prediction studies.
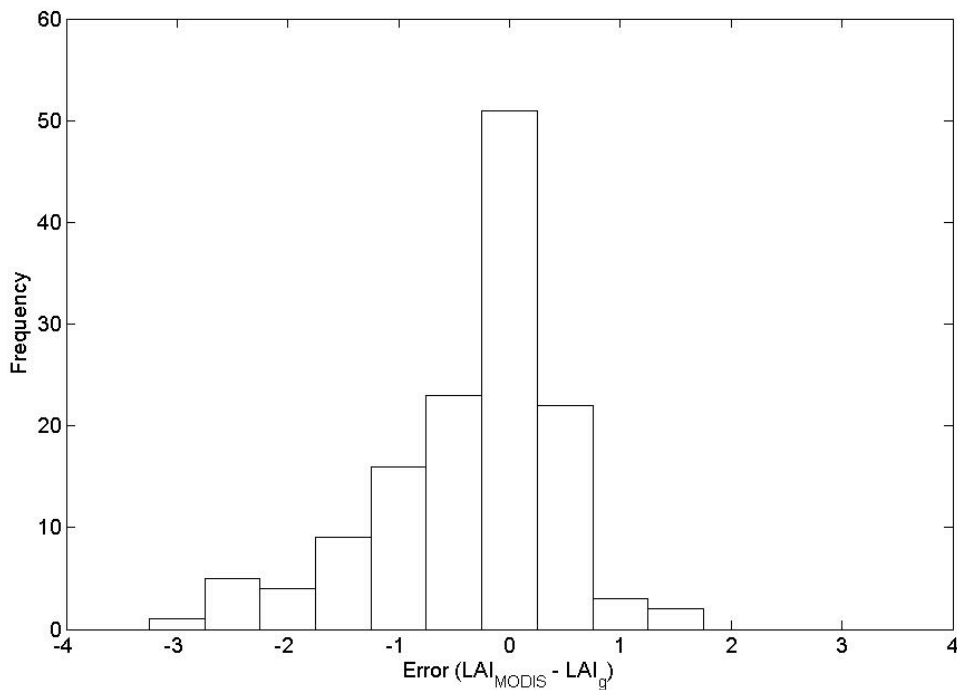
Figure 3. Histogram of errors (excluding MODIS observations that are truncated at the MODIS LAI product's minimum bound).

Apart from errors in the MODIS product itself, the error budget is also likely to include errors in the ground data, which are subject to instrument errors (e.g., eucalypt forests have different leaf angle distributions compared to northern-hemisphere woodlands for which the LiCOR is designed) and representativeness errors, particularly where only a small number of ground samples are used to estimate the true LAI value for a 1-km satellite estimate. Sampling errors may bias the estimate if ground sampling is non-random or the sample size is too small. All of these factors need to be considered in estimating the error distribution. Figure 4 (left) shows that maximum (minimum) errors decrease (increase) $\sim n^{-0.5}$ as n increases (where n is the number of ground samples in each 1-km pixel) suggesting ground sampling error has contributed to the error estimate. However, the ground sampling error should not contribute to the observed bias in ($LAI_{MODIS} - LAI_g$) since the ground observations appear to be randomly distributed within each pixel, and the sample size is relatively large.

These results suggest the spatial heterogeneity of the MODIS pixel is not fully captured by the ground data. To further investigate the dependence of errors on spatial heterogeneity of vegetation, we used the standard deviation of 250-m MODIS NDVI as a proxy for the spatial heterogeneity within a 1-km MODIS pixel. Comparing this to observed errors revealed no correlation (Figure 4, right). This implies that the 250-m MODIS NDVI did not successfully capture the spatial heterogeneity of vegetation in the sampling sites. One possible reason is that the spatial heterogeneity causing the sampling errors has a correlation length < 250 m.

Despite the larger sampling errors, these data match well the results of previous independent studies. The MODIS product is positively-biased for very low LAI values but negatively biased for higher values. We estimate the overall bias to be -0.35. This appears reasonable when compared to the -1.18 bias estimated for woodland (De Kauwe et al. 2011) and the 0.15 estimated for low vegetation savannah (Sea et al. 2011), given our observations span a range of vegetation conditions between these two extremes. Importantly, we also identify a minimum bound on the MODIS product that is partially responsible for the positive bias at low values.

The MODIS LAI shown in this work will be used in an ensemble land-surface model in subsequent studies. However, the model uses a 5-km grid, meaning the error statistics from this study cannot necessarily be applied directly. The effects of upscaling were investigated by comparing MODIS LAI error calculated at 1-

km ($\varepsilon_{1\text{-km}}$) and 5-km ($\varepsilon_{5\text{-km}}$) scales. For each 1-km cell, $\varepsilon_{1\text{-km}}$ was calculated as $\text{LAI}_{\text{MODIS}} - \text{LAI}_{\text{g}}$, and then averaged over a 5-km pixel. This value was compared to the value $\varepsilon_{5\text{-km}}$ obtained by averaging $\text{LAI}_{\text{MODIS}}$ over a 5-km pixel, and then calculating the error. The resulting errors were strongly correlated ($R^2 = 0.83$) and
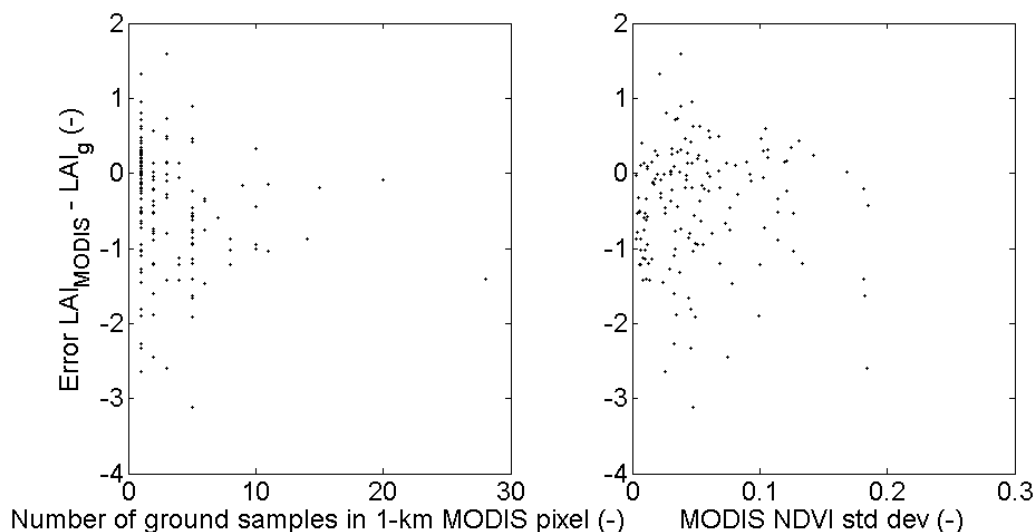

Figure 4. Observed error compared to the number of ground samples (left) and the standard deviation of 250-m MODIS NDVI observations (right) in a 1-km MODIS pixel.

almost identical (linear regression line $\varepsilon_{5\text{-km}} = 0.9209\varepsilon_{1\text{-km}} + 0.0366$). Consequently, the error quantities obtained in this study for 1-km scales can be used when perturbing MODIS LAI for the purpose of ensemble-based data assimilation applied to a land-surface model run at 5-km resolution.

4. CONCLUSIONS

Ecosystem models which represent the water, energy and carbon cycles can benefit significantly from improved observations of LAI. In this paper, we have validated the MODIS collection 5 LAI product over a wide range of seasons and agricultural landcover types. We have also characterized the MODIS product's error distribution. Reconciling conflicting results from previous studies, we find that the product overestimates LAI in lightly vegetated regions (LAI < ~0.6), but underestimates it for more vegetated regions. The error distribution is normal, with a mean of -0.35 and standard deviation of 0.82. While our estimate of the product's error is likely to include contributions from sampling errors, the observed bias is unlikely to be caused by the sampling errors. We also show that the derived error properties, at 1-km resolution, can be directly applied to characterize the LAI error for a 5-km resolution land-surface model. Understanding of the MODIS LAI product's error distribution is essential for subsequent ensemble prediction studies.

REFERENCES

De Kauwe, M.G. et al., 2011. An assessment of the MODIS collection 5 leaf area index product for a region of mixed coniferous forest. *Remote Sensing of Environment*, 115(2), pp.767-780.

Knyazikhin, Y. et al., 1999. Algorithm Theoretical Basis Document, Available at: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod15.pdf [Accessed May 11, 2011].

Kowalczyk, E.A. et al., 2006. The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model.

Leuning, R et al., 2005. Carbon and water fluxes over a temperate forest and a tropical wet/dry savanna in Australia: measurements and comparison with MODIS remote sensing estimates. *Agricultural and Forest Meteorology*, 129(3-4), pp.151-173.

Merlin, O. et al., 2008. The NAFE'06 data set: Towards soil moisture retrieval at intermediate resolution. *Advances in Water Resources*, 31(11), pp.1444-1455.

Myneni, R.B. et al., 1997. Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Transactions on Geoscience and Remote Sensing*, 35(6), pp.1380-1393.

Peischl, S. et al., 2011. Towards validation of SMOS using airborne and ground data over the Murrumbidgee Catchment. In *The 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, p. 3733–3739.

Quaife, T. et al., 2008. Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter. *Remote Sensing of Environment*, 112(4), pp.1347-1364.

Schwinger, J. et al., 2010. Sensitivity of Latent Heat Fluxes to Initial Values and Parameters of a Land-Surface Model. *Vadose Zone Journal*, 9(4), p.984.

Sea, W.B. et al., 2011. Documenting improvement in leaf area index estimates from MODIS using hemispherical photos for Australian savannas. *Agricultural and Forest Meteorology* (in press).

Sellers, P. et al., 1997. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, 275(5299), p.502.