Flux Tower Data Quality Analysis in the North American Monsoon Region

1. Motivation

The area of focus in this study is mainly Arizona, due to data richness and availability. Monsoon rains in Arizona usually start abruptly in early July and last until September. This strong precipitation pulse comprises the majority of precipitation in the region. It also influences the region's energy budget in terms of flux partitioning. The question that arises naturally is – how does the change in energy fluxes influence and propagate into the atmospheric boundary layer? Furthermore, can it influence the intensity and frequency of subsequent precipitation?
The best way to measure and analyze local energy fluxes is through flux tower data. However, before the aforementioned questions can be answered, the data quality should be tested.

We located four flux towers in the state of Arizona that have data records longer than three years. In order to check data consistency we cross check the data with the theoretical framework (sections 3 and 4), alternative measurements (section 5) and finally, within itself (section 6). The discussion of the results is given in section 7 of this paper.

2. Data

For this project, we used the data from four flux towers in Arizona (given in Table 1). All of the towers are a part of AmeriFlux network and the data is available online. The data used is "level 2", meaning that the raw voltage data have been transformed to physical units and that some quality check has been made, but no gap filling.

Flux tower data				
Location	**Tower**	**Lat, Lon**	**Temporal resolution**	**Coverage dates**
AZ	Audubon Research Ranch	31.591, -110.509	30 minutes	2003 - present
AZ	Flagstaff Managed Forest	35. 143, -111.727	30 minutes	2005 - present
AZ	Flagstaff Unmanaged Forest	35.089, -111.762	30 minutes	2005 - present
AZ	Flagstaff Wildfire	35.445, -111.772	30 minutes	2005 - present

Table 1.

All of the towers are equipped with a net radiometer, soil heat flux plate, tipping bucket rain gauge, soil moisture sensor and a 3D sonic anemometer. The anemometer is positioned at 4 meters above ground in Audubon research Ranch and in Flagstaff Wildfire location. In Flagstaff forests (both managed and unmanaged) it is located at 23 meters above ground. The four towers are shown in Figure 1.

In order to compare flux tower data with alternative measurements, satellite data was used. We used MODIS sensor products for land surface temperature and albedo. The details are given in table 2.

Satellite data				
Sensor	**Product name**	**Spatial resolution**	**Temporal resolution**	**Coverage dates**
MODIS	Land surface temperature	1 km	8 day composite	July 2000 - present
MODIS	Black sky albedo	500 m	2*16 day composite	July 2000 - present

Table 2.

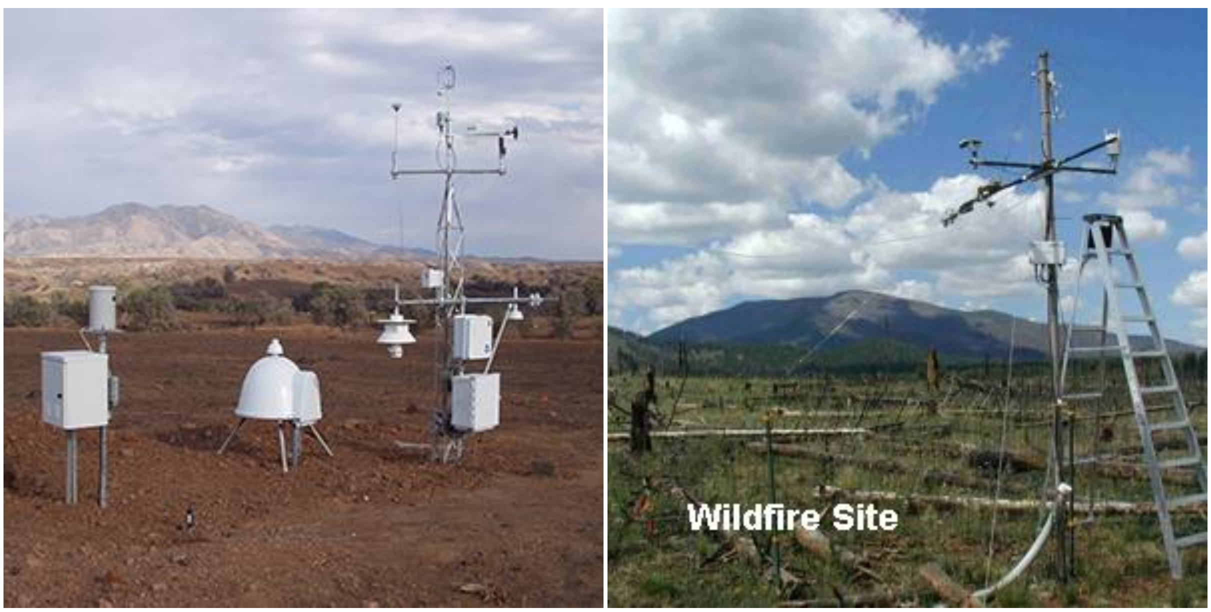

Figure 1: a) Audubon research ranch. b) Flagstaff wildfire. c) Flagstaff unmanaged forest. d) Flagstaff managed forest.

3. Energy budget

We consider the energy budget at the soil surface (equation 1). Energy flux components absorbed or emitted by the soil surface are: net radiation, latent heat flux, sensible heat flux and ground heat flux. Net radiation is comprised of solar radiation (minus the reflected fraction), longwave radiation from the atmosphere and the emitted longwave radiation from the surface (equation 2).

$$R_n - SH - LH - Fg = 0 \quad (1.)$$

$$R_n = S(1 - a) + LW\downarrow - LW\uparrow \quad (2.)$$

Where R_n denotes net radiation (positive downwards), SH is sensible heat flux, LH is latent heat flux, Fg is ground heat flux, S is the solar radiation, a is albedo, and LW denotes upward ($\uparrow$) and downward ($\downarrow$) longwave black body radiation.

When all of the radiation components are considered, the equation 1 should sum up to zero. In the following analysis, we try to close the energy budget for the four flux towers:

<u>Audubon research ranch:</u>

Measurements of net radiation, sensible heat flux, latent heat flux and ground flux are available in the level 2 data. When equation 1 is applied in practice, we end up with a residual that is not always zero. The residual is plotted below for every 30 minutes of all seven years of data. The NaN values are removed from the data, as well as the outliers.

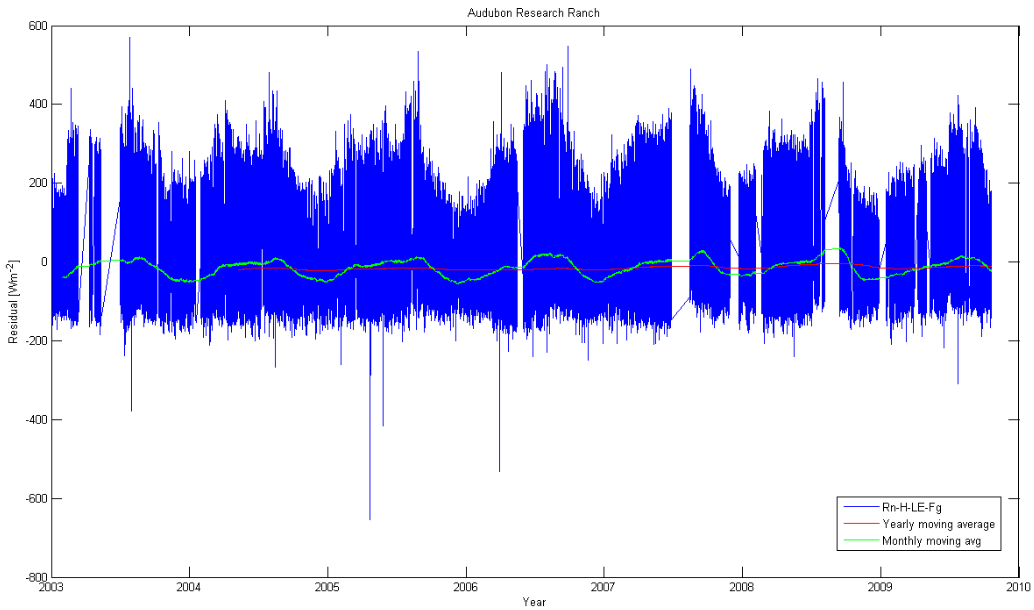

Figure 2: Heat budget residual for Audubon research ranch tower (plotted in blue). Monthly running average is plotted in green, whereas yearly running average is plotted in red.

On a 30-minute scale the residual is very large. It sometimes exceeds 400 Wm^{-2}. However, when averaged over the entire time period, the error is -15 Wm^{-2}.

Flagstaff Wildfire

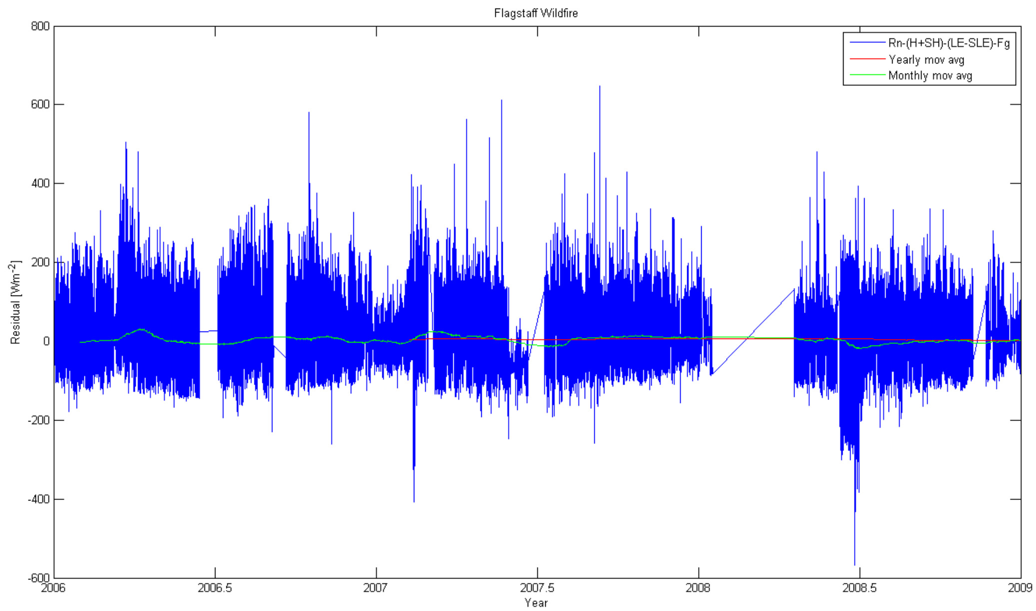

Figure 3: The residual for years 2006, 2007 and 2009. Monthly moving average is shown in green. Yearly moving average is shown in red.

The residual varies significantly on the 30-minute time scale, however, the error averaged over the entire time period is fairly small: 4.02 Wm^{-2}.

Flagstaff Managed Forest

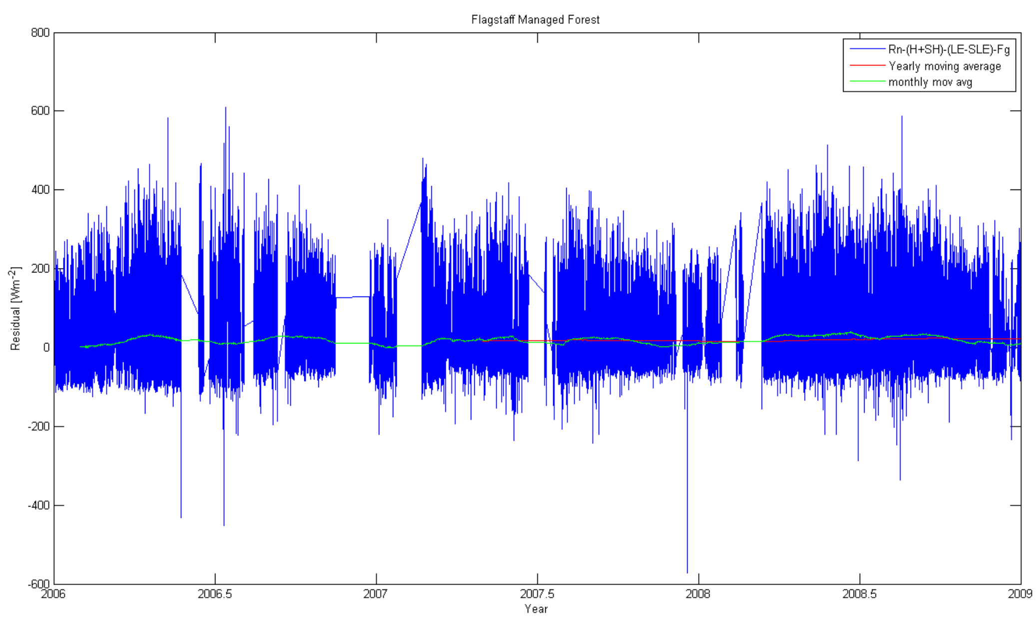

Figure 4. Energy budget residual for years 2006, 2007 and 2008. Yearly moving average is plotted in red, monthly moving average is plotted in green.

Again, error varies significantly on a short scale, but when averaged out, positive and negative anomalies cancel out. Mean error is therefore 4.4 Wm^{-2}.

Flagstaff Unmanaged Forest

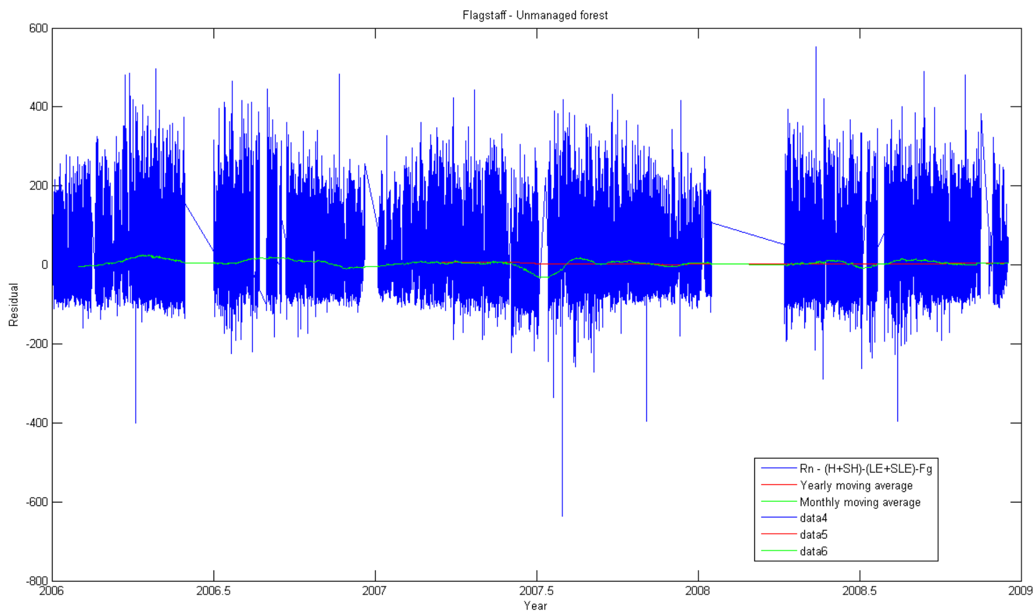

Figure 5.

For the Flagstaff forests, the energy budget also includes the canopy storage of sensible and latent heat.

Summary of results:

	Mean Residual [Wm^{-2}]
Flagstaff Unmanaged Forest	18.99
Flagstaff Managed Forest	4.42
Flagstaff Wildfire	4.02
Audubon Grasslands	-15.22

Table 3.

4. Consistency with the log law theory

In order to test the validity of wind data and the applicability of the tower data on boundary layer analysis, we checked the consistency with the log law. According to the log law, the mean wind at some height *z* has the following dependence:

$$u(z) = \frac{u^*}{K} \ln\left(\frac{z-d}{z_0}\right) \qquad (3)$$

Where u^* is friction velocity, K is Von Karman constant (equal to 0.4), d is taken to be canopy height and z_0 is roughness length. The log law describes the dependence of wind speed with height – close to the ground, the wind speed will decrease logarithmically until it reaches zero. The height at which wind speed is zero is defined as roughness length. We test the consistency of our data by calculating the roughness length and seeing if it changes over time: if it does change, our data does not obey the log law, but if it doesn't our data is consistent with the theory. The level 2 data already include calculations of u* (friction velocity). Wind speed is measured at the level of 4 meters (Audubon and Flagstaff Wildfire) and 23 meters (Flagstaff forests).

Friction velocity is defined as $u^* = \sqrt{\frac{\tau}{\rho}}$ where τ is the wind stress and ρ is air density. In our data sets it was not necessary to compute friction velocity, as level 2 data includes this quantity as well.

Audubon research ranch:

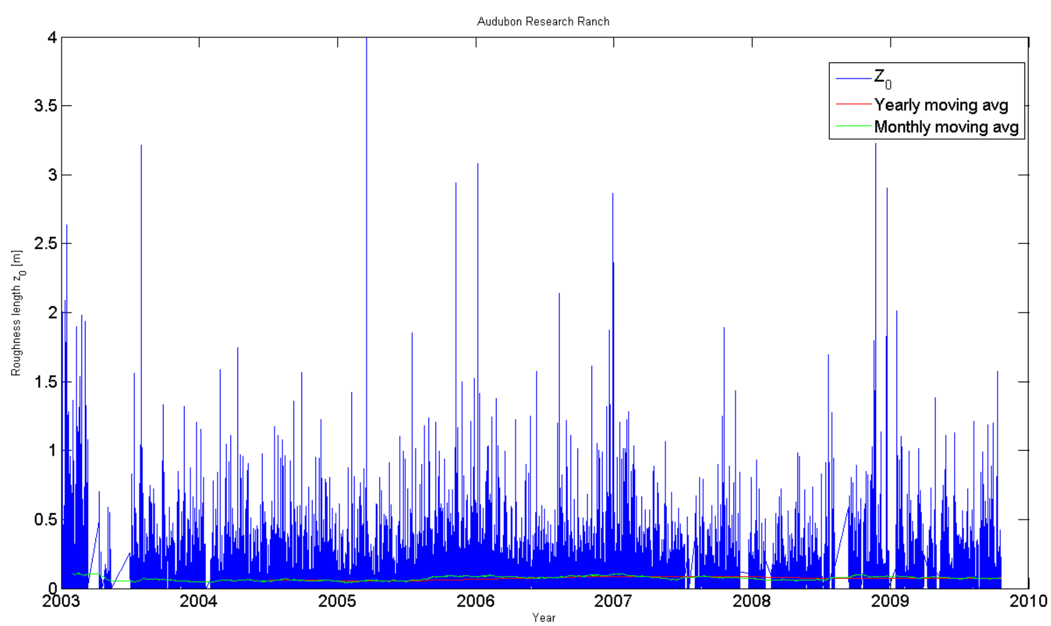

Figure 6.

Flagstaff wildfire

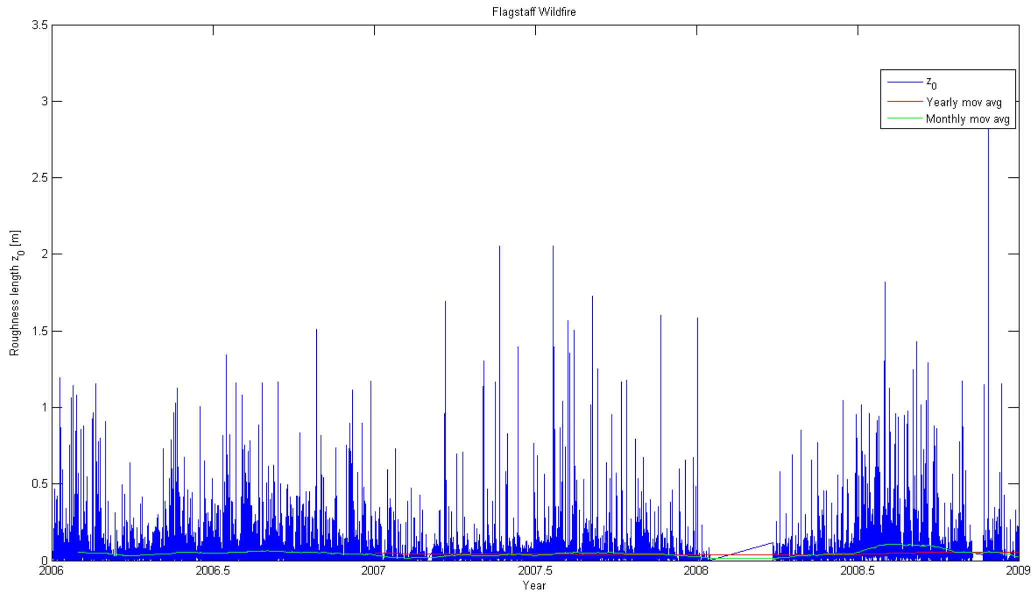

Figure 7.

Flagstaff unmanaged forest

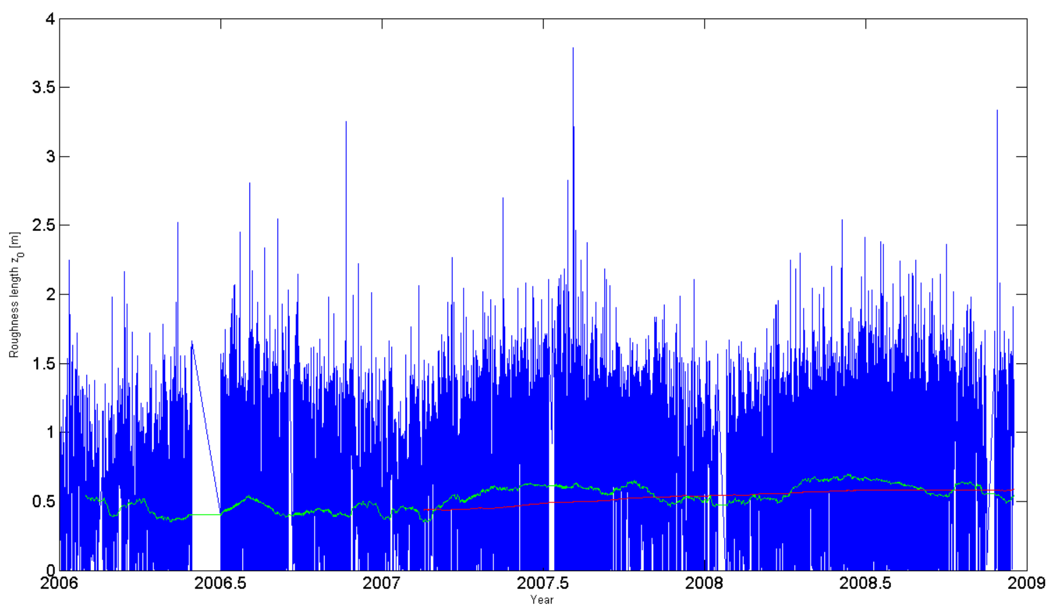

Figure 8.

Flagstaff managed forest

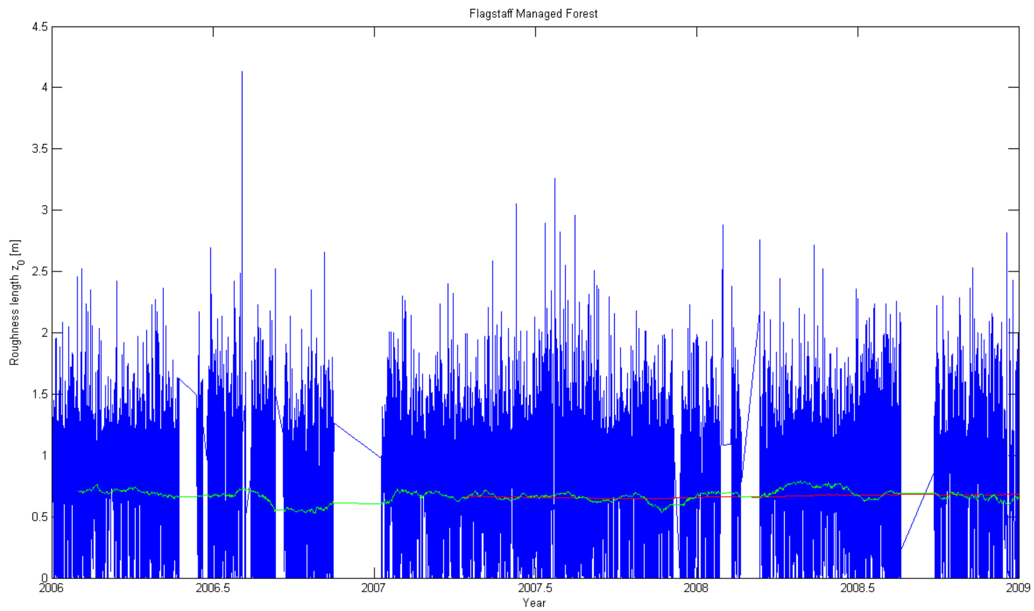

Figure 9.

In all four cases roughness length varies significantly on a short scale. This type of behavior is expected because the log law does not hold at night. However, on a longer scale the roughness length is constant, as indicated by the monthly and yearly moving averages. The magnitudes of z_0 are well within the theoretically expected range: for the wildfire site and Audubon research ranch, the roughness length is of the order of 10 centimeters. This is as expected as both locations feature short, grassy vegetation. In the forests, z_0 is of the order of 1 meter. In unmanaged forest, z_0 seems to be increasing very slowly. This could be due to vegetation growth.

5. Comparison with satellite data

In order to further test the validity of flux tower data we compared it with satellite derived quantities. Multiband satellite imagery provides the most spatially thorough picture of land surface conditions, however, it often lacks temporal resolution. Flux tower data are different – it is a point measurement and therefore spatially very limited, but with an excellent temporal resolution. With this in mind, it is clear that the two types of measurement can never give the same results, however it is interesting to see if the results are comparable.

For this purpose we used the data obtained by a MODIS sensor on board the Terra satellite. Specifically, we used land surface temperature 8-day composite with spatial resolution of 1 km and black sky albedo 8-day composite with spatial resolution of 500 meters.

A single MODIS tile swaths a large area, so each tile was subset to a region of 75 km^2 centered around the flux tower. Land surface temperature and albedo were then computed as a mean land surface temperature or mean albedo in that subset. The result was compared with tower-derived LST and albedo. This analysis was made solely for Audubon research ranch data.

Land surface temperature:

Using the Stefan Boltzmann law, we converted the tower outgoing longwave radiation. We assumed the ground surface emissivity to be $\varepsilon = 0.95$.

In order to make the temporal scale comparable, we computed 8-day moving average for tower-derived LST. MODIS LST and 8-day moving average tower-derived LST are shown in figure (10).

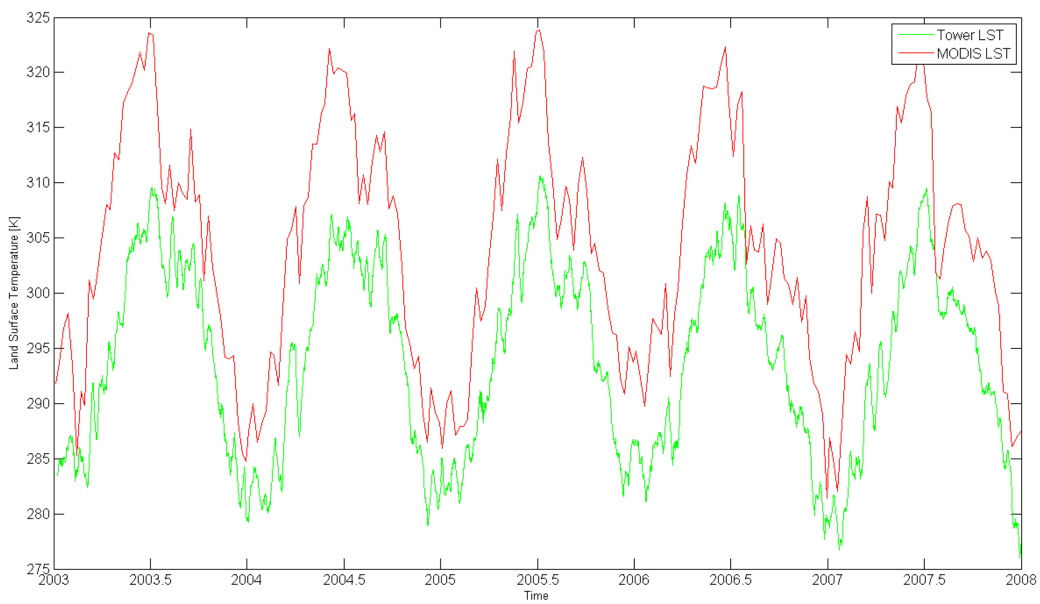

Figure 10: Land surface temperatures for Audubon research ranch. MODIS LST is shown in red, tower LST is shown in green.

The results are strikingly similar. Both temperatures rise and fall at the same time, which is especially observable on a shorter temporal scale. It is interesting to note a large dip in LST at the time of onset of monsoon rains. The tower-derived LST is constantly lower than MODIS LST. This is because MODIS LST is measured at 10:30 am, whereas the tower LST is a daily average. When only the measurements taken at 10:30am are taken into account, the temperatures match exceptionally well (Figure 11). The temperatures match up to a few degrees and generally show the same trends (both short-term and long-term). This is particularly well seen from the scatter plot (Figure 12). Correlation coefficient is found to be 0.95.

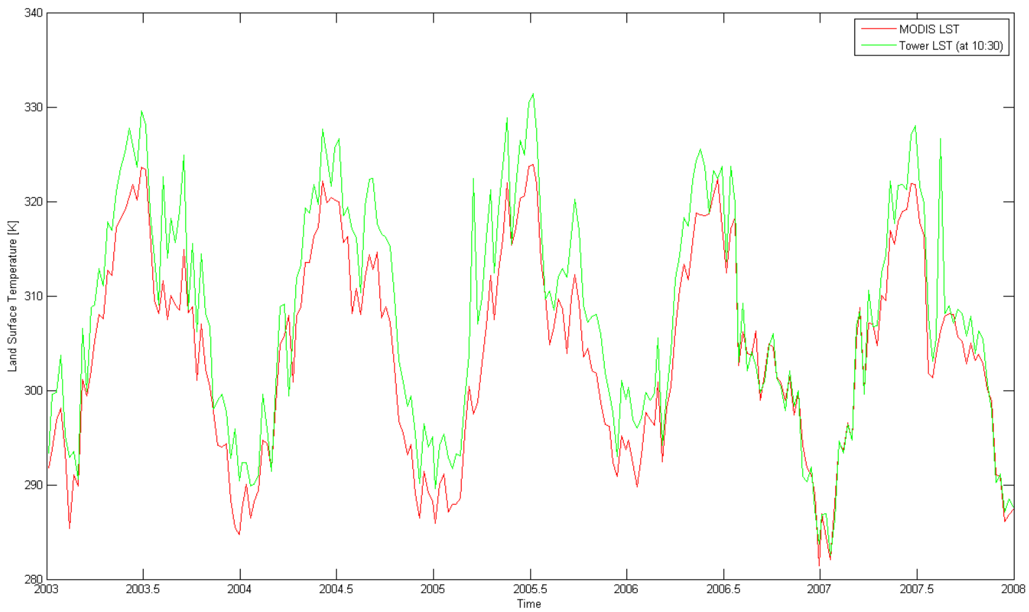

Figure 11: MODIS LST 8-day composite time series is shown in red. Tower LST measured at 10:30 am and averaged over the corresponding 8-day period is shown in green.

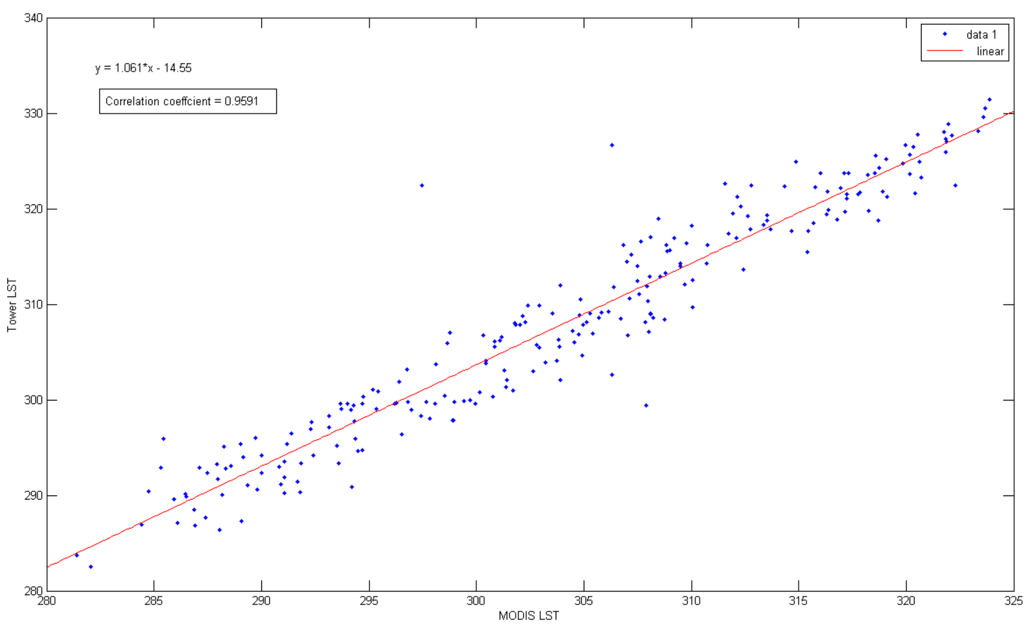

Figure 12: Scatter plot of MODIS LST 8-day composite and tower LST measured at 10:30 am and averaged over the corresponding 8-day period.

Albedo:

Audubon research ranch provides measurements of incoming and outgoing shortwave radiation. The outgoing shortwave radiation is entirely reflected solar radiation because Earth does not emit in those wavelengths. Tower albedo is therefore computed as a ratio of outgoing shortwave radiation and incoming shortwave radiation. Since there is no solar radiation at night, the night values of albedo are not physical and were discarded. In order to compare the tower albedo values with MODIS albedo values, midday albedo was computed. I.e., only albedo values from 10 am to 2 pm local time were considered. This was done because the MODIS albedo is a combined product of both Terra and Aqua satellites, and Terra and Aqua fly over the region at 10:30am and 2 pm respectively.

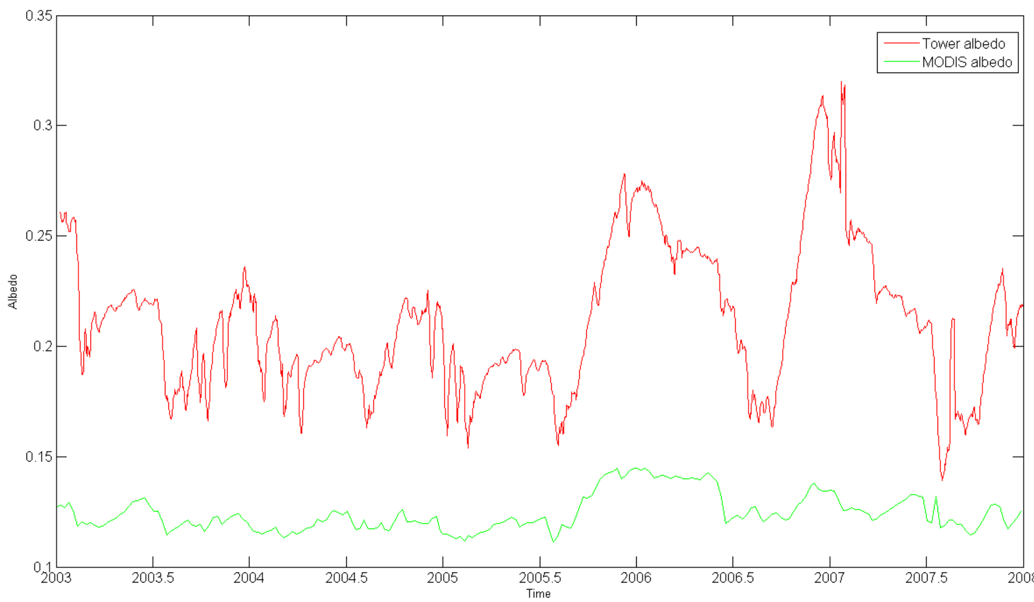

Figure 13: MODIS albedo 8-day composite is shown in green. An 8-day moving average of tower midday albedo is shown in red.

Tower albedo constantly overestimates the MODIS albedo. This could be due to the fact that MODIS albedo is "black sky albedo", meaning that it does not take into account diffuse radiation. Also, MODIS albedo gives the value for the 75 km^2 of area around the tower, whereas tower albedo measurement is much more localized.

When the two albedos are plotted on a different scale, some of the similarities are emphasized:

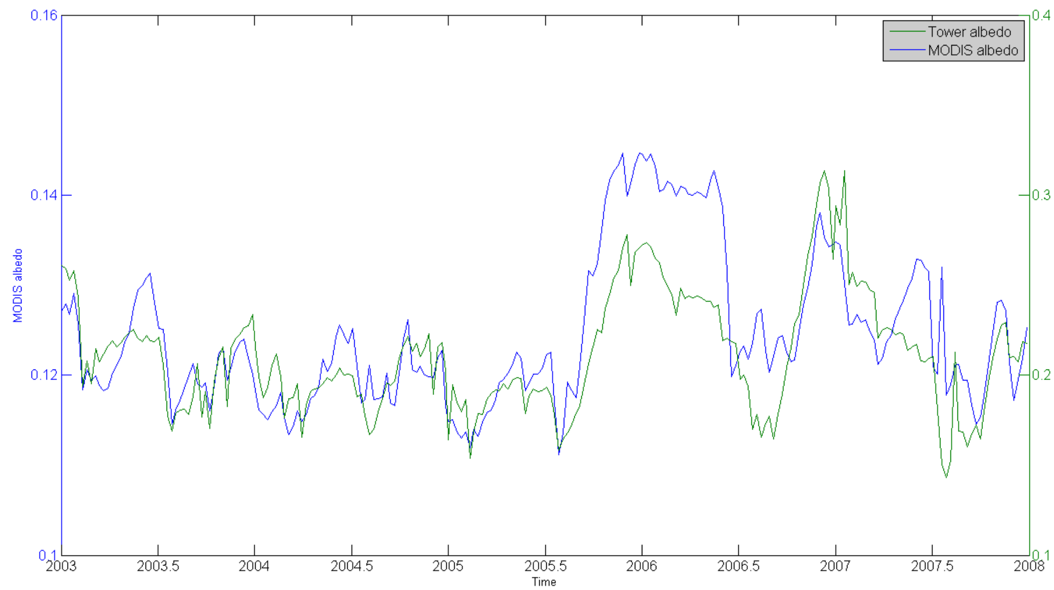

Figure 14: MODIS albedo 8-day composite is shown in blue. Tower albedo averaged over the corresponding 8-day periods is shown in green.

The albedos obtained by the two methods generally show the same features: Both time series show a hint of a seasonal cycle: Albedo generally reaches its maximum at the beginning and end of each year. It reaches the minimum in late summer, which corresponds to the onset of monsoon rains and increase in vegetation. It is interesting to note that increase in tower albedo in the winter of 2005/2006 and in the winter of 2006/2007 is also reproduced in MODIS data, suggesting it was a wider-spread phenomenon.

The scatter plot of tower albedo and MODIS albedo is shown below. The correlation coefficient is found to be 0.75.

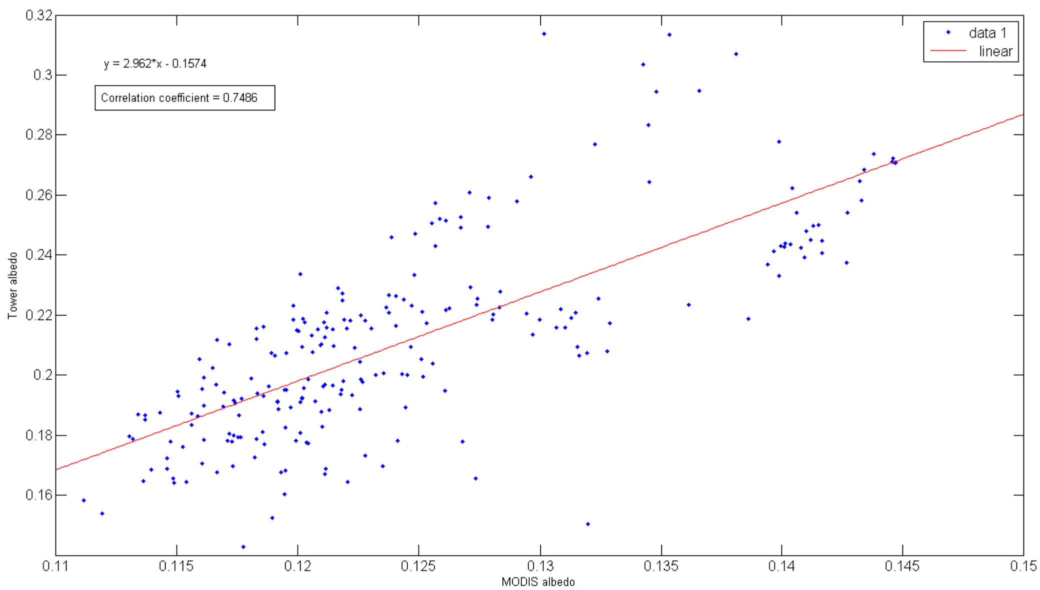

Figure 15: A scatter plot of MODIS albedo 8-day composite and tower albedo averaged over the corresponding 8-day period.

6. Precipitation and soil moisture

In order to test the consistency of the data, we compared the quantities that are related to each other. Specifically, we looked at precipitation and soil moisture. The tower at Audubon research ranch provides measurements of soil water content at two different depths – 10 cm and 20 cm. For this analysis, we used the values at 20 cm. The expected behavior is increase in soil moisture after a strong precipitation event, i.e. after monsoon rains. However, when the two quantities were compared, we observe a somewhat different pattern (figure 16).

<u>Audubon Research Ranch</u>

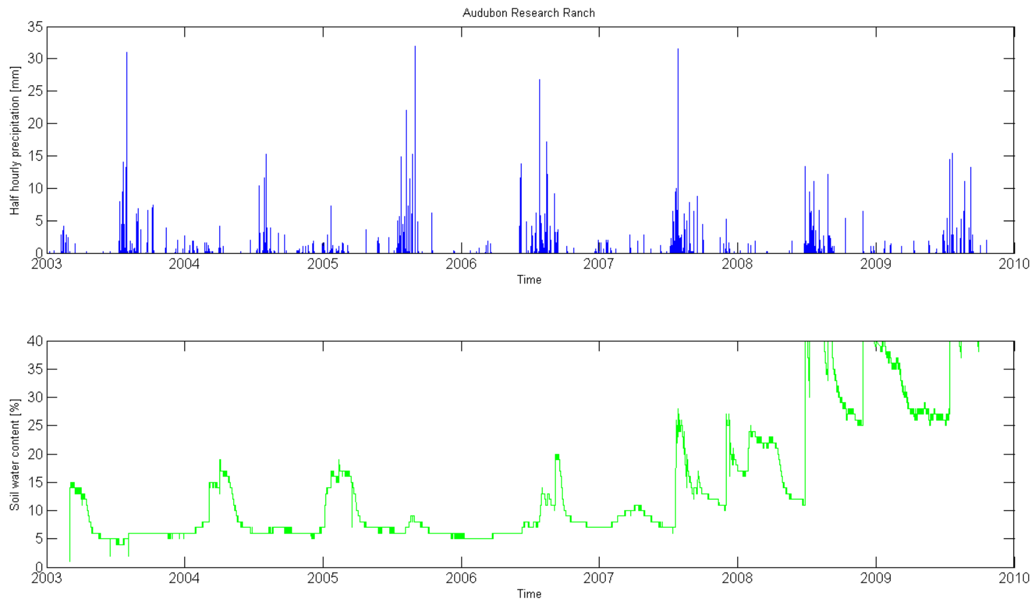

Figure 16. Audubon research ranch precipitation (upper panel) and soil water content (lower panel).

From figure 16 it is observable that soil water peaks in winter. This is rather peculiar, as winters are typically the dry season. At the time of monsoons, soil water remains virtually constant. However, in years 2006 and 2007, there is a visible increase in soil moisture after the summer rains. This type of behavior is what we expected to find in the first place. There are two possible explanations for this unexpected result: It is possible that run off is very strong in the region (this region is prone to flash floods, so it is highly plausible) so during intensive rains most of the water runs off, and very little remains in the soil. Peaks in the winter time could be explained by low evapotranspiration during the cold season. Alternatively, the instrument could be flawed and therefore shows wrong results. This assumption is somewhat justified considering a strange increase of measured soil moisture (instrument drift?) in the year 2009.

In order to test and visualize the expected results, we constructed a simple bucket model where soil moisture depends solely on precipitation and evaporation. Evaporation is modeled as a linear function of soil moisture, where α is a parameter in units of day^{-1} (equation 5):

$$\begin{aligned} \frac{dSM}{dt} &= P - E \\ E &= \alpha SM \end{aligned} \qquad (5)$$

The equation is solved in MATLAB using a Fast Fourier Transform. The results are shown in figure 17.

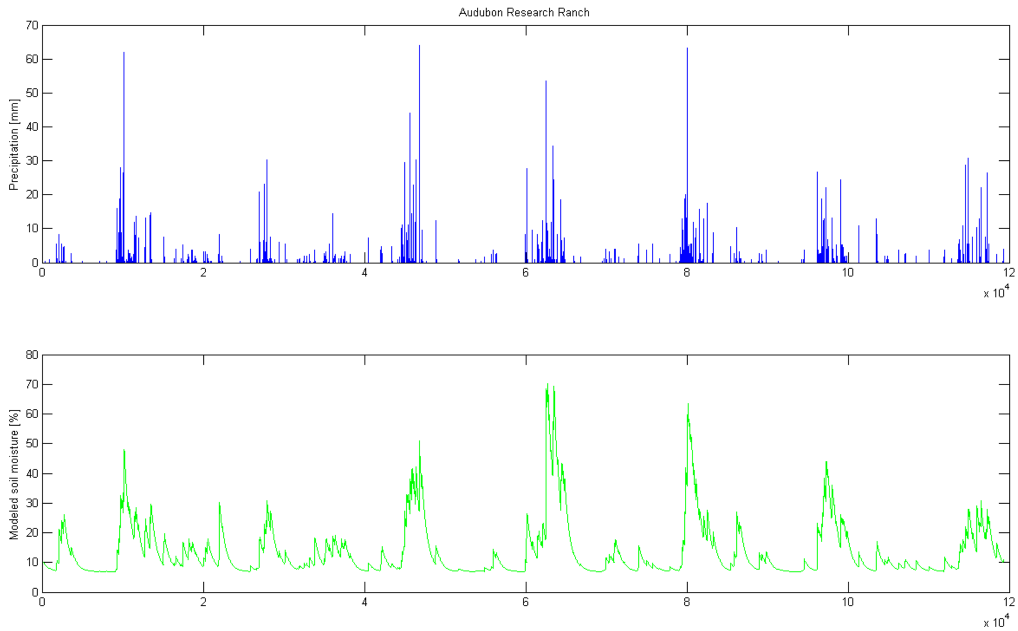

Figure 17: Upper panel: measured Audubon research ranch precipitation (in millimeters). Lower panel: Modeled soil water content (in percentages). We assumed no run-off and no temperature dependence of soil evaporation. Alpha is set to 1/(2 weeks). Note the absence of any soil moisture increase in 2009.

Assuming the measured soil moisture is correct, the model fails to capture the physics of soil wetting. This means that in this specific case the factors that were not modeled (run-off, temperature effects on evaporation…) play a significant role in the final results.

In order to further test the model and the relationship between rainfall and soil moisture, we looked at the three other data sets from the Flagstaff region. The Flagstaff data tower data is obtained by somewhat different instruments. Soil moisture is measured by a water content reflectometer that uses time domain reflectometry method. Basically, the instrument uses wave propagation to measure the permittivity of the soil, which is related to water content in the soil.

Flagstaff Managed Forest

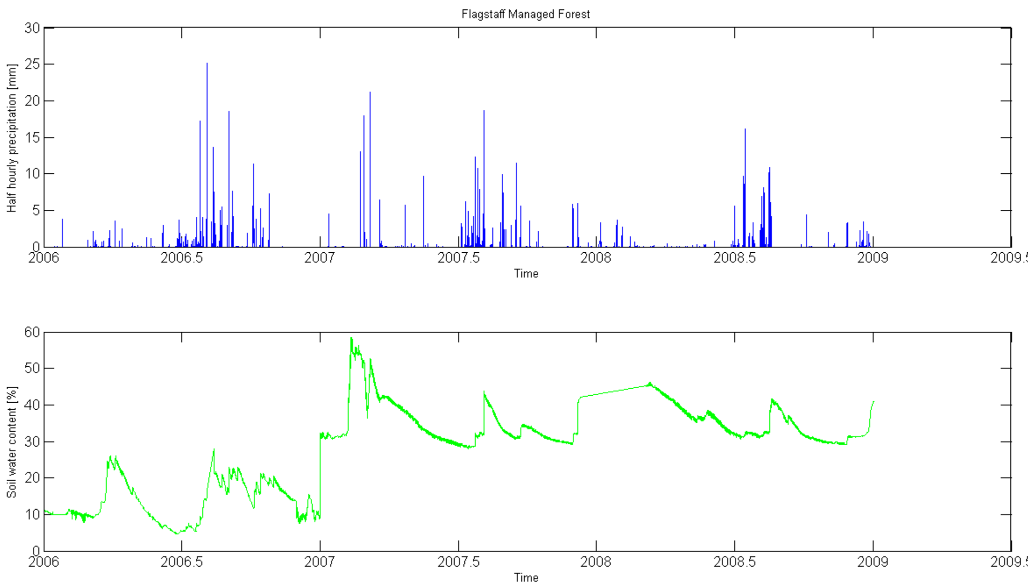

Figure 18: Measured precipitation (upper panel) and soil moisture at 20 cm (lower panel).

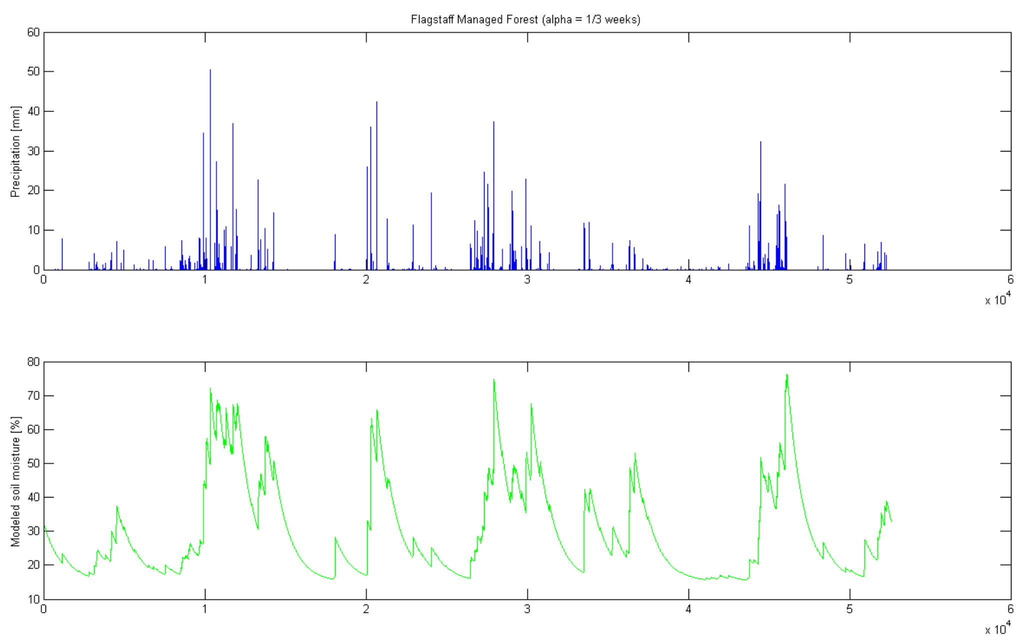

Figure 19: Measured precipitation (upper panel) and modeled soil moisture (lower panel).

This data set also shows some inconsistencies in measured soil moisture. Starting with 2007 there is a large increase in soil water content that is not reflected in an increase in precipitation. It seems that the “zeroth level” has been shifted from 0% to 30% which could be explained by instrument drift and/or lack of calibration. In the model, alpha has been set to 1/(3 weeks) in order to mimic the slope of the soil water content decrease. However, this parameterization significantly overestimates the soil water content.

Flagstaff Unmanaged Forest

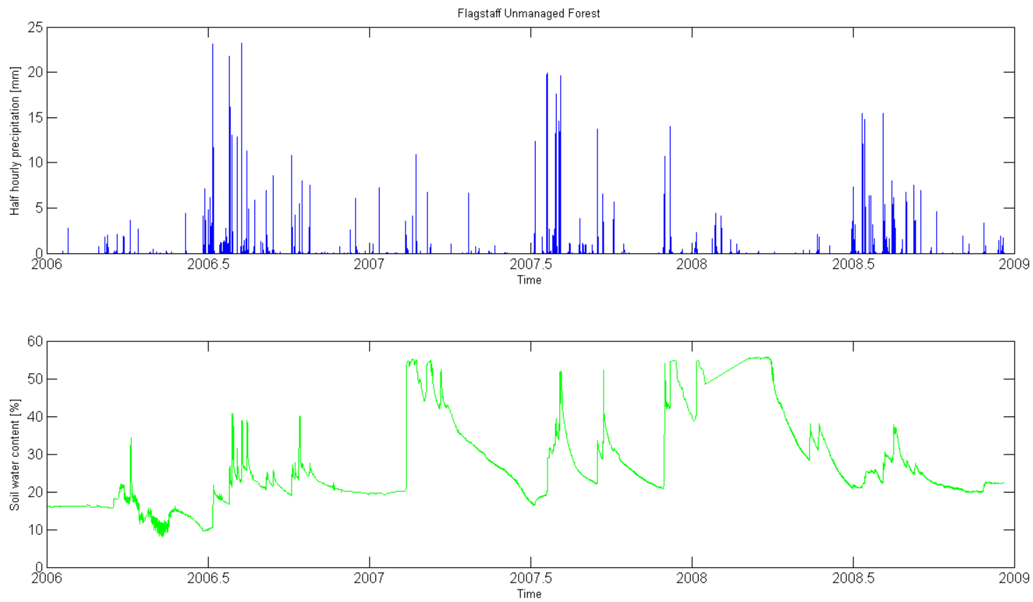

Figure 20: Measured precipitation (upper panel) and soil moisture at a depth of 20 cm (lower panel).

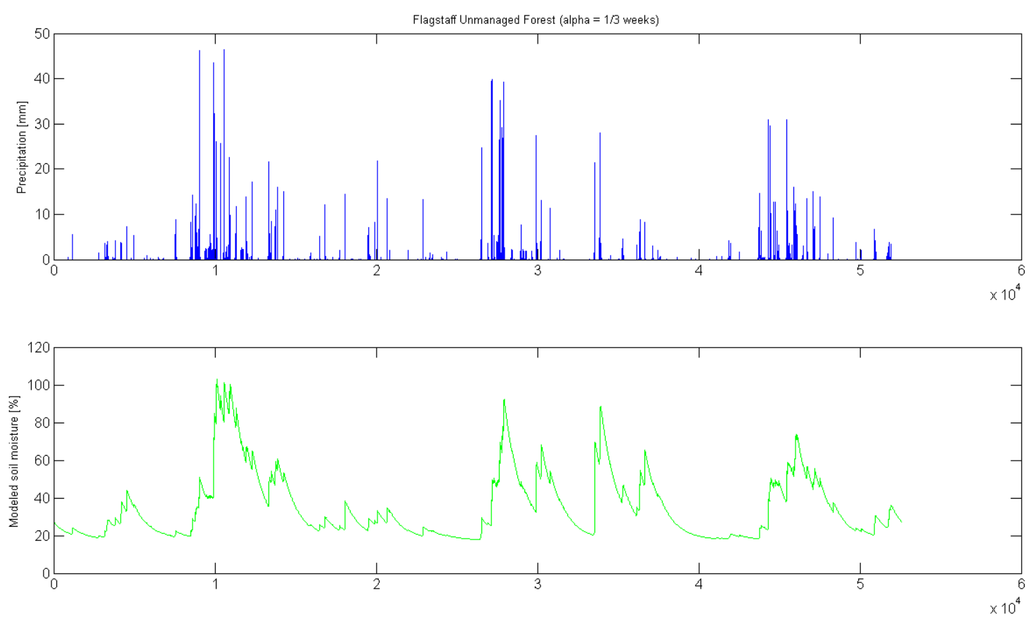

Figure 21: Measured precipitation (upper panel) and modeled soil moisture (lower panel).

There is a hint of instrument drift in this data set as well. Starting with 2007, it appears that the minimum measured soil water content shifted from 0% to about 20%. When compared to the modeled soil moisture, there is some resemblance. Alpha is set to 1/3 weeks.

Flagstaff Wildfire

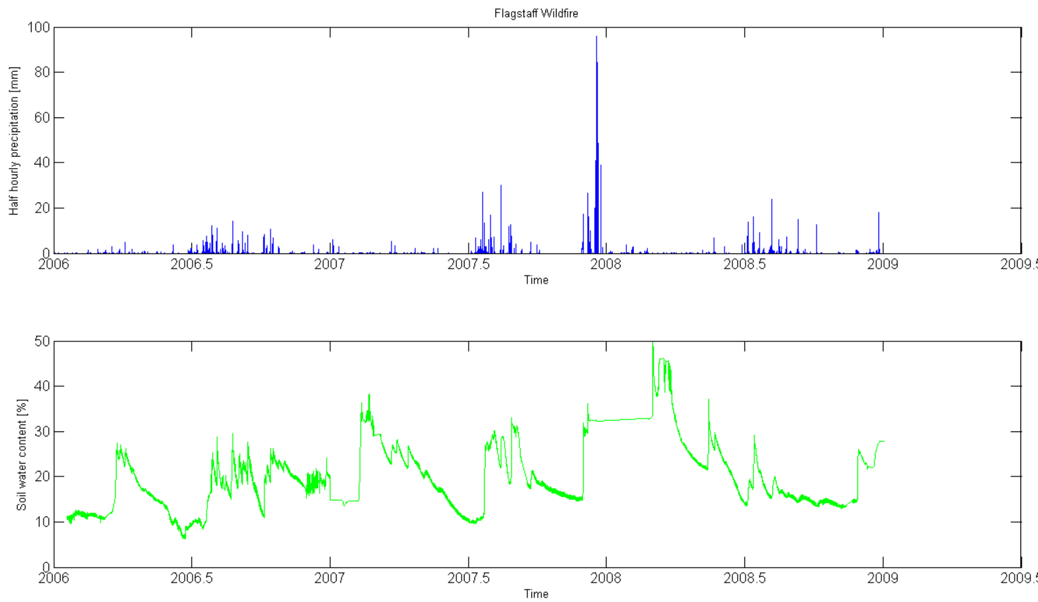

Figure 22: Measured half-hourly precipitation (upper panel) and measured soil moisture at 20 cm (lower panel).

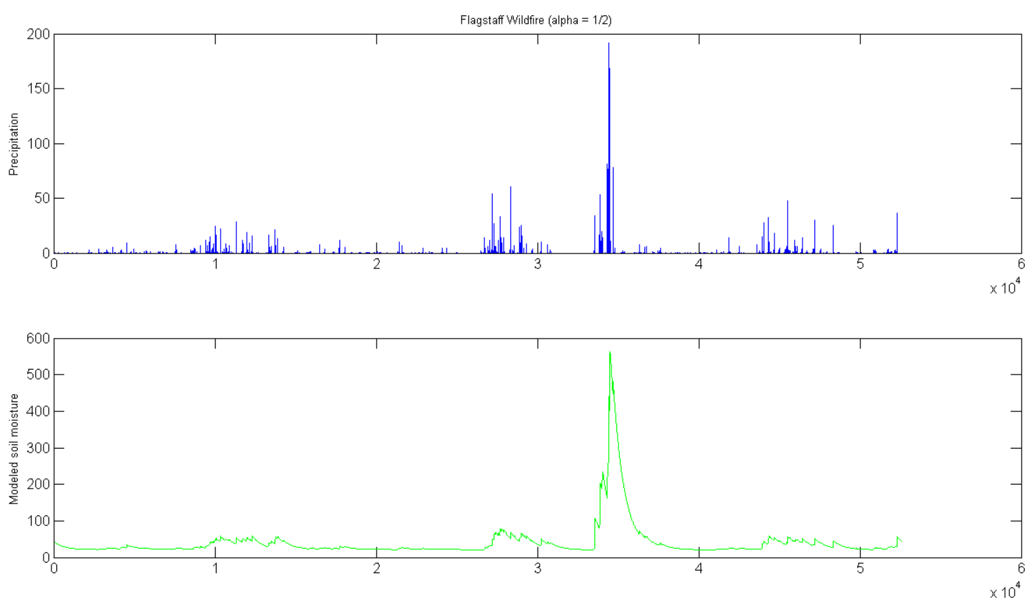

Figure 23: Upper panel: measured hourly precipitation in mm. Lowe panel: modeled soil moisture.

At the beginning of 2008, the precipitation data shows an intense precipitation event. Closer inspection reveals that there was about 200mm of rain in a single hour. This is highly unlikely especially as it was not reflected in the soil moisture measurements. However, this amount of “precipitation” skews our model results.

7. Conclusion

A very basic analysis of tower data quality is presented. In this analysis, we tried to close the heat budget, compared the data with the log law theory, compared the data with satellite measurements and tested the relations of measured variables within themselves. The heat budget analysis shows significant errors on a short temporal scale (30 minutes to a day), however, when averaged over a longer period (a month or a year) the errors cancel out and the measurements can be considered reliable. The log law analysis was performed by computing the roughness length and seeing if it remains constant. On a shorter scale (30 minutes to a day), the roughness length varies significantly, but this type of behavior is expected because the log law does not hold during nighttime conditions. On a longer scale (months to a year) 3 out of 4 towers showed no variations in roughness length. One of the towers (unmanaged forest) showed a slow increase in roughness length, but that is hypothesized to be due to vegetation growth. The two quantities compared to satellite data were land surface temperature (LST) and albedo. For this analysis we used Audubon research ranch tower data. Tower LST was computed using Stefan-Boltzmann’s law from outgoing longwave radiation. Tower albedo was computed as a ratio of outgoing shortwave and incoming shortwave radiation. Tower LST and MODIS LST showed striking similarity in trends and variations. At first glance, tower and MODIS albedo are not as similar. The tower albedo is consistently higher than the MODIS albedo. However, when plotted at a different scale, a lot of other similarities become visible. In general, both tower and MODIS albedo show the same trends and variations. For example, they both show a minimum during the monsoon season which corresponds to increase in vegetation. The possible reasons for the discordance in magnitude are plentiful: mismatch of spatial scales (MODIS albedo was computed as an average for a region of $75km^2$ around the tower), the fact that MODIS albedo does not include diffuse radiation, inconsistency in considered wavelengths, etc. Furthermore, we compared soil moisture measurements with precipitation measurements expecting to find a strong positive correlation. Instead, soil moisture seems to peak in winter time, during the dry season and only occasionally (years 2006 and 2007) in the summer, after the monsoon rains. This type of behavior was not expected given the simple bucket model we constructed. In the model soil moisture depends solely on precipitation and evaporation and there is no run-off. Also, evaporation is not a function of temperature. This leads to a conclusion that i) run-off and temperature effects are significant in the given case and therefore influence soil moisture considerably or ii) our model captures the basic physics of soil wetting, but the data we compared it to is wrong.