FINAL RESPONSE TO COMMENTS
Effect of the Revisit Interval on the Accuracy of Remote Sensing-based Estimates of Evapotranspiration at Field Scales (hess-2016-273)

Note: The responses follow the reviewers' comments and I highlighted using red text. Excepted text from the revised manuscript is shown in blue with the changes underlined when appropriate.

REVIEWER 1:
Subject of this study is the error introduced by 'temporal upscaling' of daytime evaporation as a function of the length of the interpolation interval. Temporal upscaling is used to estimate evaporation on days at which no direct evaporation data is available as E = f*X. X, the reference quantity, is known on a daily basis, the scaling factor f is assumed to vary only gradually over time and is determined through interpolation. The analysis is based on solely in situ measurements of meteorological data, turbulent fluxes and the ground heat flux, from 20 Ameriflux sites. Interpolation intervals of 1 to 32 days and five different reference quantities and associated scaling factors are considered, using three interpolation techniques and four metrics to evaluate the self-persistence of the scaling factor and error in the scaling factor and evaporation estimate. The results are analyzed per land cover class (mean of 4-6 sites per class). The authors found that the autocorrelation of the scaling factor is limited for lag times larger than a few days; that the error in both f and E quickly increases with increasing interpolation interval during the first week, more slowly afterwards; and that both the lag time for a certain autocorrelation threshold and the interpolation interval for a certain relative error in the evaporation estimate is smallest for forests, followed by cropland, grassland and open canopy. The authors relate the differences between land cover types to the mean latent heat flux per land cover type (which is largest for forest, smallest for open canopy) and define an exponential function relating mean evaporation per land cover type with a maximum interpolation interval to achieve a relative error of x%. Using the exponential relation they define a maximum return interval of 5 days to achieve a relative error in the evaporation estimate of maximally 20%.
The question the authors address in this paper is relevant for various applications of daily evaporation time series in e.g. water resources management and hydrology. The use of the relatively large set of Ameriflux sites covering different land cover types, climates, soil types, and covering many measurement years, makes the results – in potential – also more generally applicable than other literature on the subject. The authors however focus mainly on the practical implications of the results (selection of best interpolation technique and maximum return interval for a given accuracy of the evaporation estimate), than on understanding the why of the results (e.g. what actually explains the differences they find between and within land cover types and scaling approaches). This I think is a missed chance, given the data they have and the analysis they have already performed. The approach of Farah et al. (2004) for example, showing the seasonal variability of the evaporative fraction and trying to explain it, I find very useful in this respect as an additional analysis. Moreover although 20 Ameriflux sites 'distributed throughout the contiguous United States' have been used, the only information on the differences between the sites that is provided by the authors is the land cover class. The reader therefore cannot fully judge the representativeness of the results of this study, whereas given the practical focus of the analysis that should be a major benefit of this study. Of the comments below, at least comments 1,2,(3),7 and 8 I think should be addressed before publication of the manuscript.

Specific comments
1. The title of the article, mentioning the accuracy of remote sensing-based estimates of evapotranspiration, is a bit misleading. Of course the results of this study provide an indication of the effect of the revisit time on the accuracy of RS based estimates, but this study does a) not actually use remote sensing based estimates, and b) the point scale results are not one-to-one transferable to remote sensing based results. I suggest to change the title, leaving out remote sensing.

Given the significance of temporal upscaling for remote sensing-based applications, which was indeed the impetus for this study, we feel it is important to acknowledge the importance of this study to remote sensing community. However, in light of the reviewer's comments, we agree the paper's title may be confusing to some readers and have modified it to clearly indicate that surface observations were used. Based on the Editor's suggestion, the revised title reads:

Effect of the Revisit Interval and Temporal Upscaling Methods on the Accuracy of Remotely-Sensed Evapotranspiration Estimates

2.In the Introduction, p5, it is stated that perfect retrieval of the flux is assumed. If indeed the results found in this study are to be transferred to remotely sensed-based evaporation estimates, could the authors elaborate a bit on how realistic this assumption is, either in the introduction or in the discussion of the results? In other words, to what extend can the results based on the in situ, point scale measurements be applied for remotely sensed, usually more spatially aggregated, estimates?

Since any error in the modeled ET from remote sensing-based models would propagate into the estimates of the scaled quanties, *f*, used for the interpolation and then into the interpolated fluxes, the analysis conducted here represents a best-case scenario. This is now stated explicitly in the introduction. Additionally, the effects of any errors in the modeled fluxes on the maximum return interval and the conclusions drawn from the study are discussed.

Page 5, Line 11:
Since any errors in the remote sensing-based ET estimates propagate into the calculation of f and the subsequent temporal upscaling, this analysis represents the best-case scenario.

Methods
3.I understand daytime mean evaporation instead of daily mean evaporation is used because of the better representativeness of the instantaneous scaling factor during satellite overpass. For many applications however (hydrological), daily evaporation estimates are valuable. It would be interesting if the authors could show whether the results for daily evaporation are very different. (Although also EC measurements are not that perfect in measuring nighttime fluxes (Fisher et al. 2007).)

We agree that daily ET can be very useful for many applications. This decision to focus on daytime mean data is the result of a number of considerations. This research was driven by a need to characterize the effects of temporal upscaling on remote sensing-based applications. Since the internal physics of most, if not all, remote sensing-based models for estimating ET – for example, the two source energy balance models (TSEB) – to describe the moisture flux accurately at night, daytime mean data was used to emulate the model output as closely as possible. Similarly, the assumption of self-preservation that underpins temporal upscaling is not valid at night. Finally, the contribution to the total moisture flux is typically small overnight; it averages between 3% and 8% for most land cover types (Rawson and Clark 1988; Green et al. 1989; Sugita and Brutsaert 1991; Malek, 1992; Tolk et al. 2006).

4. I like that different 'reference quantities' are analyzed, but are these selected because these are the ones that are used most often (so for application purposes mainly), or could the authors also elaborate a bit more on the expected differences in self-preservation of the associated scaling factors?

The reference quantities evaluated in this study were selected because they either have been proposed for use or are commonly used by the remote sensing community for the temporal upscaling of ET derived from remotely sensed imagery. This is now stated explicitly in the manuscript.

Page 6, Line 7:
Each of the χ used is this study was selected because it is either in common usage for remote sensing-based applications or has been proposed for use in those applications.

5 (R1). I see value in using different evaluation metrics as long as the added value is clear. I wonder whether the authors could better explain the added value of D next to the RMSE and especially MAE, either in the methods section or in the results section. As it is presented now, D is just another metric, confirming the results of MAE.

5 (Ed). If D provides similar information as RMSE and MAE, but D is only scaled between 0 and 1. Then I suggest to chose one metric you prefer instead of showing 3 similar metrics.

The analyses using the index of agreement (D) have been removed from the paper. The manuscript now focuses on RMSE and MAE when describing error. The computational underpinnings and sensitivities of these two metrics are sufficiently distinct to provide unique insights into magnitude and variability of the error introduced into the ET estimates via temporal upscaling.

Page 8, Line 1 to 21:
Page 11, Line 7 to 13:
The discussion of D has been remove from sections 2.4 Statistical Metrics and
3.3 Accuracy of the Latent Heat Flux Estimates. Similarly, the figure showing how D varies with return interval has been removed.

Results
6. All results are shown as the mean value per land cover class. Although it is a logical choice to group the measurement sites per land cover class, I expect there to be considerable differences between thsites within a class reflecting differences ine.g. climate and other environmental conditions (as visible in Fig.1 as well). However, the authors do not comment on the 'within class' variability of the autocorrelation and for the remaining results the variability is not shown at all. I suggest to at least provide some additional information on the Ameriflux sites in Table 1 (at least climate and seasonality information) and to include information on the within class variability of all results, to inform the reader on a) the robustness and b) the representativeness of the analysis.

Per the reviewer's suggestion, additional information describing the climate at each of the sites has been added to Table 2. Additionally, a discussion of some of the potential reasons for the observed variability with a given land cover class has been added.

Page 9, Line 20:
    The figure also shows there was significant variability from site-to-site within a given land cover type, particularly for longer lags. Although the specific causes of these differences are not fully understood, there are number of factors that likely contribute. For example, there are difference in both species composition and climate at the various sites. Consider, as an example, the forest class which includes both coniferous and broadleaf deciduous forest. Moreover, the species composition varies even among sites of the same forest type; for example dominant species at the Niwot ridge site are Subalpine fir (Abies lasiocarpa) and Engelmann spruce (Picea engelmannii) while, as the name implies, the dominant species at the Loblolly Pine is loblolly pine (Pinus taeda). At the same time, the mean annual temperature at the forested sites ranged from 1.5 °C to 14.4 °C while the mean annual precipitation varied from 800 mm to 1372 mm. Similarly the mean annual temperature and precipitation at the cropland sites, which are all planted on a rotation of maize and soy, range between 6.4 °C and 11.0 °C and 789 mm and 991 mm, respectively.

7. In Figure 2 and Figure 9 the maximum lag time for a certain threshold of the auto- correlation and maximum return interval for a certain threshold of the relative error in evaporation estimate are plotted against the mean latent heat flux per land cover class. First, continuing on point 7, I wonder why the authors choose to plot the results for the mean latent heat flux per land cover class and not for all individual sites? Second, does there exists a similar relation between maximum lag time or return interval and latent heat flux for the data of one site? Furthermore, I would like to know whether the authors did consider other dependencies of maximum lag time (Fig.2) and maximum return interval (Fig.9) than the mean latent heat flux?

Although the same general trends are evident when the data from the individual sites are used, the relationships were rather noisy. Therefore, the means for a given land cover type was used in lieu for clarity. The aim of this paper was to explore the effects of temporal upscaling on the accuracy of ET estimates and, with that in mind, the relationships between the mean flux and both the autocorrelation or maximum return interval is shown as a guideline for selecting an appropriate return interval. They are not intended to suggest causality. Indeed, the underlying cause of the relationships is unclear.

Additionally, although it falls outside the intent of the paper, several analyses – for example, an investigation of atmospheric coupling (McNaughton and Jarvis, 1983) - were conducted in an effort to discern the fundamental cause of the relationships between the magnitude of the moisture flux and both the autocorrelation and maximum return interval. Unfortunately, these did not yield fruitful results and the data available at all of the sites is insufficient to conduct the extensive earth system modeling efforts needed to ascertain the underlying cause of the relationships.

8. Page 8-9. The authors show and describe the results concerning the RMSE and MAE of the scaling factor as a function of itself (75% of its maximum reached after xx days). Could these values be related to the absolute value of the scaling factor as well? Or in other words, how does the relative error in f develop as a function of the lag? Idem for the evaporation estimates (Page 9).

This comment is unclear. Both the text and associated figures describe the error as a function of the lag (return interval). The reference to the amount of time needed to reach 75% of the maximum is given simply to highlight how quickly the error increase for short lag times. To improve the clarity, the mean maximum RMSE and MAE is given for each land use type.

Page 10, Line 10:
For comparison, the mean maximum RMSE for each land cover type was 0.26, 0.28, and 0.17 for croplands, grasslands, forest, and open canopies, respectively. Although it also increased logarithmically, the amount of time needed for MAE to reach 75% of the peak value was more variable, ranging between 5 to 10 days. Again, for purposes of comparison, the mean maximum MAE was 0.22, 0.14, 0.16, and 0.10, respectively, for croplands, grasslands, forest, and open canopies.

9. I think that it may not be necessary to show Figures 5 and 6. The results for the scaling factor already provide enough ground to prefer linear interpolation above spline interpolation methods, so Figure 7 summarizes the relevant information on the development of the error in evaporation estimates with increasing interpolation interval. (Although it is interesting why the MAE in the evaporation estimate for forested sites is lowest for Rn as reference quantity (compared to the other reference quantities), whereas the MAE for the associated scaling factor is largest of all scaling factors?)

While we appreciate the reviewer's suggestion, we respectfully disagree. The additional figures not only reinforce the benefit of using linear interpolation, they also show how the errors associated with the scaled quantities are propagated into the final flux estimates. By comparing the figures, the reader can clearly see that the errors in the interpolated values of the scaled quantities are mirrored in the errors in flux estimates.

10. As I understand, in the analysis as presented here, all measurement days are grouped without differentiating between e.g. seasons, phenological stages, sowing and harvest dates, cloudiness (apart from on days that are used for interpolation) and moisture availability. I'm curious whether the authors did consider doing the analysis for e.g. the growing season only, since the variability of the evaporative fraction most likely changes with the seasons (water availability, phenology (e.g. Farah et al. 2004). For applications in water resources management one might be especially interested in evaporation during the growing season and I wonder whether the maximum interval would be different.

Preliminary analyses indicate there is a seasonal pattern in the magnitude of the error with the greatest discrepancy occurring during growing season when the magnitude of the fluxes is also greatest. However, when the error is considered in relative terms, the seasonal effect vanishes. As a result, the maximum return interval would not change with the time of year.

11. It would be informative if the authors could include some figures showing the actual latent heat flux for all land cover types or preferably sites (e.g. the average annual cycle per land cover type/site), so that the reader gets some more insight into the daily and seasonal variability of the latent heat flux and evaporative fraction itself, before showing the thereof derived analysis of self-preservation of the scaling factor and interpolation errors.

Per the reviewer's suggestion a figure showing the annual pattern of the mean daytime latent heat flux for each land cover type has been added to the manuscript.
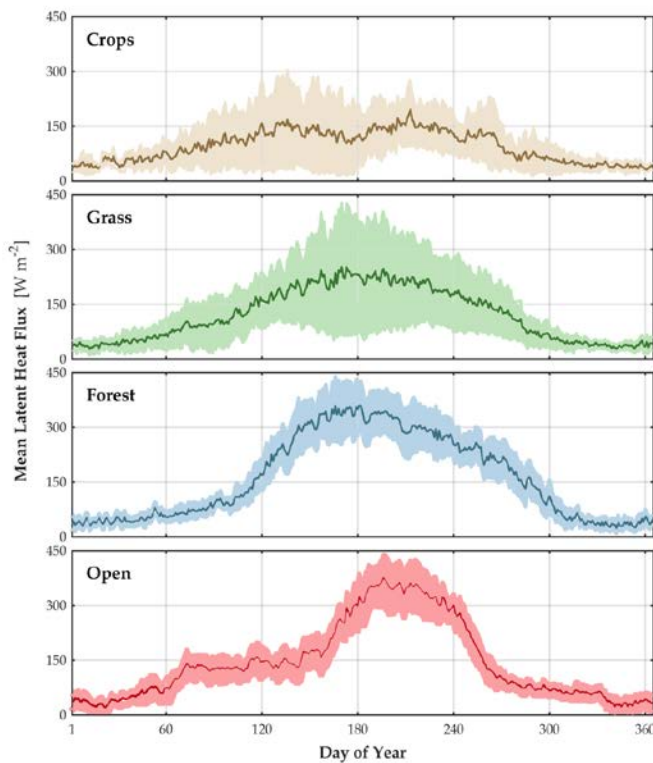

Figure 2 The mean daytime latebt heat flux is shown for each of the land cover types. The mean flux was calculated using the daytime mean flux data for all of years considered at each site. The shaded area represents one standard deviation about the mean.

Minor and technical comments

12. Although one can figure out what is meant, the authors may want to go through the introduction again and clarify / restructure the text here and there. For example page 6, lines 19-26 could be rewritten a bit more clearly. I understand it like this: Based on the daily time series time series with intervals of 2 to 32 days are generated. Per interval of length n we generate the n possible unique time series starting at the first n days of the daily series, to maximize the robustness of the statistical analysis. The generated time series were subsequently filtered for clouded days.

Although the reviewer has it exactly right, we agree that the paragraph referred to above is awkward. It has been rewritten to improve clarity. Additionally, the entire paper has been carefully proofread and additional revisions were made when needed to improve clarity or correct typographic errors.

The text the reviewer refers to now reads:

Page 7, Line 13:
For this analysis, temporal upscaling was conducted at each of the Ameriflux sites using all possible combinations of f and interpolation methods. Specifically, it was conducted with data representing return intervals of up to 32 day generated from the daytime mean data at each site. In order to maximize the robustness of the statistical analysis, all possible realizations - the unique yet equivalent time series that can be generated from the data collected at a particular site while maintaining constant return interval – were considered in the analysis. The total number of possible realizations for a given return interval is equal to the length of the return interval. The individual realizations were generated by beginning the time series on consecutive days.

13: I would include the variable and unit on the y-axis, at least in the top left panel of the figure (in stead of fx, everywhere).

Figures 1, 2, 5, and 6 indicate the variable used by the scaling function as a subscript; for example, the scaling metric associated with the available energy is denoted f_A while the one associated with net radiation is denoted f_{Rn}. In all cases, the scaled metric is dimensionless.

14. Page 4, line 15: I think Lhomme and Elguero (1999) were one of the first to describe the influence of clouds on the self-preservation of the evaporative fraction? I suggest to include the reference and maybe skip the word recent.

Per the reviewer's suggestion, the reference has been added. The sentence was also revised to improve clarity.

Page 4, Line 29:
Similarly, Lhomme and Elguero (1999) and later Van Niel et al. (2012) showed that the degree of self-preservation can be influenced by cloud cover.

15. For clarity, could the formulae of the five scaling factors and reference quantities be provided? Page 5, line 26: how is the available energy defined? I suppose you use A = Rn-G? Or are other (e.g. change of energy storage in the canopy) terms involved?

Per the reviewer's suggestion, available energy is now explicitly defined as the net radiation less the soil heat flux in the manuscript. The calculation of the remaining reference quantities that are derived from other measurements, i.e. λE_0 and λE_{eq}, are formally defined in section 2.2.

Page 4, Line 18:
For example, it is quite common to estimate ET expressed in terms of the latent heat flux (λE) using the available energy (A), here defined as the net radiation less the soil heat flux, as the reference quantity and evaporative fraction (f_A) as the scaled metric (e.g. Crago and Brutsaert , 1996; Bastiaanssen et al., 1998; Suleiman and Crago, 2004; Colaizzi et al., 2006; Hoedjes et al. 2008; van Niel et al., 2011; Delogu et al., 2012).

16: Page 6, line 28. The threshold for a cloud free day seems to tolerate quite some cloudiness? On what is the 25% based?

Preliminary analyses comparing the relatively simple model output with the observed incident solar radiation using known clear sky data showed agreement to within ±10%. To be confident that cloud cover was present, the 25% threshold was selected. This is now stated in the paper.

Clear-sky days were identified as those where the daytime mean of the measured K↓ was within 25% of the predicted value from a simple radiation model; this threshold was selected based on a preliminary analyses comparing the model results with observations on known clear-sky days.

17. Page 7, line 1: What is the total length of the observation time series after removal of the clouded days?

A sentence has been added indicating that a minimum of 1200 days was considered the analyses at each site.

Page 7, Line 27:
Although the number of days flagged due to cloudy conditions and omitted from subsequent analyses varied depending on the site and the return interval being modelled, at least 1200 days were considered for each of the analyses at each site.

18: Page 7, line 22-26. The authors could better explain the use of the 'necessary transforms'.

The section has been revised to better explain why the transforms are needed and how they were conducted. In part, it now reads:

Page 8, Line 24:
The aggregation was accomplished by calculating the arithmetic means after conducting any necessary transform. For example, both the auto-correlation and RMSE are non-additive quantities that cannot be averaged directly; instead, they must first be transformed into an additive quantity. In the case of the former, the auto-correlation was aggregated by averaging the results for the individual analysis periods at each of the sites after applying a Fisher z-transformation (Burt and Barber, 1996). Similarly, the RMSE data was averaged after first transforming it to the mean square error.

20: Typos:
Page 3, Line 31: presence of
Corrected

Page 4, Line 8:
This infrequent acquisition of imagery is due to both lengthy return intervals and the presence of cloud cover (Ryu et al., 2012; van Niel et al., 2012; Cammalleri et al., 2013).

Line 33: referred to as

Corrected

Page 4, Line 11:
To provide temporally continuous ET estimates, the moisture flux during the period between data acquisitions is often estimated using an interpolation technique commonly referred to as temporal upscaling.

Page 4, Line 22: dot behind (Van Niel et al. 2012)

This sentence has been deleted in response to a suggestion from Reviewer 2.

Line 26: For example, [something missing?] and Tasumi et al,

The spurious word "and" has been deleted.

Page 4, Line 34:
For example, Tasumi et al. (2005) proposed using the reference ET for alfalfa (ET_r) as χ; later, Allen et al. (2007) proposed using the standardized reference evapotranspiration (ET_0) as χ.

Line 30: no dot behind time

The period has been deleted.

Page 5, Line 2:
As a result, f derived from ET_r or ET_0 can be treated in much the same fashion as a crop coefficient and assumed to be nearly constant changing only slowly with time (Colaizzi et al., 2006; Chavez et al., 2009).

Page 5, Line 13: Delete Namely

Deleted per the reviewer's request

Page 5, Line 21:
These are *i.* croplands (maize (Zea mays)/soy (Glycine max) rotation); *ii.* grasslands; *iii.* forests (evergreen needleleaf and broadleaf deciduous); and, *iv.* open-canopy (shrubland and woody savanna).

Line 18: while/by maintaining

Corrected

Page 5, Line 26:
After forcing closure of the energy balance <u>while</u> maintaining a constant Bowen ratio (Twine et al. 2000) in order to more closely match the characteristics of the output from the models, the 30-minute measurements were used to calculate the various χ and f.

Page 7, Line 22: Once calculated for the individual the sites

Corrected

Page 8, Line 20:
Once calculated for the individual the sites, the statistics were aggregated to represent the typical results for a given land cover type.

Page 8. Line 13: Derived from meteorological data

Corrected

Page 9, Line 12:
Nonetheless, there were statistically significant, albeit modest, differences between the auto-correlation functions associated with f derived <u>from</u> evaporative fraction analogues and those derived from meteorological data.

Line 14: associated with

Corrected

Page 9, Line 14:
Regardless of land cover, ρ associated <u>with</u> $f_{K\downarrow}$, f_{Rn}, and f_A, tended to be greater than ρ associated with either $f_{\lambda E0}$ or $f_{\lambda Eeq}$.

Line 28. 'Interpolated estimates of each X', must be 'each f' I think?

Corrected

Page 10, Line 5:
Both RMSE and MAE of the interpolated estimates of each f were calculated for all land cover types and return intervals up to 32 days.

References
Farah, H. O., W. G. M. Bastiaanssen, and R. A. Feddes (2004). Evaluation of the temporal variability of the evaporative fraction in a tropical watershed. International Journal of Applied Earth Observation and Geoinformation 5.2 (2004): pp. 129-140.
Fisher, J.B., D. D. Baldocchi, L. Misson, T. E. Dawson, and A. H. Goldstein (2007). What the towers don't see at night: nocturnal sap flow in trees and shrubs at two Ameriflux sites in California. Tree Physiology, 27 (4) (2007): pp. 597–610. doi: 10.1093/treephys/27.4.597.
Lhomme and Elguero (1999). Examination of evaporative fraction diurnal behaviour using a soil-vegetation model coupled with a mixed-layer model. Hydrol. Earth Syst. Sci., 3 (2) (1999), pp. 259–270

Green, SR, McNaughton, K, Clothier, BE. 1989: Observations of night-time water use in kiwifruit vines and apple trees. Agric. For. Meteorol. 48,251–261.
Malek, E. 1992: Night-time evapotranspiration vs. daytime and 24 h evapotranspiration. J. Hydrol. 138, 119-129.
McNaughton, KG, Jarvis, PG, 1983: Predicting effects of vegetation changes on transpiration and evaporation. In: Kozlowski, TT (Ed.), Water Deficits and Plant Growth, vol. VII. Academic Press, pp. 1–47.
Rawson, HM, Clarke, JM. 1988: Nocturnal transpiration in wheat. Aust. J. Plant Physiol. 15, 397–406.
Sugita, M, Brutsaert, W. 1991: Daily evaporation over a region from lower boundary layer profiles measured with radiosondes. Water Resour. Res. 27, 747–752.
Tolk, JA, Howell, TA, Evett, SR. 2006: Nighttime evapotranspiration from alfalfa and cotton in a semiarid climate. Agron. J. 98, 730–736.

REVIEWER 2:
- The manuscript addresses a key challenge in the estimation of evapotranspiration from satellite data, during the interval of satellite overpasses, i.e., the temporal interpolation question. This is an important topic and has many practical implication, including water management.

The authors would like to thank the reviewer for the kind words and suggestions.

- The main shortcoming of this paper it limits the analysis to the statistical approach only without clear discussion of the physical interpretation of (cli- mate and surface) processes involved. E.g., how will ET varies between two satellite overpasses, and why?

Although a comprehensive review of the physical processes and factors controlling ET is be the scope of this paper, the introduction has been expanded to outline the key physical processes and factors controlling evapotranspiration as a foundation for understanding how ET varies over time and why temporal upscaling is needed.

Page 3, Line 2:
        As one component of a complex network of interconnected processes, evapotranspiration (ET) is influenced by numerous factors such as available energy, soil moisture, vegetation density, and humidity (Farquhar and Sharkey 1982; van de Griend and Owe 1994; Alves and Pereira 2003; Alfieri et al. 2007). For example, the amount energy available to drive ET depends on atmospheric properties, such as humidity and aerosol content, which influence atmospheric transmissivity (Brutseart 1975; Bird and Riordan 1986). The available energy is also controlled by surface properties, such as the type and density of vegetation cover and soil moisture, which influence not only the surface albedo and emissivity (Wittich 1997; Asner etal. 1998; Myneni et al.1989; Song et al. 1999; Lobell and Asner 2002), but also impact the amount of energy conducted into the ground (Friedl and Davis 1994; Kustas et al. 2000; Abu-Hamadeh 2003; Santanell and Friedl 2003). Moreover, the magnitude of the moisture flux can vary over a range of timescales in response to changes in the environmental conditions influencing ET. One example of this, which has been pointed out by Williams et al. (1998), Scott et al. (2014), and others, is the rapid and often persistent change in ET in response to a rain event.

Such discussions may allow for more physically based upscaling methods than statistical. E.g., it is well possible to compute daily ET0 using climate data measured at ground stations, which then can be used to upscale ETa (actual) evapotranspiration derived from satellite data, e.g., see Allen et al. (2001), Mohamede al. (2004).

The authors agree Reference ET (ETo) may be a useful quantity for temporal upscaling and a close parallel to the approach the reviewer suggests here is discussed in the manuscript; Reference ET is one of the quantities evaluated in the study. In contrast to Allen et al. (2001) and Mohamede al. (2004), however, ETo was not scaled by a crop coefficient. This was done in recognition that a crop coefficient, which have been derived for only a limited number of plant species, may not be available even when the vegetation composition of the remotely-sensed scene is known.

- P4, l20, a bit lengthy paragraph to confirm a known fact that it is erroneous to assume clear sky condition, if it is not!

Per the reviewer's suggestion, the paragraph, which repeats information provided previously, has been deleted for the sake of conciseness.

- P5, l13, would be good to give map of sites to easily see different climate zones

A figure showing the location of the Ameriflux sites used in this study has been added.

Figure 1 Map showing the location of the Ameriflux sites used in this study.

- P5, l15, give year from .... to .....

The study period for each of the sites is now included as a part of Table 2.

- (R2a) P6, l3, Eq. 2, would have been easier to follow if the Penman-Monrtieth equation is written complete as given in the reference (Eq. 6, p24 of Allen et al., 1998), and then give the new derivation.

  (Ea) I agree with the reviewer to change equation 2 to the way the Penman-Monteith equation is denoted in Allen et 1998 and add the density of water to the left hand site of the equation so $\lambda = \lambda_v$,

- (R2b) P6, l3, Eq. 2, what is the difference between λv and λ

  (Eb) There is no difference between λ and λ_v. Both are the latent heat of vaporization. I think the confusion is caused by the fact that the authors call the 'latent heat flux' λE instead of $\rho \lambda E$, with ρ the density of water.

The discussion of the reference latent heat flux (λE_o) has been modified to improve its clarity while maintaining λE as symbolic abbreviation for the latent heat flux. While the authors recognize in the context of an equation that λE might be misinterpreted by some as the product of two terms, i.e. λ and E, the use of λE to represent the latent heat flux is well-established and commonplace in the literature. The authors are concerned that the usage of a nonstandard symbol would not only create an inconsistency with the symbols use for other related quantities, e.g. the equilibrium latent heat flux (λE_{eq}), it would also be a potential source of confusion to readers.

Although it nearly identical the equation given in the FAO56 documentation, the reference ET (ET_o) calculated for this study was determined following the updated version of the relationship given Walter et al. (2005). This is now clarified in the paper and that relationship is shown. The conversion from ETo to λE_o is then described. In this manner, the potentially confusing situation where both λ and λ_v appear in a single equation can be avoided.

Other errors identified in the description were also correct.

Page 6, Line 1 to 15:

The first of the χ derived from meteorological data, λE_0, is derived from ET_0 which is described by Allen et al. (1998) as the hypothetical ET (or λE) from a well-watered grass surface with an assumed height of 0.12 m and albedo of 0.23. It is calculated using a simplified form of the Penman-Monteith equation. For this study, the updated relationship given by Walter et al. (2005) was used:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\frac{C_n}{(T+273)}U(e_s - e_a)}{\Delta + \gamma(1 + UC_d)} \tag{2}$$

where Δ is the slope of the saturation vapor pressure-temperature curve (kPa K^{-1}), R_n is the net radiation (W m^{-2}), G is the soil heat flux (W m^{-2}), γ is the psychrometric constant (kPa K^{-1}), C_n is a constant (37 °C s^2 m^{-2}), T is the air temperature (°C), U is the wind speed (m s^{-1}), e_s is the saturation water vapor pressure (kPa), e_a is the actual water vapor pressure (kPa), and C_d is a constant (0.24 s m^{-1}). This relationship is nearly identical to the one given in Allen et al. (1998); the two formulae differ only with regard to the assumed surface resistance. While the surface resistance is assumed to be 70 s m^{-1} by Allen et al. (1998), it is assumed to be 50 s m^{-1} in the later work. While modest, this modification yields improved results when the daytime moisture flux is calculated on an hourly basis (Walter et al. 2005). The result is converted to λE_0 by multiplying by the product of the density of water and the latent heat of vaporization. Similarly, λE_{eq}, which can be thought of as the energy-driven moisture flux that is independent of surface resistance, can be expressed according to:

$$\lambda E_{eq} = A\frac{\Delta}{\Delta + \gamma} \tag{3}$$

with the variables defined as above (McNaughton, 1976; Raupach, 2001).

- P6, l14, would be good to briefly describe the interpolation methods, linear, spline, and hermite. Also the description of some parameters seems very short in some places, e.g., the autocorrelation calculation.

Per the reviewer's suggestion, additional information describing the algorithms and implementation of the interpolation methods has been added to the manuscript.

Page 8, Line 28:

As the name implies, the piecewise linear interpolation estimates f using a family of $n - 1$ linear relationships defined such that the linearly-interpolated f ($\hat{f}_L$) at time t is determined according to:

$$\hat{f}_{L_i}(t) = f_i + (t_{i+1} - t_i)m_i h \quad t_i \leq t \leq t_{i+1} \tag{4}$$

where n is the number of observed data points, f_i is the known f at time t_i, m_i is the slope of straight line relationship for the period between t_i and t_{i+1} defined as $m_i = (f_{i+1} - f_i)/(t_{i+1} - t_i)$, and h is the time normalized between 0 and 1 and is defined as $h = (t - t_i)/(t_{i+1} - t_i)$. The piecewise cubic spline interpolation function is family of $n - 1$ cubic polynomials defined such that the interpolated f ($\hat{f}_S$) at time t is determined according to:

$$\hat{f}_{S_i}(t) = f_i + a_i[(t_{i+1} - t_i)h]^3 + b_i[(t_{i+1} - t_i)h]^2 + c_i[(t_{i+1} - t_i)h] \quad t_i \leq t \leq t_{i+1} \tag{5}$$

where the coefficients a_i, b_i, and c_i are determined by simultaneously solving the series of $n - 1$ equations with the constraints that the interpolation function, as well as its first and second derivatives, must be continuous and pass exactly through the known values of f. Similarly, the final interpolation technique, piecewise hermite cubic spline, defines the

$$\hat{f}_{H_i}(t) = (2h^3 - 3h^2 + 1)f_i + (-2h^3 - 3h^2)f_{i+1} + \cdots$$
$$h(h^2 - 2h + 1)(t_{i+1} - t_i)s_i + h(h^2 - h)(t_{i+1} - t_i)s_{i+1} \quad t_i \leq t \leq t_{i+1} \tag{6}$$

where s_i is the slope of the curve at time t_i (De Boor, 1994). For this study, it is calculated according to:

$$s_i = \frac{1}{2}\left(\frac{f_{i+1} - f_i}{t_{i+1} - t_i} + \frac{f_i - f_{i-1}}{t_i - t_{i-1}}\right) \tag{7}$$

and the variables are defined as above (Moler, 2004).

- P 7, l10, would also be good to give the % of the RMSE.

It is unclear what is meant here. The text referred to by the reviewer is describes the calculation of RMSE.

-P10, l25, it would have been interesting to test the statistical results derived from the analysis against actual upscaling of satellite-based ET results - There is some spelling and grammar errors

While we agree with the reviewer in principle that such an analysis has the potential to provide many useful insights, there is no remote sensing-based data set with a sufficiently high temporal resolution available. Thus, as is discussed in the manuscript, as well as other such as Ryu et al. (2012) and Cammalleri et al. (2014), in-situ observations are a necessary proxy for studies investigating the temporal upscaling of remote sensing-based flux estimates.

References:
- Allen, R.G., Bastiaanssen, W.G.M., Tasumi, M., Morse, A., 2001. Evap- otranspiration on the watershed scale using the SEBAL model and Landsat images, ASAE Meeting Presentation, Paper Number 01-2224, Sacramento, California, USA, July 30–August 1, 2001,.
-Mohamed, Y.A., Bastiaanssen, W.G.M., Savenije, H.H.G., 2004, Spatial variability of evaporation and moisture storage in the swamps of the upper Nile studied by remote sensing techniques. J. Hydrology 289:145–164.

Walter, I. A., Allen, R. G., Elliott, R. L., Itenfisu, D., Brown, P., Jensen, M. E., Mecham, B., Howell, T. A., Snyder, R., Echings, S., Spofford, T., Hattendorf, M., Martin, D. L., Cuenca, R. H., and Wright, J. L.: The ASCE Standardized Reference Evapotranspiration Equation. Technical Committee report to the Environmental and Water Resources Institute of the American Society of Civil Engineers from the Task Committee on Standardization of Reference Evapotranspiration, 173 pp., Reston, VA, 2005.

Effect of the Revisit Interval and Temporal Upscaling Methods on the Accuracy of Remotely-Sensed Evapotranspiration Estimates

Joseph G. Alfieri[1], Martha C. Anderson[1], William P. Kustas[1], Carmelo Cammalleri[2]

[1] US Department of Agriculture, Agricultural Research Service,
   Hydrology and Remote Sensing Laboratory, Beltsville, MD, USA
[2] Joint Research Centre, European Commission, Ispra, Italy

Correspondence to: Joseph G. Alfieri (joe.alfieri@ars.usda.gov)

1

Abstract. Accurate spatially distributed estimates of actual evapotranspiration (ET) derived from remotely sensed data are critical to a broad range of practical and operational applications. However, due to lengthy return intervals and cloud cover, data acquisition is not continuous over time, particularly for satellite sensors operating at medium (~100 m) or finer resolutions. To fill the data gaps between clear-sky data acquisitions, interpolation methods that take advantage of the relationship between ET and other environmental properties that can be continuously monitored are often used. This study sought to evaluate the accuracy of this approach, which is commonly referred to as temporal upscaling, as a function of satellite revisit interval. Using data collected at 20 Ameriflux sites distributed throughout the contiguous United States and representing 4 distinct land cover types (cropland, grassland, forest, and open canopy) as a proxy for perfect retrievals on satellite overpass dates, this study assesses daily ET estimates derived using 5 different reference quantities (incident solar radiation, net radiation, available energy, reference ET, and equilibrium latent heat flux) and 3 different interpolation methods (linear, cubic spline, and hermite spline). Not only did the analyses find that the temporal autocorrelation, i.e. persistence, of all of the reference quantities was short, it also found that those land cover types with the greatest ET exhibited the least persistence. This carries over to the error associated with both the various scaled quantities and flux estimates. In terms of both the root mean square error (RMSE) and mean absolute error (MAE), the errors increased rapidly with increasing return interval following a logarithmic relationship. Again, those land cover types with the greatest ET showed the largest errors. Moreover, using a threshold of 20% relative error, this study indicates that a return interval of no more than 5 days is necessary for accurate daily ET estimates. It also found that the spline interpolation methods performed erratically for long return intervals and should be avoided.

1 Introduction

As one component of a complex network of interconnected processes, evapotranspiration (ET) is influenced by numerous factors such as available energy, soil moisture, vegetation density, and humidity (Farquhar and Sharkey 1982; van de Griend and Owe 1994; Alves and Pereira 2003; Alfieri et al. 2007). For example, the amount energy available to drive ET depends on atmospheric properties, such as humidity and aerosol content, which influence atmospheric transmissivity (Brutseart 1975; Bird and Riordan 1986). The available energy is also controlled by surface properties, such as the type and density of vegetation cover and soil moisture, which influence not only the surface albedo and emissivity (Wittich 1997; Asner etal. 1998; Myneni et al.1989; Song et al. 1999; Lobell and Asner 2002), but also impact the amount of energy conducted into the ground (Friedl and Davis 1994; Kustas et al. 2000; Abu-Hamadeh 2003; Santanell and Friedl 2003). Moreover, the magnitude of the moisture flux can vary over a range of timescales in response to changes in the environmental conditions influencing ET. One example of this, which has been pointed out by Williams et al. (1998), Scott et al. (2014), and others, is the rapid and often persistent change in ET in response to a rain event.

Because it is a fundamental linkage between numerous biogephysical and biogeochemical processes, accurate information regarding evapotranspiration (ET) is critical for a broad range of scientific and practical applications with significant social, economic, and environmental impacts. For example, reliable information about ET is essential for accurately forecasting weather and assessing the impacts of changing climate (Katul et al., 2012; Wang and Dickinson, 2012); monitoring and mitigating the adverse effects of extreme weather events such as drought (Anderson et al., 2007, 2011, 2016; Otkin et al. 2016); and, identifying and predicting the changes in both the biogeographical characteristics of ecosystems and the services they provide in response to changing environmental conditions (Hawkins and Porter, 2003; Kreft and Jetz, 2007; Midgley et al. 2002). However, as pointed out by Seguin and Itier (1983), Abdelghabi et al. (2008), and Anderson et al. (2012), among others, perhaps the most important application of ET data is providing information critical to satisfying the competing demands for scare water resources.

Already, the competing demands for fresh water by agricultural, industrial, and urban consumers exceed the available supply for nearly one-third of the world's population (Qadir et al. 2003) and it is predicted that number will increase to more than two-thirds of the global population in the coming decades (Wallace, 2008; Vörösmarty et al., 2010). To meet the current and future demand for water, resource managers and other policymakers must make informed decisions regarding the needs of competing stakeholders when allocating limited water resources in order to maximize their effective use. In the case of irrigated agriculture, which is the largest consumer of fresh water and accounts for 1200 km^3 or approximately 85% of annual current water use (Drooger et al., 2010; Thenkabail et al., 2010), the need for water is largely driven by evaporative loss. Thus, ET measurements are needed not only to monitor evaporative water loss and determine crop irrigation needs, it is also needed to develop the irrigation techniques and management practices necessary to ensure the efficient use of water in agricultural environments (Howell, 2001; Schultz and Wrachien, 2002; Gordon et al., 2010; de Fraiture and Wichelns, 2010).

3

While in situ observations are invaluable for some of these applications, many of them require spatially distributed measures of ET at field to continental scales that cannot be supplied by the existing flux measurement infrastructure. Remote sensing-based approaches are the only viable mean for monitoring ET over this continuum of scales (McCabe et al., 2008; Kalma et al., 2008; Gonzalez-Dugo et al., 2009). As discussed by Anderson et al. (2012), any comprehensive program for monitoring water resources will by necessity use remote sensing data collected by multiple platforms at a range of spatial and temporal scales.

Nonetheless, remote sensing is not without limitations. Chief among these is the infrequent acquisition of the medium to high-resolution imagery needed as input for remote sensing-based models to determine ET. This infrequent acquisition of imagery is due to both lengthy return intervals and the presence of cloud cover (Ryu et al., 2012; van Niel et al., 2012; Cammalleri et al., 2013). To provide temporally continuous ET estimates, the moisture flux during the period between data acquisitions is often estimated using an interpolation technique commonly referred to as temporal upscaling. This well-established approach, which can be applied at either sub-daily or daily time steps, estimates the moisture flux as the product of some reference quantity (χ) and its associate scaled metric (f) according to:

$$\widehat{ET}_t = \chi_t f_t \tag{1}$$

where $\widehat{ET}$ is the estimated ET and t indicates the time period of the estimate. While it is typically related to the moisture flux, χ is a quantity that can be measured or estimated more readily than the moisture flux itself. The scaled metric is the ratio between χ and the moisture flux. For example, it is quite common to estimate ET expressed in terms of the latent heat flux (λE) using the available energy (A), here defined as the net radiation less the soil heat flux, as the reference quantity and evaporative fraction (f_A) as the scaled metric (*e.g.* Crago and Brutsaert , 1996; Bastiaanssen et al., 1998; Suleiman and Crago, 2004; Colaizzi et al., 2006; Hoedjes et al. 2008; van Niel et al., 2011; Delogu et al., 2012).

For the periods between data retrievals, f is estimated via interpolation. As a result, this approach is predicated on the assumption that f is self-preserving, *i.e.* it is constant or nearly constant, and thus varies only slowly over time (Brutsaert and Sugita, 1992; Nichols and Cuenca, 1993; Crago, 1996). In order to conform to this assumption, the components of the radiation or energy budget are often selected as χ such that f is an analogue of evaporative fraction. Examples of these quantities include the incident solar radiation ($K{\downarrow}$; Jackson et al. 1983; Zhang and Lemeur, 1995), or extraterrestrial solar radiation ($K{\downarrow}_{TOA}$; Ryu et al. 2012). However, a number of recent studies indicate the assumption of self-preservation is only approximate for these quantities. For example, both Gentine et al. (2007) and Hoedjes et al. (2008) showed that the self-preservation of evaporative faction is sensitive to soil moisture conditions and fractional vegetation cover. Similarly, Lhomme and Elguero (1999) and later Van Niel et al. (2012) showed that the degree of self-preservation can be influenced by cloud cover. As such, the assumption of clear-sky conditions is a significant potential source of error in the ET estimates that must be considered when utilizing or evaluating temporally upscaled moisture flux data (van Niel et al., 2012; Peng et al., 2013; Cammalleri et al., 2014). Other studies have focused on using a quantity derived from the local meteorological conditions as χ because it would consider many of the meteorological factors that influence the moisture flux. For example, Tasumi et al. (2005) proposed using the reference ET for alfalfa (ET_r) as χ; later, Allen et al. (2007) proposed using the

4

standardized reference evapotranspiration (ET_0) as χ. In both cases, the resulting f is equivalent to a crop coefficient and would share its characteristics. As a result, f derived from ET_r or ET_0 can be treated in much the same fashion as a crop coefficient and assumed to be nearly constant changing only slowly with time (Colaizzi et al., 2006; Chavez et al., 2009).

By assessing the error introduced into ET estimates by temporal upscaling under realistic conditions, this study sought to achieve two goals: i. to provide insights into the relative strengths of the differing temporal upscaling approaches, and *ii*. to determine the maximum return interval threshold for obtaining acceptable estimates of daily ET. Specifically, this study uses in situ measurements collected over a variety of land cover types as a proxy for remotely-sensed data to evaluate the impact of multiple reference quantities, interpolation techniques, and revisit intervals on the estimated daily moisture flux. The study focuses on daytime mean data to evaluate temporal upscaling at a daily time step. It also assumes perfect retrieval of the flux; in other terms, no error was introduced into ET data to approximate the error or uncertainty in the estimates of ET from the remote sensing-based models. Since any errors in the remote sensing-based ET estimates propagate into the calculation of f and the subsequent temporal upscaling, this analysis represents the best-case scenario. The following section provides an overview of the field measurements along with the reference quantities, interpolation techniques, and evaluation methods used in this study. Section 3 provides a discussion of the results of this study while the final section encapsulates the conclusions that can be drawn from those results.

2. Methods

2.1. Datasets

Data, including local meteorological conditions, surface fluxes, and surface conditions (Table 1), collected at numerous sites within the Ameriflux network (Baldocchi et al., 2001) were used for this study. Specifically, the data were collected at 20 Ameriflux sites (Fig. 1 and 2; Table 2) distributed across the contiguous United States and representing four distinct land cover types. These are *i*. croplands (maize (Zea mays)/soy (Glycine max) rotation); *ii*. grasslands; *iii*. forests (evergreen needleleaf and broadleaf deciduous); and, *iv*. open-canopy (shrubland and woody savanna). Measurements were collected for a minimum of five years at each of the sites selected. Further information regarding the field sites, measurement procedures, and post-processing protocols for Ameriflux is presented in Baldocchi et al., 2001; the data are archived at the Oak Ridge National laboratory and available at http://ameriflux.ornl.gov/.

After forcing closure of the energy balance while maintaining a constant Bowen ratio (Twine et al. 2000) in order to more closely match the characteristics of the output from the models, the 30-minute measurements were used to calculate the various χ and f. Finally, the daytime mean of the fluxes and other necessary quantities were calculated for use in the subsequent analyses. Although it can be taken as nominally as the period between 0800 and 1800 LST, daytime is defined herein as the period between the first and last measurements during a given day when the incident solar radiation exceeded 100 W m^{-2}.

Deleted: .

Deleted: is

Deleted: Namely, t

Deleted: by

Deleted: energy balance

Deleted: Next, t

Deleted: these

Deleted: and the fluxes

5

2.2. Reference Quantities and Scaled Metrics

The first of the χ derived from meteorological data, λE_0, is derived from ET_0 which is described by Allen et al. (1998) as the hypothetical ET (or λE) from a well-watered grass surface with an assumed height of 0.12 m and albedo of 0.23. It is calculated using a simplified form of the Penman-Monteith equation. For this study, the updated relationship given by Walter et al. (2005) was used:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\frac{C_n}{(T+273)}U(e_s - e_a)}{\Delta + \gamma(1 + UC_d)} \tag{2}$$

where Δ is the slope of the saturation vapor pressure-temperature curve (kPa K^{-1}), R_n is the net radiation (W m^{-2}), G is the soil heat flux (W m^{-2}), γ is the psychrometric constant (kPa K^{-1}), C_n is a constant (37 °C s^2 m^{-2}), T is the air temperature (°C), U is the wind speed (m s^{-1}), e_s is the saturation water vapor pressure (kPa), e_a is the actual water vapor pressure (kPa), and C_d is a constant (0.24 s m^{-1}). This relationship is nearly identical to the one given in Allen et al. (1998); the two formulae differ only with regard to the assumed surface resistance. While the surface resistance is assumed to be 70 s m^{-1} by Allen et al. (1998), it is assumed to be 50 s m^{-1} in the later work. While modest, this modification yields improved results when the daytime moisture flux is calculated on an hourly basis (Walter et al. 2005). The result is converted to λE_0 by multiplying by the product of the density of water and the latent heat of vaporization. Similarly, λE_{eq}, which can be thought of as the energy-driven moisture flux that is independent of surface resistance, can be expressed according to:

$$\lambda E_{eq} = A\frac{\Delta}{\Delta + \gamma} \tag{3}$$

with the variables defined as above (McNaughton, 1976; Raupach, 2001).

2.3. Interpolation Techniques

In addition to piecewise linear interpolation, two piecewise spline interpolation methods were evaluated as a part of this study, namely cubic and hermite spline interpolation. (These are indicated when necessary hereafter by the subscripts L, S, H, respectively.) In contrast to linear interpolation, which tends to yield accurate results only when the underlying data vary smoothly over time, the splining methods are less prone to error when the observed data change abruptly (Trefethen 2013). Similarly, the more computational complex hermite spline method typically yields more accurate results when the gaps between observed data points are large (DeBoor, 1994).

As the name implies, the piecewise linear interpolation estimates f using a family of $n - 1$ linear relationships defined such that the linearly-interpolated f ($\hat{f}_L$) at time t is determined according to:

$$\hat{f}_{L_i}(t) = f_i + (t_{i+1} - t_i)m_i h \quad t_i \le t \le t_{i+1} \tag{4}$$

where n is the number of observed data points, f_i is the known f at time t_i, m_i is the slope of straight line relationship for the period between t_i and t_{i+1} defined as $m_i = (f_{i+1} - f_i)/(t_{i+1} - t_i)$, and h is the time normalized between 0 and 1 and is

6

Deleted: s

Deleted: For this study, five χ, along with their associated f, discussed in the literature were evaluated. The first three of these quantities ($K\downarrow$, R_n, and A) yield analogues of evaporative fraction and require incrementally greater information regarding the surface radiation and energy budgets. (Hereafter, the associated scaled metrics are denoted as f_χ where the subscript represents the reference quantity; for example, the ratio of λE and A is indicated as f_A.) The remaining χ are derived from estimates of the moisture flux calculated based on local meteorological conditions. These are ET_0 expressed in terms of energy (λE_0) and the equilibrium latent heat flux (λE_{eq}). The first of the, λE_0, is described by Allen et al. (1998) as the hypothetical ET (or λE) from a grass reference surface with an assumed height of 0.12 m, albedo of 0.23, and surface resistance of 70 s m^{-1} of a moderately moist soil. It is defined using a simplified form of the Penman-Monteith relationship as:¶

$\lambda E_0 = \lambda_v \frac{a\,\Delta A + \gamma\frac{C_n}{T_K}UD}{\Delta + \gamma(1 + UC_d)}$ — — — — — — — (2)¶

where λ_v is the latent heat of vaporization (J kg^{-1}), a is a constant (1.1333×10^{-4} kg J^{-1}), Δ is the slope of the saturation vapor pressure-temperature curve (kPa K^{-1}), A is the available energy (W m^{-2}), γ is the psychrometric constant (kPa K^{-1}), C_n is a constant (1.1333×10^{-5} K s^2 m^{-2}), T_K is the air temperature (K), U is the wind speed (m s^{-1}), D is the water vapor pressure deficit (kPa), and C_d is a constant (0.25 s m^{-1}; Allen et al., 1998).

Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1

defined as $h = (t - t_i)/(t_{i+1} - t_i)$. The piecewise cubic spline interpolation function is family of n - 1 cubic polynomials defined such that the interpolated f $(\hat{f}_S)$ at time t is determined according to:

$$\hat{f}_{S_i}(t) = f_i + a_i[(t_{i+1} - t_i)h]^3 + b_i[(t_{i+1} - t_i)h]^2 + c_i[(t_{i+1} - t_i)h] \quad t_i \leq t \leq t_{i+1} \tag{5}$$

where the coefficients a_i, b_i, and c_i are determined by simultaneously solving the series of $n - 1$ equations with the

5 constraints that the interpolation function, as well as its first and second derivatives, must be continuous and pass exactly through the known values of f. Similarly, the final interpolation technique, piecewise hermite cubic spline, defines the

$$\hat{f}_{H_i}(t) = (2h^3 - 3h^2 + 1)f_i + (-2h^3 - 3h^2)f_{i+1} + \cdots \\ h(h^2 - 2h + 1)(t_{i+1} - t_i)s_i + h(h^2 - h)(t_{i+1} - t_i)s_{i+1} \quad t_i \leq t \leq t_{i+1} \tag{6}$$

where s_i is the slope of the curve at time t_i (De Boor, 1994). For this study, it is calculated according to:

$$s_i = \frac{1}{2}\left(\frac{f_{i+1} - f_i}{t_{i+1} - t_i} + \frac{f_i - f_{i-1}}{t_i - t_{i-1}}\right) \tag{7}$$

10 and the variables are defined as above (Moler, 2004).

For this analysis, temporal upscaling was conducted at each of the Ameriflux sites using all possible combinations of f and interpolation methods. Specifically, it was conducted with data representing return intervals of up to 32 day generated from the daytime mean data at each site. In order to maximize the robustness of the statistical analysis, all possible realizations - the unique yet equivalent time series that can be generated from the data collected at a particular site while

15 maintaining constant return interval – were considered in the analysis. The total number of possible realizations for a given return interval is equal to the length of the return interval. The individual realizations were generated by beginning the time series on consecutive days.

Again, to emulate the temporal upscaling of flux data derived from remotely sensed-data as closely as possible, efforts were made to ensure that the observations used for the interpolation were collected on clear-sky days. Clear-sky days

20 were identified as those where the daytime mean of the measured $K\downarrow$ was within 25% of the predicted value from a simple radiation model; this threshold was selected based on a preliminary analyses comparing the model results with observations on known clear-sky days. The incident solar radiation was estimated as the product of $K\downarrow_{\text{TOA}}$ calculated following Meeus et al. (1991) and atmospheric transmissivity calculated according to Brutsaert (1975). In order to ensure a constant return interval for a given interpolation, if a day was judged to be cloudy, both the observed flux on that day and the estimated flux

25 for those subsequent days derived from it were omitted from the statistical analysis. Although the number of days flagged due to cloudy conditions and omitted from subsequent analyses varied depending on the site and the return interval being modelled, at least 1200 days were considered for each of the analyses at each site.

2.4. Statistical Metrics

As discussed by Wilks (2006), persistence, *i.e.* the degree of self-preservation, can be assessed via auto-correlation

30 (ρ). For a given lag (h), *i.e.* the offset between measurements pairs, the auto-correlation is defined according to:

$$\rho = \frac{\sum_{i=1}^{n-h}[(x_i - \bar{x}_-)(x_{i+h} - \bar{x}_+)]}{\sqrt{\sum_{i=1}^{n-h}(x_i - \bar{x}_-)^2 \sum_{i=1}^{n-h}(x_{i+h} - \bar{x}_+)^2}} \tag{8}$$

where n is the number of data points, $\bar{x}_-$ is the mean of the first $n - h$ data points and $\bar{x}_+$ is the mean of the final $n - h$ data points.

A pair of statistics are used to evaluate the accuracy of the temporal upscaling. The first of these is the root mean
5   square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2} \tag{9}$$

where n is the number of data points, x is the observed flux, and $\hat{x}$ is the flux predicted by temporal upscaling. However, because the squared difference term in RMSE tends to overemphasize the effects of large errors (Legates and McCabe, 1999; Willmott and Matsuura, 2005; Willmott et al., 2012), the mean absolute error (MAE) was also calculated as follows:

10    $$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i| \tag{10}$$

with the variables defined as above.

Once calculated for the individual the sites, the statistics were aggregated to represent the typical results for a given land cover type. The aggregation was accomplished by calculating the arithmetic means after conducting any necessary transform. For example, both the auto-correlation and RMSE are non-additive quantities that cannot be averaged directly;
15   instead, they must first be transformed into an additive quantity. In the case of the former, the auto-correlation was aggregated by averaging the results for the individual analysis periods at each of the sites after applying a Fisher z-transformation (Burt and Barber, 1996). Similarly, the RMSE data was averaged after first transforming it to the mean square error.

3. Results and Discussion

20   **3.1 Persistence of Scaled Quantities**

Due to its importance in determining the accuracy of the estimates, the persistence or degree of self-preservation exhibited by the various f used in this study was evaluated by determining its autocorrelation function. For each site, the autocorrelation was calculated for each contiguous segment of daytime mean data that was at least 48 days in length (1.5 times the maximum return interval considered herein).

25   As can be seen in Fig. 3, which shows the mean auto-correlation function for each f and land cover type, all f performed similarly. In all cases, ρ decreased in inverse proportion to h, dropping to less than 0.50 within three to ten days. Also, for any given land cover, the mean auto-correlation functions for the analogues of evaporative fraction, namely $f_{K\downarrow}$, f_{Rn}, and f_A, were statistically indistinguishable from one another based on t-tests conducted at the 95% confidence level. Similarly, no statistically significant difference between the mean auto-correlation functions of $f_{\lambda E0}$ and $f_{\lambda Eeq}$ was found.
30   Nonetheless, there were statistically significant, albeit modest, differences between the auto-correlation functions associated

8

Deleted: 4

Deleted: Three

Deleted: 5

Deleted: 6

Deleted: The third metric is the modified index of agreement (D) defined by Willmott et al. (1985) as:¶
$D = 1 - \frac{\sum_{i=1}^{n}|x_i - \hat{x}_i|}{\sum_{i=1}^{n}|x_i - \bar{x}| + |\hat{x}_i - \bar{x}|}$ · · · · · · · · · (7)¶
where $\bar{x}$ is the mean of the observational data and the other variables are defined as above. Like MAE, the use of the absolute difference for this metric circumvents the tendency of similar statistics based on squared errors to be overly influenced by outliers or large errors. However, since it bounded between zero, which indicates no agreement, and unity, which indicates perfect agreement, D both indicates the relative magnitude of the error and facilitates the comparison of the error from differing upscaling methods.

Deleted: For example, the auto-correlation was aggregated by averaging the results for the individual analysis periods and sites after applying a Fisher z transformation (Burt and Barber, 1996).

Deleted: 1

with f derived from evaporative fraction analogues and those derived from meteorological data. Regardless of land cover, ρ associated with $f_{K\downarrow}$, f_{Rn}, and f_A, tended to be greater than ρ associated with either $f_{\lambda E0}$ or $f_{\lambda Eeq}$. On average, the difference was approximately 0.03.

The results of this analysis, which are consistent with results of other studies (Farah et al. 2004; Lu et al., 2013) that found significant day-to-day and seasonal variations in evaporative fraction, indicates the long-term persistence of f is very limited. This result also suggests that interpolated values of f may not accurately reflect the actual values and, as a result, may be a key source of error when using temporal upscaling to estimate the moisture flux between image retrievals.

The figure also shows there was significant variability from site-to-site within a given land cover type, particularly for longer lags. Although the specific causes of these differences are not fully understood, there are number of factors that likely contribute. For example, there are difference in both species composition and climate at the various sites. Consider, as an example, the forest class which includes both coniferous and broadleaf deciduous forest. Moreover, the species composition varies even among sites of the same forest type; for example dominant species at the Niwot ridge site are Subalpine fir (Abies lasiocarpa) and Engelmann spruce (Picea engelmannii) while, as the name implies, the dominant species at the Loblolly Pine is loblolly pine (Pinus taeda). At the same time, the mean annual temperature at the forested sites ranged from 1.5 °C to 14.4 °C while the mean annual precipitation varied from 800 mm to 1372 mm. Similarly the mean annual temperature and precipitation at the cropland sites, which are all planted on a rotation of maize and soy, range between 6.4 °C and 11.0 °C and 789 mm and 991 mm, respectively.

Further analysis shows differences in the mean autocorrelation functions exist between land cover types. Regardless of the scaled quantity considered, the mean autocorrelation function decreases most rapidly over forested sites and the most slowly over the open canopy sites. Indeed, if the lag where the mean autocorrelation function reaches some threshold value, e.g. 0.50, is plotted as a function of the mean daytime latent heat flux (Fig. 4), it can be seen that persistence decreases exponentially with the increasing moisture flux. While the underlying cause of this relationship is unclear, it suggests the return interval necessary to achieve accurate estimates of ET via temporal upscaling will be longer over relatively dry regions with a low moisture flux than over regions where ET is high.

3.2 Accuracy of the Interpolated Scaled Quantities

Both RMSE and MAE of the interpolated estimates of each f were calculated for all land cover types and return intervals up to 32 days. As can be seen in Fig. 5 and Fig. 6, both metrics behaved similarly; regardless of the land cover type, scaled quantity, or interpolation method considered, the error increased rapidly with increasing return interval until a plateau was reached. In all cases, the RMSE, which increased according to a logarithmic function of return interval, reached 75% of its peak value within five days. For comparison, the mean maximum RMSE for each land cover type was 0.26, 0.28, and 0.17 for croplands, grasslands, forest, and open canopies, respectively. Although it also increased logarithmically, the amount of time needed for MAE to reach 75% of the peak value was more variable, ranging between 5 to 10 days. Again, for purposes of comparison, the mean maximum MAE was 0.22, 0.14, 0.16, and 0.10, respectively, for croplands, grasslands,

9

Deleted: 2

Deleted: This

Deleted: χ

Deleted: 3

Deleted: 4

forest, and open canopies. Further, MAE increased most rapidly for those land cover types that exhibited the highest moisture flux. The largest error, whether measured in terms of RMSE or MAE, also tended to be associated with the forest and cropland sites where the mean ET was largest.

The results also show that all of the interpolation methods yielded similar results for short return intervals of less than eight days. In contrast, for longer return intervals, both RMSE and MAE of the estimates using the spline interpolation methods were greater than when linear interpolation is used (Fig.5 and 6). Moreover, the error of the estimates tended to much noisier for the spline techniques, particularly the cubic spline method which exhibited periods of very large errors. These large noisy errors, which are most evident in RMSE – perhaps because it is more sensitive to outliers than MAE – are indicative of "overshoot" errors by the spline interpolation. The large errors are also most pronounced for those land cover types that also demonstrated the highest average ET and the lowest autocorrelation

3.3 Accuracy of the Latent Heat Flux Estimates

Not unexpectedly, the accuracy of the moisture fluxes estimated via temporal upscaling closely mirrors the accuracy of the interpolated f. As was the case with f, both the RMSE and MAE of the flux estimates increase rapidly with return interval to a maximum value following a logarithmic function (Fig. 7 and Fig. 8). In the case of RMSE, the maximum error ranged between 31 W m^{-2} and 66 W m^{-2}. In the case of MAE, it ranged between 22 W m^{-2} and 54 W m^{-2}. Again, the greatest error is associated with the land cover with the highest ET, i.e. forest and cropland.

These plots, like those for f, show little difference among the interpolation techniques when the return interval is short. For return intervals longer than about 8 days, however, the spline interpolation techniques, and especially the cubic spline method, can introduce large errors into the flux estimates due to the "overshoot" errors in the interpolation of f. These large noisy errors are most evident in the RMSE of forested sites (Fig. 8), but may also be seen to a lesser extent at the cropland sites. Overall, this suggests there is no substantive advantage of using the more computational complex spline techniques over linear regression; rather, the propensity of spline methods to introduce large errors due to interpolation "overshoot" indicates these techniques should be avoided.

The accuracy, and thus utility, of the various f was evaluated while focusing specifically on the results when linear interpolation was used. Regardless of f, an intercomparison of the estimated fluxes using t-tests conducted at the 95% confidence level, indicated there was no statistically significant difference in either the flux estimates or the error due to temporal upscaling when the return interval was less than eight days. For longer return intervals, analyses using RMSE (Fig. 9) and MAE (not shown), which yielded similar results, indicated the error due temporal upscaling was very similar when f_{Rn}, f_A, or $f_{\lambda Eeq}$ was used. Indeed, the error introduced using any of these three quantities was statistically identical based on t-tests conducted at the 95% confidence level. Moreover, with the exception of the forest sites, where the error due to temporal upscaling using $f_{K\downarrow}$ was the same as the error introduced by using f_{Rn}, f_A, or $f_{\lambda Eeq}$, temporal upscaling using f_{Rn}, f_A, and $f_{\lambda Eeq}$ consistently introduced the least error. For a 10-day return interval, as an example, the percent error introduced by these quantities ranges between 21% and 23% depending on land cover. In contrast, temporal upscaling using $f_{\lambda E0}$ introduced the

10

Deleted: 3
Deleted: 4
Deleted: for
Deleted: 2
Deleted: 5
Deleted: 6.
Deleted: 5
Deleted: d
Deleted: based on
Deleted: ,
Deleted: there
Deleted: short,
Deleted: 7

greatest error. Again, for a 10-day return interval, the percent error associated with $f_{\lambda E0}$ ranges between 24% and 30% depending on land cover.

The analysis of D reinforces the earlier results. Initially, D decreases rapidly with increasing return interval to less than 0.75 within 3 days and less than 0.50 within 4 to 7 days. This sharp decline in D, which is consistent for all land cover types and f (Fig. 10), indicates there is only moderate agreement between the observed flux and that estimated via temporal upscaling for all but the shortest return intervals. Also, while the moisture flux estimated using f_{Rn} tends to maintain the highest degree of agreement, followed closely by estimates using f_A, and $f_{\lambda Eeq}$, the variability in D tends to be modest; for any given land cover type and return interval, D varies by 0.024, on average. These results reconfirm the earlier results by indicating that the accuracy of temporal upscaling is greatest for f_{Rn}, f_A, and $f_{\lambda Eeq}$.

3.4 Estimating Optimal Return Interval Thresholds

Again focusing on the flux estimates when linear interpolation was used, the return interval threshold yielding errors of less than 20% in the daily ET estimates was identified (Table 3). The 20% threshold was selected because it is the nominal uncertainty commonly associated with in situ observations such as those collected via eddy covariance. While the return interval associated with the 20% threshold varied depending on land cover type and f, the longest return intervals are associated with $f_{\lambda Eeq}$ followed by f_{Rn}, and f_A, which yield statistically identical results, and finally $f_{K\downarrow}$ and $f_{\lambda E0}$, which also yield statistically indistinguishable results based on t-tests at the 95% confidence level. The range of values among the various f was 2 days, on average. This again indicates that the accuracy of temporal upscaling is greatest for f_{Rn}, f_A, and $f_{\lambda Eeq}$.

By plotting the average threshold return interval for each land cover class against its corresponding mean latent heat flux for that class (Fig. 10), it can be seen that length of the return interval that will result in no more than 20% error decreases with the increasing moisture flux. Like ρ, the relationship follows an exponential decay function. In this case, however, the curve has a lower bound of five days. Based on this, the maximum return interval that can be expected to introduce less than 20% error to the flux estimates via temporal upscaling for all land cover classes is 5 days. If a threshold of 10% relative error is used, the threshold falls to only 3 days. Importantly, since the determination of the maximum return interval was made assuming there is no error in the moisture flux used to calculate f, they represent the best-case scenario. In practice, any error in the flux retrieval will propagate into the interpolated flux. As a result the maximum return interval would be somewhat shorter.

4. Conclusions

The results of this study indicate that the day-to-day persistence of χ typically used in the temporal upscaling of satellite-based ET retrievals is quite limited. The autocorrelation of daytime means of these quantities decreases to less than 0.5 within 10 day and to less than 0.25 in 7 to 24 days depending on land cover class. More generally, it was found that the number of days for ρ to reach to a given threshold decreases with increasing λE following a well-defined exponential decay

Deleted: 8

Deleted: 3

Deleted: 9

function. This suggests that the utility of temporal upscaling is limited to short return intervals, especially for land covers such as forest and croplands, which are characterized by large moisture fluxes. The analyses of RMSE and MAE confirm this inference; in both cases the magnitude of the error increases rapidly with increasing return interval and typically reaches 75% of the maximum error within 3 to 7 days. Again, the magnitude of the error due to temporal upscaling was greatest over those land cover types with the highest ET. Using 20% relative error as the threshold, the maximum return interval ranged between five and eight days, on average, depending on land cover type. However, since the maximum return interval decreases to a minimum of five days following an exponential decay function of the mean moisture flux, five days is the longest return interval that would allow for accurate ET estimates over all land cover types assuming perfect retrieval. While the study found that using λE_{eq}, R_n, or A as χ tended to produce the most accurate estimates of λE for longer return intervals, for return intervals of five days or less, there was no statistically significant difference in the flux estimates. Finally, the comparison of interpolation methods indicated there is no advantage to using the more computationally complex spline interpolation methods.

Acknowledgements:

Deleted: also

Deleted:

Deleted: to

5. References

Abdelghani, C., Hoedjes, J. C. B., Rodriquez, J. - C., Watts, C. J., Garatuza, J., Jacob, F., and Kerr, Y. H.: Using remotely sensed data to estimate area-averaged daily surface fluxes over a semi-arid mixed agricultural land, Agric. Forest Meteorol., 148: 330-342, 2008.

Abu-Hamdeh, N.H.: Thermal Properties of soils as affected by density and water content. Biosystems. Eng., 86, 97-102, 2003.

Alfieri, J.G., Blanken, P.D., Yates, D.N., and Steffen, K.: Variability in the Environmental Factors Driving Evapotranspiration from a Grazed Rangeland during Severe Drought Conditions. J. Hydrometeorol., 8, 207-220, 2007.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration, Guideline for computing crop water requirements, FAO irrigation and drainage paper n. 56, 326 pp., Rome, Italy, 1998.

Allen, R. G., Tasumi, M., and Trezza, R.: Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-model, J. Irrig. Drain. Eng., 133, 380-394, 2007.

Alves, I., and Pereira, L.S: Modeling surface resistance from climatic variables? Agric. Water Manage., 42, 371-385, 2000.

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. A., and Kustas, W. P.: A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 2. Surface moisture climatology, J. Geophys. Res., doi:10.1029/2006JD007507, 2007.

Anderson, M. C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J. R., and Kustas, W.P.: Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States, J. Climate, 15, 2025-2044, 2111.

Anderson, M. C., Allen, R. G., Morse, A., and Kustas, W. P.: Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources, Remote Sens. Environ., 122, 56-65, 2012.

Anderson, M. C., Zolin, C. A, Sentelhas, P. C., Hain, H. R., Semmens, K., Tugrul Yilmaz, M., Gao, F., Otkin J. A., Tetrault, R.: The evaporative stress index as an indicator of agricultural drought in Brazil: An assessment based on crop yield impacts, Remote Sens. Environ., 174, 82-99.

Asner, G.: Biophysical and biochemical sources of variability in canopy reflectance. Remote Sens. Environ., 64, 234-253, 1998.

Baldocchi, D., Falge, E., Gu, L. H., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T.,Wilson, K., and Wolfsy, S.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities, Bull. Amer. Meteorol. Soc., 82, 2415–2434, 2001.

Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., and Holtslag, A. A. M.: A remote sensing surface energy balance algorithm for land (SEBAL): 1. Formulation. J. Hydrol., 212-213, 198-212, 1998.

Bird, R.E., and Riordan C.: Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth's surface for cloudless atmosphere. J. Climate Appl. Meteorol., 25, 87-96, 1986.

Brutsaert, W.: On a derivable formula for longwave radiation from clear skies. Water Resour. Res., 11, 742-744, 1975.

Brutsaert, W. and Sugita, M.: Application of self-preservation in the diurnal evolution of the surface energy budget to determine daily evaporation, J. Geophys. Res., 97, 18377–18382, 1992.

Burt, J. E., and Barber, G. M.: Elementary statistics for Geographers. Guilford Press, New York, New York, USA, 1996.

Cammalleri, C., Ciraolo, G., La Loggia, G., and Maltese, A.: Daily evapotranspiration assessment by means of residual surface energy balance modeling: A critical analysis under a wide range of water availability. J. Hydrol., 452–453, 119–129, 2012.

Cammalleri, C., Anderson, M. C., Gao, F., Hain, C. R., and Kustas, W.P.: A data fusion approach for mapping daily evapotranspiration at field scale, Water Resour. Res., 49, 4672–4686, 2013.

Cammalleri, C., Anderson, M.C., and Kustas, W.P.: Upscaling of evapotranspiration fluxes from instantaneous to daytime scales for thermal remote sensing applications. Hydrol. Earth Syst. Sci., 18, 1885-1894, 2014.

Chavez, J. L., Neale, C. M. U., Prueger, J. H., and Kustas, W. P.: Daily evapotranspiration estimates from extrapolating instantaneous airborne remote sensing ET values, Irrig. Sci., 27, 67–81, 2008.

Colaizzi, P. D., Evett, S. R., Howell, T. A., and Tolk, J. A.: Comparison of five models to scale daily evapotranspiration from one-time-of-day measurements, Trans. ASABE, 49, 1409–1417, 2006.

Crago, R.: Conservation and variability of the evaporative fraction during the daytime, J. Hydrol., 180, 173–194, 1996.

Crago, R. D., and Brutsaert, W.: Daytime evaporation and the self-preservation of the evaporative fraction and the Bowen ratio, J. Hydrol., 178, 241-255, 1996.

De Boor, C.: A Practical Guide to Splines. Springer-Verlag, New York, New York, USA, 1994.

de Fraiture, C., and Wichelns, D.: Satisfying future water demands for agriculture, Agric. Water. Manag., 97, 502–511, 2010.

Delogu, E., Boulet, G., Olioso, A., Coudert, B., Chirouze, J., Ceschia, E., Le Dantec, V., Marloie, O., Chehbouni, G., and Lagouarde, J.-P.: Reconstruction of temporal variations of Evapotranspiration using instantaneous estimates at the time of satellite overpass, Hydrol. Earth Syst. Sci., 16, 2995–3010, 2012.

Droogers, P., Immerzeel, W. W., and Lorite, I. J.: Estimating actual irrigation application by remotely sensed evapotranspiration observations, Agric. Water Manag., 97, 1351-1359, 2010.

Farah, H. O., Bastiaanssen, W. G. M., and Feddes, R. A.: Evaluation of the temporal variability of the evaporative fraction in a tropical watershed, Int. J. Appl. Earth Obs. Geoinfo., 5, 129-140, 2004.

Farquhar, G.D., Sharkey, T.D.: Stomatal conductance and photosynthesis. Ann. Rev. Plant Phys., 33, 317-345, 1982.

Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G., and Duchemin, B.: Analysis of evaporative fraction diurnal behavior, Agric. Forest Meteorol., 143, 13–29, 2007.

Gonzalez-Dugo, M. P., Neale, C. M. U., Mateos, L., Kustas, W. P., Prueger, J. H., Anderson, M. C., and Li, F.: A comparison of operational remote sensing-based models for estimating crop evapotranspiration, Agric. Forest Meteorol., 149, 1843-1853, 2009.

Gordon L. J., Finlayson, C., and Falkenmark M.: Managing water in agriculture for food production and other ecosystem
5 services, Agric. Water. Manag., 97: 512–519, 2010.

Hawkins, B. A., and Porter, E. E.: Relative influences of current and historical factors on mammal and bird diversity patterns in deglaciated North America, Global Ecol. Biogeo., 12, 475–481, 2003.

Hoedjes, J. C. B., Chehbouni, A., Jacob, F., Ezzahar, J., and Boulet, G.: Deriving daily evapotranspiration from remotely sensed instantaneous evaporative fraction over olive orchard in semi-arid Morocco. J. Hydro., 354, 53-64.

10 Howell, T. A..: Enhancing water use efficiency in irrigated agriculture, Agron. J., 93: 281–289, 2001.

Jackson, R. D., Hatfield, J. L., Reginato, R. J., Idso, S. B., and Pinter Jr., P. L.: Estimation of daily evapotranspiration from one time-of-day measurements, Agric. Water Manage., 7: 351–362, 1983.

Kalma, J. D. McVicar, T. R., and McCabe, M. D.: Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data, Surv. Geophys., 29: 421–469, 2008.

15 Katul, G. G., Oren, R., Manzano, S., Higgins, C., and Parlange, M. B.: Evapotranspiration: A process driving mass transport and energy exchange in the soil-plant-atmosphere-climate system, Rev. Geophys., 50, doi:10.1029/2011RG000366, 2012.

Kreft, H., and Jetz, W.: Global patterns and determinants of vascular plant diversity, Proc. Nat. Acad. Sci. USA, 104: 5925–5930, 2007.

Kustas, W.P., Prueger, J.H., Hatfield, J.L., Ramalingan, K., and Hipps, L.E.: Variability in soil heat flux from a mesquite
20 dune site. Agric. Forest Meteorol., 103, 249-264, 2000.

Lobell, D.B., and Asner, G.P: Moisture effects on soil reflectance. Soil Sci. Soc. Amer. J., 66, 722– 727, 2002.

Legates, D. R., and McCabe, G. R.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res.,35, 233-241, 1999.

Lhomme, J. P., and Elguero, E.: Examination of evaporative fraction diurnal behaviour using a soil-vegetation model
25 coupled with a mixed-layer model. Hydrol. Earth Syst. Sci., 3, 259–270, 1999.

Lu, J., Tang, R., Tang, H., and Li, Z.-L.: Derivation of daily evaporative fraction based on temporal variations in surface temperature, air temperature, and net radiation, Remote Sens., 5, 5369-5396, 2013.

McCabe, M. F., Wood, E., Wojcik, R., Pan, M., Sheffeld, J., Gao, H., and Su, H.: Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, Remote Sens. Environ., 112, 430-444, 2008.

30 McNaughton, K. G.: Evaporation and advection I: evaporation from extensive homogeneous surfaces, Quart. J. R. Met. Soc., 102, 181-191, 1976.

Meeus, J.: Astronomical algorithms, Willmann-Bell, Richmond , Virginia, USA, 1999.

Moler, C.: Numerical Computing with Matlab, SIAM Publishing, Philadelphia, 2004.

Midgley, G. F., Hannah, L., Millar, D., Rutherford, M. C., and Powrie, L. W.: Assessing the vulnerability of species richness to anthropogenic climate change in a biodiversity hotspot, Global Ecol. Biogeo., 11, 445–451, 2002.

Myneni, R., Ross, J., Asner, G.: A review on the theory of photon transport in leaf canopies. Agric. Forest Meteorol., 45, 1-153, 1989.

5   Nichols, W., and Cuenca, R. H.: Evaluation of the evaporative fraction for parameterization of the surface, energy-balance, Water Resour. Res., 29: 3681–3690, 1993.

Otkin, J. A., Anderson, M. C., Hain, C., Svoboda, M., Johnson, D., Mueller, R., Tadess, T., Wardlow, B., and Brown, J.: Assessing the evolution of soil moisture and vegetation conditions during the 2012 United States flash drought, Agric. Forest Meteorol., 208-219, 230-242.

10   Peng, J., Borsche, M., Liu, Y., and Loew, A.: How representative are instantaneous evaporative fraction measurements of daytime fluxes? Hydrol. Earth Syst. Sci., 17, 3913–3919, 2013.

Qadir, M., Boers, T.M., Schubert, S., Ghafoor, A., and Murtaza, G,: Agricultural water management in water-starved countries: challenges and opportunities, Agric. Water Manag., 62, 165–85, 2003.

Raupach, M: Combination theory and equilibrium evaporation. Quart. J. R. Met. Soc., 127, 1149-1181.

15   Ryu, Y., Baldocchi, D. D., Black, T. A., Detto, M., Law, B. E., Leuning, R., Miyata, A., Reichstein, M., Vargas, R., Ammann, C., Beringer, J., Flanagan, L. B., Gu, L., Hutley, L. B., Kim, J., McCaughey, H., Moors, E. J., Rambal, S., and Vesala, T.: On the temporal upscaling of evapotranspiration from instantaneous remote sensing measurements to 8-day mean daily-sums, Agric. Forest Meteorol., 152, 212–222, 2012.

Santanello, J.S., and Friedl, M.A.: Diurnal covariation in soil heat flux and net radiation. J. Appl. Meteorol., 42, 851-862,
20   2003.

Schultz, B., and De Wrachien, D.: Irrigation and drainage systems research and development in the 21st century, Irrig. Drain. Sci., 51, 311–327, 2002.

Scott, R.L., Edwards, E.A., Shutleworth, W.J., Huxman, T., Watts, C., and Goodrich, D.C.: Interannual and seasonal variation in fluxes of water and carbon dioxide from a riparian woodland ecosystem. Agric. Forest. Meteorol., 122, 65-84,
25   2014.

Seguin, B., and Itier, B.: Using midday surface-temperature to estimate daily evaporation from satellite thermal IR data, Int. J. Remote Sens., 4, 371–383, 1983.

Song, J.: Phenological influences of the albedo of prairie grasslands and crop fields. Int. J. Biometeorol., 42, 153-157, 1999.

Suleiman, A., and Crago, R. D.: Hourly and daytime evapotranspiration from grassland using radiometric surface
30   temperatures. Agron. J., 96, 384-390, 2004.

Tasumi, M., Allen, R. G., Trezza, and Wright, J. L.: 2005. Satellite-based energy balance to assess within-population variance of crop coefficient curves, J. Irrig. Drain., 131, 94-109.

Thenkabail, P. S., Hanjra, M. A., Dheeravath, V., and Gumma, M.: Global croplands and their water use from remote sensing and nonremote sensing perspectives, in Advances in environmental remote sensing: sensors, algorithms and applications, Q. Weng (Ed.), Taylor and Francis, CRC Press, Boca Raton, Florida, USA, 383-420, 2010.

Trefethen, L. N.: Approximation theory and approximation practice. SIAM books, Philadelphia, Pennsylvania, USA, 2013.

5 Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P.R., Meyer, T. P., Prueger, J. H., Starks, P. J., and Wesely, M. L.: Correcting eddy covariance flux underestimates over a grassland, Agric. Forest Meteorol., 103, 279–300, 2000.

Van de Griend, A.A., and M. Owe.: Bare soil surface resistance to evaporation by vapor diffusion under semiarid conditions. Water Resourc. Res., 30, 181-188, 1994.

10 Van Niel, T. G., McVicar, T. R., Roderick, M. L., van Dijk, A. I. J. M., Renzullo, L. J., and van Gorsel, E.: Correcting for systematic error in satellite-derived latent heat flux due to assumptions in temporal scaling: assessment from flux tower observations. J. Hydrol. 409, 140–148, 2011.

Van Niel, T. G., McVicar, T. R., Roderick, M. L., van Djik, A. I. J. M., Beringer, J., Hutley, L. B., and van Gorsel, E.: Upscaling latent heat flux for thermal remote sensing studies: Comparison of alternative approaches and correction of bias, J.

15 Hydrol., 468–469, 35–46, 2012.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., and Davies, P. M.: Global threats to human water security and river biodiversity, Nature, 467: 555–561, 2010.

Wallace, J. S.: Increasing agricultural water use efficiency to meet future food production. Agric. Ecosyst. Environ., 82,

20 105–19, 2000.

Walter, I. A., Allen, R. G., Elliott, R. L., Itenfisu, D., Brown, P., Jensen, M. E., Mecham, B., Howell, T. A., Snyder, R., Echings, S., Spofford, T., Hattendorf, M., Martin, D. L., Cuenca, R. H., and Wright, J. L.: The ASCE Standardized Reference Evapotranspiration Equation. Technical Committee report to the Environmental and Water Resources Institute of the American Society of Civil Engineers from the Task Committee on Standardization of Reference Evapotranspiration, 173

25 pp., Reston, VA, 2005.

Wang, K., and Dickinson R. E.: A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, Rev. Geophys., 50, doi: 10.1029/2011RG000373, 2012.

Williams, M., Malhi, Y., Nobre, A. D., Rastetter, B., Grace, J, Pereira, P.: Seasonal variation in net carbon exchange and evapotranspiration in a Brazilian rain forest: a modelling analysis. Plant Cell Environ., 21, 953-968, 2003.

30 Willmott, C. J., Ackleson, S. G., Davis, R.E., Feddema, J.J., Klink K.M., Legates, D.R., O'Donnell, J., and Rowe, C.M.: Statistics for the evaluation of model performance. J. Geophys. Res., 90, 8995–9005, 1995.

Willmott, C., and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate Res., 30, 79-82, 2005.

Willmott, C., Robeson, S. M., Matsuura, K., A refined index of model performance, Int. J. Climitol., 32, 2088-2094, 2012.
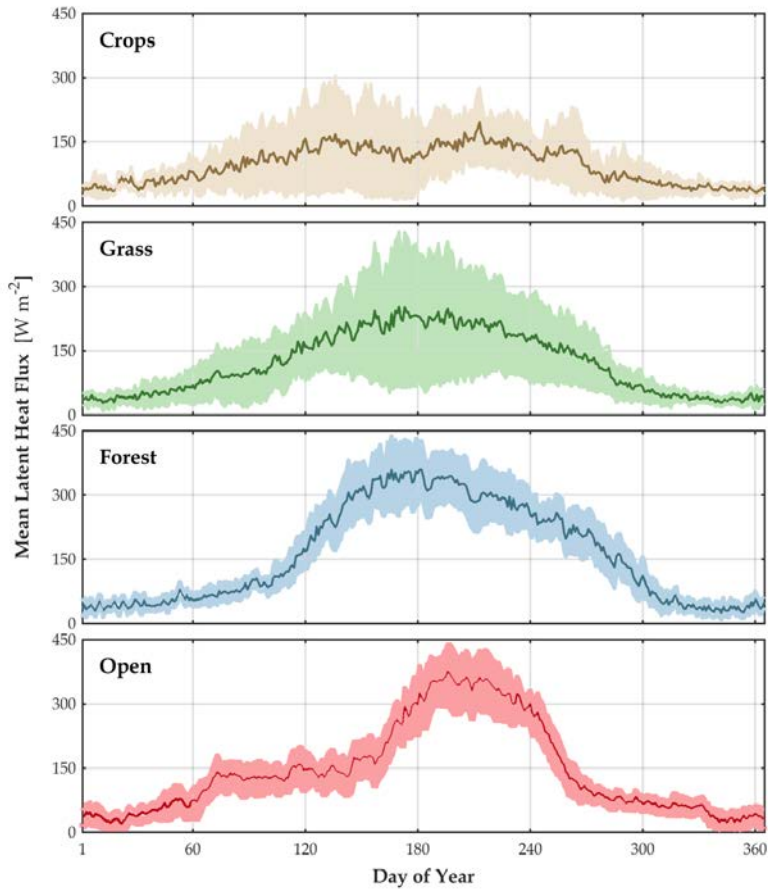
17

Wilks, W. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, Burlington, Massachusetts, USA, 2006.

Zhang, L. and Lemeur, R.: Evaluation of daily evapotranspiration estimates from instantaneous measurements, Agric. Forest Meteorol., 74, 139–154, 1995.

Figures

Figure 1 Map showing the location of the Ameriflux sites used in this study.

Figure 2 The mean daytime latebt heat flux is shown for each of the land cover types. The mean flux was calculated using the daytime mean flux data for all of years considered at each site. The shaded area represents one standard deviation about the mean.
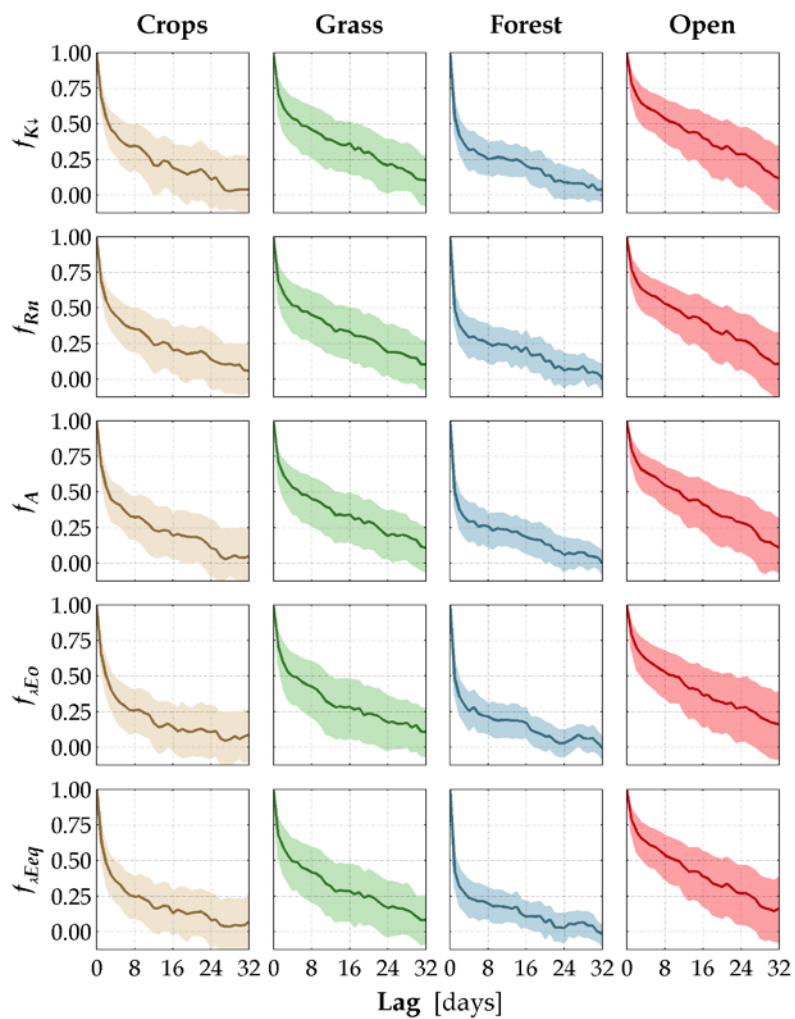
Figure 3 The representative autocorrelation function derived for each land cover type and scaled metric used in this study is shown. The shaded area represents one standard deviation about the mean.
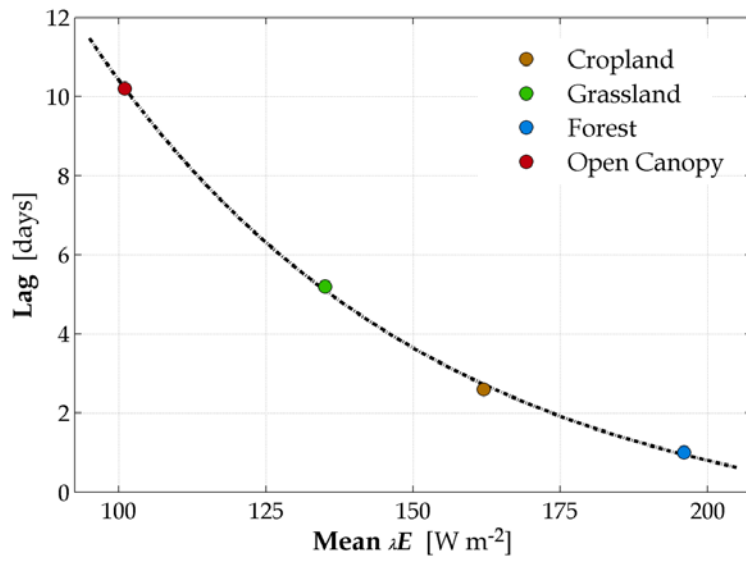
Figure 4 The maximum lag where the autocorrelation function exceeds 0.50 plotted as a function of the mean daytime latent heat flux is shown.

Deleted: 2

22
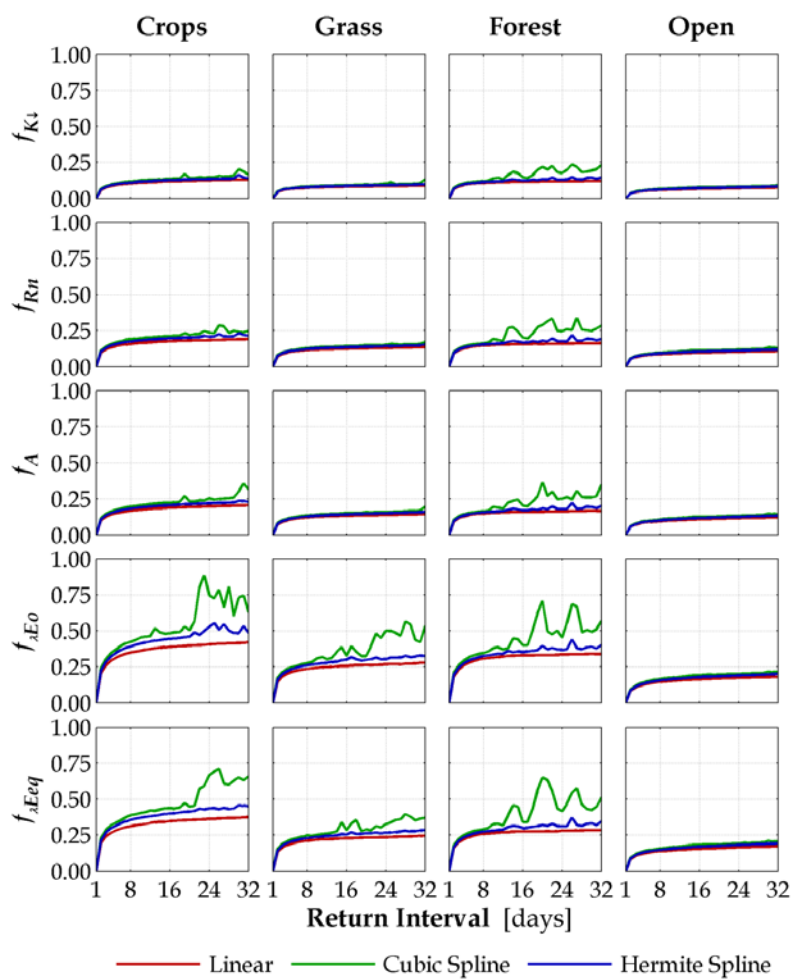
Figure 5 The root mean square error (RMSE) of the estimates of the scaled quantities is shown for each land cover type and interpolation scheme.

5
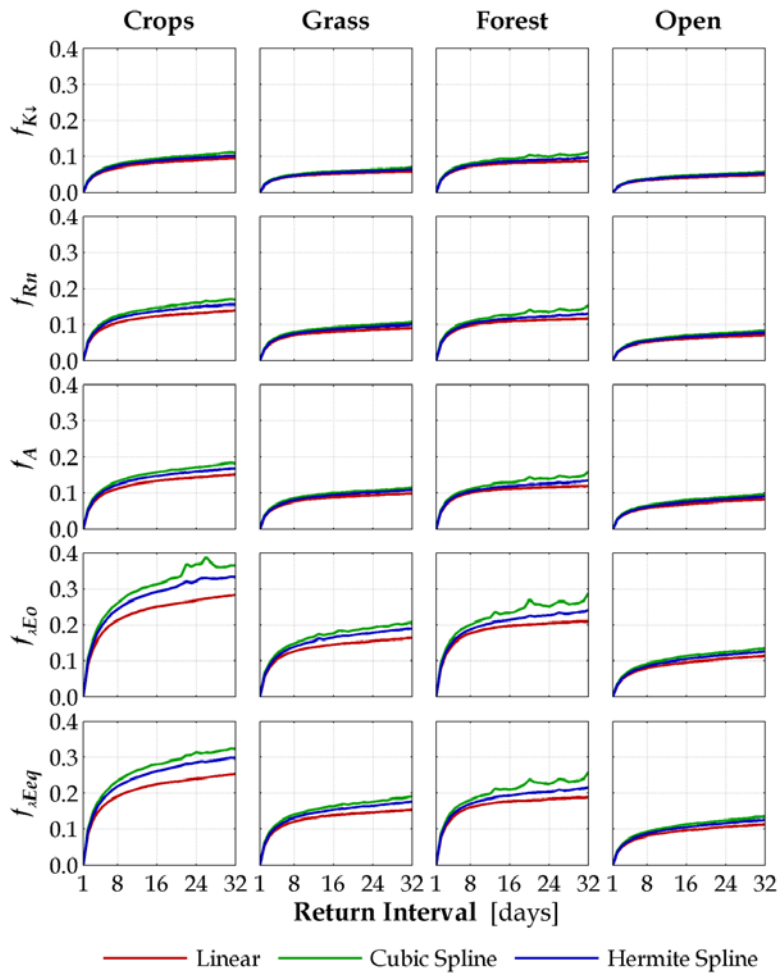
Figure 6 The mean absolute error (MAE) of the estimates of the scaled quantities is shown for each land cover type and interpolation scheme.

Deleted: 4
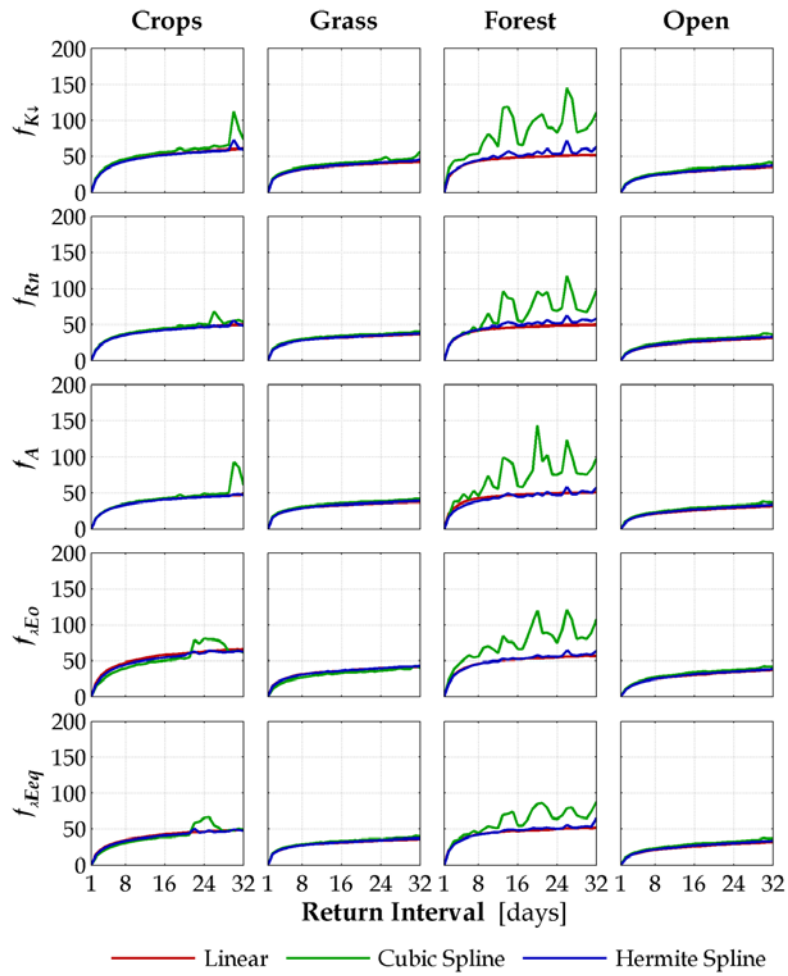
Figure 7 The root mean square error (RMSE) of the latent heat flux derived from each of the scaled quantities is shown for each land cover type and interpolation scheme.
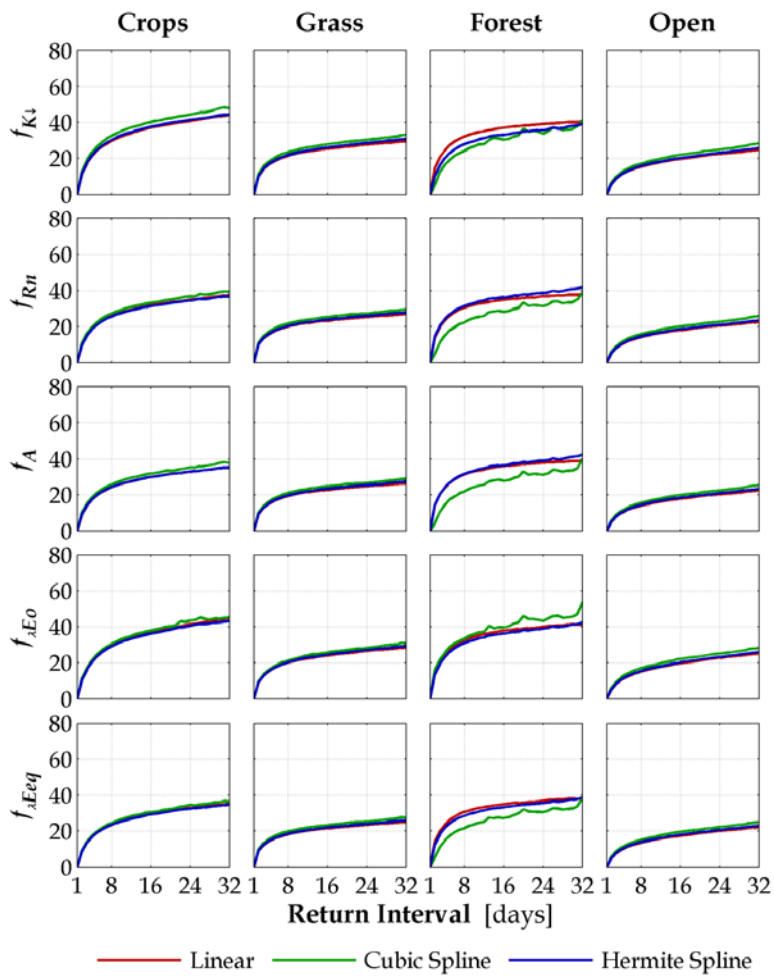
Figure 8 The mean absolute error (MAE) of the latent heat flux derived from each of the scaled quantities is shown for each land cover type and interpolation scheme.

Deleted: 6
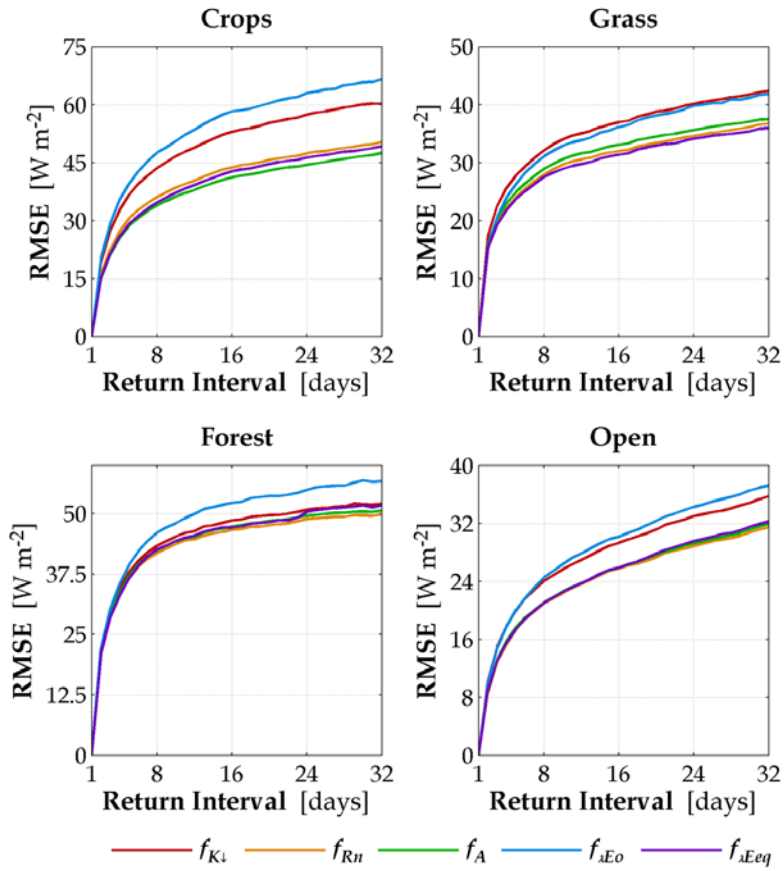
Figure 9 The root mean square error (RMSE) of the latent heat flux derived from each of the scaled quantities is shown for each land cover type when linear interpolation is used.
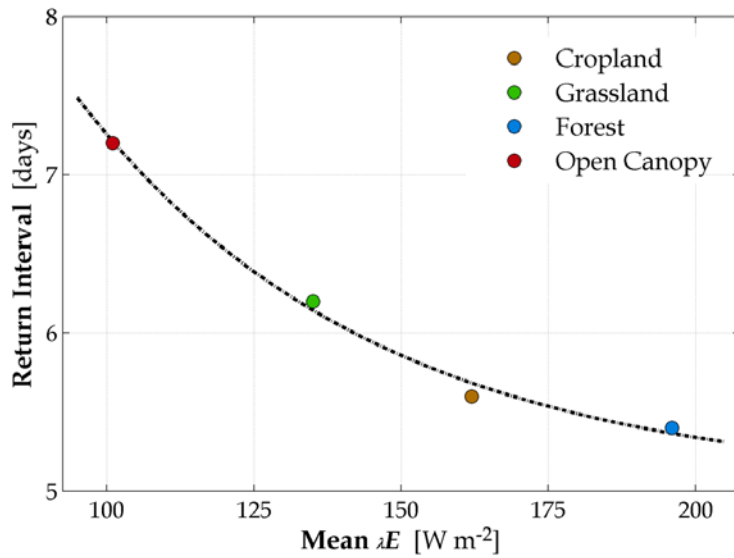
5

Figure 10 The maximum return interval where the relative error is less than 20% plotted as a function of the mean daytime latent heat flux.

Tables:

Table 1 Summary of the 30-minute measurements collected at each Ameriflux site used in this study.

Meteorological Conditions		
Wind Speed	Wind Direction	Air Temperature
Water Vapor Density	Vapor Pressure Deficit	Relative Humidity
Atmospheric Pressure	Precipitation	
Radiation and Energy Budget		
Incident Solar Radiation	Reflected Solar Radiation	Incident Longwave Radiation
Terrestrial Longwave Radiation	Net Radiation	Soil Heat Flux
Sensible Heat Flux	Latent Heat Flux	Friction Velocity

Table 2 Summary of Ameriflux sites used in this study.

Site	Location	Land Cover	Mean Annual Temp.	Mean Annual Rainfall	Study Period	Site	Location	Land Cover	Mean Annual Temp.	Mean Annual Rainfall	Study Period
Bondville	40.006 °N 88.290 °W	Cropland (maize/soy)	11.0°C	991 mm	2000-2008	Lucky Hills	31.744 °N 110.052 °W	Shrubland	17.6°C	320 mm	2007-2012
Brookings	44.345 °N 96.836 °W	Woody Savanna	6.0°C	586 mm	2004-2010	Mead	41.165 °N 96.477 °W	Cropland (maize/soy)	10.1°C	789 mm	2001-2012
Brooks Field	41.692 °N 93.691 °W	Cropland	8.9°C	847 mm	2005-2011	Morgan Monroe	39.323 °N 86.413 °W	Broadleaf Deciduous Forest	10.9°C	1032 mm	2004-2014
Chestnut Ridge	35.931 °N 84.332 °W	Broadleaf Deciduous Forest	13.9°C	1359 mm	2005-2010	Niwot Ridge	40.033 °N 105.546 °W	Evergreen Needleleaf Forest	1.5°C	800 mm	2001-2012
Fermi Cropland	41.859 °N 88.223 °W	Cropland (maize/soy)	9.2°C	929 mm	2005-2011	Missouri Ozarks	38.744 °N -92.200 °W	Broadleaf Deciduous Forest	12.1°C	986 mm	2004-2013
Fermi Grassland	41.841 °N 88.241 °W	Grassland	9.2°C	929 mm	2005-2011	Rosemount	44.714 °N 93.090 °W	Cropland (maize/soy)	6.4°C	879 mm	2004-2012
Freeman Ranch	29.940 °N -97.990 °W	Woody Savanna	19.5°C	864 mm	2005-2009	Santa Rita Mesquite	31.821 °N 110.866 °W	Woody Savanna	17.9°C	380 mm	2004-2012
Kendall Grassland	31.737 °N 109.942 °W	Grassland	15.6°C	407 mm	2004-2012	Tonzi Ranch	38.432 °N 120.966 °W	Woody Savanna	15.8°C	559 mm	2001-2012
Konza Prairie	39.082 °N 96.560 °W	Grassland	12.8°C	867 nn	2006-2012	Vaira Ranch	38.407 °N 120.910 °W	Grassland	15.8°C	559 mm	2001-2012
Loblolly Pine	35.978 °N 79.094 °W	Evergreen Needleleaf Forest	14.4°C	1170 mm	2001-2008	Walker Branch	35.959 °N 84.787 °W	Broadleaf Deciduous Forest	13.7°C	1372 mm	2001-2007

Table 3 The maximum return interval with a relative error of less than 20% is given for each reference quantity and LULC when linear interpolation was used.

		Reference Quantity				
		Incident Solar Radiation	Net Radiation	Available Energy	Reference Latent Heat Flux	Equilibrium Latent Heat Flux
Land Cover	**Cropland**	4	6	7	4	7
	Grassland	5	7	6	5	8
	Forest	5	6	5	5	6
	Open Canopy	6	8	8	7	8