JAN. 30, 09

UP-SCALING OF SEBAL DERIVED EVAPOTRANSPIRATION MAPS FROM LANDSAT (30m) TO MODIS (250m) SCALE

Sung-ho Hong, Jan M.H. Hendrickx and Brian Borchers

New Mexico Tech, Socorro, NM 87801

ABSTRACT

Remotely sensed imagery of the Earth's surface via satellite sensors provides information to estimate the spatial distribution of evapotranspiration (ET). The spatial resolution of ET predictions depends on the sensor type and varies from the 30 – 60 m Landsat scale to the 250 – 1000 m MODIS scale. Therefore, for an accurate characterization of the regional distribution of ET, scaling transfer between images of different resolutions is important. Scaling transfer includes both up-scaling (aggregation) and down-scaling (disaggregation). In this paper, we address the up-scaling problem.

The Surface Energy Balance Algorithm for Land (SEBAL) was used to derive ET maps from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Moderate Resolution Imaging Spectroradiometer (MODIS) images. Landsat 7 bands have spatial resolutions of

30 to 60 m, while MODIS bands have resolutions of 250, 500 and 1000 m. Evaluations were conducted for both “output” and “input” up-scaling procedures, with aggregation accomplished by both simple averaging and nearest neighboring resampling techniques. Output up-scaling consisted of first applying SEBAL and then aggregating the output variable (daily ET). Input up-scaling consisted of aggregating 30 m Landsat pixels of the input variable (radiance) to obtain pixels at 60, 120, 250, 500 and 1000 m before SEBAL was applied. The objectives of this study were first to test the consistency of SEBAL algorithm for Landsat and MODIS satellite images and second to investigate the effect of the four different up-scaling processes on the spatial distribution of ET.

We conclude that good agreement exists between SEBAL estimated ET maps directly derived from Landsat 7 and MODIS images. Among the four up-scaling methods compared, the output simple averaging method produced aggregated data and aggregated differences with the most statistically and spatially predictable behavior. The input nearest neighbor method was the least predictable but was still acceptable. Overall, the daily ET maps over the Middle Rio Grande Basin aggregated from Landsat images were in good agreement with ET maps directly derived from MODIS images.

1. INTRODUCTION

Remote sensing data from satellite-based sensors have the potential to provide detailed information on land surface properties and parameters at local to regional scales. Perhaps

one of the most important land surface parameters that can be derived from remote sensing is actual ET. The spatio-temporal distribution of ET is needed for sustainable management of water resources as well as for a better understanding of water exchange processes between the land surface and the atmosphere. However, ground measurements of ET over a range of space and time scales are very difficult to obtain due to the time and cost involved. Remotely sensed imagery with numerous spatial and temporal resolutions is therefore an ideal solution for determination of the spatio-temporal distribution of ET.

Today, large amounts of remotely sensed data with variable spatial, temporal, and spectral resolutions are available. A number of studies have attempted to estimate ET from different satellite sensors, including the Land remote sensing satellite Enhanced Thematic Mapper Plus (Landsat ETM+) (Bastiaanssen et al., 2005; Hendrickx and Hong, 2005; Allen et al., 2007; Hong, 2008), the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (French et al., 2002), the Advanced Very High Resolution Radiometer (AVHRR) (Seguin et al., 1991), the Moderate Resolution Imaging Spectroradiometer (MODIS) (Nishida et al., 2003; Hong et al., 2005) and the Geostationary Orbiting Environmental Satellite (GOES) (Mecikalski et al., 1999).

We employ the Surface Energy Balance Algorithm for Land (SEBAL) that is one of several remote sensing algorithms used to extract information from raw satellite data. It estimates various land surface parameters, including surface albedo, normalized difference vegetation index (*NDVI*), surface temperature, and energy balance parameters

from the remotely sensed radiance values obtained from satellite sensors. Since satellite sensors have different spatial, spectral and radiometric resolutions, the consistency of ET estimates from different satellites by SEBAL needs to be certified.

The validation of products of remote sensing algorithms is dependent upon the spatial resolution (Liang, 2004). Fine resolution products (< 100 m) such as Landsat can be validated with ground measurements. However, validating coarse resolution products, such as MODIS (1000 m in thermal band), using ground measurements is very difficult because of the scale disparity between ground "point" measurements and the coarse spatial resolution imagery. Therefore, for validation of MODIS products, the products of high resolution remotely sensed imagery such as Landsat 7 (30 to 60 m resolution) need to be first validated with ground point measurements. MODIS products can then be compared against up-scaled (aggregated) Landsat product. A comparison of SEBAL ET estimates against independent ground based measurements typically yields accuracies of about ±15% and ± 5% for, daily and seasonal evaporation estimates, respectively (Bastiaanssen et al., 2005). In the southwestern USA, daily SEBAL ET estimates agreed with ground observation with an accuracy of ±10% (Hendrickx and Hong, 2005; Hong, 2008). Similar results have been reported by Morse et al. (2000) and Allen et al. (2007).

Many studies regarding the effect of up-scaling data sets have been reported (Mark and Aronson, 1984; Nellis and Briggs, 1989; Turner et al., 1989; Lam and Quattrochi, 1992; Stoms, 1992; Brown et al., 1993; Vieux, 1993; De Cola, 1994; Wolock and Price, 1994; Zhang and Montgomery, 1994; Bian et al., 1999). During an aggregation process, the

92 raster spatial data are reduced to a smaller number of data pixels covering the same
93 spatial extent. It is generally recognized that data aggregation modifies the statistical and
94 spatial characteristics of the data (Bian et al., 1999). Since the total number of pixels is
95 reduced, the variance and frequency distribution of the sampled data may deviate from
96 the original data set and tends to reduce spatial autocorrelation at coarser resolutions
97 (Bian, 1997). Some studies have pointed out that data accuracy is enhanced significantly
98 by reduction of spatial resolution (Townshend et al., 1992; Dai and Khorram, 1998; Van
99 Rompaey et al., 1999; Carmel, 2004). Several studies have also argued that aggregation
100 to a coarser resolution reveals certain spatial patterns which are not shown until the data
101 are presented at a coarser scale (Zhang and Montgomery, 1994; Seyfried and Wilcox,
102 1995). On the other hand, the decrease in spatial resolution possibly results in a loss of
103 information that may be valuable for particular applications (Carmel et al., 2001).
104
105 The methodology for aggregating simple rectangular grid data is well developed (Bian,
106 1997; Bian et al., 1999; Mengelkamp et al., 2006). In this study, the simple averaging and
107 nearest neighbor resampling methods were selected for the data aggregation scheme,
108 since these methods have been the most popular and convenient to use (Atkinson, 1985;
109 Liang, 2004). The simple averaging method calculates the average value over an area of
110 interest to produce a new coarser resolution data set. Nearest neighbor sampling produces
111 a subset of the original data; the extremes and subtleties of the data values are therefore
112 preserved.
113

For the up-scaling scheme, numerous studies have used the assumption that surface fluxes can be expressed as direct area averages of the surface fluxes (Shuttleworth, 1991; Lhomme, 1992; Li and Avissar, 1994). Liang (2000) simply averaged the remotely sensed reflectance values from 30 m to 1 km and explored the aggregation effect. He concluded that the spectral reflectance was basically linear from 30 m resolution to 1000 m resolution. More recently, Mengelkamp et al (2006) mentioned that area averaged small scale ET calculated from local measurements was in good agreement with the area represented regional values. Nevertheless, few papers have examined the effect of different up-scaling schemes on the relative accuracy of the aggregated data despite its practical importance. A spatial resolution gap exists between the data requirements of regional-scale models and the output data from remote sensing energy balance algorithms such as SEBAL. For example, general global circulation models or regional weather prediction models need input data with a spatial resolution of hundreds of kilometers which is much larger than the spatial resolution of most satellite sensors (Liang, 2004). Therefore, an up-scaling (data aggregation) procedure is needed to fill the scale gap between satellite measurements and input requirements for large scale models. Increasing spatial resolution by data aggregation has shown the potential to generate observed or modeled surface flux estimates over a range of different spatial resolutions (Gupta et al., 1986; Lhomme, 1992; Ebleringer and Field, 1993).

In this study, high quality scenes of two different dates of Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Moderate Resolution Imaging Spectroradiometer (MODIS) imagery were selected and SEBAL was applied to estimate daily ET. Landsat scale pixels

(30 m) were aggregated to larger scale (60, 120, 250, 500 and 1000 m). The objectives of this study were first to test the consistency of the SEBAL algorithm for Landsat 7 and MODIS images, and second to investigate the effects of four different up-scaling processes on the spatial distribution of ET, especially how the relative accuracy of ET changes with different up-scaling processes.

2. METHOD AND MATERIALS

2.1. STUDY AREA AND SATELLITE IMAGERY

Landsat 7 and Terra MODIS images (Figure 1) on two different dates during the growing season (September 14, 2000 and June 16, 2002) were used to examine the effect of aggregation processes. On these two dates, high quality Landsat 7 and MODIS images were available. The June 16 images are representative for conditions of full vegetative cover at the height of the growing season, while the September 14 images represent somewhat drier conditions towards the end of the growing season. Four satellite images used in this study were georeferenced to match the spatial coordinates as closely as possible. This was done by identifying the several accurate Ground Control Points (e.g. road intersections and agricultural field boundaries) on the images and aligning them to fit on between images. The image used in this study is the subset of the Middle Rio Grande Basin that covers an area of 18 by 90 km. The Middle Rio Grande setting is mainly composed of agricultural fields, riparian forests and surrounding desert areas (Figure 1).

2.2. SURFACE ENERGY BALANCE ALGORITHM FOR LAND (SEBAL)

SEBAL is a physically based analytical image processing method that evaluates the components of the energy balance and determines the ET rate as the residual. SEBAL is based on the computation of energy balance parameters from multi spectral satellite data (Bastiaanssen et al., 1998; Morse et al., 2000; Allen et al., 2007). To implement SEBAL, images are needed with information on reflectance in the visible, near-infrared and mid-infrared bands, as well as emission in the thermal infrared band. To account for the influence of topographical variations on the energy balance components, a digital elevation model (DEM) with the same spatial resolution as the satellite imagery is also required. The slope and aspect were calculated from DEM using models provided in ERDAS IMAGINE software (ERDAS, 2002).

The energy balance equation is

$$R_n - G - H = \lambda ET \quad (1)$$

where R_n is the net incoming radiation flux density (Wm^{-2}), G is the ground heat flux density (Wm^{-2}), H is the sensible heat flux density (Wm^{-2}), λET is the latent heat flux density (Wm^{-2}), and parameter λ is the latent heat of vaporization of water (Jkg^{-1}).

The net radiation (R_n) was computed for each pixel from the radiation balance using surface albedo obtained from short-wave radiation and using emissivity estimated from

183 the long-wave radiation (Allen et al., 1998; Bastiaanssen et al., 1998; Morse et al., 2000).
184 Soil heat flux (*G*) was estimated from net radiation together with other parameters such
185 as normalized difference vegetation index (*NDVI*), surface temperature and surface
186 albedo (Clothier et al., 1986; Choudhury et al., 1987; Daughtry et al., 1990; Bastiaanssen,
187 2000). Sensible heat flux (*H*) was calculated from wind speed, estimated surface
188 roughness for momentum transport, and air temperature differences between two heights
189 (0.1 and 2 m) using an iterative process based on the Monin-Obukhov similarity theory
190 (Brutsaert, 1982; Morse et al., 2000; Tasumi, 2003).
191
192 The spatial resolutions of the Landsat 7 bands are 30 and 60 m, compared with 250, 500
193 and 1000m for the MODIS bands (Table 1). Besides the difference in the spatial
194 resolution between Landsat 7 and MODIS, a difference in radiance measurements
195 between the two sensors is expected as a result of slightly different band widths for each
196 sensor. Table 1 also shows the spectral bands of Landsat 7 and MODIS in the visible,
197 near infrared and thermal infrared wavelength regions used for SEBAL application.
198 MODIS bands 1, 2, 3, 4, 6 and 7 are compatible with Landsat 7 bands 3, 4, 1, 2, 5 and 7,
199 respectively. The band widths of MODIS in the visible and near infrared, with the
200 exception of Band 3, are narrower than those of Landsat. This results in different
201 responses from the surface, which in turn may alter the computed surface albedo and
202 vegetation index.
203
204 **2.2.1. Brightness temperature**

205 The major difference in the ET derivation from Landsat and MODIS images was in the
206 surface temperature calculations. SEBAL used one thermal band for surface temperature
207 estimation for Landsat 7 data while two thermal bands were used with MODIS data.
208
209 The temperature detected by a thermal sensor is called the brightness temperature.
210 Radiance data from Landsat 7 and MODIS thermal infrared bands were first converted to
211 brightness temperatures with an inversion of Planck's equation:

212
$$T_b = \frac{hc/k\lambda}{\ln\left(\frac{2hc^2\lambda^{-5}}{L_\lambda}+1\right)} = \frac{K_2}{\ln\left(\frac{K_1}{L_\lambda}+1\right)} \quad (2)$$

213 T_b is the brightness temperature in Kelvin [K], c is the speed of light (2.998 x 10^8) [ms^{-1}],
214 h is the Planck's Constant (6.626 x 10^{-34}) [Js], k is the Boltzmann constant (1.3807 x 10^{-23})
215 [JK^{-1}], L_λ is the spectral radiance [$Wm^{-2}\mu m^{-1}sr^{-1}$], λ is the band effective center
216 wavelength [μm] and K_1 and K_2 are calibration coefficients [$Wm^{-2}sr^{-1}\mu m^{-1}$] (Table 2).
217
218 **2.2.2. Surface temperature**
219 For Landsat images the surface temperature (T_s) is estimated using T_b and ε_0 with the
220 following empirical relationship (Morse et al., 2000).
221
222
$$T_s = \frac{T_b}{\varepsilon_0^{0.25}} \quad (3)$$

223
224 where $\varepsilon_0 = 1.009 + 0.47\ ln(NDVI)$.

For MODIS images the split window technique is used. Split window algorithms take advantage of the differential absorption in two close infrared bands to account for the effects of absorption by atmospheric gases. Several split window algorithms are currently available to derive surface temperature from brightness temperature when multiple thermal bands are available. In this study, the algorithm developed by Price (1984) was applied since Vazquez et al. (1997) determined that it performed better than other algorithms. T_s is given by

$$T_s = T_{31} + 1.8(T_{31} - T_{32}) + 48(1 - \varepsilon) - 75\Delta\varepsilon \quad (4)$$

where T_{31} is the brightness temperature obtained from band31 [K], T_{32} is the brightness temperature obtained from band 32 [K], $\varepsilon = (\varepsilon_{31} + \varepsilon_{32})/2$, $\Delta\varepsilon = \varepsilon_{31} - \varepsilon_{32}$, ε_{31} is the surface emissivity in band 31 and ε_{32} is the surface emissivity in band 32.

Cihlar et al. (1997) developed an algorithm to calculate the surface emissivity from *NDVI*.

$$\Delta\varepsilon = \varepsilon_{31} - \varepsilon_{32} = 0.01019 + 0.01344 \ln(NDVI) \quad (5)$$

where $\varepsilon_{31} = 0.9897 + 0.029 \ln(NDVI)$.

2.2.3. Daily evapotranspiration

In SEBAL, daily ET was interpolated by assuming the instantaneous evaporative fraction (*EF*) when the satellite was passing over is approximately equal to the daily mean value

246 (Shuttleworth et al., 1989; Brutsaert and Sugita, 1992; Crago, 1996; Farah et al., 2004;
247 Gentine et al., 2007). The soil heat flux is assumed to be zero on a daily basis (Kustas et
248 al., 1993). Based on the known value of the instantaneous EF, the daily-averaged net
249 radiation flux, and the soil heat flux over a daily period, daily ET (ET_{24}) can be computed
250 by (Bastiaanssen et al., 1998):
251
252 $$ET_{24} = \frac{86400 EF(R_{n24} - G_{24})}{\lambda} \quad (6)$$
253
254 where $EF = \lambda E/(\lambda E + H)$, 86400 is a constant for time scale conversion, ET_{24} is daily
255 ET [mmd^{-1}], R_{n24} is daily-averaged net radiation [Wm^{-2}] and G_{24} is daily-averaged soil
256 heat flux [Wm^{-2}].
257
258 **2.3. UP-SCALING (AGGREGATION) PROCESS**
259 In the up-scaling process, two different procedures were evaluated. The first consisted of
260 applying SEBAL first and then aggregating the output variable (daily ET). The second
261 consisted of aggregating Landsat pixels of input variable (radiance) to obtain pixels at the
262 MODIS scale before SEBAL was applied (Figure 2). If the model is insensitive to an
263 input parameter, aggregating the value with increasing scale will have little influence on
264 model predictions. However, when the model does not operate linearly, the change in
265 data aggregation could increase or decrease model predictions (Quattrochi and Goodchild,
266 1997; French, 2001; Liang, 2004).
267

Aggregation imagery was obtained by simple averaging and by nearest neighbor selection, and done with ERDAS IMAGINE (Leica Geosystems LLC). The simple averaging resampling method calculated the arithmetic mean over an n by n window. Since a pixel value of satellite imagery is considered to be the integrated value over the corresponding area on the ground, simple averaging is considered appropriate for aggregating remotely sensed images. The simple averaging method smoothes the original data values and therefore produces a “tighter” histogram than the original data set. Furthermore, aggregating a data set by simple averaging generally decreases the variance and also increases the spatial autocorrelation (Anselin and Getis, 1993).

The nearest neighbor approach uses the value of the input pixel closest to the center of the output pixel. To determine the nearest neighbor, the algorithm uses the inverse of the transformation matrix to calculate the image file coordinates of the desired geographic coordinate. The pixel value occupying the closest image file coordinate to the estimated coordinate will be used for the output pixel value in the georeferenced image. Unlike simple averaging, nearest neighbor is appropriate for thematic files having data file values based on a qualitative system. One advantage of the nearest neighbor method is that, unlike the simple averaging resampling method, its output values are original input values. The other advantage is that it is easy to compute and therefore fastest to use. However, the disadvantage is that nearest neighbor generates a choppy, "stair-stepped" effect. The image tends to have a rough appearance relative to the original data (Cover and Hart, 1967; Atkinson, 1985; Dodgson, 1997; Bian et al., 1999).

The aggregation was operated at six levels: 30, 60, 120, 250, 500 and 1000 m pixel sizes. At each level, Landsat scale 30 by 30 m pixels were broken into 10 by 10 m pixels with the same pixel values; the data were then aggregated directly from the 10 m resolution instead of from a previous aggregation. This procedure made it easier to aggregate from the Landsat 30 m pixel size to MODIS 250, 500 and 1000 m pixel sizes.

3. RESULTS AND DISCUSSION

3.1. SEBAL CONSISTENCY BETWEEN LANDSAT AND MODIS

The SEBAL algorithm was applied to both Landsat 7 and MODIS images acquired on September 14, 2000 and June 16, 2002 and estimated their daily ET rates. In order to check the consistency of SEBAL performance for the different satellite sensors, SEBAL estimated ET from Landsat and MODIS images were compared each other. Spatial distribution of ET maps for visual verification and histograms and basic statistics for quantitative examination were selected. Two approaches were used to inspect the ET estimation difference between two different satellite sensors: one is a difference image (pixel-by-pixel difference between Landsat and MODIS estimates), while the other was a relative difference image (absolute value of the pixel difference was divided by the MODIS derived pixel value). Basic statistics of the difference and relative difference images were also computed to quantify the discrepancy between Landsat and MODIS estimates.

3.1.1 Comparison between Landsat (30m) and MODIS (250m) estimated ET

Figure 3 shows that the June image taken during the summer has significantly higher ET rates than the September image taken in the early fall. All of the ET images clearly show high ET rates in the irrigated fields and riparian areas along the Rio Grande Valley, while low ET rates are shown in the surrounding desert areas and bare soils. The city of Albuquerque has a somewhat higher ET rate than surrounding desert areas due to urban and residential vegetations.

The disparate spatial resolutions of Landsat- and MODIS-based ET images result in some differences in ET distribution, as may be expected. Many small areas (length scale on the order of 10 to 100 m) with high ET rates along the river are captured well in the Landsat-based ET map with a spatial resolution of 30 m. These peak ET rates are averaged out, however, on the MODIS derived ET map with a spatial resolution of 250 m. Figure 3 shows that MODIS derived ET distributions have a tighter and taller histogram and fewer pixels have close to zero (0.0 to 0.5) ET than the histogram from Landsat imagery. In the table of basic statistics in Figure 3, the ET map derived from the Landsat 7 image shows a higher maximum and standard deviation than the one derived from the MODIS images. However, the mean values of Landsat- and MODIS-based ET images are very similar. The minimum value of ET in each image equals to zero.

Difference images between the Landsat-based ET at 30m resolution and MODIS-based ET at 250m resolution show how these products are dissimilar to each other (Figure 4). Each difference image was produced by subtracting MODIS-based ET from Landsat-based ET [$ET_{Landsat} - ET_{MODIS}$], with brown-colored pixels in the difference map in Figure

335 4 representing points where the MODIS-based ET is significantly higher than Landsat-
336 based ET. Blue-colored pixels represent points where the ET from Landsat is
337 significantly higher than the ET from MODIS imagery. Areas with apparently high ET
338 differences (> +2.0 or < -2.0 mm/d) shown as brown or blue, are observed along the
339 boundary between Rio Grande River riparian areas and surrounding deserts. These high
340 differences are mostly due to (1) disagreement in image georeferencing between the
341 Landsat and MODIS imagery and (2) differences resulting from subtracting the ET value
342 of a large (250m) MODIS based pixel from that of a small (30m) Landsat based pixel.
343
344 It is not trivial to generate georeferenced imagery with error of less than one pixel
345 (Eugenio and Marqués, 2003). The georeferencing of two maps with spatial resolutions
346 differing an order of magnitude is especially difficult (Liang et al., 2002). One or two
347 pixels of georeferencing disagreement can cause abrupt ET changes at the boundaries
348 between riparian (high ET) and desert (low ET) areas. The effect of different pixel sizes
349 is clearly demonstrated with the brown and blue pixels located along the sudden
350 transition from riparian area to desert. The brown-colored pixels (ET difference < -2
351 mm/d) are located in the desert and result from subtracting a large MODIS pixel located
352 partially in the riparian area with relatively high ET from a small Landsat pixel located in
353 the desert with zero ET. The blue-colored pixels (ET difference > 2 mm/d) are located in
354 the riparian area and result from subtracting a large MODIS pixel located partially in the
355 desert from a riparian area located small Landsat pixel.
356

Basic statistics (mean and standard deviation) allow a quantitative means of comparison and evaluation. Positive and negative differences due to georeferencing disagreement between two images tend to cancel each other in these calculations since they occur in opposite directions at both sides of the transgression from riparian to desert area. Therefore the mean and standard deviation of each difference image were calculated based on the "absolute" difference between Landsat- and MODIS-based ET images. For both study dates, the mean and standard deviation of difference between the Landsat and MODIS-based ET are within 1.0 mm/day. Basic statistics in Figure 4 show that the September images have a slightly lower mean difference and standard deviation than the June images. However, this does not imply that the September Landsat- and MODIS-based ET images agree better than June images. The difference in basic statistics is caused by the smaller values of the mean and standard deviation of ET rates in the September images.

Relative difference images were produced as well by dividing the absolute difference image by the MODIS derived ET image [$|(ET_{Landsat} - ET_{MODIS})|$ /ET_{MODIS}] (Figure 5). The relative difference value ranges from zero to infinity. The infinity values occur when the MODIS-based ET is much smaller than the Landsat-based ET. The infinity values were constrained to 1.0 and pixels having zero values either in the MODIS-based ET or in the Landsat-based ET image are also assigned to 1.0 as relative difference. Most of the pixels having 1.0 (red-colored) relative difference are located in the desert area. One interesting point is that the quite a few pixels having 1.0 as relative difference are found along the transition zone between riparian and desert areas. Those pixels result from 30 m Landsat

pixels having high ET inside 250 m coarse resolution of MODIS pixels having low ET (Figure 5).

Figure 6 presents three dimensional graphs of the relationship between relative difference and daily ET rate on both June and September images. Both graphs in Figure 6 show that large relative difference predominantly occur in areas having low ET while areas having ET such as greater than 3 mm/d exhibit relative differences of about less than 0.4. However, there are some points having 1.0 relative difference with daily ET greater than 2.0 mm/d. These points are resulted from pixels having significant difference between Landsat and MODIS derived ET and mainly due to georeferencing disagreement between Landsat and MODIS satellite images. These questionable points are mostly located in the boundary area between riparian and surrounding desert.

3.2. ANALYSIS OF UP-SCALING EFFECTS

The spatial distribution and its statistical features were evaluated and compared among the four different up-scaling methods across the five aggregation levels. Output up-scaling aggregated the SEBAL estimated daily ET rates either with simple averaging or the nearest neighbor resampling method. The resultant aggregated ET map may represent the best estimate of ET at the coarser resolution, since the aggregated ET was derived directly by aggregation of the fine resolution ET data. For input up-scaling, since the radiometric observations (radiance) or SEBAL model inputs were aggregated, one expects to retrieve the best estimate of a radiometric observation at the coarser

resolutions. These aggregated data were used as input to the SEBAL model and calculated daily ET.

The different up-scaling methodologies were evaluated by: (1) spatial distribution of aggregated imagery by four different schemes at each aggregation level to evaluate the changes in spatial pattern after aggregation, and (2) histograms and basic statistics of the aggregated data for different up-scaling schemes at all levels. The spatial details lost during aggregation were considered to be the difference between original image and up-scaled image. In this study difference images were created by subtracting the up-scaled pixels from the original pixels of the Landsat- or MODIS-based ET estimates. While relative difference images were produced by dividing the absolute difference by the original Landsat- and MODIS-based ET images. The statistical and spatial characteristics of differences were evaluated by analyzing the spatial distribution of differences as well as the mean and standard deviation of absolute differences.

3.2.1. Effect of aggregation

Spatial and statistical characteristics of up-scaled products from June and September Landsat-based ET maps at 30m resolution to five aggregation levels are presented in Figures 7 –10. Figure 7 presents ET maps from output up-scaling using simple averaging resampling on June 16, 2002, at spatial resolutions of 60, 120, 250, 500 and 1000m. This method produces the most statistically and spatially predictable behavior. The least predictable – but still acceptable – behavior is produced by input up-scaling using nearest neighbor resampling. An example for June 16, 2002 is presented in Figure 8. Figures 9

and 10 present the histograms and statistics for the different up-scaling methods on, respectively, June 16, 2002 and September 14, 2000. Although spatial detail was lost with the increase in pixel size, the overall spatial distribution of ET of each aggregated map (for example Figures 7 and 8) was in agreement with the original ET maps in Figure 3.

All histograms of ET distribution (Figures 9 and 10) show the dominance of close to zero ET values and this frequency decreases a few percent (3.4 to 1.3%) with pixel size only when output up-scaling with simple averaging was applied. This feature might be explained by the observation that desert areas along the riparian corridors are classified to have zero ET in fine resolution of 30m. However, these desert areas are easily mixed with riparian areas when applying simple averaging, while nearest neighbor resampling schemes hardly affect the frequencies in the histogram since nearest neighbor produces a subset of the original data. The 60 and 120m pixel sized histograms in Figures 9 – 10 exhibit an almost constant frequency occurrence of 2.0% for June imagery and 3.0% for September imagery over ET rates ranging from 2.5 to 7.5 mm/d and from 1.0 to 5.0 mm/d, respectively. This constant frequency changes into a concave down shape as pixel size is increased further with simple averaging resampling in both output and input up-scaling. That is, the frequency of pixels having 5 – 6 mm/d ET increases but the frequency of pixels having 3 – 4 mm/d decreases with simple averaging is applied. Pixels having 5 – 6 mm/d of ET in this study area are mainly surface water, agricultural fields and riparian vegetation pixels located along the Rio Grande riparian corridor. There are pixels having 3 – 4 mm/d of ET located inside of the riparian corridor as well as in the

448 transition zone between riparian and surrounding desert. These pixels are mostly located
449 along the transition zone between riparian areas and surrounding deserts areas and
450 adjacent to the Rio Grande River. These pixels have low ET, but when averaged with
451 adjacent higher ET pixels the contrast disappears. However, histograms from nearest
452 neighbor resampling stayed rather consistent in shape at each resolution.
453
454 The basic statistics and histograms also show the statistical changes through aggregation.
455 With either output up-scaling or input up-scaling, the mean values of the simple
456 averaging and nearest neighbor images remain essentially constant across all aggregation
457 levels in both days. However, ET maps derived using nearest neighbor show a more
458 "blocky" pattern than those derived using from simple averaging (for example Figures 7
459 and 8). This difference in spatial distribution is due to the fact that simple averaging
460 decreases the standard deviation with increasing pixel size, while the standard deviation
461 from nearest neighbor aggregation stays fairly constant across all aggregation levels.
462
463 The differences in aggregation procedures between simple averaging and nearest
464 neighbor cause the fundamental difference in statistics of the aggregated data. The simple
465 averaging method aggregates based on data values, and the resulting values are confined
466 to the mid range. However, the nearest neighbor resampling is based on location, its pixel
467 value varying with the location of central pixels in new coordinates as the pixel size
468 changes. Therefore, the aggregated results are a systematically sampled subset of the
469 original data, and their values are expected to be less confined. This explains the
470 somewhat larger data ranges for the nearest neighbor resampling method, but the mean of

471 the data does not change significantly. Many regional-scale hydrological process models
472 require input parameters over a large area. Direct area averaging technique has often
473 been used to generate the regional-scale model input parameters (Shuttleworth, 1991;
474 Chehbouni and Njoku, 1995; Croley et al., 2005; Maayar and Chen, 2006). For example,
475 direct averaged values of air temperature, precipitation, humidity, surface roughness
476 length and so on were used as input parameters in hydrologic models (Brown et al., 1993;
477 Maayar and Chen, 2006). However, the standard deviation of the data set decreases as the
478 aggregation level increases, therefore users need to check the sensitivity of the range of
479 the variable of the model prior to applying direct averaging for data aggregation.
480
481 In fact, the SEBAL algorithm is nonlinear; that is the mean aggregated ET (output up-
482 scaled) at any given resolution does not equal the modeled ET value of an aggregated
483 input value (input up-scaled). However, as demonstrated by visual examination of the
484 spatial distribution of ET in Figures 7 – 10, the contrast as well as the basic patterns (high
485 and low values and their relative locations) of ET between output up-scaling and input
486 up-scaling show a slight disagreement. A slightly higher mean and standard deviation
487 was found in the results from input up-scaling with simple averaging than from output
488 up-scaling with simple averaging; however there is almost no difference between input
489 and output up-scaling when applying the nearest neighbor method. Overall, statistical and
490 spatial characteristics produced by input up-scaling show relatively good agreement with
491 those of the output up-scaling method.
492

493 **3.2.2. Difference of aggregated data versus original Landsat (30m) and MODIS**
494 **(250m) estimated ET**
495 First, aggregation difference was examined by comparing aggregated maps with the
496 original ET map at 30m resolution derived from Landsat imagery. Tables 3 and 4 present
497 the basic statistics of difference and relative difference against original Landsat derived
498 ET on June 16, 2002 and September 14, 2000 produced by four different up-scaling. The
499 mean values of absolute difference and relative difference range from 0.14 to 0.63 mm/d
500 and from 0.55 to 0.82, respectively.
501 The mean and standard deviation values of absolute difference from September image are
502 smaller than those from June image. The smaller mean difference and standard deviation
503 is explained by the smaller values of the ET rates in the September image. Mean values
504 of absolute difference from output up-scaled maps are similar with those from input up-
505 scaled maps; however consistently higher standard deviations are found in input up-
506 scaled maps (Tables 3 – 4). This result confirms that aggregated model output data
507 provide the best estimate of model output at the coarser resolution.
508
509 The mean and standard deviation of the absolute differences also increase with pixel size.
510 This is mainly due to the mixed pixel effect. Since aggregation tends to average out the
511 small surface features, the difference between aggregated imagery and the original fine
512 resolution imagery increases with aggregation levels. One interesting note is that the
513 mean of the relative difference increase with pixel size, however standard deviation
514 actually slightly decreases with pixel size. In this study relative difference is bounded to
515 be not greater than 1.0. Therefore, as mean values increase to approach 1.0, the standard

516 deviation of absolute difference actually decreases with increasing relative difference.
517 Based on the mean and standard deviation of the absolute difference and relative
518 difference, although the difference increases with aggregation levels, the ET of the
519 original images seems to be better preserved from the output up-scaling than input up-
520 scaling.
521
522 Both of the 3D frequency plots in Figure 11 between up-scaled ET and its relative
523 difference against Landsat-based ET show patterns similar to those in Figure 6. That is,
524 relative difference decreases with ET. However, points having 1.0 relative difference
525 with daily ET greater than 1.0 mm/d are greatly diminished in Figure 11. In particular,
526 the top portion of Figure 11, which shows the relative difference between the output up-
527 scaled ET and the ET obtained from simple averaging, shows very few of these
528 questionable points. In the bottom portion of Figure 11, which shows the relative
529 difference between the input up-scaled ET and the nearest neighbor up-scaled ET, there
530 are some points with a relative difference of 1.0, but there are far fewer such points than
531 in Figure 6. This indicates that there are fewer georeferencing disagreements between
532 Landsat-derived ET and output up-scaled ET than the one between Landsat and MODIS
533 images.
534
535 Next, we compare aggregation differences by comparing up-scaled maps at 250m
536 resolution with the original ET map from MODIS. This requires that we first examine
537 which aggregation scheme produces the best match with the original MODIS-based ET
538 and then check the quality of the different aggregation schemes. Landsat-based ET maps

539 at 30 m resolution were aggregated into 250 m resolution maps by applying the four
540 different aggregation schemes already presented in Figure 2. Table 5 show the basic
541 statistics of the absolute difference and relative difference of images from the four
542 different up-scaling schemes at 250m resolution compared with MODIS-based ET of
543 June and September.
544
545 The mean and standard deviation of absolute difference and relative difference from
546 output up-scaling with the simple averaging map are smaller than the one from input up-
547 scaling (Table 5). Also the simple averaging method generates smaller absolute
548 difference and relative difference than the nearest neighboring method. This implies that
549 output up-scaling with simple averaging map has best agreement with MODIS derived
550 ET. No difference between output and input up-scaling is found from the nearest
551 neighbor aggregation method. As shown in the previous section, the maximum and
552 standard deviation of the ET maps produced by simple averaging are decreased as data
553 were aggregated to 250m resolution. However, the nearest neighbor aggregation method
554 generated images having a similar maximum and standard deviation to the original image
555 (Figures 9 – 10). This explains why the mean and standard deviation of absolute
556 difference between aggregated Landsat ET image using simple averaging and MODIS-
557 based ET are smaller than from nearest neighbor (Table 5).
558
559 Although the difference increases with aggregation levels, the ET of
560 the original images seems to be better preserved with output up-scaling than

561 with input up-scaling. Out of the four different up-scaling procedures, output up-scaling
562 with simple averaging performs best. However, all four aggregation schemes are still
563 acceptable since the mean and standard deviation values of absolute difference are all less
564 than those from the original Landsat ET imagery in Figure 4.
565
566 **3.3. COMPARISON OF SEBAL UNCERTAINTY WITH UP-SCALING EFFECTS**
567
568 As mentioned in Section 1, it has been reported that SEBAL daily ET estimates agree
569 with ground observation with an accuracy of ±10%.The great strength of the SEBAL is
570 due to its internal calibration procedure that eliminates most of the bias in latent heat flux
571 at the expense of increased bias in sensible heat flux (Allen et al., 2007; Hong, 2008).
572
573 In order to examine the difference among up-scaling schemes, the relative difference
574 between up-scaled ET images at 120 and 1000m resolutions are calculated and histogram
575 and descriptive statistics are shown in Figure 12. Relative difference is calculated
576 between Upscaling2, 3 or 4 against Upscaling1 (Figure 2) as [|($ET_{upscaling2,\ 3\ or\ 4}$ −
577 $ET_{upscaling1}$)| /$ET_{upscaling1}$]. Up-scaling1 (output simple averaging) is taken as reference
578 since output simple averaging generates the best matched up-scaled map with respect to
579 the MODIS-derived ET map (Table 5). As shown in Figure 1, the study area includes
580 surrounding desert where soil moisture is little thus ET is very small. Since area having
581 very low ET can easily introduce very high relative difference and moreover it is difficult
582 to precisely estimate ET in desert area anyhow, the area having less than 1mm/day is

583 excluded in this analysis. The portion of area having less than 1mm/day covers about
584 50% of whole study area.
585
586 As shown in Figure 12, first, mean relative difference increases with pixel spatial
587 resolution, and second, relative difference between simple averaging and nearest
588 neighboring resampling (Upscaling1-Upscaling2 and Upscaling1-Upscaling4) has a lot
589 higher mean and standard deviation than the one between two simple averaging schemes
590 (Upscaling1-Upscaling3). This simply indicates that as pixel size increases, the difference
591 between simple averaging and nearest neighboring resampling increases. However, mean
592 relative difference between Upscaling1 and Upscaling3 (input simple averaging) in both
593 120m and 1000m resolution are all less than 10% which is smaller than the magnitude of
594 SEBAL uncertainty. A little difference between output and input up-scaling implies that
595 SEBAL is close to linearity model and that is due to its internal calibration procedure
596 (dT-T_s relationship). Another interesting point is that for the 1000m resolution histograms
597 of Upscaling1-Upscaling2 and Upscaling1-Upscaling3, considerable data points have
598 relative difference greater than 10% and especially lots of pixels (15% frequency) have
599 relative difference greater than 90%. Those areas having >90% relative difference are
600 mainly located along the boundary between riparian and desert areas. These pixels in the
601 boundary area are mixed with riparian (high ET) and desert (low ET), thus the difference
602 between up-scaled ET map by simple averaging and nearest neighbor resampling is
603 significant and causes very high relative difference.
604

605 Based on the result of relative difference analysis, the difference between simple
606 averaging and nearest neighbor is a lot bigger than the uncertainty of the SEBAL
607 procedure. Therefore, users have to aware of the difference and are careful to select
608 appropriate up-scaling scheme for their research.
609
610
611 ## 4. CONCLUSIONS
612
613 Daily evapotranspiration rates were predicted using the SEBAL algorithm from Landsat
614 7 and MODIS imagery. The objectives of this study were to test the consistency of the
615 SEBAL algorithm for the different satellite sensors and to investigate the effect of
616 various proposed aggregation procedures.
617
618 Although ET maps derived from the Landsat 7 images showed higher maximum and
619 standard deviation values than those derived from the MODIS images, the mean values of
620 Landsat- and MODIS-based ET images were very similar. Discrepancy in direct pixel-
621 by-pixel comparison between Landsat- and MODIS-based ET was due to mainly
622 georeferencing disagreement as well as the inherent differences in spatial, spectral and
623 radiometric resolutions between imagery from the different satellite sensors.
624
625 The output up-scaling scheme produced slightly better ET maps than the input up-scaling
626 scheme. Both simple averaging and nearest neighbor resampling methods can preserve
627 the mean values of the original images across aggregation levels. However, the simple

averaging resampling method resulted in decreasing standard deviation values as the resolution coarsened, while the standard deviation did not change across aggregation levels with the nearest neighbor resampling method. For difference analysis, large relative differences predominantly occur in areas having low ET (desert and bare soil) while areas having high ET (agricultural field and riparian vegetation) exhibit small relative differences. Out of the four different up-scaling procedures proposed in this study, output up-scaling with simple averaging performs best. However, other aggregation schemes are still acceptable.

Results of the relative difference analysis among up-scaling schemes show that a little difference between output and input up-scaling is found. However, there significant difference exists between simple averaging and nearest neighbor and its difference is a lot bigger than the uncertainty of the SEBAL procedure.

5. ACKNOWLEDGEMENT

The following sponsors have contributed to this study: U.S. Department of Agriculture, CSREES grant No.: 2003-35102-13654 and NSF EPSCoR grant EPS-0132632.

6. REFERENCES

Allen, R.G., M. Tasumi, and R. Trezza. 2007. Satellite-based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC) – Model. Journal of Irrigation and Drainage Engineering, ASCE 133:380-394.

652 Allen, R.G., L.S. Pereira, D. Raes, and M. Smith. 1998. Crop evapotranspiration. FAO
653 Irrigation and drainage paper 56, FAO. Rome.
654 Anselin, L., and A. Getis. 1993. Spatial statistical analysis and geographic information
655 systems Springer-Verlag, New York.
656 Atkinson, P.M. 1985. Preliminary Results of the Effect of Resampling on Thematic
657 Mapper Imagery. ACSM-ASPRS Fall Convention Technical Papers:929-936.
658 Bastiaanssen, W.G.M. 2000. SEBAL-based sensible and latent heat fluxes in the Irrigated
659 Gediz Basin, Turkey. Journal of Hydrology 229:87-100.
660 Bastiaanssen, W.G.M., M. Menenti, R.A. Feddes, and A.A.M. Holtslag. 1998. A remote
661 sensing surface energy balance algorithm for land (SEBAL). Part 1: Formulation.
662 Journal of Hydrology 212-213:198-212.
663 Bastiaanssen, W.G.M., E.J.M. Noordman, H. Pelgrum, G. Davids, B.P. Thoreson, and
664 R.G. Allen. 2005. SEBAL model with remotely sensed data to improve water-
665 resources management under actual field conditions. Journal of Irrigation and
666 Drainage Engineering 131:85-93.
667 Bian, L. 1997. Multiscale nature of spatial data in scaling up environment models CRC
668 Press, Inc.
669 Bian, L., R. Butler, D.A. Quattrochi, and P.M. Atkinson. 1999. Comparing effects of
670 aggregation methods on statistical and spatial properties of simulated spatial data.
671 Photogrammetry and Remote Sensing 65:73-84.
672 Brown, D.G., L. Bian, and S.J. Walsh. 1993. Response of a distributed watershed erosion
673 model to variations in input data aggregation levels. Computers and Geosciences
674 19:499-509.
675 Brutsaert, W. 1982. Evaporation into the atmosphere D. Reidel Pub. Co., Dordrecht, The
676 Netherlands.
677 Brutsaert, W., and M. Sugita. 1992. Application of self-preservation in the diurnal
678 evolution of the surface energy budget to determine daily evaporation. Journal of
679 Geophysical Research 97:18,377-18,382.
680 Carmel, Y. 2004. Controlling data uncertainty via aggregation in remotely sensed data.
681 IEEE Transaction on Geoscience and Remote Sensing Letters 1:39-41.
682 Carmel, Y., J. Dean, and C.H. Flather. 2001. Combining location and classification error
683 sources of estimating multi-temporal database accuracy. Photogrammetric
684 Engineering and Remote Sensing 67:865-872.
685 Chehbouni, A., and E.G. Njoku. 1995. Approaches for averaging surface parameters and
686 fluxes over heterogeneous terrain. Journal of Climate 8:1386-1393.
687 Choudhury, B.J., S.B. Idso, and R.J. Reginato. 1987. Analysis of an empirical model for
688 soil heat flux under a growing wheat crop for estimating evaporation by an
689 infrared-temperature-based energy balance equation. Agriculture and Forest
690 Meteorology 39:283-297.
691 Cihlar, J., H. Ly, Z. Li, J. Chen, H. Pokrant, and F. Hung. 1997. Multi-temporal, Multi-
692 channel AVHRR data sets for land biosphere studies – Artifacts and corrections.
693 Remote Sensing of Environment 60:35-57.
694 Clothier, B.E., K.L. Clawson, J. P.J. Pinter, M.S. Moran, R.J. Reginato, and R.D. Jackson.
695 1986. Estimation of soil heat flux from net radiation during growth of alfalfa.
696 Agricultural and Forest Meteorology 37:319-329,.

697 Cover, T., and P. Hart. 1967. Nearest neighbor pattern classification. IEEE Transactions
698 on Information Theory 13:21- 27.
699 Crago, R.D. 1996. Conservation and variability of the evaporative fraction during the
700 daytime. Journal of Hydrology 180:173-194.
701 Croley, I., T. E., C. He, and D.H. Lee. 2005. Distributed-Parameter Large Basin Runoff
702 Model. II: Application. Journal of Hydrologic Engineering 10:182-191
703 Dai, X.L., and S. Khorram. 1998. The effects of image misregistration on the accuracy of
704 remotely sensed change detection. IEEE Transaction on Geoscience and Remote
705 Sensing 36:1566-1577.
706 Daughtry, C.S., W.P. Kustas, M.S. Moran, R.D. Jackson, and J. Pinter. 1990. Spectral
707 estimates of net radiation and soil heat flux. Remote Sensing of Environment
708 32:111-124.
709 De Cola, L. 1994. Simulating and mapping spatial complexity using multi-scale
710 techniques. International Journal of Geographical Information Systems 8:411-427.
711 Dodgson, N.A. 1997. Quadratic interpolation for image resampling. IEEE Transactions
712 on Image Processing 6:1322-1326.
713 Ebleringer, J.R., and C.B. Field. 1993. Scaling Physiological Processes, Leaf to Globe
714 Academic Press, Inc., New York.
715 ERDAS. 2002. Field Guide 6th ed. Atlanta, Georgia, ERDAS Inc.
716 Eugenio, F., and F. Marqués. 2003. Automatic Satellite Image Georeferencing Using a
717 Contour-Matching Approach. IEEE Transactions on Geoscience and Remote
718 Sensing 41:2869-2880.
719 Farah, H.O., W.G.M. Bastiaanssen, and R.A. Feddes. 2004. Evaluation of the temporal
720 variability of the evaporative fraction in a tropical watershed. International
721 Journal of Applied Earth Observation and Geoinformation 5:129-140.
722 French, A.N. 2001. Scaling of surface energy fluxes using remotely sensed data. PhD
723 Thesis, University of Maryland, College Park.
724 French, A.N., T.J. Schmugge, and W.P. Kustas. 2002. Estimating evapotranspiration over
725 El Reno, Oklahoma with ASTER imagery. Agronomie 22:105-106.
726 Gentine, P., D. Entekhabi, A. Chehbouni, G. Boulet, and B. Duchemin. 2007. Analysis of
727 evaporative fraction diurnal behaviour. Agricultural and Forest Meteorology
728 143:13-29.
729 Gupta, V.K., I. Rodriguez-Iturbe, and E.F. Wood. 1986. Scale Problems in Hydrology D.
730 Reidel Publishing Company, Boston.
731 Hendrickx, J.M.H., and S.-H. Hong. 2005. Mapping sensible and latent heat fluxes in arid
732 areas using optical imagery. Proceedings of International Society for Optical
733 Engineering, SPIE 5811:138-146.
734 Hong, S.-H. 2008. Mapping regional distributions of energy balance components using
735 optical remotely sensed imagery. PhD. Dissertation. Dept. of Earth &
736 Environmental Science (Hydrology Program), New Mexico Institute of Mining
737 and Technology, Socorro, NM, USA.
738 Hong, S.-H., J.M.H. Hendrickx, and B. Borchers. 2005. Effect of scaling transfer
739 between evapotranspiration maps derived from LandSat 7 and MODIS images.
740 Proceedings of International Society for Optical Engineering, SPIE 5811:147-158.

741 Kustas, W.P., C.S.T. Daughtry, and P.J.V. Oevelen. 1993. Analytical treatment of the
742 relationships between soil heat flux/net radiation ratio and vegetation indices.
743 Remote Sensing and Environment 46:319-330.
744 Lam, N., and D.A. Quattrochi. 1992. On the issues of scale, resolution, and fractal
745 analysis in the mapping sciences. Professional Geographer 44:88-98.
746 Lhomme, J.-P. 1992. Energy balance of heterogeneous terrain: Averaging the controlling
747 parameters. Agricultural and Forest Meteorology 61:11-21.
748 Li, B., and R. Avissar. 1994. The impact of spatial variability of land-surface
749 characteristics on land-surface fluxes. Journal of Climate 7:527-537.
750 Liang, S. 2000. Numerical experiments on spatial scaling of land surface albedo and leaf.
751 area index. Remote Sensing Reviews 19:225–242.
752 Liang, S. 2004. Quantitative Remote Sensing of Land Surfaces John Wiley and Sons, Inc.
753 Liang, S., H. Fang, M. Chen, C.J. Shuey, C. Walthall, C. Daughtry, J. Morisette, C.
754 Schaaf, and A. Strahler. 2002. Validating MODIS land surface reflectance and
755 albedo products: methods and preliminary results. Remote Sensing of
756 Environment 83:149-162.
757 Maayar, M.E., and J.M. Chen. 2006. Spatial scaling of evapotranspiration as affected by
758 heterogeneities in vegetation, topography, and soil texture. Remote Sensing of
759 Environment 102:33-51.
760 Mark, D.M., and P.B. Aronson. 1984. Scale dependent fractal dimensions of topographic
761 surfaces: An empirical investigation with applications in geomorphology and
762 computer mapping. Mathmatical Geology 16:671-683.
763 Mecikalski, J.R., G.R. Diak, M.C. Anderson, and J.M. Norman. 1999. Estimating fluxes
764 on continental scales using remotely-sensed data in an atmospheric-land exchange
765 model. Journal of Applied Meteorology 38:1352-1369.
766 Mengelkamp, H.-T., F. Beyrich, G. Heinemann, F. Ament, J. Bange, F. Berger, J.
767 Bösenberg, T. Foken, B. Hennemuth, C. Heret, S. Huneke, K.-P. Johnsen, M.
768 Kerschgens, W. Kohsiek, J.-P. Leps, C. Liebethal, H. Lohse, M. Mauder, W.
769 Meijninger, S. Raasch, C. Simmer, T. Spieß, A. Tittebrand, J. Uhlenbrock, and P.
770 Zittel. 2006. Evaporation Over A Heterogeneous Land Surface. Bulletin of the
771 American Meteorological Society 87:775-786.
772 Morse, A., M. Tasumi, R.G. Allen, and W.J. Kramer. 2000. Application of the SEBAL
773 methodology for estimating consumptive use of water and streamflow depletion
774 in the Bear river basin of Idaho through remote sensing. Final report submitted to
775 the Raytheon Systems Company, Earth Observation System Data and Information
776 System Project, by Idaho Department of Water Resources and University of Idaho.
777 Nellis, M.D., and J.M. Briggs. 1989. The effect of spatial scale on Konza landscape
778 classification using textural analysis. Landscape Ecology 2:93-100.
779 Nishida, K., R.R. Nemani, S.W. Running, and J.M. Glassy. 2003. An operational remote
780 sensing algorithm of land surface evaporation. Journal of Geophysical Research
781 108 (D9):doi:10.1029/2002JD002062.
782 Price, J.C. 1984. Land surface temperature measurements from the split window channel
783 of the NOAA 7 Advanced Very High Resolution Radiometer. Journal of
784 Geophysical Research 89:7231-7237.
785 Quattrochi, D.A., and M.F. Goodchild. 1997. Scale, multiscaling, remote sensing, and
786 GIS Lewis Publishers, New York.

787 Seguin, B., J.-P. Lagouarde, and M. Saranc. 1991. The assessment of regional crop water
788 conditions from meteorological satellite thermal infrared data. Remote Sensing of
789 Environment 35:141-148.
790 Seyfried, M.S., and B.P. Wilcox. 1995. Scale and the nature of spatial variability: Field
791 examples having implications for hydrologic modeling. Water Resources
792 Research 31:173 - 184.
793 Shuttleworth, W.J. 1991. The modllion concept. Review of Geophysics 29:585-606.
794 Shuttleworth, W.J., R.J. Gurney, A.Y. Hsu, and J.P. Ormsby. 1989. The variation in
795 energy partition at surface flux sites. Proceedings of the IAHS Third International
796 Assembly, Baltimore, MD. 186:67-74.
797 Stoms, D. 1992. Effects of habitat map generalization in biodiversity assessment.
798 Photogrammetric Engineering and Remote Sensing 58:1587-1591.
799 Tasumi, M. 2003. Progress in operational estimation of regional evapotranspiration using
800 satellite imagery. Ph.D., University of Idaho, Moscow, Idaho.
801 Townshend, J.R.G., C.O. Justice, C. Gurney, and J. McManus. 1992. The impact of
802 misregistration on change detection. IEEE Transaction on Geoscience and
803 Remote Sensing 30:1054-1060.
804 Turner, M.G., R.V. O'Neil, R.H. Gardner, and B.T. Milne. 1989. Effects of changing
805 spatial scale on the analysis of landscape pattern. Landscape Ecology 3:153-162.
806 Van Rompaey, A.J.J., G. Govers, and M. Baudet. 1999. A strategy for controlling error of
807 distributed environmental models by aggregation. International Journal of
808 Geographical Information Science 13:577-590.
809 Vazquez, D.P., F.J. Olmo Reyes, and L.A. Arboledas. 1997. A comparative study of
810 algorithms for estimating land surface temperature from AVHRR data. Remote
811 Sensing of Environment 62:215-222.
812 Vieux, B.E. 1993. DEM Aggregation and smoothing effects on surface runoff modeling.
813 Journal of Computing in Civil Engineering 7:310-338.
814 Wolock, D.M., and C.V. Price. 1994. Effects of digital elevation model map scale and
815 data resolution on a topography-based watershed model. Water Resources
816 Research 40:3041-3052.
817 Zhang, W., and D.R. Montgomery. 1994. Digital elevation model grid size, landscape
818 representation, and hydrologic simulations. Water Resources Research 30:1019-
819 1028.
820
821

Table 1. Band spatial resolutions (m] and wavelengths (μm) of Landsat 7 and MODIS sensors.

Sensors		Band number								
		1	2	3	4	5[#]	6	7	31	32
Landsat 7	Pixel size [m]	30	30	30	30	30	60	30	NA*	NA*
	Band width [μm]	0.45 – 0.51	0.52 – 0.60	0.63 – 0.69	0.75 – 0.9	1.55 – 1.75	10.4 – 12.5	2.09 – 2.35	NA*	NA
MODIS	Pixel size [m]	250	250	500	500	500	500	500	1000	1000
	Band width [μm]	0.62 – 0.67	0.84 – 0.87	0.46 – 0.48	0.54 – 0.56	1.23 – 1.25	1.63 – 1.65	2.11 – 2.15	10.8 – 11.3	11.8 – 12.3

[#]MODIS band5 is not used in this study because of streaking noise,
*Not available

Table 2. Constants K_1 and K_2 [$Wm^{-2}ster^{-1}\mu m^{-1}$] for Landsat 7 ETM+ (NASA, 2002) and MODIS (http://modis.gsfc.nasa.gov/).

	K_1	K_2
Landsat 7	666.09	1282.71
MODIS (band 31)	730.01	1305.84
MODIS (band 32)	474.99	1198.29

Table 3. Basic statistics of difference [mm/d] between up-scaled ET and original Landsat-based ET (30m). (note: statistics are calculated from absolute value of the difference)

Up-scaling approach	Up-scaling operation	June 16, 2002		September 14, 2000	
		Mean difference	Standard deviation	Mean difference	Standard deviation
Output	AVG_60[1]	0.20	0.34	0.14	0.24
	NN_60[2]	0.18	0.30	0.14	0.28
	AVG_120	0.30	0.48	0.17	0.27
	NN_120	0.32	0.35	0.23	0.33
	AVG_250	0.51	0.79	0.25	0.35
	NN_250	0.32	0.38	0.23	0.35
	AVG_500	0.54	0.81	0.27	0.36
	NN_500	0.38	0.41	0.27	0.36
	AVG_1000	0.63	0.90	0.30	0.38
	NN_1000	0.43	0.43	0.33	0.42
Input	AVG_60	0.28	0.50	0.14	0.25
	NN_60	0.18	0.31	0.15	0.28
	AVG_120	0.29	0.51	0.16	0.28
	NN_120	0.28	0.36	0.23	0.34
	AVG_250	0.53	0.85	0.24	0.36
	NN_250	0.32	0.38	0.23	0.35
	AVG_500	0.54	0.87	0.25	0.37
	NN_500	0.38	0.41	0.28	0.39
	AVG_1000	0.62	0.95	0.28	0.39
	NN_1000	0.43	0.43	0.32	0.42

[1]Aggregated to 60m by simple averaging, [2] Aggregated to 60m by nearest neighbor

Table 4. Basic statistics of relative difference [-] between up-scaled ET and original Landsat-based ET (30m). (note: statistics are calculated from absolute value of the relative difference)

Up-scaling approach	Up-scaling operation	June 16, 2002		September 14, 2000	
		Mean relative difference	Standard deviation	Mean relative difference	Standard deviation
Output	AVG_60[1]	0.55	0.44	0.68	0.42
	NN_60[2]	0.56	0.45	0.69	0.43
	AVG_120	0.58	0.42	0.70	0.40
	NN_120	0.60	0.42	0.73	0.40
	AVG_250	0.64	0.40	0.74	0.37
	NN_250	0.65	0.41	0.75	0.38
	AVG_500	0.64	0.39	0.74	0.37
	NN_500	0.69	0.38	0.78	0.35
	AVG_1000	0.65	0.39	0.76	0.36
	NN_1000	0.72	0.37	0.82	0.32
Input	AVG_60	0.60	0.44	0.70	0.41
	NN_60	0.56	0.45	0.70	0.42
	AVG_120	0.61	0.42	0.71	0.40
	NN_120	0.62	0.42	0.73	0.39
	AVG_250	0.66	0.40	0.76	0.37
	NN_250	0.65	0.41	0.75	0.38
	AVG_500	0.67	0.40	0.76	0.37
	NN_500	0.69	0.39	0.78	0.35
	AVG_1000	0.68	0.39	0.77	0.36
	NN_1000	0.73	0.36	0.82	0.32

[1]Aggregated to 60m by simple averaging, [2] Aggregated to 60m by nearest neighbor

Table 5. Basic statistics of difference [mm/d] and relative difference [-] of up-scaled ET against original MODIS-based ET (250m). (note: statistics are calculated from absolute value of the difference)

Up-scaling approach	Up-scaling operation	June 16, 2002		September 14, 2000	
		Mean difference	Standard deviation	Mean difference	Standard deviation
Output	AVG_250[1]	0.41	0.39	0.31	0.37
	NN_250[2]	0.46	0.41	0.36	0.41
Input	AVG_250	0.43	0.40	0.32	0.38
	NN_250	0.46	0.41	0.36	0.41
		Mean relative difference	Standard deviation	Mean relative difference	Standard deviation
Output	AVG_250	0.60	0.38	0.71	0.36
	NN_250	0.67	0.37	0.78	0.34
Input	AVG_250	0.65	0.38	0.75	0.36
	NN_250	0.67	0.37	0.78	0.34

[1]Aggregated to 250m by simple averaging, [2] Aggregated to 250m by nearest neighbor

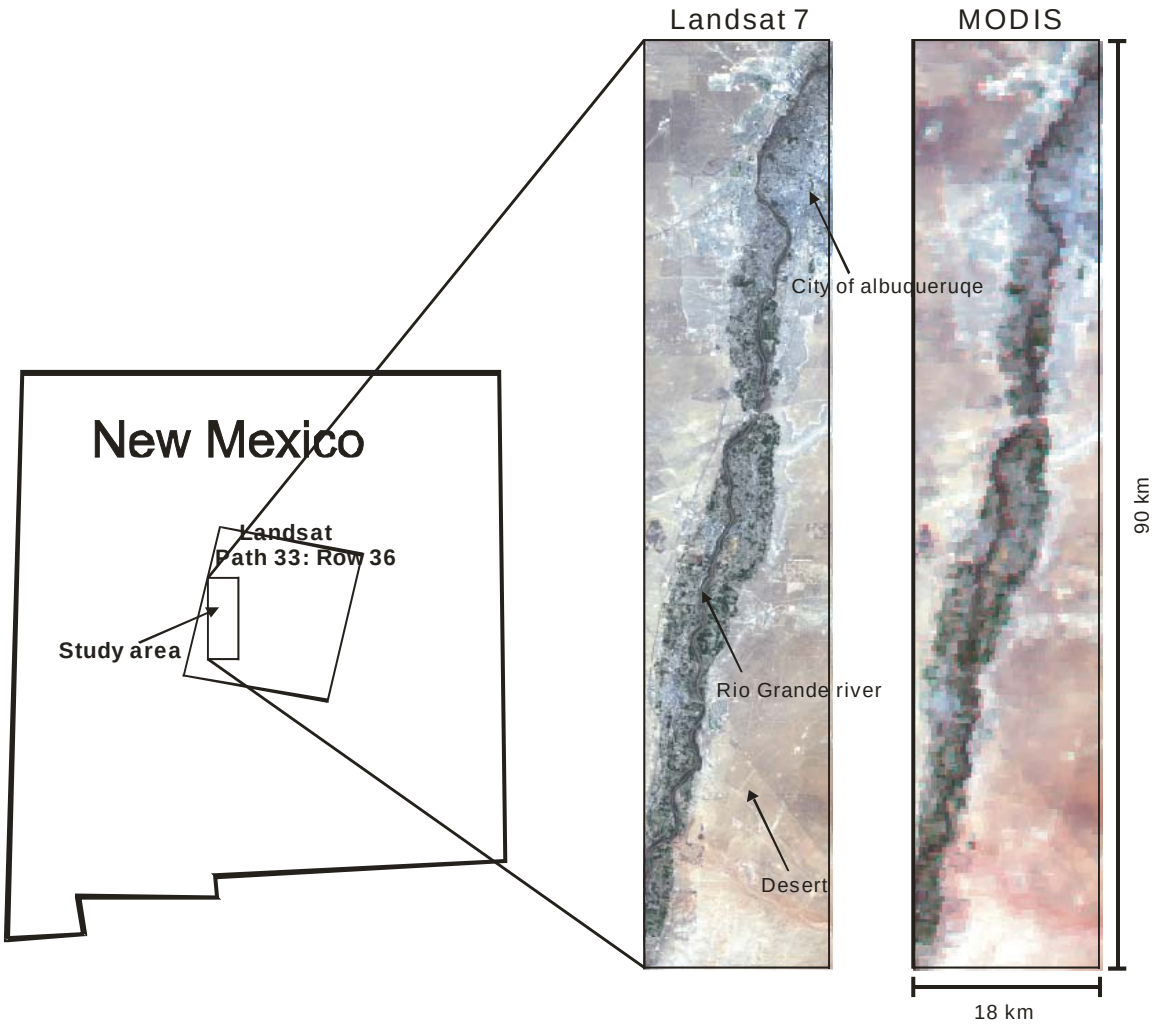


Figure 2. Location of the study area (18km by 90km). True color Landsat 7 (30 m by 30 m resolution) and MODIS (250 m by 250 m resolution) images on June 16, 2002.

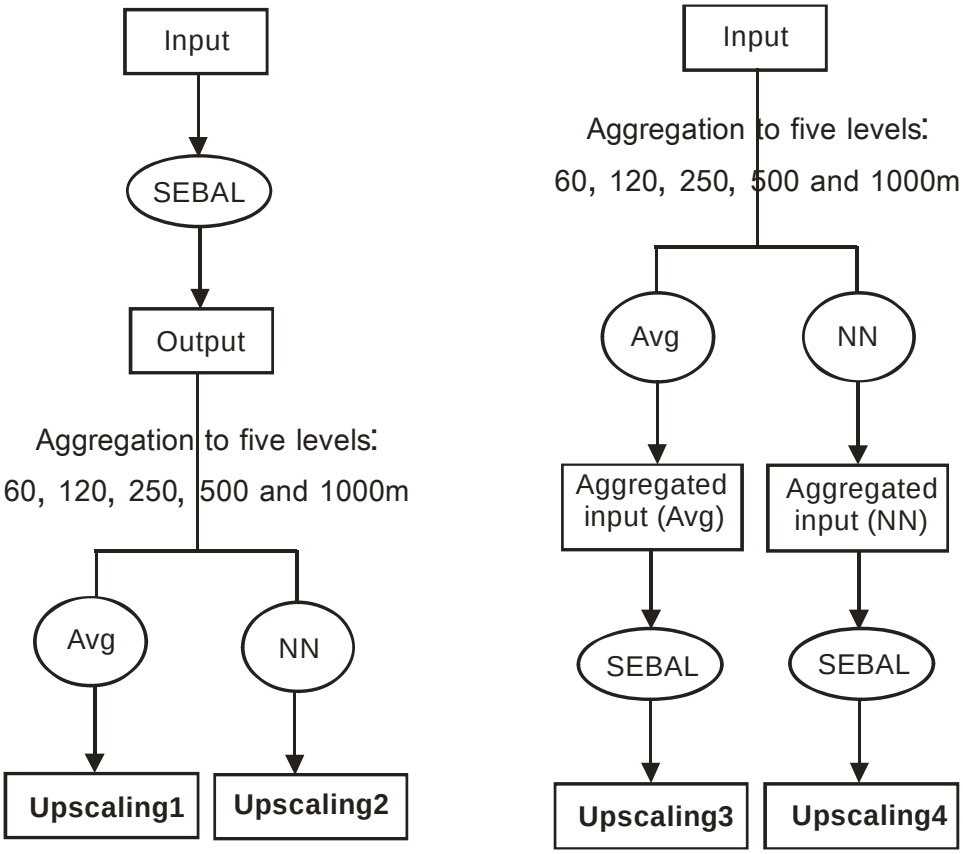
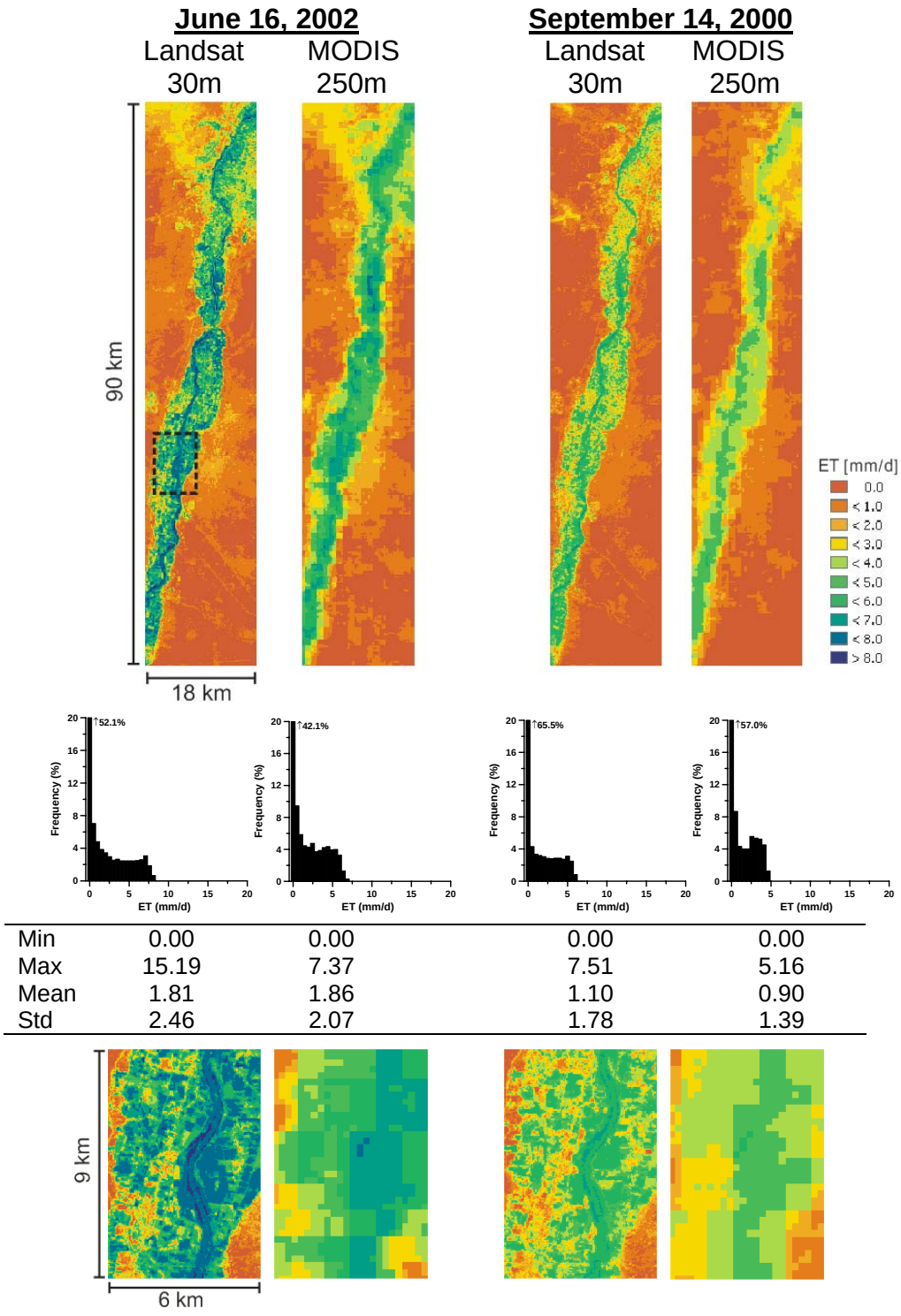
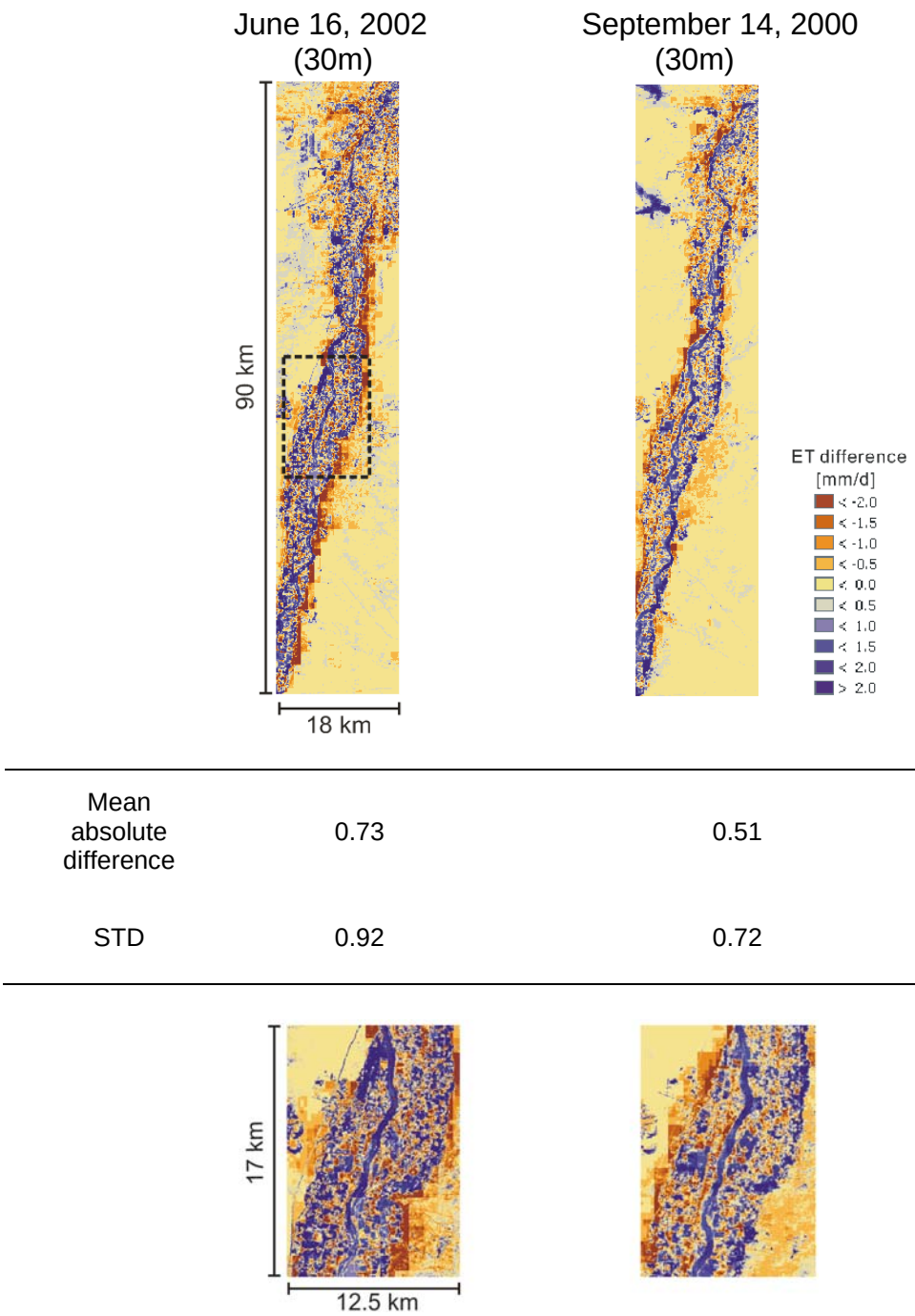


Figure 2. Schematic of up-scaling schemes applied in this study. (Upscaling1: output up-scaling with simple averaging, Upscaling2: output up-scaling with nearest neighbor, Upscaling3: input up-scaling with simple-averaging and Upscaling4: input-up-scaling with nearest neighbor).



Min	0.00	0.00	0.00	0.00
Max	15.19	7.37	7.51	5.16
Mean	1.81	1.86	1.10	0.90
Std	2.46	2.07	1.78	1.39

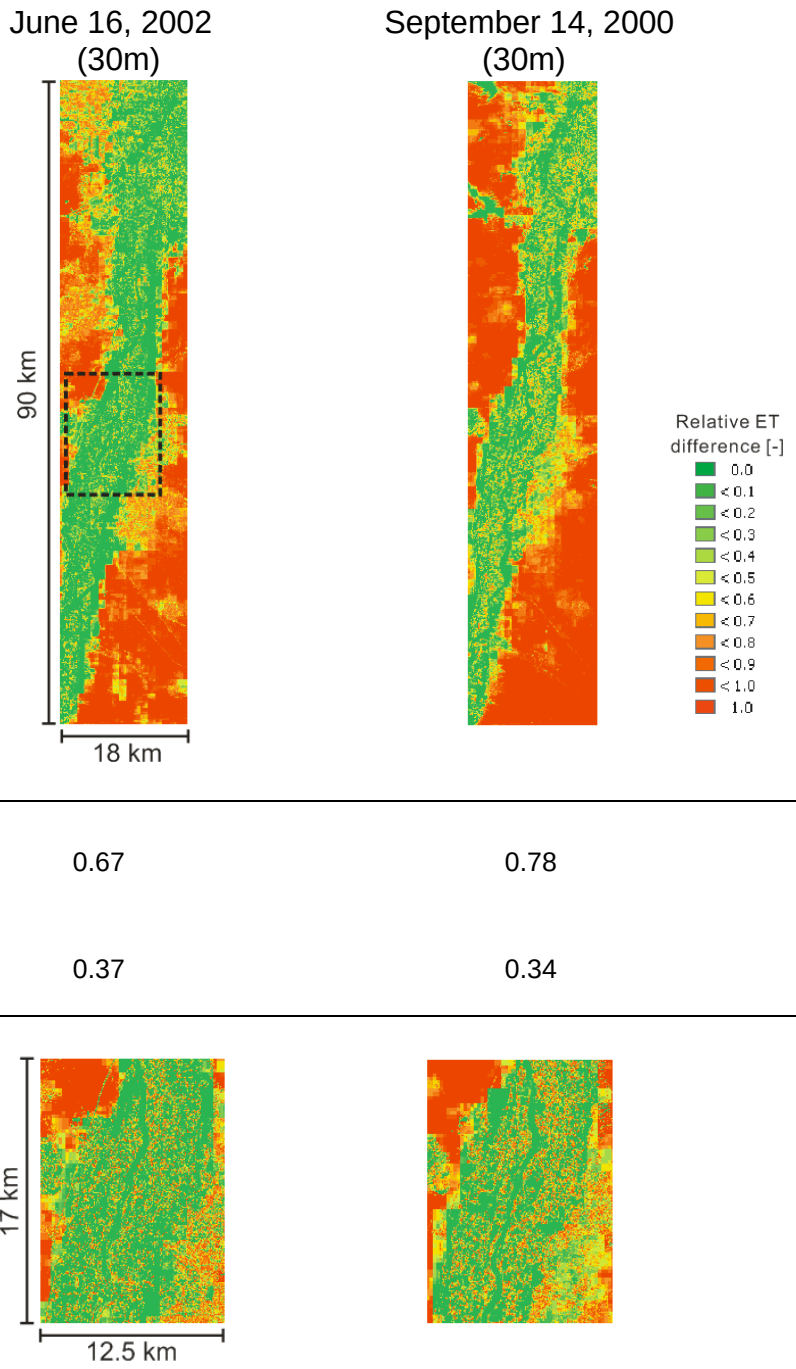
Figure 3. Landsat (30 m) and MODIS (250 m) derived ET by SEBAL of June and September. Bin size of the histogram is 0.5 mm/d and frequency occurrence exceeding 20% marked next to the arrow. The histograms and basic statistics are based on the entire maps (18 km x 90 km). Enlarged areas (9 by 6 km) shown at the bottom correspond to the dotted square of the upper images



	June 16, 2002 (30m)	September 14, 2000 (30m)
Mean absolute difference	0.73	0.51
STD	0.92	0.72

Figure 4. ET difference map (30 m) between the Landsat estimated ET (30m) and the MODIS estimate ET (250m). (note: mean and standard deviation (STD) are calculated with the absolute difference). Enlarged areas (12.5 by 17 km) shown at the bottom correspond to the dotted square of the upper images.

Relative difference between Landsat estimated ET (30m) and MODIS estimate ET (250m): $[|(ET_{Landsat} - ET_{MODIS})| / ET_{MODIS}]|$



	June 16, 2002 (30m)	September 14, 2000 (30m)
Mean relative difference	0.67	0.78
STD	0.37	0.34

Figure 5. Relative difference (30 m) between the Landsat estimated ET (30m) and the MODIS estimate ET (250m) on June 16, 2002. Enlarged areas (12.5 by 17 km) shown at the bottom correspond to the dotted square of the upper images.

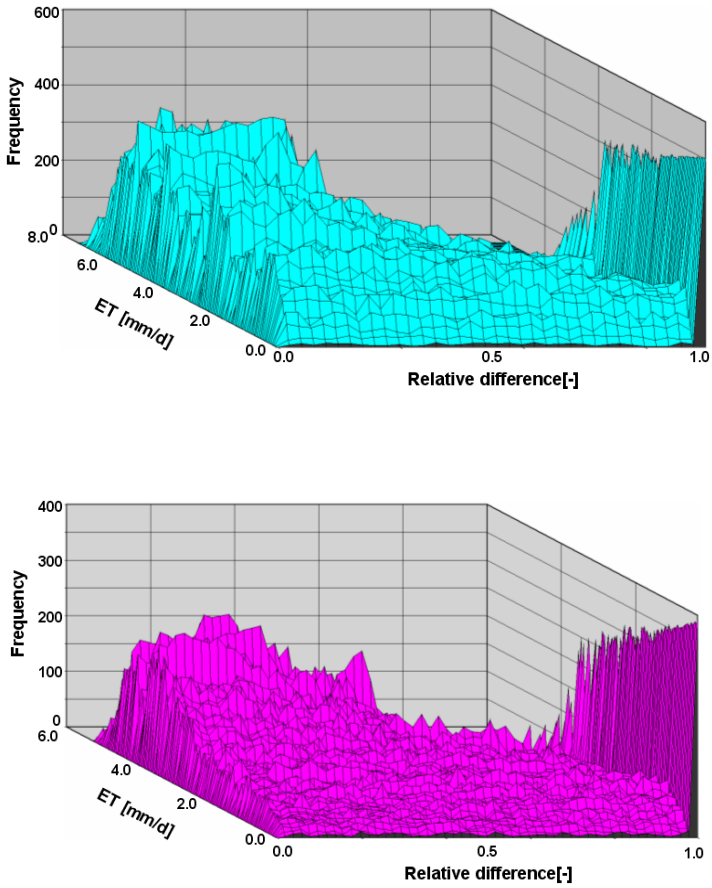


Figure 6. 3D frequency plot of the relative difference between Landsat drived ET (30m) and MODIS derived ET (250m) against MODIS derived ET (250m) (top: June 16, 2002 and bottom: September 14, 2000).

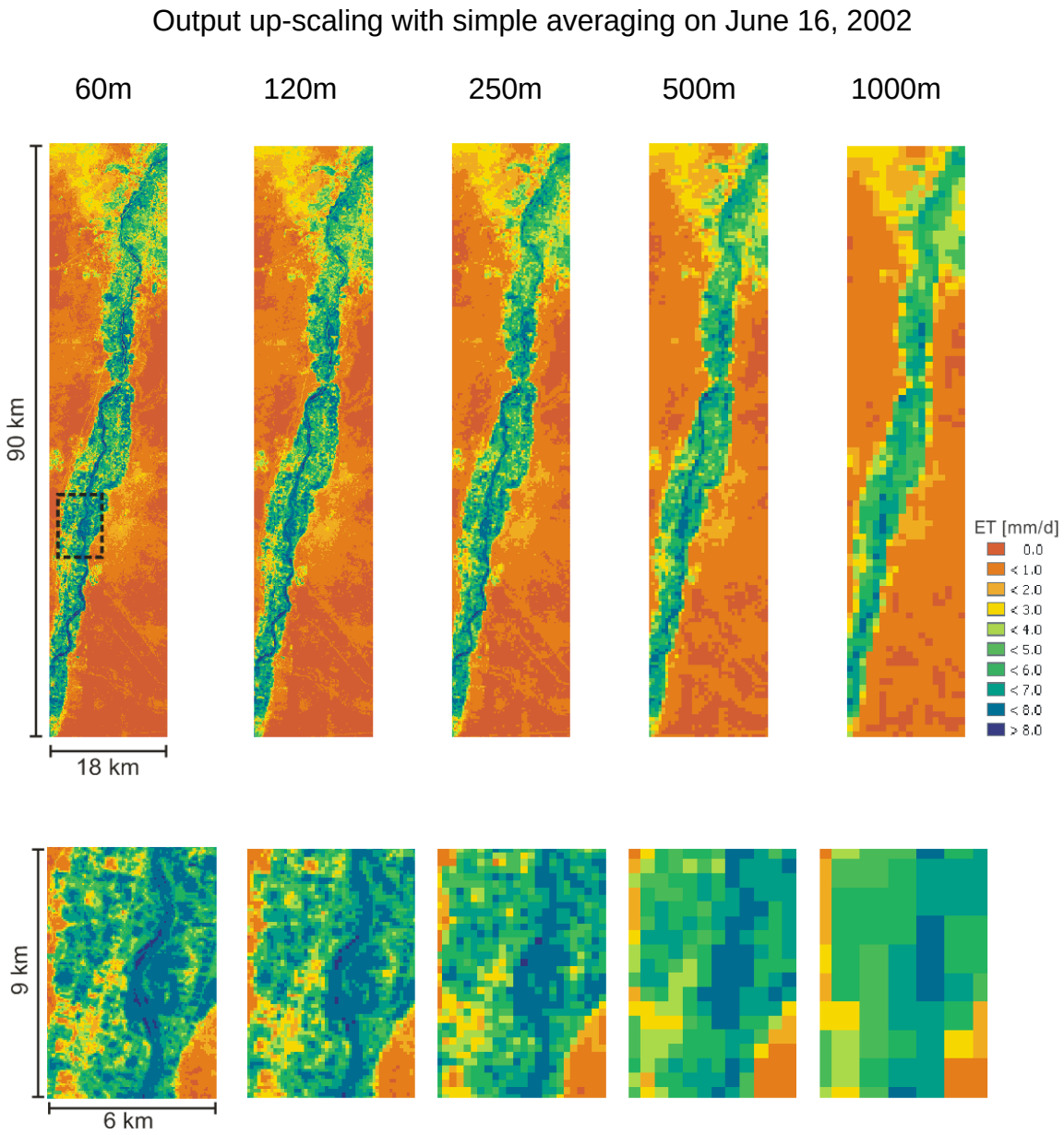


Figure 7. ET maps from output up-scaling using simple averaging resampling on June 16, 2002. Spatial resolutions are 60, 120, 250, 500 and 1000 m from the left. This method produces the most statistically and spatially predictable behavior.

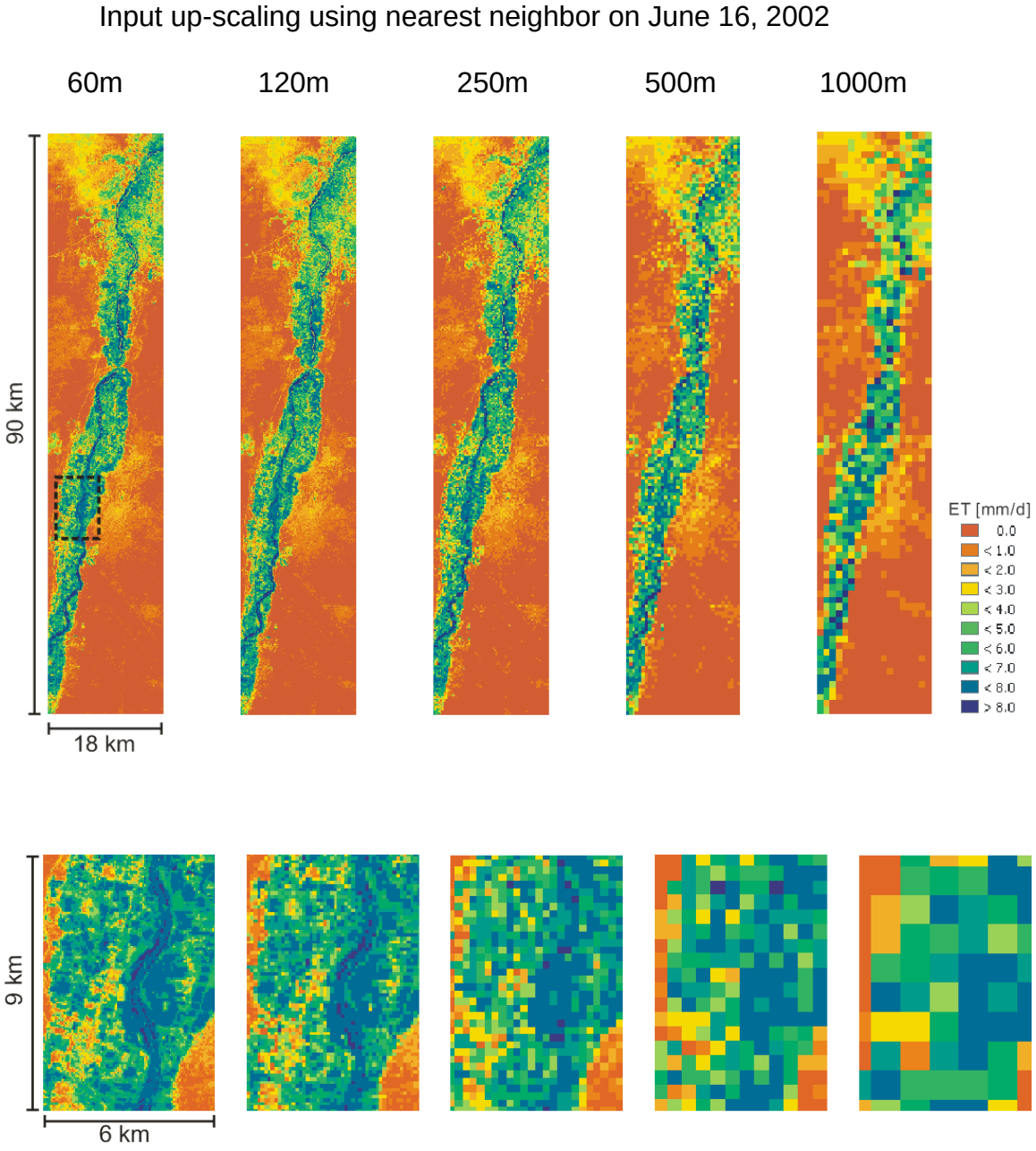
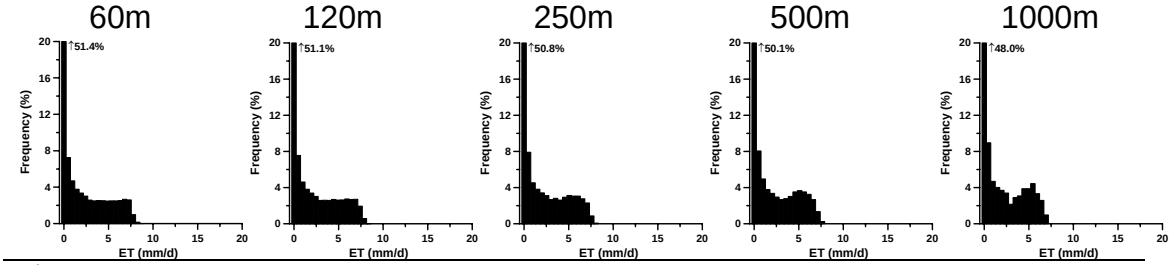


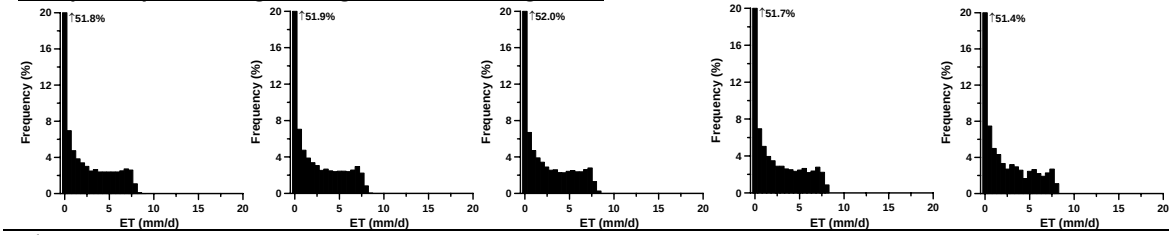
Figure 8. ET maps from input up-scaling using nearest neighbor resampling on June 16, 2002. Spatial resolutions are 60, 120, 250, 500 and 1000 m from the left. This method produces the best predictable behavior but is still acceptable.

Output up-scaling with simple averaging on **June 16, 2002**



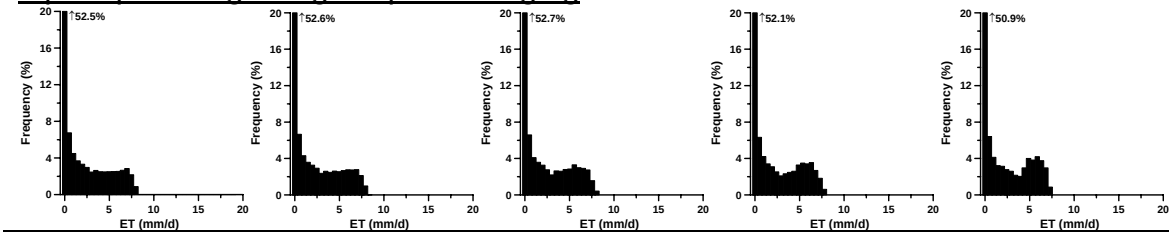
	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	9.25	8.90	8.60	8.17	7.44
Mean	1.81	1.81	1.81	1.81	1.81
Std	2.43	2.40	2.35	2.30	2.22

Output up-scaling using nearest neighbor



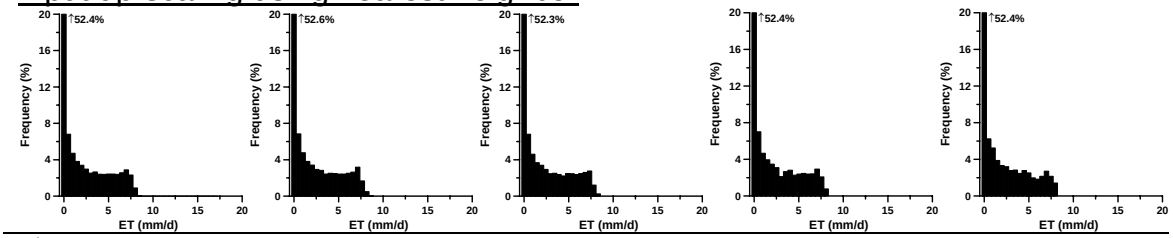
	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	9.38	9.02	8.78	8.56	8.53
Mean	1.81	1.81	1.81	1.81	1.79
Std	2.46	2.46	2.46	2.46	2.45

Input up-scaling using simple averaging



	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	9.26	8.92	8.57	8.28	7.81
Mean	1.81	1.82	1.84	1.84	1.85
Std	2.47	2.47	2.47	2.46	2.44

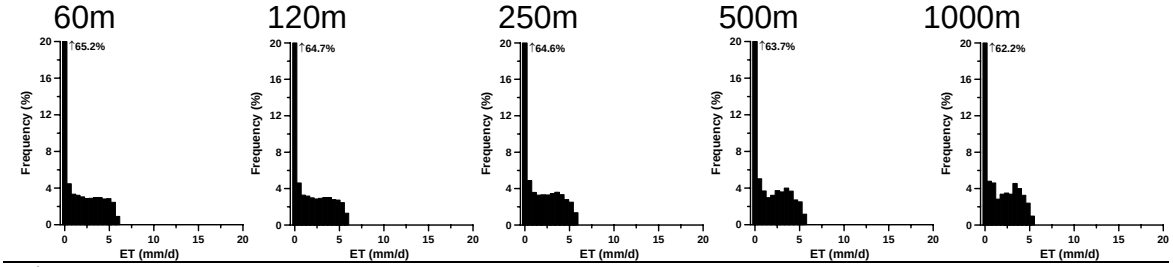
Input up-scaling using nearest neighbor



	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	9.26	8.92	8.57	8.28	7.81
Mean	1.81	1.82	1.84	1.84	1.85
Std	2.47	2.47	2.47	2.46	2.44

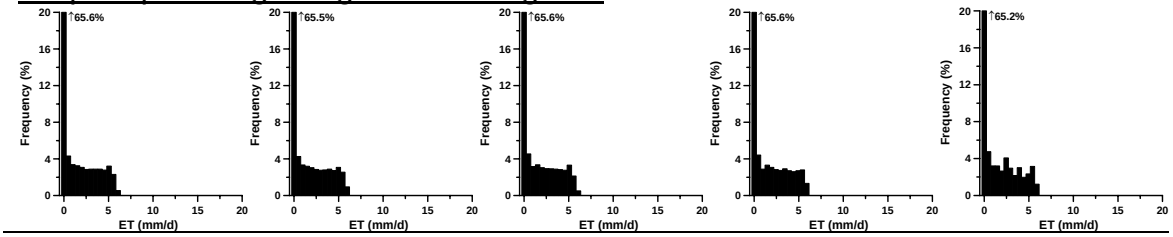
Figure 9. Frequency distribution and basic statistics of up-scaled maps on June 16, 2002. Bin size of the histogram is 0.5 mm/d and frequency occurrence exceeding 20% marked next to the arrow.

Output up-scaling using simple averaging on **September 14, 2000**



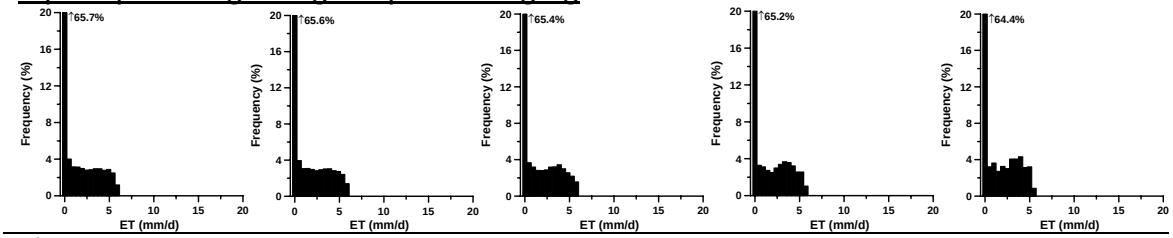
	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	6.43	6.29	6.21	6.05	5.78
Mean	1.10	1.10	1.10	1.10	1.10
Std	1.76	1.74	1.71	1.66	1.60

Output up-scaling using nearest neighbor



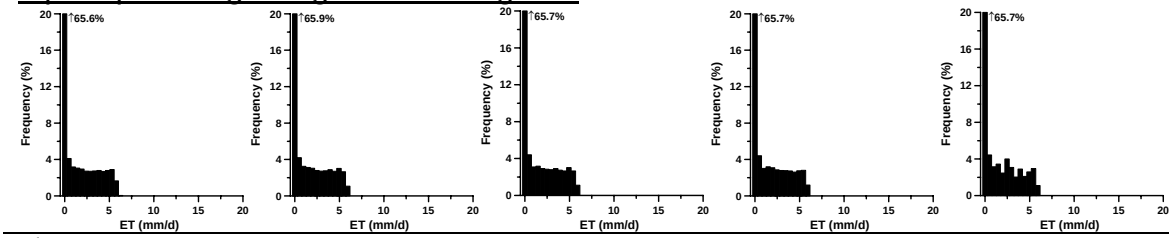
	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	7.10	6.61	6.48	6.40	6.40
Mean	1.10	1.10	1.10	1.10	1.08
Std	1.78	1.78	1.78	1.78	1.75

Input up-scaling using simple averaging



	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	6.40	6.35	6.27	6.19	5.99
Mean	1.10	1.10	1.11	1.11	1.12
Std	1.78	1.77	1.76	1.75	1.71

Input up-scaling using nearest neighbor



	60m	120m	250m	500m	1000m
Min	0.00	0.00	0.00	0.00	0.00
Max	6.75	6.44	6.43	6.42	6.42
Mean	1.10	1.10	1.09	1.09	1.07
Std	1.78	1.78	1.78	1.78	1.75

Figure 10. Frequency distribution and basic statistics of up-scaled maps on September 14, 2000. Bin size of the histogram is 0.5 mm/d and frequency occurrence exceeding 20% marked next to the arrow.

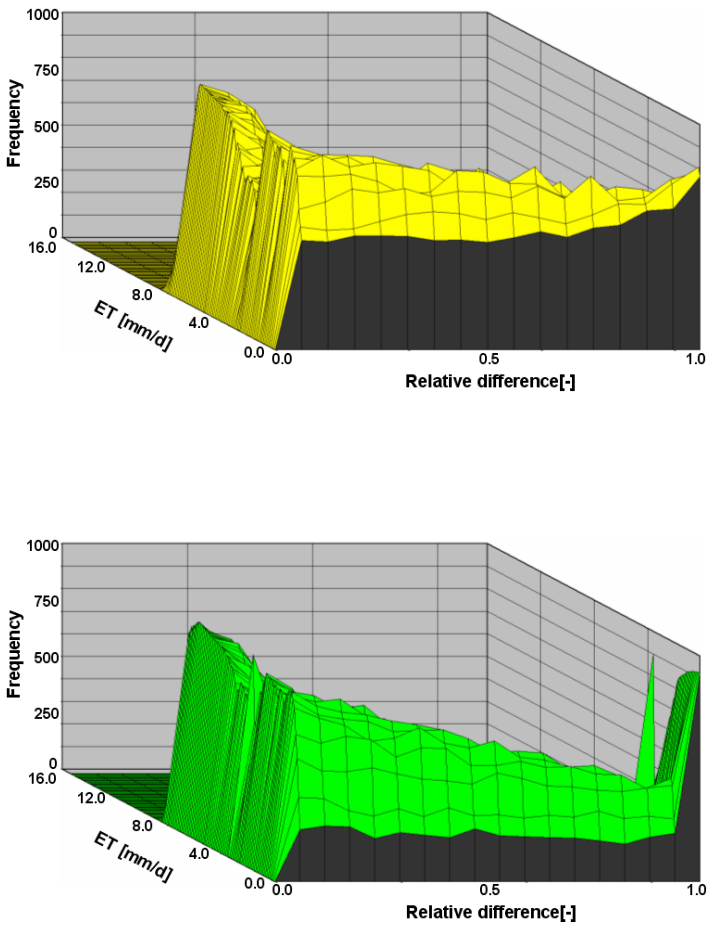


Figure 11. 3D frequency plot of the relative difference between up-scaled daily ET (250 m) and Landsat derived ET (30m) on June 16, 2002 against Landsat derived ET (30m) (top: up-scaling output with simple averaging and bottom: up-scaling input with nearest neighbor).

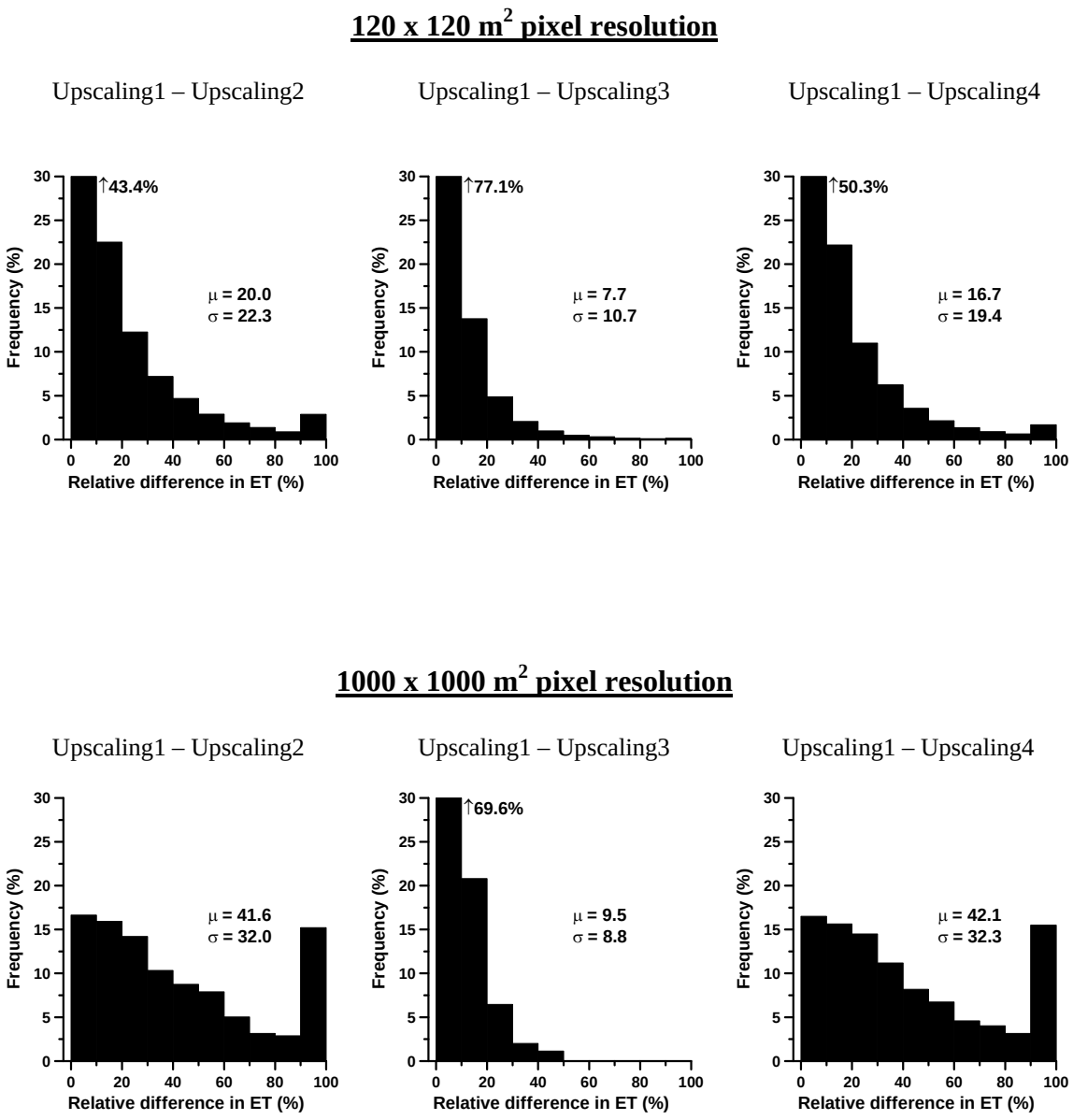


Figure 12. Frequency of the relative difference among up-scaling schemes.