ELSEVIER

A vegetation index based technique for spatial sharpening of thermal imagery

Nurit Agam [a,*], William P. Kustas [a], Martha C. Anderson [a], Fuqin Li [a], Christopher M.U. Neale [b]

[a] Hydrology and Remote Sensing Laboratory, ARS, USDA, Beltsville, Maryland, USA
[b] Biological and Irrigation Engineering Department, Utah State University, Logan, USA

Abstract

High spatial resolution (~ 100 m) thermal infrared band imagery has utility in a variety of applications in environmental monitoring. However, currently such data have limited availability and only at low temporal resolution, while coarser resolution thermal data (~ 1000 m) are routinely available, but not as useful for identifying environmental features for many landscapes. An algorithm for sharpening thermal imagery (TsHARP) to higher resolutions typically associated with the shorter wavebands (visible and near-infrared) used to compute vegetation indices is examined over an extensive corn/soybean production area in central Iowa during a period of rapid crop growth. This algorithm is based on the assumption that a unique relationship between radiometric surface temperature (T_R) relationship and vegetation index (VI) exists at multiple resolutions. Four different methods for defining a VI $- T_R$ basis function for sharpening were examined, and an optimal form involving a transformation to fractional vegetation cover was identified. The accuracy of the high-resolution temperature retrieval was evaluated using aircraft and Landsat thermal imagery, aggregated to simulate native and target resolutions associated with Landsat, MODIS, and GOES short- and longwave datasets. Applying TsHARP to simulated MODIS thermal maps at 1-km resolution and sharpening down to ~ 250 m (MODIS VI resolution) yielded root-mean-square errors (RMSE) of 0.67– 1.35 °C compared to the 'observed' temperature fields, directly aggregated to 250 m. Sharpening simulated Landsat thermal maps (60 and 120 m) to Landsat VI resolution (30 m) yielded errors of 1.8–2.4 °C, while sharpening simulated GOES thermal maps from 5 km to 1 km and 250 m yielded RMSEs of 0.98 and 1.97, respectively. These results demonstrate the potential for improving the spatial resolution of thermal-band satellite imagery over this type of rainfed agricultural region. By combining GOES thermal data with shortwave VI data from polar orbiters, thermal imagery with 250-m spatial resolution and 15-min temporal resolution can be generated with reasonable accuracy. Further research is required to examine the performance of TsHARP over regions with different climatic and land-use characteristics at local and regional scales.
© 2006 Published by Elsevier Inc.

Keywords: Land surface temperature; Spatial resolution; Image sharpening; Thermal remote sensing; Radiometric temperature

1. Introduction

High spatial resolution ($\leq 10^2$ m) thermal infrared (TIR) band imagery has utility in a variety of applications in environmental monitoring, which include detecting conditions conducive to wildfire, assessing ecosystem health and drought severity (Quattrochi & Luvall, 2004), monitoring volcanic eruptive activity (Pieri & Abrams, 2005), and exploring urban heat island effects (Voogt & Oke, 2003). A common use of thermal data is to derive surface energy budgets (Diak et al.,

2004), with high-resolution thermal providing assessments of evapotranspiration (ET) down to scales of individual agricultural fields (Norman et al., 2003) and evaporative losses along canals and riparian corridors (Loheide Ii & Gorelick, 2005). This type of information is needed to reliably plan water distribution in the western U.S. as well as in other arid and semiarid regions around the world.

Satellite-based thermal datasets currently available are summarized in Table 1. These datasets reflect a tradeoff between temporal and spatial resolution such that the systems have either high-spatial/low-temporal resolution (e.g., Landsat Thematic Mapper — TM; and Landsat Enhanced Thematic Mapper Plus — ETM+) or low-spatial/high-temporal resolution (e.g., National Oceanic and Atmospheric Administration-Advanced Very High Resolution Radiometer — NOAA-AVHRR; Terra/Aqua-Moderate Resolution Imaging Spectrometer —

* Corresponding author. Nurit Agam USDA-ARS Hydrology and Remote Sensing Laboratory Bldg. 007, BARC-WEST Beltsville, MD 20705. Tel.: +1 301 504 5825; fax: +1 301 504 8931.
E-mail addresses: agam@hydrolab.arsusda.gov, nurit.agam@gmail.com (N. Agam).

Table 1
Resolutions of vis/NIR and thermal bands of currently available spaceborne systems

Sensor	Satellite	vis/NIR bands spatial resolution	Thermal band spatial resolution	Revisit time
ETM+	Landsat 7	30 m	60 m	16 days
ASTER	Terra	15 m	90 m	On demand
TM	Landsat 5	30 m	120 m	16 days
MODIS	Terra, Aqua	250 m/1 km*	1 km	1–2/day
AVHRR	NOAA	1 km	1 km	2/day
GOES imager	GOES	–	5 km	15 min

ETM+ — Enhanced Thematic Mapper Plus.
TM — Thematic Mapper.
ASTER — Advanced Spaceborne Thermal Emission and Reflection Radiometer.
MODIS — Moderate Resolution Imaging Spectroradiometer.
GOES — Geostationary Operational Environmental Satellites.
*MODIS NDVI products are 16 day composite (MOD13Q1 and MOD13A2 for 250 m and 1 km, respectively).

MODIS; Geostationary Operational Environmental Satellite — GOES). ASTER data are only available by demand and therefore provide only sporadic temporal coverage at a given site. The Landsat 7 satellite has been experiencing technical difficulties with its Scan Line Corrector since May 31, 2003, resulting in striping that makes the imagery difficult to use for research and operational applications. At the time of writing, Landsat 5 has been operating for 22 years, and its lifetime is uncertain. The continuation of thermal band imaging on Landsat Data Continuity Mission platforms is currently under debate, and such high-resolution thermal data may soon be unavailable at any temporal resolution.

A technique to derive higher resolution land surface temperature (LST) from other available data is therefore highly desirable. For many landscapes, variability in LST is driven primarily by variations in vegetation cover amount. The functional relationship between the spaceborne-derived radiometric surface temperature (T_R) and vegetation indices (VIs) such as the Normalized Difference Vegetation Index (NDVI), which are generally available at finer spatial resolution than thermal band data (see Table 1), has been exploited by various remote sensing based energy balance modeling schemes for constraining/defining model variables/parameters (e.g., Gillies & Carlson, 1995; Moran et al., 1994; Price, 1990). A technique for using empirically derived NDVI-T_R relationship to disaggregate T_R to the shortwave band resolution has been described and evaluated for study sites in the Southern Great Plains (Anderson et al., 2004a; Kustas et al., 2003).

Kustas et al. (2003) used a 2nd degree polynomial least squares regression to describe the relationship between NDVI and T_R, and obtained errors of ∼1.5 °C when disaggregating ∼1550 and 770 m resolution surface temperature maps to 250 m, simulating an application to MODIS-resolution datasets. However, no additional information was gained in sharpening thermal imagery at resolutions finer than 100 m, in comparison with assuming a uniform sub-pixel temperature distribution. These results are consistent with an independent spatial scaling study showing that the greatest loss of T_R spatial variability came at pixel resolutions larger than 200–400 m, the typical

dimension of the agricultural field boundaries (French, 2001). However, it is hypothesized that by better describing the NDVI−T_R relationship the errors that result from sharpening T_R may decrease and some value may be gained also at finer resolutions.

In this paper, the original sharpening algorithm presented by Kustas et al. (2003) is refined, exploring alternative sharpening basis functions and evaluating their performance over an extensive corn/soybean production area in central Iowa during a period of rapid crop growth. Sharpening as applied to TM, ETM+, MODIS and GOES thermal imagery is simulated using high-resolution aircraft and Landsat imagery aggregated to coarser resolutions, and expected errors in retrieved temperature are evaluated over this agricultural landscape at watershed and regional scales.

2. Methodology

2.1. TsHARP algorithm

The TsHARP technique (previously referred to as disTrad), developed by Kustas et al. (2003) and utilized by Anderson et al. (2004a), is based on the assumption that a unique NDVI−T_R relationship exists within a sensor scene at multiple spatial resolutions, largely related to fractional vegetation cover. This assumption arises from the well-documented observation of an inverse relationship between land surface temperature and vegetation cover (e.g., Badeck et al., 2004; Running et al., 1995; Tucker et al., 2001; White et al., 1997). The slope of this relationship, however, varies with land cover and climate (Karnieli et al., 2006), and therefore is site- and scene-specific. In TsHARP, the parameters of the sharpening function are determined within the context of the scene to be sharpened, and thus reflect the local relationship between surface temperature and vegetation cover fraction.

The sharpening methodology is as follows. A least-squares regression is performed between T_R and some function of NDVI, aggregated to the coarser thermal resolution (NDVI$_{low}$):

$$\hat{T}_R(\text{NDVI}_{low}) = f(\text{NDVI}_{low}) \tag{1}$$

and then this regression relationship is applied to the NDVI data at their finer, native resolution (NDVI$_{high}$). Here the "hat" symbol indicates a temperature value predicted using the VI regression equation. The divergence of the retrieved temperatures from the observed temperature field is due to spatial variability in LST driven by factors other than vegetation cover fraction (for example, soil moisture variations), and can be assessed at the coarse scale:

$$\Delta\hat{T}_{R\ low} = T_{R\ low} - \hat{T}_R(\text{NDVI}_{low}) \tag{2}$$

This coarse-scale residual field is added back into the sharpened map such that the original temperature field is recovered through re-aggregation. The sharpened sub-pixel temperatures within each coarse pixel are therefore computed via:

$$\hat{T}_{R\ high} = \hat{T}_R(\text{NDVI}_{high}) + \Delta\hat{T}_{R\ low} \tag{3}$$

Table 2
Starting and target resolutions used for evaluating TsHARP

Starting resolution (m)	4800		960				120	60
Target resolution (m)	960	240	240	120	60	30	60	30	30
Landsat data	X	X	X	X	X				
Aircraft data			X	X	X	X	X	X	X

X denotes the resolution pairs to which TsHARP was applied.

where the first term of the right-hand side is evaluated using the regression function, f, determined at the coarse scale (Eq. (1)), and the second residual term from Eq. (2) is constant over the coarse pixel area.

In developing the scene-dependent regression function (Eq. (1)), coarse pixels containing water bodies must be excluded based on a land cover classification or an NDVI threshold. Water bodies tend to have both low temperature and low NDVI and do not conform to the inverse T_R−NDVI trend defined by vegetated land pixels. Similarly, it is useful to apply a selection criteria screening out coarse pixels with high sub-pixel variability in NDVI. These pixels tend to be outliers in T_R−NDVI distributions, representing strongly heterogeneous patches in the landscape. Kustas et al. (2003) recommend stratifying pixels in bins of NDVI, and selecting only the pixels with the lowest 25% coefficient of variation in NDVI to form the regression equations. This conservative measure will also tend to screen out pixels with unresolved water bodies.

2.2. Alternative sharpening basis functions

Kustas et al. (2003) and Anderson et al. (2004a) applied this sharpening technique to scenes over range and agricultural land in Oklahoma using a 2nd-degree polynomial regression between T_R and NDVI:

$$f(\text{NDVI}) = a_0 + a_1\text{NDVI} + a_2\text{NDVI}^2 \qquad (4)$$

with a_0, a_1, and a_2 being the scene-specific parameters derived from the regression analysis. In those studies, a polynomial function was selected over a simpler linear fit ($a_2 = 0$ in Eq. (4)) because it yielded a higher coefficient of determination, r^2. For general applications, however, a polynomial function may be unduly influenced by outliers at high and low NDVI values caused, for example, by unresolved water bodies or urban structures, potentially yielding poor retrievals in some cases. Furthermore, there is no physical basis for assuming the 2nd order relationship between T_R and NDVI described by Eq. (4). In this work, we seek a more robust and justifiable form for the sharpening basis function, f.

A linear function has the advantage of being less sensitive to outliers in the upper and lower tails of the NDVI distribution. Physically, however, T_R is expected to be more linearly correlated with fractional vegetation cover (f_C) than with NDVI, as soil and vegetation contribute to the composite surface radiometric temperature in proportion to the fraction of the radiometer view that is occupied by each component (Norman et al., 1995). Consequently, we also explored a transformation of NDVI into values associated with cover fraction, as a possible basis function.

Fractional vegetation cover (f_C) can be estimated from NDVI, for example using the form suggested by Choudhury et al. (1994):

$$f_C = 1 - \left(\frac{\text{NDVI}_{max} - \text{NDVI}}{\text{NDVI}_{max} - \text{NDVI}_{min}}\right)^{0.625} \qquad (5)$$

where NDVI_{min} and NDVI_{max} nominally describe the minimum and maximum NDVI values expected within the scene, and are
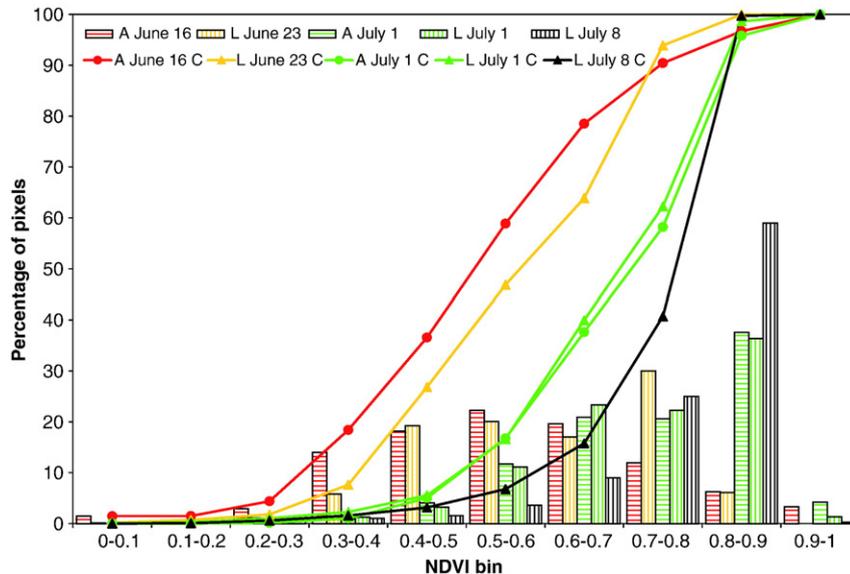
Fig. 1. The development of the agricultural crops in the Walnut Creek watershed during SMEX02 study period, demonstrated by the NDVI distributions (percentages are presented by bars, cumulative percentage by lines) for each satellite and aircraft scene (noted by L and A, respectively). Note the gradual shift of NDVI distribution from a centrally peaked in mid-June to a distribution skewed toward high NDVI in July, reflecting the corn and soybean rapid growth.

determined here by the lower and upper 3% tails of the NDVI distribution, chosen to exclude outliers. Pixels with NDVI outside these limits are reset to the limit values, and therefore there is loss of sensitivity to NDVI in 6% of the scene using this function. Combining Eq. (4) ($a_2=0$) with Eq. (5) yields

$$f(NDVI) = a_0 + a_1 \left(1 - \left(\frac{NDVI_{max} - NDVI}{NDVI_{max} - NDVI_{min}} \right)^{0.625} \right). \quad (6)$$

Since a_0, a_1, $NDVI_{max}$, and $NDVI_{min}$ are all constant for a given scene, Eq. (6) may be rewritten in the form

$$f(NDVI) = a_0' - a_1' (NDVI_{max} - NDVI)^{0.625} \quad (7)$$

in which

$$\begin{aligned} a_0' &= a_0 + a_1 \\ a_1' &= a_1 (NDVI_{max} - NDVI_{min})^{-0.625}. \end{aligned} \quad (8)$$

Aiming for a simple operational technique requiring a minimal amount of scene-specific parameterization, an addi-

tional NDVI transformation was examined, in which $NDVI_{max}$ is set to 1 and $NDVI_{min}$ is set to 0, yielding

$$f(NDVI) = a_0 - a_1 (1 - NDVI)^{0.625}. \quad (9)$$

While this simplified fraction vegetation cover basis function (referred to as f_Cs) does not yield an actual estimate of f_C, it appears to yield a linear relationship between T_R and $f(NDVI)$, without the need for defining $NDVI_{max}$ or $NDVI_{min}$, or excluding pixels in the tails of the NDVI distribution (except for those associated with water bodies).

In summary, the following five forms of $f(NDVI)$ have been evaluated in terms of their utility as sharpening basis functions:

$$f(NDVI) = \begin{cases} (a)\, a_0 + a_1 NDVI & \text{linear fit} \\ (b)\, a_0 + a_1 NDVI + a_2 NDVI^2 & \text{polynomial fit} \\ (c)\, a_0 + a_1 \left(1 - \left(\dfrac{NDVI_{max} - NDVI}{NDVI_{max} - NDVI_{min}} \right)^{0.625} \right) & f_C \text{ transformation} \\ (d)\, a_0 - a_1 (1 - NDVI)^{0.625} & \text{simplified } f_C \text{ trans.} \\ (e)\, 0 & \text{no sharpening} \end{cases} \quad (10)$$

Note that the fifth form (e) is included as a baseline case, representing no sharpening, i.e., no incorporation of high-resolution NDVI information. This case, where $\hat{T}_R = T_{R\,low}$ uniformly (referred
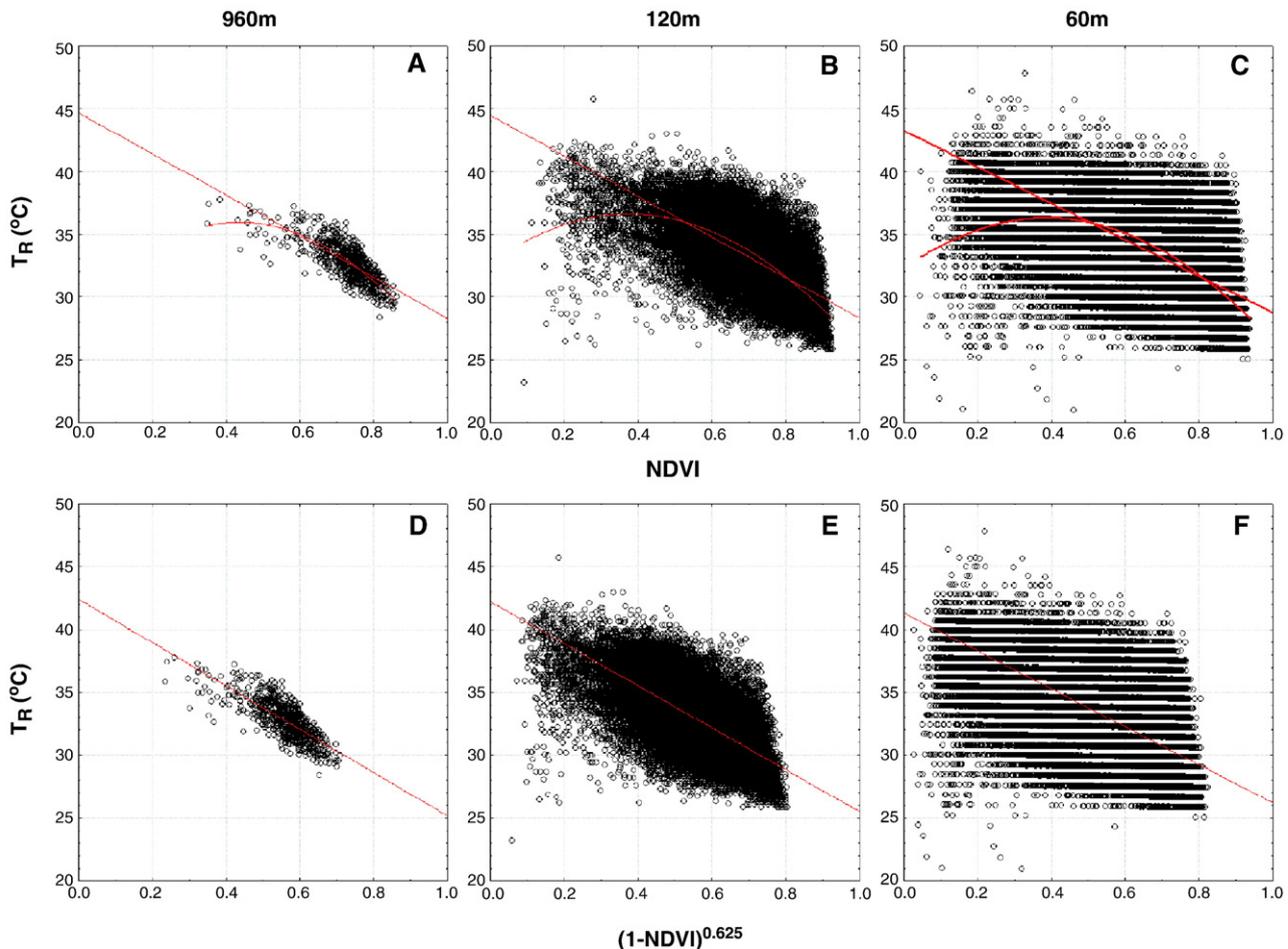
Fig. 2. The regression functions, describing the NDVI−T_R relationships, developed for forms A, B (panels A–C) and D (panels D–F) of Eq. (10) at the 960 m, 120 m, and 60 m scales for the Landsat scene acquired on July 1, 2002.

to as uniT_R), served as a reference against which the improvement due to sharpening using the first four forms (a–d) was assessed.

2.3. Multi-platform simulations

To evaluate the utility of the thermal sharpening procedure as applied to data from various satellite platforms, Landsat and high-resolution aircraft imagery were aggregated to simulate coarser-scale datasets representative of the GOES/AVHRR, MODIS, and Landsat systems. Landsat 7 data at 60-m native resolution were aggregated to source resolutions of 4800 and 960 m (simulating the native resolutions of the GOES and MODIS/AVHRR thermal bands, respectively), and these coarse data were then sharpened to target resolutions of 240, 120 and 60 m. Aircraft imagery at 6-m resolution was aggregated to source resolutions of 960, 120 and 60 m (simulating MODIS, Landsat 5 and Landsat 7, respectively), and sharpened to target resolutions of 240, 120, 60, and 30 m. The different combinations of initial and target resolutions examined in this study are listed in Table 2.

3. Data

Data used in this study were collected during the Soil Moisture–Atmosphere Coupling Experiment (SMACEX), con-ducted in an upper Midwest corn and soybean production region in central Iowa during the growing season of 2002. The study focused on the Walnut Creek watershed south of Ames, IA, which is representative of the Des Moines lobe occupying approximately one third of the Iowa state area. A detailed description of the site and the SMACEX experiment is provided by Kustas et al. (2005).

Both airborne and spaceborne remote sensing imagery were acquired over the SMACEX study area during periods of peak corn and soybean growth, in late June to early July. Airborne images were collected on June 16 and July 1, while Landsat data were available on June 23 (Landsat 5) and July 1 and 8 (Landsat 7). The combination of these two sources provided nearly weekly coverage through the peak growing season, with an overlap on July 1 that was used for comparing the two datasets.

3.1. Aircraft data

Multispectral imagery was obtained during SMACEX from the Utah State University (USU) twin engine Seneca Piper II aircraft. Airborne images in the optical range were acquired by Nikon digital cameras with Thematic Mapper filters (0.545–0.560, 0.665–0.680 and 0.795–0.809 μm) at 1.5 m resolution.
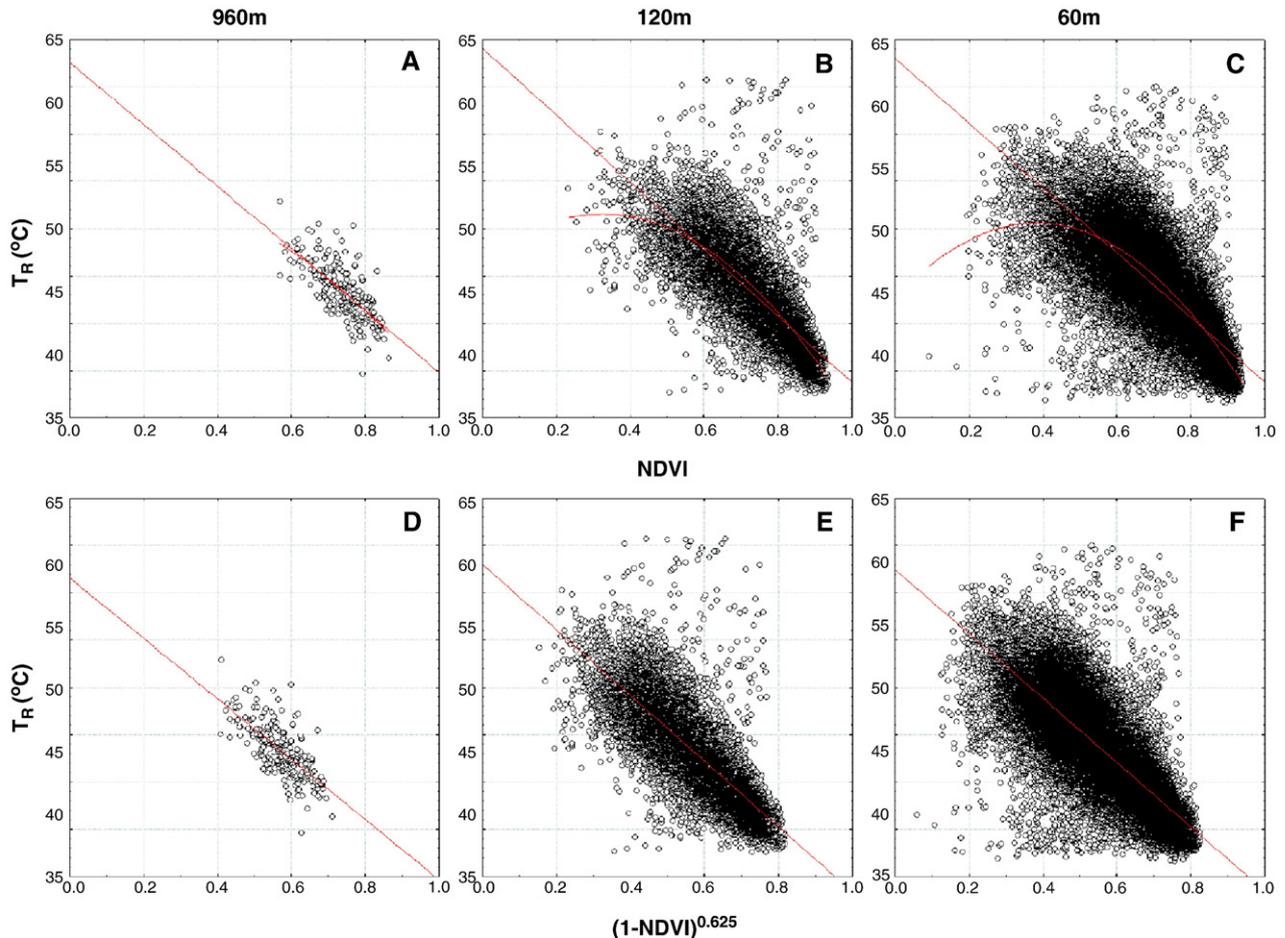
Fig. 3. The regression functions, describing the NDVI − T_R relationships, developed for forms A, B (panels A–C) and D (panels D–F) of Eq. (10) at the 960 m, 120 m, and 60 m scales for the aircraft scene acquired on July 1, 2002.

Individual scenes were first geometrically and atmospherically corrected, geo-registered, and then mosaicked along the flight lines to generate a large calibrated image covering an area of ~210 km² (30 km east–west; 7 km north–south, centered on the Walnut Creek watershed). Thermal-infrared (10.5–12.5 μm) imagery was acquired at 6 m resolution using an Inframetric 760 scanner. A line-shift correction was applied to remove effects of aircraft movement, followed by a geo-rectification of individual images to a base map. The individual images were then mosaicked and "at sensor" brightness temperature was atmospherically corrected using MODTRAN (Berk et al., 1998) to

obtain surface brightness temperature. The latter was corrected for surface emissivity according to the procedure suggested by Hipps (1989). The image strips were acquired from north to south during ~1 h flights between 11:15 AM and 12:15 PM on both dates. A detailed description of the aircraft data acquisition and analysis is provided by Chávez et al. (2005).

3.2. Satellite data

Spaceborne multispectral imagery from the Landsat TM (5) and ETM+ (7) provided vis/NIR data at 30 m resolution and
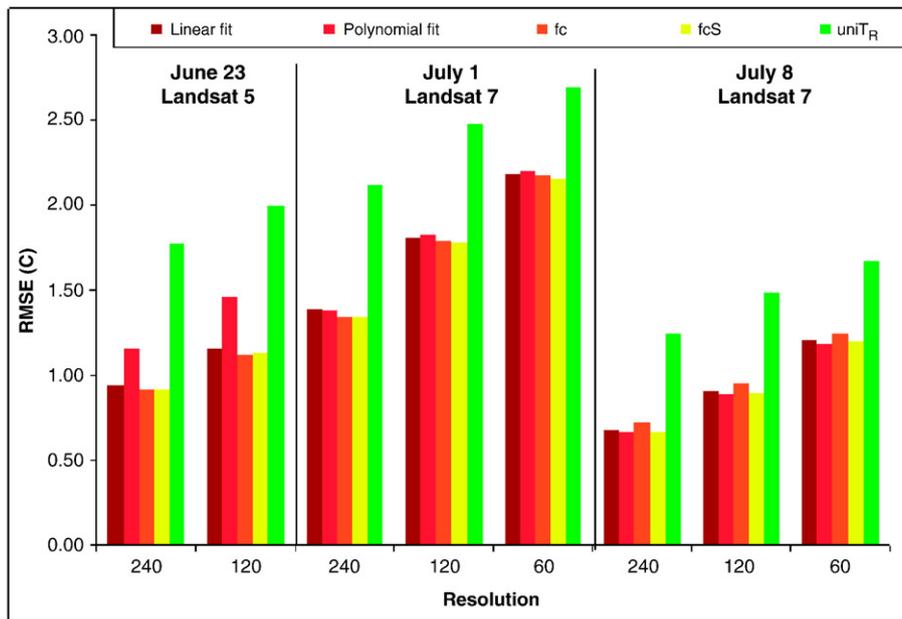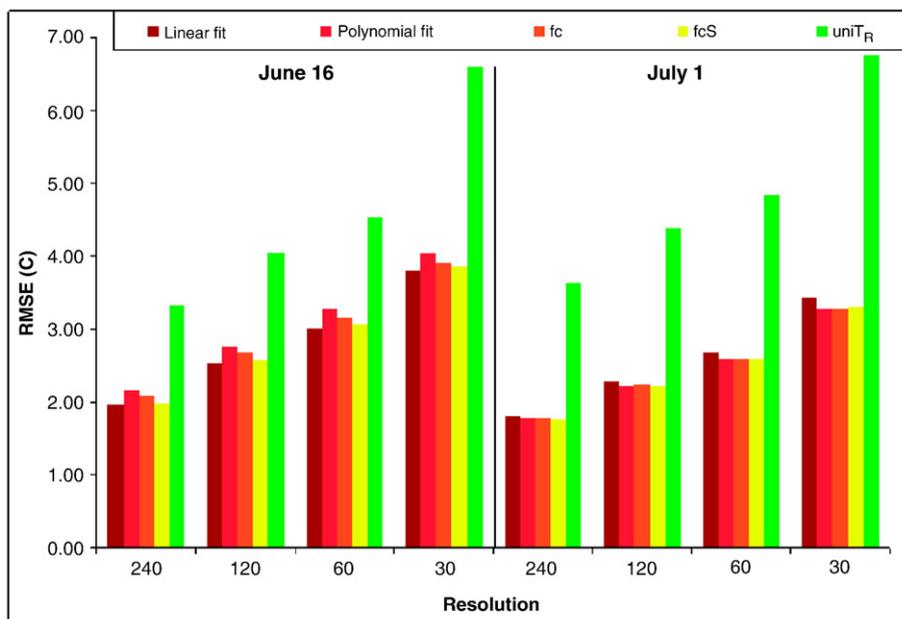
Fig. 4. Root Mean Square Errors (RMSE) associated with applying TsHARP to simulated native resolution of 960 m using the five forms of Eq. (10) at each date to obtain thermal maps at the different target resolutions.

TIR data at 120 and 60 m for Landsat 5 and 7, respectively. The study site falls in an area overlapped by two adjacent Landsat scenes, thus resulting in near weekly coverage. The processing and analysis required for obtaining T_R from the single thermal band on Landsat are fully detailed by Li et al. (2004). In short, the at-sensor brightness temperatures were atmospherically corrected using MODTRAN (Berk et al., 1998), and then corrected for emissivity using a fractional cover mixture model (Sobrino et al., 2001) assuming a soil emissivity based upon laboratory analysis of soil samples collected in the watershed (an average value of 0.978 for the Landsat thermal band) and a vegetation emissivity of 0.985 (Sobrino et al., 2001).

Watershed-scale evaluations of the sharpening algorithm were performed over a subset (~10 km north–south by 30 km east–west, centered at 41.96°N, 93.6°W) of the full Landsat scene, encompassing the Walnut Creek watershed. For regional-scale analyses, the full Landsat scenes were used.

3.3. Image aggregation procedure

For this study, low resolution images used in the sharpening procedure were derived by aggregating high-resolution T_R and NDVI data to the coarser source scales listed in Table 2. The T_R reference images ($T_{R\ ref}$) at the target resolutions, used for validation of the retrieval accuracy, were also generated through aggregation. NDVI maps were aggregated by simple areal average. Note that many studies have shown that NDVI is relatively scale-invariant (e.g., Anderson et al., 2004b; De Cola, 1997; Friedl et al., 1995; Hall et al., 1992). In aggregating the thermal images, the original (finer resolution) T_R data were first converted to radiance values using the Stephan–Boltzmann law ($R = \varepsilon \sigma T_R^4$ in which R is the radiance, σ is Stephan–Boltzmann constant, and ε is the emissivity). The emissivity was set to a constant value of 0.98 based on Li et al. (2004) who found it to be fairly constant over the entire scene for the duration of SMACEX. The derived radiances were then aggregated, and converted back to temperature.

3.4. Evaluation of retrieval accuracy

The level of agreement between the reference ($T_{R\ ref}$) and sharpened ($\hat{T}_{R\ high}$) temperature fields was assessed by means of the Root-Mean-Square-Error (RMSE) and Mean-Absolute-Error (MAE) computed according to (Willmott, 1982):

$$\text{RMSE} = \left[n^{-1} \sum_{i=1}^{n} \left(\hat{T}_{R\ high} - T_{R\ ref} \right)^2 \right]^{1/2}$$

$$\text{MAE} = \left[n^{-1} \sum_{i=1}^{n} \left| \hat{T}_{R\ high} - T_{R\ ref} \right| \right] \quad (11)$$

Recently, Willmott and Matsuura (2005) compared these two measures of error and concluded that MAE is the preferable

metric, because RMSE is sensitive not only to error size, but also to error variability and number of data points. In this study, RMSE was found to be well correlated with MAE, with $\text{RMSE} \approx 1.5\text{MAE}$ ($r^2 = 0.91$). Therefore, error is reported as RMSE to maintain comparability with previous studies, although nominal MAE is easily estimated from the RMSE and will have $\approx 45\%$ lower values.

4. Results and discussion

4.1. Sharpening at the watershed scale

4.1.1. Evaluation of NDVI−T_R regression basis functions

The development of the agricultural crops in the Walnut Creek watershed during the SMEX02 study period is demonstrated in Fig. 1, showing the distribution in NDVI for each satellite and aircraft scene collected during the experiment. As the growing season progressed, the NDVI values over the watershed evolved from a centrally peaked distribution to a distribution skewed toward high NDVI. These changes reflect the corn and soybean rapid growth that occurred during SMEX02, with canopy heights starting at nominally 0.15 and 0.75 m, for soybean and corn, respectively, reaching heights of ~0.5 and 2 m by mid-July (Anderson et al., 2004b). Both Landsat and aircraft imagery yielded similar distributions in NDVI on July 1, the day when both datasets were available.

To ascertain the optimal regression function for this landscape, the five functional forms in Eq. (10a−e) were applied to each source−target resolution pair in Table 2, and the characteristics of retrieval errors were examined in comparison

Table 3
Comparison of regression function performance in sharpening thermal Landsat and aircraft images from 960 m resolution to various target resolutions

Date	Target resolution (m)	$\bar{T}_R$ A (°C)	σT_R (°C)	RMSE (°C)				
				Linear	Polynomial	f_C	f_Cs	uniT_R
Landsat scenes								
6/23/2002	240	34.47	2.36	0.94	1.15	0.91	0.92	1.77
6/23/2002	120	34.47	2.53	1.16	1.46	1.12	1.13	2.00
7/1/2002	240	34.59	2.66	1.39	1.38	1.35	1.35	2.12
7/1/2002	120	34.58	2.95	1.81	1.82	1.79	1.78	2.48
7/1/2002	60	34.54	3.13	2.18	2.20	2.17	2.16	2.69
7/8/2002	240	31.96	2.15	0.67	0.67	0.72	0.67	1.25
7/8/2002	120	31.95	2.29	0.91	0.89	0.95	0.90	1.49
7/8/2002	60	32.07	2.42	1.21	1.18	1.25	1.20	1.68
Aircraft scenes								
6/16/2002	240	34.21	5.21	1.97	2.17	2.10	2.00	3.33
6/16/2002	120	34.18	5.70	2.53	2.76	2.68	2.57	4.05
6/16/2002	60	34.16	6.05	3.00	3.27	3.16	3.07	4.53
6/16/2002	30	34.03	6.45	3.81	4.02	3.90	3.87	6.58
7/1/2002	240	39.51	4.43	1.81	1.78	1.78	1.77	3.63
7/1/2002	120	39.49	5.07	2.28	2.23	2.24	2.22	4.39
7/1/2002	60	39.47	5.47	2.69	2.59	2.60	2.60	4.85
7/1/2002	30	39.34	5.88	3.42	3.28	3.27	3.29	6.75

Statistics show average temperature ($\bar{T}_R$) and standard deviation (σT_R) in the 'observed' fields, and RMSE between the 'observed' and sharpened images as computed using the regression functions in Eq. (10a−d). The fifth method (uniT_R; Eq. (10e)) represents the case where no sharpening is applied, and serves as a reference to which TsHARP performance is compared.

with reference images at the target resolutions. Optimally, a function that minimizes overall retrieval errors, and yields errors with minimal dependence on local temperature/vegetation cover conditions is desired.

Figs. 2 and 3 demonstrate regression functions developed for the 1 July Landsat and aircraft scenes, respectively, using forms a, b and d of Eq. (10) at the 960, 120, and 60 m scales. Note that regression results using Eqs. (10c) and (10d) were very similar, and therefore not shown. The fitting parameters varied only slightly with resolution, reinforcing the assumption that the NDVI−T_R relationship is relatively scale independent. The coefficient of determination (r^2) ranged between 0.4 and 0.7 (all statistically significant), being generally similar for all methods. The 2nd degree polynomial (Eq. (10b)) is potentially most

problematic since it can be highly sensitive to outliers, yielding a non-physical convex shaped curve in temperature at low NDVI. Note that the aggregation process served to narrow the dynamic range in both T_R and NDVI. These plots show all pixels within the scene. In practice, only a subset of points showing low sub-pixel heterogeneity in NDVI would be used to develop the regression equations, serving to filter out many of the outliers apparent in these distributions. For this landscape, however, such a filter did not significantly impact the derived fits.

Errors associated with sharpening Landsat and aircraft scenes from 960 m to a variety of target resolutions using the 5 different regression functions are presented in Fig. 4 and Table 3. Fig. 4 demonstrates that the four sharpening
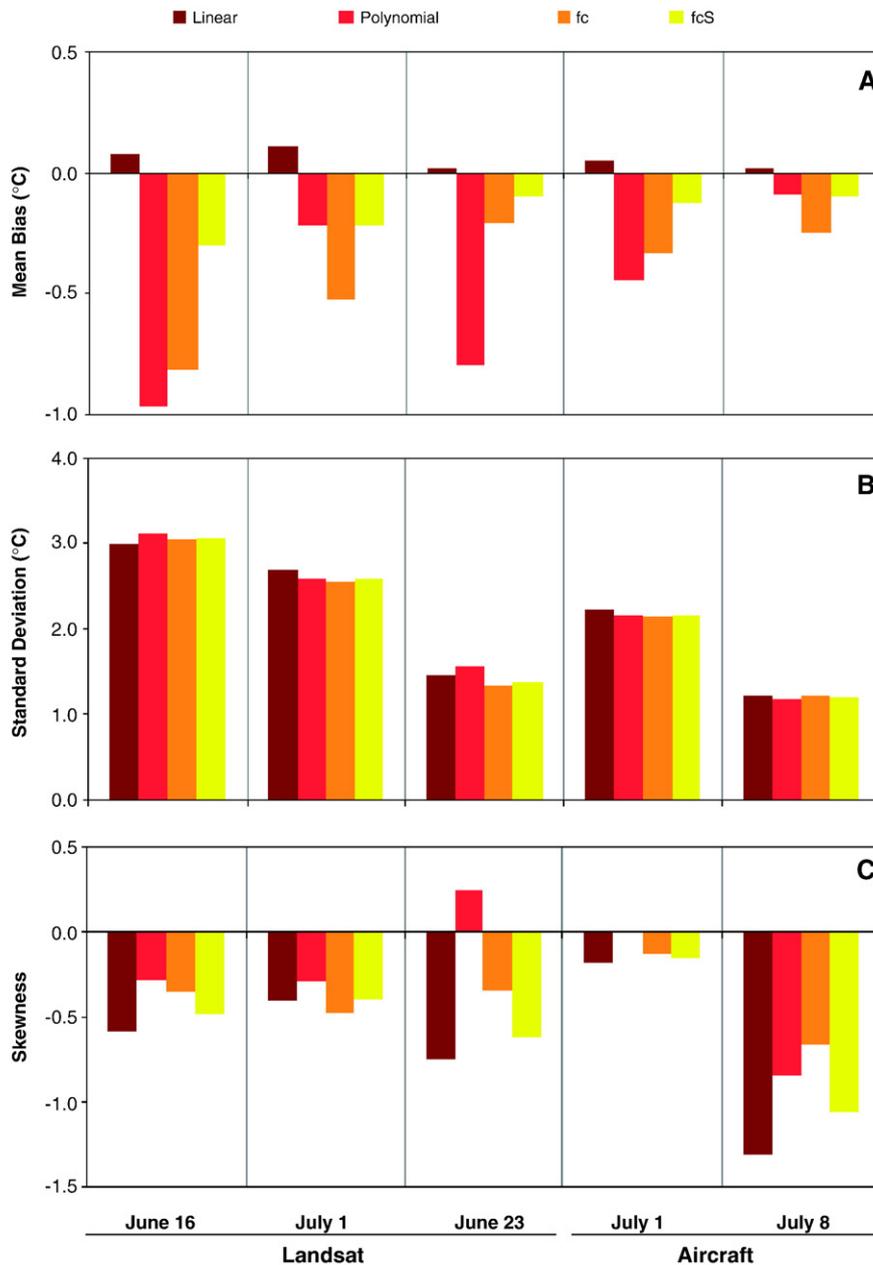
Fig. 5. Statistic characteristics of the distributions of the errors yielded by TsHARP applied to simulated native resolution of 960 m to obtain 120 m thermal pixels.

functions perform similarly and significantly better than uniT_R (Eq. (10e) — no sharpening applied). Generally, errors increased at finer target resolutions. The RMSE for sharpening to 240 m was less than 1.5 and 2.2 °C for the Landsat and aircraft imagery, respectively. For target resolutions of 60 m, RMSE for Landsat and aircraft did not exceed 2.2 and 3.3 °C, respectively. This decrease in accuracy with finer target resolutions is expected, considering the increasing degree of sub-pixel variability that must be retrieved.

Error characteristics also differ between scenes. Generally, the aircraft data yielded higher RMSEs than did the Landsat images. However, the aircraft scenes also showed larger intrinsic variability in observed temperature (reflected by the standard deviation in Table 3). This may be due to the fact that, whereas Landsat images represent snapshots in time, the individual aircraft images creating the mosaic were acquired over a 1 h interval, causing a temporally-induced temperature gradient over the scene from north to south. Therefore, the errors associated with the two datasets are not directly comparable. It does appear, however, that the retrieval errors are related to overall scene heterogeneity, expressed in terms of the standard deviation in T_R(σT_R).

Of the four sharpening functionals (Eq. (10a–d)), the polynomial fit (Eq. (10b)) had the largest RMSE value for three out of the five scenes studied here, particularly earlier in the growing season. Moreover, the polynomial functional fit
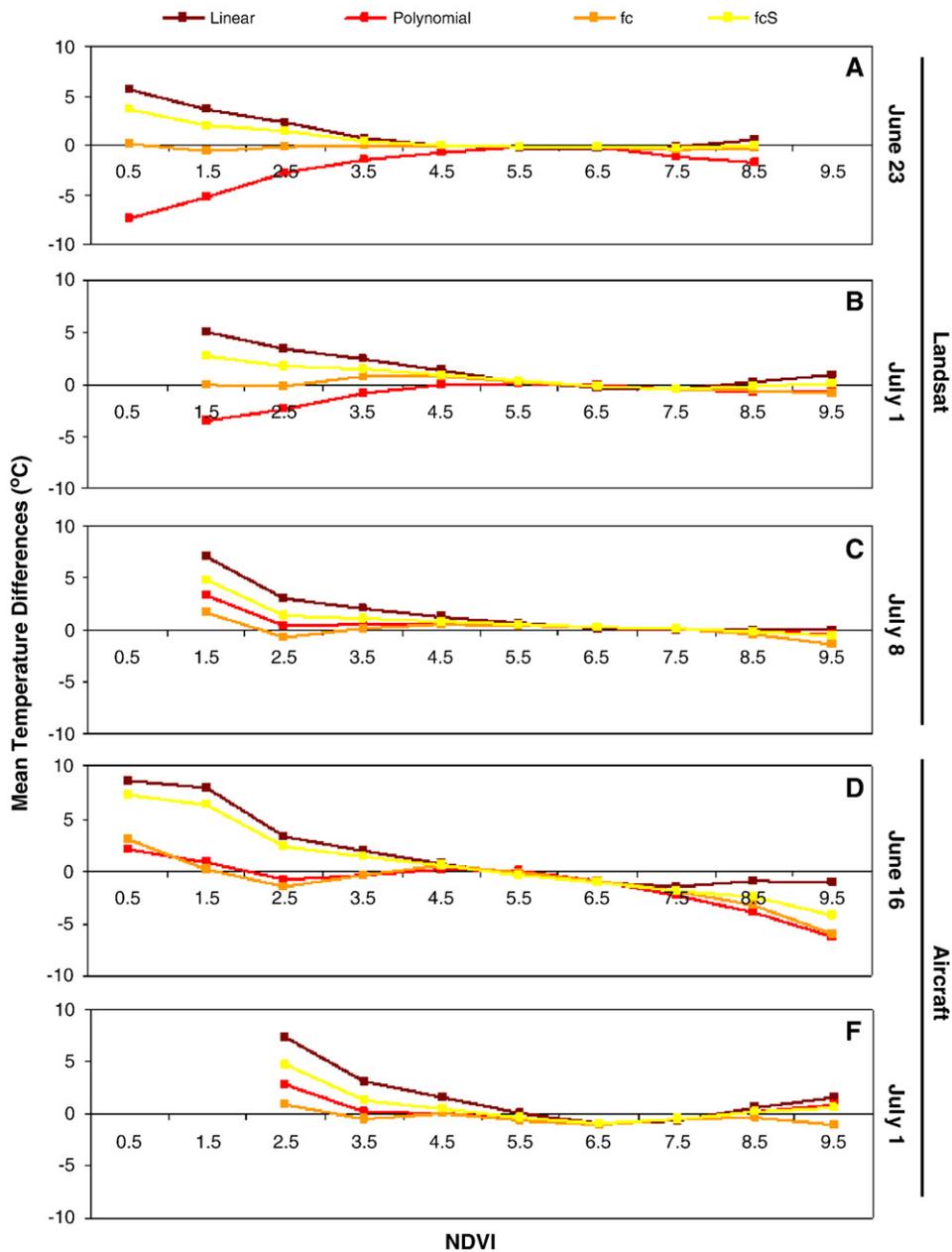
Fig. 6. The dependence of retrieval error on vegetation cover amount, demonstrated by plotting the mean temperature bias, for sharpening from 960 to 120 m, at each NDVI-bin for the five scenes.

Fig. 7. The potential utilization of TsHARP$_{fcS}$ on 'simulated' MODIS imagery demonstrated using the Landsat 7 data acquired on July 1. The upper panel (A) is the simulated MODIS temperature field (960 m pixel size); the next (B) is the 'observed' temperature field (240 m pixel size), and the last (C) is the temperature field created by TsHARP starting at 960 m and sharpening to 240 m.

linear function, however, generally showed the most strongly skewed error distributions, relating to the non-linear relationship between NDVI and T_R.

The dependence of retrieval error on vegetation cover amount is demonstrated in Fig. 6, showing the mean bias for each regression method as a function of NDVI in bins of width 0.1. In general, all methods yielded small biases at the mid NDVI range, increasing at both high and low NDVI. The linear method yielded the largest errors at low NDVI for all dates, while the fraction vegetation cover based methods generated errors that are most uniformly small over the full range in NDVI, reflecting the fact that T_R is expected to be linear in vegetation cover fraction. It should be noted that only a small number of pixels fall within the very low and very high NDVI bins, meaning that statistics in these bins have less significance.

These statistical analyses suggest that the polynomial fit between T_R and NDVI that was originally used by Kustas et al. (2003) did not perform as well as the other sharpening formulations for this agricultural landscape, being overly sensitive to outliers and producing a non-physical turnover at low NDVI. Retrieval errors from a linear regression showed sensitivity to fractional vegetation cover, an undesirable characteristic. The transformation to fraction vegetation cover improved the linear correlation with surface temperature. Of the two f_C-based methods tested, the simplified method performed similarly to the Choudhury et al. (1994) relationship and has the benefit of not requiring specification of NDVI$_{max}$ and NDVI$_{min}$ scaling parameters, and is applicable over the entire range in NDVI present in the scene. The simplified fraction vegetation cover method (hereafter TsHARP$_{fcS}$ method) is therefore used for the rest of the analyses in this paper.

Note that these conclusions may be scene dependent and further examination under different climate and land cover conditions is required to decisively determine which of these relationships universally yields better results.

4.1.2. Applications to simulated MODIS and Landsat imagery

Applying TsHARP to MODIS imagery can potentially increase the spatial resolution of the thermal band data by a

generally yielded the highest mean bias errors $[n^{-1}\sum(\hat{T}_R-T_R)]$ while the linear (Eq. (10a)) and the simplified fraction-cover (Eq. (10d)) methods generated the smallest biases (Fig. 5). The

Table 4
Sharpening error statistics associated with various combinations of simulated spaceborne systems

Date	Observing instrument	Simulated thermal	Simulated NDVI	Native resolution (m)	Target resolution (m)	RMSE (°C)	
						TsHARPf_CS	uniT_R
6/16/2002	Aircraft	Landsat 7	Landsat 7	60	30	1.80	2.02
7/1/2002	Aircraft	Landsat 7	Landsat 7	60	30	1.80	2.01
6/16/2002	Aircraft	Landsat 5	Landsat 5	120	30	2.39	3.28
7/1/2002	Aircraft	Landsat 5	Landsat 5	120	30	2.35	3.31
6/23/2002	Landsat 5	MODIS	MODIS	960	240	0.92	1.77
7/1/2002	Landsat 7	MODIS	MODIS	960	240	1.35	2.12
7/8/2002	Landsat 7	MODIS	MODIS	960	240	0.67	1.25
6/16/2002	Aircraft	MODIS	MODIS	960	240	2.00	3.33
7/1/2002	Aircraft	MODIS	MODIS	960	240	1.77	3.63
7/1/2002	Landsat 7	GOES	MODIS/AVHRR	4800	960	0.98	1.42
7/1/2002	Landsat 7	GOES	MODIS	4800	240	1.97	2.34

RMSE values were computed by comparing the 'observed' temperature fields with TsHARP$_{fcS}$ and uniT_R fields.
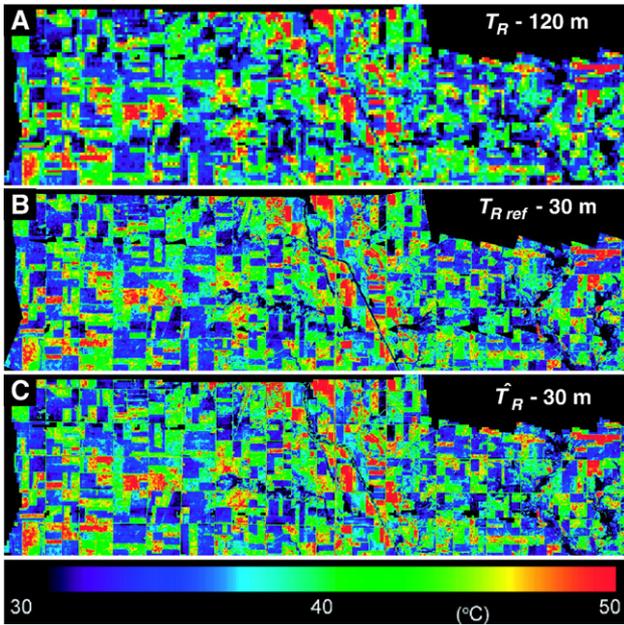
Fig. 8. Application of TsHARP$_{fcS}$ with 'simulated' Landsat 5 imagery derived from aircraft data acquired on July 1. The upper panel (A) is the simulated Landsat TM temperature field (120 m pixel size); the next (B) is the 'observed' temperature field (30 m pixel size), and the last (C) is the temperature field created by TsHARP starting at 120 m and sharpening to 30 m.

factor of 4, from 1 km to 250 m. This spatial resolution may in some cases be fine enough to monitor individual agricultural fields. As the temporal resolution of MODIS Terra and Aqua

provide 1–2 images per day, depending on local cloud-cover conditions, a day-to-day monitoring of individual fields may become possible using sharpened MODIS thermal images. Fig. 7 demonstrates the potential utility of TsHARP$_{fcS}$ as applied to MODIS resolution data, simulated with the Landsat 7 scene acquired on July 1. The upper Fig. 7A approximates the MODIS temperature field (960 m pixel size); the middle panel (Fig. 7b) shows the original 60 m Landsat 7 temperature field aggregated to 240 m (the "observed" scene, $T_{R\,ref}$), and the bottom panel (Fig. 7c) shows the 960 m field sharpened to 240 m resolution. Moisture and vegetation conditions were strongly variable across the watershed on this date, and the RMSE of the sharpened image was 1.35 °C (Table 4), slightly lower than that found by Kustas et al. (2003). In comparison, the uniT_R method yielded an RMSE of 2.12 °C. Although there are detectable differences between the observed and sharpened fields, in general the sharpening algorithm recovered much of the detail apparent in Fig. 7B. Clearly the sharpened image (Fig. 7C) provides significantly better visual information content than does the original MODIS-resolution image in terms of interpreting spatial variability in ET at this watershed scale. Note, however, that the 250 m resolution of the MODIS shortwave bands is not adequate to clearly separate individual fields.

A similar experiment was conducted to examine the utility of TsHARP$_{fcS}$ as applied to Landsat 5, sharpening 120-m thermal fields to the 30-m resolution of the vis/NIR bands. The fields in Fig. 8 were created from aircraft imagery on July 1, showing temperature fields aggregated directly to 120 m and 30 m, and the
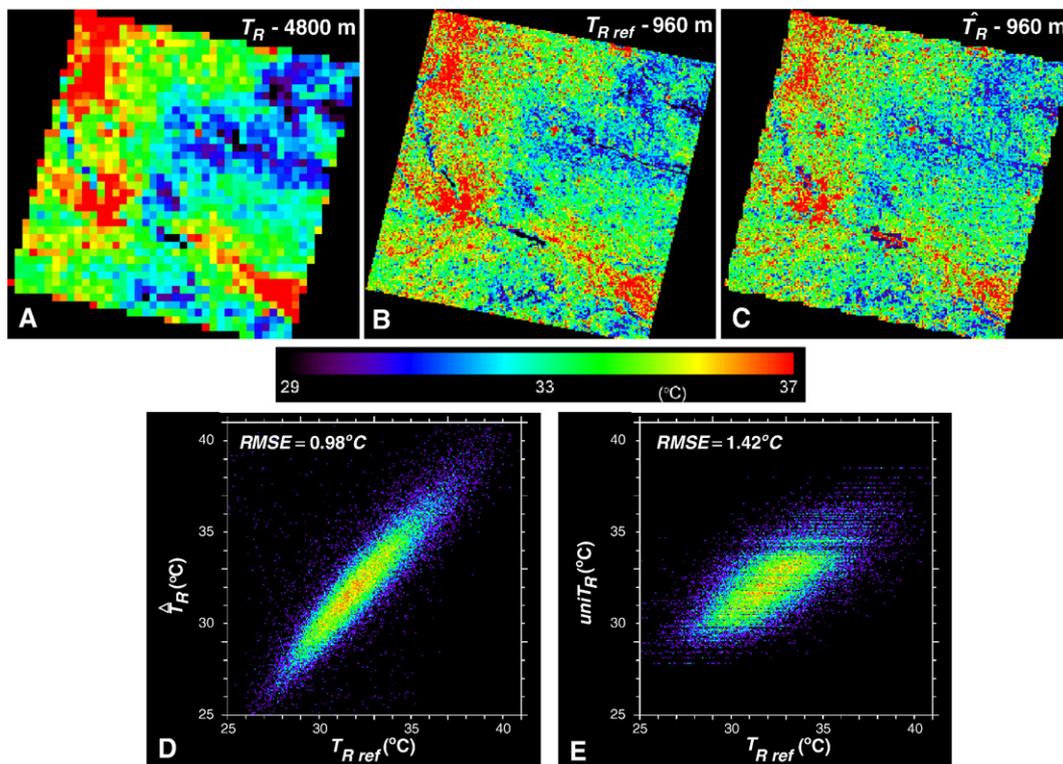
Fig. 9. The utility of TsHARP$_{fcS}$ applied to 'simulated' GOES thermal imagery with 1-km NDVI products of MODIS derived from the Landsat 7 data acquired on July 1. The upper-left panel (A) is the simulated GOES temperature field (4800 m pixel size); the next (B) is the 'observed' temperature field (960 m pixel size), and the last (C) is the temperature field created by TsHARP$_{fcS}$ starting at 4800 m and sharpening to 960 m. Panels (D) and (E) show the scatter plot of the modeled vs. the 'observed' temperatures, for TsHARP and uniTrad, respectively.

results of sharpening from 120 to 30 m. Qualitatively, good similarity was observed between the two 30 m resolution images, with TsHARP$_{fcS}$ providing considerably more detail than is present in the 120 m image. Field edges are sharper, and even subfield variability is reproduced with good fidelity. The RMSE of the sharpened image was 2.35 °C, while uniT_R yielded 3.31 °C (Table 4). In contrast, Kustas et al. (2003), using a polynomial regression function, found little improvement over uniT_R when sharpening images at native resolutions finer than 200 m. This disparity could be due to the improved performance of the f_C-based regression function, but may also be related to spatial characteristics of landscape features for the two study areas.

A third test (not shown), simulating sharpening of Landsat 7 data from 60 to 30 m resolution resulted in smaller errors (RMSE=1.80 °C) in comparison with the 120 to 30 m test case (RMSE=2.37±0.02 °C). This is expected due to the smaller sharpening factor (×2 vs. ×4). However, uniT_R provided similar RMSE at these resolutions. While the quantitative impact of sharpening is marginal in this case, Anderson et al. (2004a) demonstrate that sharpened Landsat 7 thermal images have significant utility in surface energy balance modeling in terms of improved visual information content and model convergence rate.

4.2. Sharpening at the regional scale — applications to simulated GOES data

By combining geostationary thermal data from the GOES satellite (nominal pixel size over the study area of ∼ 5 km) with moderate resolution NDVI data from MODIS, thermal maps at 250 m spatial resolution and 15 min temporal resolution can potentially be achieved. Figs. 9 and 10 present examples of utilizing TsHARP$_{fcS}$ to sharpen simulated GOES thermal data (Landsat 7 data aggregated to 4800 m resolution) over Iowa to 960 and 240 m, respectively, for July 1.

In this example, the RMSE of sharpened images were 0.98 and 1.97 °C for target resolutions 960 and 240 m, respectively, while uniT_R yielded errors of 1.42 and 2.34 °C (Table 4). While the differences in RMSE between TsHARP$_{fcS}$ and uniT_R were relatively small in comparison with the earlier sharpening exercises, the density scatter plots of sharpened vs. observed temperature in Figs. 9 and 10 show clear benefit of TsHARP$_{fcS}$ over uniT_R.

Due to the assumption that vegetation cover amount is the primary driver of temperature variations, the sharpening algorithm described here cannot be applied to water bodies. Coarse pixels within resolved water bodies therefore should be left unsharpened. This is not greatly problematic, as contiguous water pixels tend to have relatively uniform surface temperature. However, as evident in Figs. 9 and 10, sub-pixels in close vicinity to unresolved water bodies show large boxy anomalies, resulting from the reintroduction of the coarse-scale residual ΔT_R. There may be instances where these waterside sub-pixels are of interest, for example in studying evapotranspiration from riparian corridors. Future papers will detail simple techniques for addressing unresolved water bodies. Note also that urban areas should be carefully considered, again because their NDVI−T_R
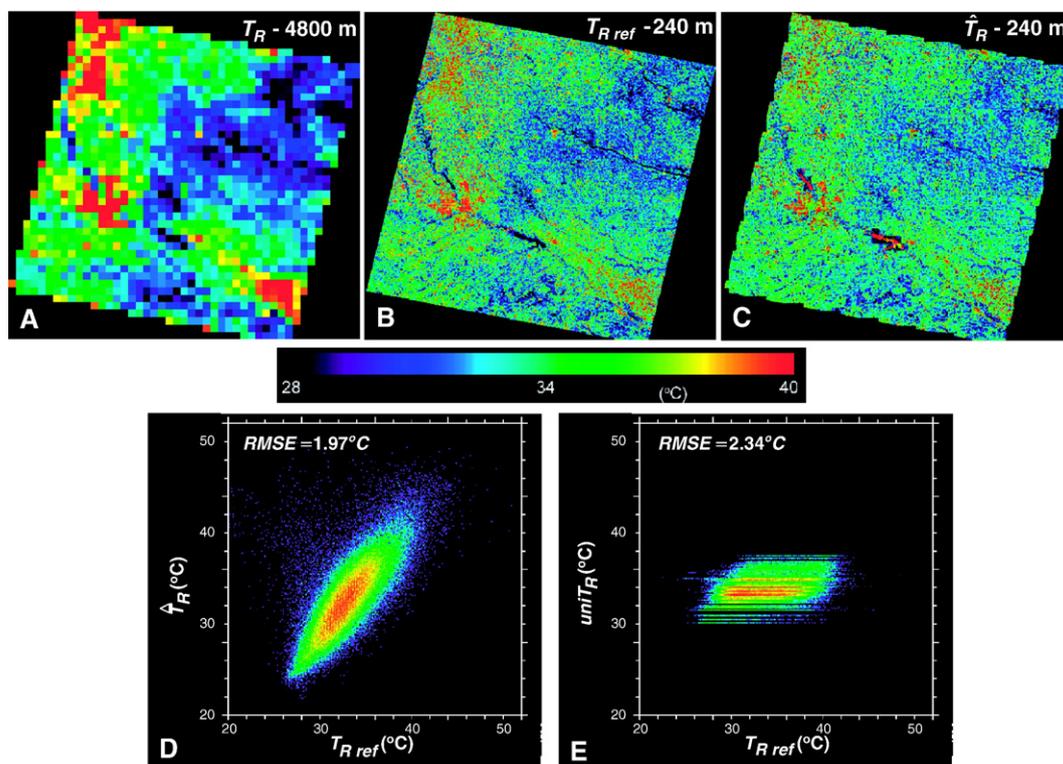
Fig. 10. The result of applying TsHARP$_{fcS}$ to 'simulated' GOES thermal imagery in conjunction with 250-m NDVI product from MODIS demonstrated using the Landsat 7 data acquired on July 1. The upper-left panel (A) is the simulated GOES temperature field (4800 m pixel size); the next (B) is the 'observed' temperature field (240 m pixel size), and the last (C) is the temperature field created by TsHARP$_{fcS}$ starting at 4800 m and sharpening to 240 m. Panels (D) and (E) show the scatter plot of the modeled vs. the 'observed' temperatures, for TsHARP and uniT_R, respectively.

relationship may significantly differ from adjacent natural or agricultural areas, and hence not be applicable. It may be that pre-stratification of the image by land-cover class may provide better results over strongly heterogeneous landscapes (Friedl & Davis, 1994).

5. Summary and conclusions

The need for high-spatial/high-temporal resolution thermal data has led to a further examination of the sharpening procedure first developed by Kustas et al. (2003). Four different methods for determining the $NDVI - T_R$ relationship were examined over the agricultural fields of Walnut Creek watershed in Iowa during the growing season, using Landsat and aircraft data, and the potential utilization of TsHARP at regional scales with GOES and MODIS imagery was examined.

TsHARP was found to be fairly insensitive to the choice of $NDVI - T_R$ relationship under the conditions of this study. However, a 2nd-order polynomial regression tended to yield the largest errors. The simplified fraction vegetation cover method (TsHARP$_{fcS}$) performed slightly better that the other methods, likely due in part to the linearity between f_C and T_R. It is recommended to use TsHARP$_{fcS}$ for sharpening surface temperature imagery, although further investigation, under various vegetation and climatic conditions, is required to support this conclusion. Applying TsHARP to simulated MODIS 1 km thermal maps to sharpen down to ~250 m (MODIS VI resolution) yielded RMSE of 0.67–1.35 °C, compared to the 'observed' temperatures, while RMSE results for uniT_R (no sharpening) were 1.25–2.12 °C. TsHARP decreased the errors by ~0.8 °C for all three Landsat scenes and by 1.3–1.8 °C for the aircraft scenes. By sharpening MODIS, 250 m thermal maps could be available daily, and more frequently if applied to both Terra and Aqua overpasses. The discrepancies in surface energy budget computations, caused by the sharpening, will be the subject of a future investigation.

Sharpening simulated Landsat thermal maps (60 and 120 m) to Landsat VI resolution (30 m) was found to yield slightly larger errors compared to those found for the MODIS resolutions. RMSE for TsHARP$_{fcS}$ and uniT_R were 2.4 vs. 3.3 °C, respectively, for 120 m starting resolution, and 1.8 vs. 2.0 °C for starting resolution of 60 m. Contrary to the findings of Kustas et al. (2003) these results show significant improvement using TsHARP$_{fcS}$ versus uniT_R at 120 m resolution, although no real advantage was found for sharpening from 60 to 30 m. However, there still may be significant utility sharpening to 30 m for surface energy balance modeling in terms of improved visual information content and model convergence rate (Anderson et al., 2004a).

Lastly, sharpening simulated GOES thermal maps (derived from a complete Landsat scene) from 5 km to 1 km and 250 m yielded RMSEs of 0.98 and 1.97, respectively. These results prove a great potential in significantly enhancing the thermal information available over this agricultural region. Notwithstanding the uncertainties, 15 min thermal maps at field scale resolution may be possible, permitting routine monitoring of evaporative fluxes and stress conditions for the two important crops (corn and soybean) produced in this region. In fact the corn and soybean production in this study area is indicative of a much larger agricultural region of the United States, namely, the upper Midwest corn–soybean region, which encompasses over 60 million ha and represents approximately 60% of all U.S. cultivated cropland. Further research is planned to examine the general applicability of the TsHARP algorithms for different climatic regions and land-use characteristics.

Although TsHARP may provide T_R estimates at resolutions more useful for environmental monitoring of different land cover conditions, this thermal image sharpening technique is unable to recreate land-surface temperature variations due to unresolved soil-moisture anomalies. This technique therefore does not preclude the need for high-resolution (<100 m) satellite-based thermal imaging platforms for accurately discriminating land cover having significantly different vegetation/crop cover and moisture conditions. In many instances this will include environmentally sensitive areas such as riparian corridors and wetland–dryland interfaces and other types of transitional areas or ecotones.

References

Anderson, M. C., Neale, C. M. U., Li, F., Norman, J. M., Kustas, W. P., Jayanthi, H., et al. (2004). Upscaling ground observations of vegetation water content, canopy height, and leaf area index during SMEX02 using aircraft and Landsat imagery. *Remote Sensing of Environment, 92*, 447−464.

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Torn, R. D., Kustas, W. P., & Basara, J. B. (2004). A multi-scale remote sensing model for disaggregating regional fluxes to micrometeorological scales. *Journal of Hydrometeorology, 5*, 343−363.

Badeck, F. W., Bondeau, A., Bottcher, K., Doktor, D., Lucht, W., Schaber, J., et al. (2004). Responses of spring phenology to climate change. *New Phytologist, 162*, 295−309.

Berk, A., Bernstein, L. S., Anderson, G. P., Acharya, P. K., Robertson, D. C., Chetwynd, J. H., et al. (1998). MODTRAN cloud and multiple scattering upgrades with application to AVIRIS. *Remote Sensing of Environment, 65*, 367−375.

Chávez, J. L., Neale, C. M. U., Hipps, L. E., Prueger, J. H., & Kustas, W. P. (2005). Comparing aircraft-based remotely sensed energy balance fluxes with Eddy covariance tower data using heat flux source area functions. *Journal of Hydrometeorology, 6*, 923−940.

Choudhury, B. J., Ahmed, N. U., Idso, S. B., Reginato, R. J., & Daughtry, C. S. T. (1994). Relations between evaporation coefficients and vegetation indices studied by model simulations. *Remote Sensing of Environment, 50*, 1−17.

De Cola, L. (1997). Multiresolution covariation among Landsat and AVHRR vegetation indices. In D. A. Quattrochi, & M. F. Goodchild (Eds.), *Scale in Remote Sensing and GIS* (pp. 73−91). CRC Lewis.

Diak, G. R., Mecikalski, J. R., Anderson, M. C., Norman, J. M., Kustas, W. P., Torn, R. D., et al. (2004). Estimating land-surface energy budgets from space: Review and current efforts at the University of Wisconsin—Madison

and USDA-ARS. *Bulletin of the American Meteorological Society, 85*, 65−78.

French, A. 2001. Scaling of surface energy fluxes using remotely sensed data, PhD thesis, University of Maryland.

Friedl, M. A., & Davis, F. W. (1994). Sources of validation in radiometric surface temperature over a Tallgrass Prairie. *Remote Sensing of Environment, 48*, 1−17.

Friedl, M. A., Davis, F. W., Michaelsen, J., & Moritz, M. A. (1995). Scaling and uncertainty in the relationship between NDVI and land surface biophysical variables: An analysis using a scene simulation model and data from FIFE. *Remote Sensing of Environment, 54*, 233−246.

Gillies, R. R., & Carlson, T. N. (1995). Thermal remote-sensing of surface soil–water content with partial vegetation cover for incorporation into climate models. *Journal of Applied Meteorology, 34*, 745−756.

Hall, F. G., Huemmrich, K. F., Goetz, S. J., Sellers, P. J., & Nickerson, J. E. (1992). Satellite remote sensing of surface energy balance: Success, failures and unresolved issues in FIFE. *Journal of Geophysical Research, 97*, 19,061−19,089.

Hipps, L. E. (1989). The infrared emissivity of soil and *Artemisia tridenata* and subsequent temperature corrections in a shrub-steppe ecosystem. *Remote Sensing of Environment, 27*, 337−342.

Karnieli, A., Bayasgalan, M., Bayarjargal, Y., Agam, N., Khundulmur, S., & Tucker, C. J. (2006). Comments on the use of the Vegetation Health Index over Mongolia. *International Journal of Remote Sensing, 27*, 2017−2024.

Kustas, W. P., Hatfield, J., & Prueger, J. H. (2005). The Soil Moisture Atmosphere Coupling Experiment (SMACEX): Background, hydrometeorological conditions and preliminary findings. *Journal of Hydrometeorology, 6*, 791−804.

Kustas, W. P., Norman, J. M., Anderson, M. C., & French, A. N. (2003). Estimating subpixel surface temperatures and energy fluxes from the vegetation index-radiometric temperature relationship. *Remote Sensing of Environment, 85*, 429−440.

Li, F., Jackson, T. J., Kustas, W. P., Schmugge, T. J., French, A. N., Cosh, M. H., & Bindlish, R. (2004). Deriving land surface temperature from Landsat 5 and 7 during SMEX02/SMACEX. *Remote Sensing of Environment, 92*, 521−534.

Loheide Ii, S. P., & Gorelick, S. M. (2005). A local-scale, high-resolution evapotranspiration mapping algorithm (ETMA) with hydroecological applications at riparian meadow restoration sites. *Remote Sensing of Environment, 98*, 182−200.

Moran, M. S., Clarke, T. R., Inoue, Y., & Vidal, A. (1994). Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. *Remote Sensing of Environment, 49*, 246−263.

Norman, J. M., Anderson, M. C., Kustas, W. P., French, A. N., Mecikalski, J. R., Torn, R. D., et al. (2003). Remote sensing of surface energy fluxes at 10[1]-m pixel resolutions. *Water Resources Research, 39*. doi:10.1029/2002WR001775.

Norman, J. M., Kustas, W. P., & Humes, K. S. (1995). A two-source approach for estimating soil and vegetation energy fluxes from observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology, 77*, 263−293.

Pieri, D., & Abrams, M. (2005). ASTER observations of thermal anomalies preceding the April 2003 eruption of Chikurachki volcano, Kurile Islands, Russia. *Remote Sensing of Environment, 99*, 84−94.

Price, J. C. (1990). Using spatial context in satellite data to infer regional scale evapotranspiration. *IEEE Transactions on Geosciences and Remote Sensing, GE-28*, 940−948.

Quattrochi, D. A. & Luvall, J. (Eds.). (2004). *Thermal remote sensing in land surface processes*: CRC.

Running, S. W., Loveland, T. R., Pierce, L. L., Nemani, R. R., & Hunt, E. R. (1995). A remote sensing based vegetation classification logic for global land cover analysis. *Remote Sensing of Environment, 51*, 39−48.

Sobrino, J. A., Raissouni, N., & Li, Z. -L. (2001). A comparative study of land surface emissivity retrieval from NOAA data. *Remote Sensing of Environment, 75*, 256−266.

Tucker, C. J., Slayback, D. A., Pinzon, J. E., Los, S. O., Myeni, R. B., & Taylor, M. G. (2001). Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999. *International Journal of Biometeorology, 45*, 184−190.

Voogt, J. A., & Oke, T. R. (2003). Thermal remote sensing of urban climates. *Remote Sensing of Environment, 86*, 370−384.

White, M. A., Thornton, P. E., Running, S. W., & Global (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Biogeochemical Cycles, 11*, 217−234.

Willmott, C. J. (1982). Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society, 63*, 1309−1313.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research, 30*, 79−82.