

# Estimating near-surface air temperature with NOAA AVHRR

James P. Riddering and Lloyd P. Queen

**Abstract.** A technique is presented for producing estimates of near-surface air temperature ( $T_a$ ) in complex terrain based on biweekly composite data from the National Oceanic and Atmospheric Administration (NOAA) advanced very high resolution radiometer (AVHRR). Results are tested against independently derived DAYMET gridded meteorological data. The model utilizes radiant surface temperature ( $T_s$ ) extrapolations as a surrogate for  $T_a$ , and digital terrain information is used to adjust temperatures as a function of environmental adiabatic lapse rates. The Earth–Sun–sensor geometry information is used to constrain air temperature estimates in three different implementations. Increasing the complexity of geometric constraints on the model resulted in increased accuracies of the predictions, and yet caused a decrease in the number of days  $T_a$  estimates could be made. When comparing DAYMET surfaces to the satellite estimates of  $T_a$ , correlation coefficients ( $R$ ) as high as 0.742 and standard errors of the estimate of 2.73 were observed in the final model implementation. The logic presented here serves as a useful technique for relatively simple derivation of near-surface air temperatures in a variety of remote sensing applications.

**Résumé.** On présente une technique pour produire des estimations de la température de l'air au niveau de la surface ( $T_a$ ) dans des paysages complexes basée sur les images synthèses temporelles bihebdomadaires NOAA (« National Oceanic and Atmospheric Administration ») AVHRR (« advanced very high resolution radiometer »). Les résultats sont testés par rapport aux données météorologiques spatialisées DAYMET indépendantes. Le modèle utilise des extrapolations de température radiante de surface ( $T_s$ ) à titre de substitut pour  $T_a$ , alors que l'information dérivée du modèle numérique d'altitude est utilisée pour ajuster les températures en tant que fonction du gradient adiabatique environnemental. L'information sur la géométrie terre–soleil–capteur est utilisée pour contraindre les estimations de température dans trois versions différentes. L'accroissement de la complexité des contraintes géométriques dans le modèle s'est soldé par une amélioration des précisions de prévision tout en causant une diminution dans le nombre de jours où des estimations de  $T_a$  pouvaient être réalisées. En comparant les surfaces DAYMET aux estimations satellitaires de  $T_a$ , on a pu observer des coefficients de corrélation ( $R$ ) aussi élevés que 0,742 et des erreurs types de 2,73 dans la version finale du modèle. La logique présentée ici a démontré son utilité comme technique pour dériver de façon relativement simple la température de l'air au niveau de la surface pour une variété d'applications en télédétection.

[Traduit par la Rédaction]

## Introduction

Gridded near-surface air temperature ( $T_a$ ) has been a much sought after variable in the realm of earth-system science due to the importance  $T_a$  plays in regulating many terrestrial processes. Approaches to deriving these surfaces have resulted in varying techniques, ranging from spatial interpolations of meteorological station data to satellite and aircraft remote sensing methods of predicting  $T_a$ . The various station-based interpolation techniques suffer from an arbitrary location of weather stations that often lack near real time data accessibility. This has driven many researchers to look for satellite-based methods of filling the gaps.

The relationship between vegetation indices and radiant surface temperature ( $T_s$ ) has been well established and exploited for a variety of means. Examples include surface resistance and vegetation water balance (Nemani and Running, 1989; Nemani et al., 1993; Carlson et al., 1994), atmospheric water vapor (Prince et al., 1998), and predictive fire danger (Vidal and Devaux-Ros, 1995). One of the more interesting applications of the  $T_s$  and vegetation index ( $T_s$ –VI) relationship,

however, has been the derivation of  $T_a$  (Goward et al., 1994; Prihodko and Goward, 1997; Czajkowski et al., 2000).

The more common satellite-based techniques for estimating  $T_a$  rely on the negative relationship between a vegetation index (VI), often the normalized difference vegetation index (NDVI), and a calculated  $T_s$  from thermal channel data. A relationship is defined through regression analysis, and extrapolations occur to some theoretical VI value that represents a full canopy as viewed by the sensor. The theoretical full canopy is important because of the low thermal mass of the canopy and the notion that this canopy  $T_s$  will likely be near the local  $T_a$  (Prihodko and Goward, 1997). Indeed, microclimate studies have shown this to be the case (Geiger, 1965; Aston and van Bavel, 1972; Gates, 1980). Many examples of satellite-based  $T_a$  estimation follow this concept and rely on a neighborhood analysis for defining the  $T_s$ –VI relationship. Goward et al. (1994), Prihodko and Goward (1997), Prince et al. (1998), and Lakshmi et al. (2001)

Received 28 February 2005. Accepted 20 January 2006.

**J.P. Riddering<sup>1</sup> and L.P. Queen.** College of Forestry and Conservation, University of Montana, Missoula, MT 59812, USA.

<sup>1</sup>Corresponding author (e-mail: jrp@ntsg.umt.edu).

all use a neighborhood of advanced very high resolution radiometer (AVHRR) pixels to define the relationship and extrapolate accordingly. Comparisons are typically made with weather station data (Goward et al., 1994; Prihodko and Goward, 1997; Prince et al., 1998) or with other remote sensing derived air temperature estimates (Lakshmi et al., 2001).

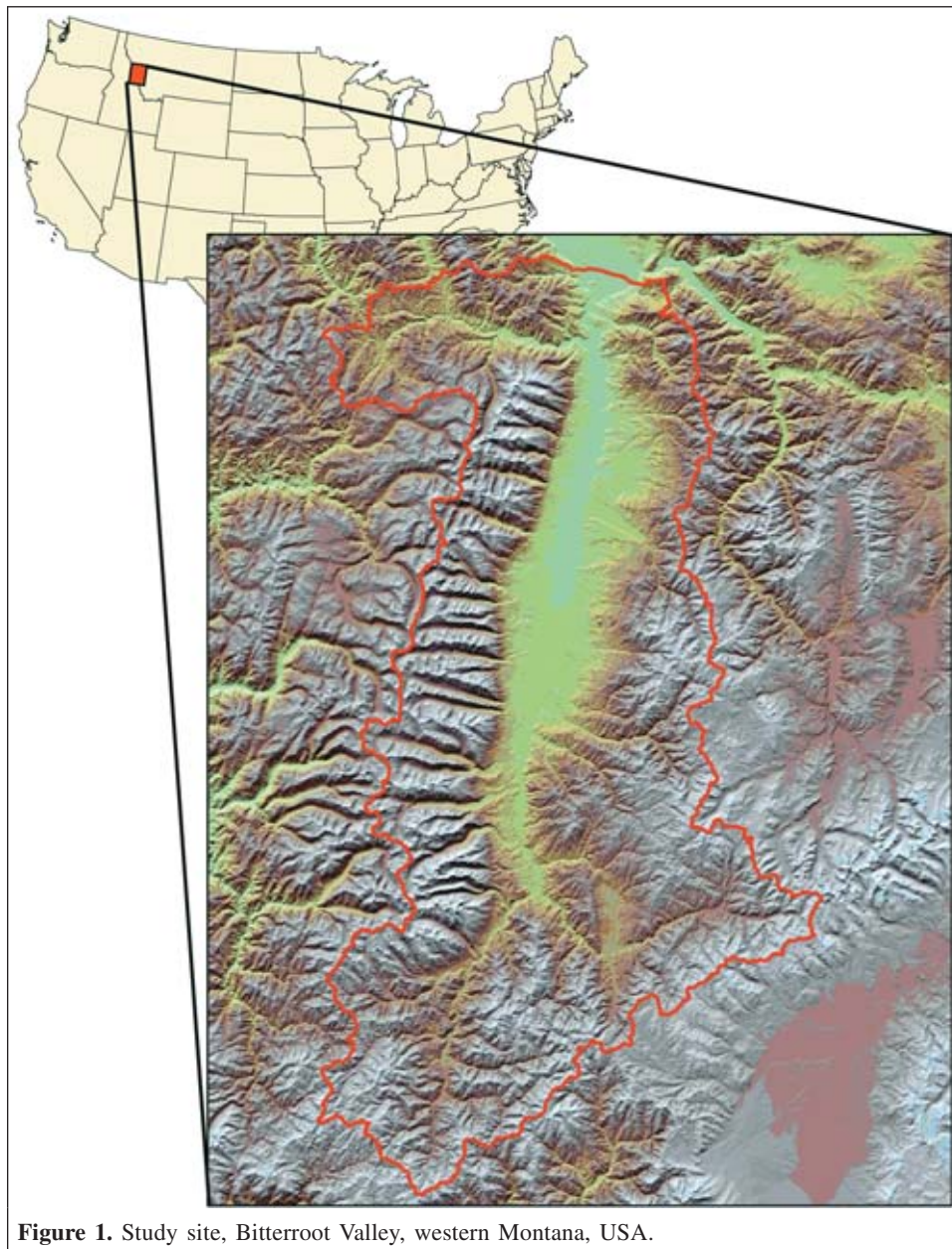
In this paper we present a method of estimating near-surface air temperature using National Oceanic and Atmospheric Administration (NOAA) advanced very high resolution radiometer (AVHRR) *composite* data in an area of complex terrain. Composite AVHRR data were chosen because they are readily available and relatively efficient to work with. Since composite data are being used, a *non-neighborhood* technique is presented. The goal is to achieve robust estimates of air temperature from satellite data while maintaining computational efficiency. This study was conducted over one

complete growing season and utilizes independently derived gridded air temperature surfaces for direct comparison at the time of satellite overpass.

## Methods

### Study area

This study was conducted in the Bitterroot Valley of western Montana, USA (**Figure 1**). The geographic extent was defined by the 1 : 100 000 scale 11 digit fourth hydrologic unit code (HUC) developed by the US Department of Agriculture Natural Resources Conservation Service, Montana State office. The fourth HUC defines major river drainage basins and in this case delineates the Bitterroot River and major tributaries. The Bitterroot Valley is characterized by a major mountain range on



**Figure 1.** Study site, Bitterroot Valley, western Montana, USA.

its western border and a lesser range on the eastern border. Vegetation is typically agriculture and grasslands, with riparian vegetation along river corridors at low elevations and mixed conifers at middle to higher elevations. The mixed conifers transition from *Pinus ponderosa* – *Pseudotsuga menziesii* through *Pinus contorta* and into *Picea–Abies* mixes at higher elevations. The highest peaks are characterized by alpine vegetation on the western edge of the study area. Elevations range from approximately 976 m where the Bitterroot River joins the Clark Fork of the Columbia River to 3098 m at Trapper peak. The Bitterroot HUC stretches approximately 157 km north–south and approximately 58 km east–west and contains approximately 743 000 ha.

### Satellite data

Since 1987, the US Geological Survey EROS Data Center (EDC) has been receiving, processing, and archiving AVHRR data. For this study, biweekly maximum NDVI composites derived from afternoon overpasses of NOAA-14 were obtained. The maximum NDVI compositing technique relies on 14 daily calculations of NDVI. The maximum NDVI observed per pixel in that period is retained, and all individual channel data (channels 1–5) from the day of maximum NDVI observation are also retained in the dataset. Additionally, there are three geometry files included, namely satellite zenith, solar zenith, and relative zenith angles, as well as a date file referencing the day of data retention within the composite period. The channel data are converted to percent reflectance (channels 1 and 2) and brightness temperature (channels 3–5). The maximum NDVI composite technique is used to minimize cloud contamination (Holben, 1986); however, recent studies indicate a preferential selection of off-nadir pixels (Stoms et al., 1997). In addition to calibration and compositing, all data are registered and projected at the EDC into Lambert's azimuthal equal area projection to facilitate expeditious use in geographic studies. The 1997 composite periods P10–P19 (9 May – 25 September) were used in this study. No radiometric or geometric postprocessing of the EDC data was done.

### Processing stream

**Figure 2** describes a conceptual layout of the processing stream used to estimate air temperature. Part 1 refers to the calculation of surface temperature and emissivity estimates, part 2 discusses the site mask imposed on the analysis, and part 3 explains the central processing loop. The core process relies on inputs from parts 1 and 2 and produces *day*-specific output from the composite data for all pixels not removed by the masking criteria.

#### Surface temperature (flow chart part 1)

The split-window technique of Ulivieri et al. (1994) was used in this study to calculate surface temperature. The primary reason for using this technique is its simple formulation, accurate results (Vazquez et al., 1997; Ouaidrari et al., 2002), applicability across multiple AVHRR sensors (Ulivieri et al.,

1994; Ouaidrari et al., 2002), and relative resistance to emissivity errors (Vazquez et al., 1997). The surface temperature ( $T_s$ ) formulation is as follows:

$$T_s = T_4 + 1.8(T_4 - T_5) + 48(1 - \epsilon) - 75(\epsilon_4 - \epsilon_5) \quad (1)$$

where  $T_4$  and  $T_5$  are the brightness temperatures of AVHRR channels 4 and 5, respectively;  $\epsilon_4$  and  $\epsilon_5$  are emissivity estimates for channels 4 and 5, respectively; and  $\epsilon$  is the mean emissivity of channels 4 and 5.

The choice was made to estimate emissivity in an attempt to reduce the associated error in  $T_s$  calculation, which can be quite large. The method of van de Griend and Owe (1994) was used for estimating  $\epsilon_4$  (Equation (2)), and the method of Thornton (1998), which was developed over a region that includes this study area, was used to estimate the difference in channel 4 and 5 emissivities ( $\Delta\epsilon$ ). Average emissivity ( $\epsilon$ ) and  $\epsilon_5$  were then calculated from the estimated parameters. Equations (2) and (3) are empirical relationships, and Equations (4) and (5) are simply algebraic expressions. The formulations are as follows:

$$\epsilon_4 = -0.99 - 0.09 \frac{(0.7 - \text{NDVI})}{0.6} \quad (2)$$

$$\Delta\epsilon = -0.02938 + 0.04957 \times \text{NDVI} \quad (3)$$

$$\epsilon_5 = \epsilon_4 - \Delta\epsilon \quad (4)$$

$$\epsilon = \frac{\epsilon_4 + \epsilon_5}{2} \quad (5)$$

#### Site mask (flow chart part 2)

There were a series of constraints imposed on the estimation of  $T_a$  to remove pixels that contained cloud contamination, undesirable land cover classes, extreme satellite zenith angles, or relative azimuth angles that would likely lead to poor predictions of air temperature. Any pixel in the data that failed to meet the listed criteria was removed from further analysis. Cloud contamination, undesirable land cover classes, and satellite zenith angles are discussed in the following sections, and the addition of the relative azimuth constraint is discussed in the Results section.

The maximum NDVI compositing technique assumes clouds are removed through the simple selection of those pixels which have the highest NDVI for the composite period. Although this is generally an acceptable notion by virtue of reflectance functions of clouds, we chose to embed a more explicit cloud-screening algorithm to minimize the impact subpixel cloud contamination might have in attenuating the responses in the thermal channels. The resulting technique is based on the methods of Flasse and Ceccato (1996), Thornton (1998), and Seielstad et al. (2002). Pixels were considered cloud contaminated if they met any of the following criteria:

$$\text{channel 1 reflectance (red)} > 0.2 \quad (6)$$

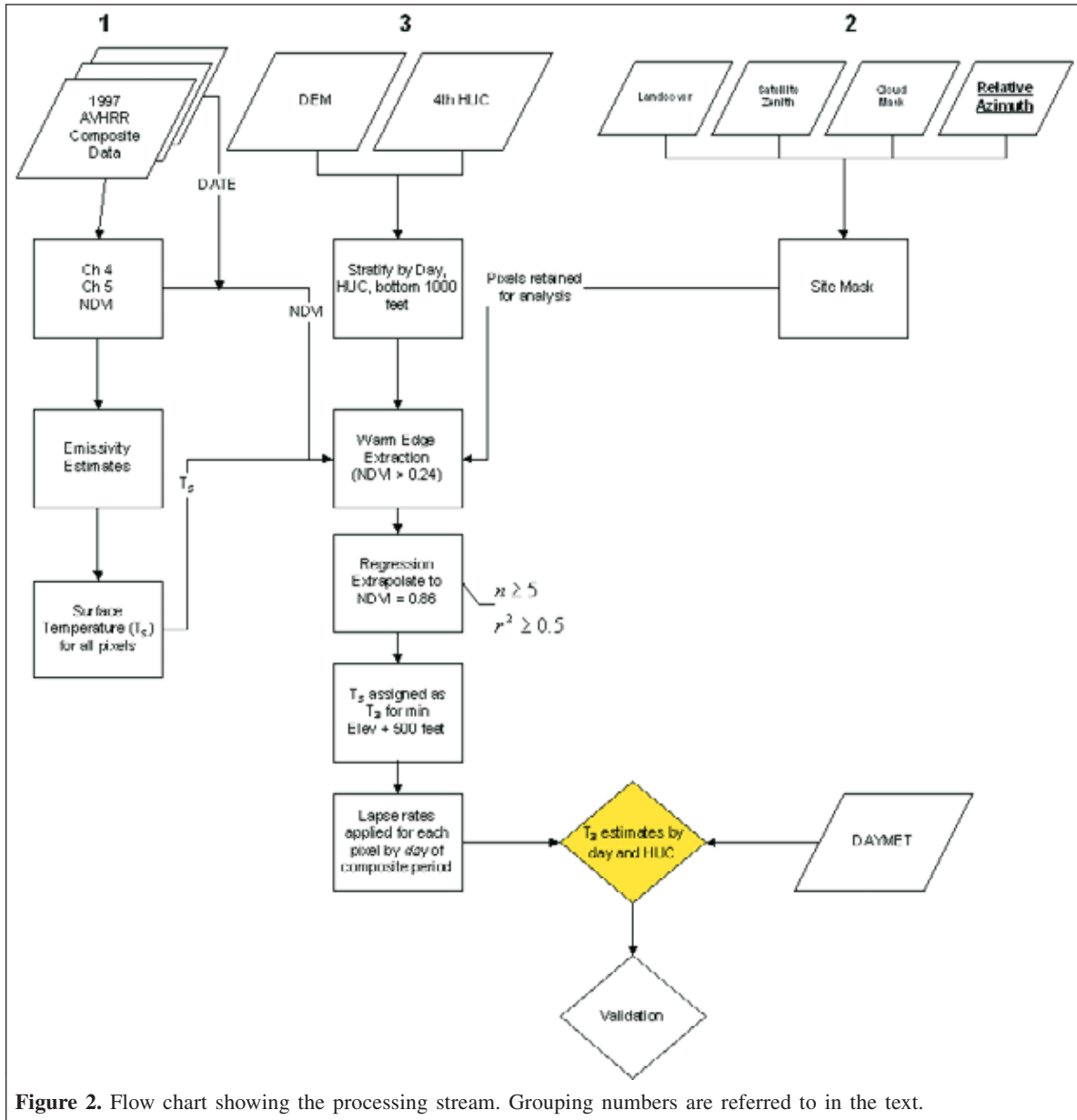


Figure 2. Flow chart showing the processing stream. Grouping numbers are referred to in the text.

$$Q < 1.20 \quad (7)$$

$$\Delta T_{45} > 4.5 \text{ (K)} \text{ or } \Delta T_{45} < -1.5 \text{ (K)} \quad (8)$$

$$\Delta T_{34} > 15 \text{ (K)} \quad (9)$$

where  $Q$  is the ratio of near-infrared (NIR) reflectance (channel 2) to red reflectance (channel 1),  $\Delta T_{45}$  is the difference between channel 4 and channel 5 brightness temperature, and  $\Delta T_{34}$  is the difference between channel 3 and channel 4 brightness temperature. If the pixel was shown to be cloud contaminated in a given composite period, that pixel was flagged and removed from subsequent analysis.

The land cover constraint was based on the *North American land cover characteristics data base* (US Geological Survey 2000). This is a generalized land cover map showing such

classes as shrubland, irrigated cropland, and evergreen needleleaf forest. The primary goal was to remove pixels that contained potentially confounding land cover practices such as agriculture, water, or urban-developed or those which were potentially mislabeled (e.g., evergreen broadleaf forest in Montana). The resulting mask retained only those pixels from the following land cover classes: (i) grassland, (ii) shrubland, (iii) deciduous broadleaf forest, (iv) evergreen needleleaf forest, and (v) mixed forest.

A simple satellite zenith angle truncation was included to remove extreme look angles from the dataset. As suggested by Prihodko and Goward (1997), a zenith angle greater than  $40^\circ$  was removed in the initial implementation to minimize the thermal attenuation caused by long atmospheric path lengths. Subsequent examination of the estimates, however, resulted in the satellite zenith angle constraint being changed to  $30^\circ$ . The

impacts of zenith angles are discussed in greater detail in the Results section.

#### *Warm edge extraction: $T_a$ technique (flow chart part 3)*

Many implementations of the  $T_s$ -VI technique for estimating  $T_a$  rely on a contextual approach where a neighborhood of pixels is used for the regression analysis. The “window” is then shifted (usually one row or column) and the calculation is repeated. This technique relies on spatial autocorrelation of adjacent pixels for the solution of  $T_a$ . Unfortunately, it results in an effective reduction in the spatial resolution of the solution itself (Prince et al., 1998). Other studies have exploited the  $T_s$ -VI relationship in a different manner. Although not used explicitly for  $T_a$  estimation, some authors chose combinations of contextual analysis with a data-reduction technique. For example, Nemani and Running (1989) and Carlson et al. (1995) limited their analysis to only those NDVI- $T_s$  pairs which showed maximum  $T_s$  for each unique NDVI, the so-called warm-edge extraction. The logic behind this selection suggests that those pixels are most likely well illuminated, near nadir, and in thermal equilibrium at the time of satellite overpass.

Since one of the primary interests is in a relatively simple and computationally efficient technique of estimating air temperature at landscape scales, a method was chosen that combines both contextual and noncontextual elements. A stepwise explanation of the process follows.

First, a digital elevation model (DEM) was used to build a mask of the lowest 305 m (1000 ft) of the Bitterroot HUC. The assumption here is that an estimate of mean valley temperature could be used with a simple adiabatic lapse rate to calculate temperatures at higher elevations. The bottom 305 m was chosen because it gave an adequate number of  $T_s$  and NDVI values for use in the air temperature estimates and provided enough of the required diversity in land cover (hence NDVI scores; see Riddering et al., 1999) to define the NDVI- $T_s$  relationship. For this bottom 1000 ft, the data were then stratified into individual year-day bins so that only data from the same day were used in the individual initial air temperature estimate for the valley floor.

Once sorted by day, NDVI scores of 0.24 and below (and the associated  $T_s$ ) were removed from the analysis. This minimum value was based on empirical observation of NDVI- $T_s$  pairs and reported minimum NDVI values for vegetated surfaces (Carlson et al., 1994), effectively removing confounding pixels while reducing the size of the dataset. Next, the “warm edge” was extracted, resulting in greater data reduction. Lastly, a least squares regression was fit to the data. Regression constraints were imposed on the data such that any line composed of less than five pixels ( $n = 5$ ) or exhibiting a coefficient of determination ( $r^2$ ) less than 0.5 was removed from further consideration. This has the effect of maintaining a minimal quality control standard in the data. Lastly, the equation describing this line was extrapolated to NDVI = 0.86 to estimate the associated full-canopy  $T_s$ , which was used as a surrogate for the mean valley temperature. A value of 0.86 as the full-canopy NDVI was chosen, following the reports of

Prihodko and Goward (1997) and based on long-term maximum NDVI scores (Burgan and Chase, 1998) for known full-canopy vegetation typical of western Montana forests. The extrapolated value of  $T_s$  (which becomes the initial  $T_a$  value) was assumed to be the mean value of the bottom 305 m and was therefore assigned to the mean valley-bottom elevation (1128 m or 3700 ft for the Bitterroot Valley). A simple environmental lapse rate of 1.98 K per 305 m (Barry and Chorley, 1998) was used to adjust temperatures according to DEM derived elevations. The result is pixel-based  $T_a$  estimates that are unique for each day of data within the composite period.

### DAYMET

Most previous studies have used point weather station data for comparisons with satellite estimates of  $T_a$  (e.g., Prihodko and Goward, 1997; Florio et al., 2004). Our goal, however, was to test the  $T_a$  logic across a gridded, topographically complex landscape. Therefore, we required an independently derived gridded air temperature surface to test the satellite  $T_a$  estimates against. As a result, the DAYMET model was chosen to derive the independent air temperature grids for model comparison.

DAYMET has been described in detail elsewhere (Thornton et al., 1997). Briefly, DAYMET is a computer program that produces daily gridded meteorological variables (maximum and minimum temperature, precipitation, radiation, and humidity) over complex terrain from weather station observations. It utilizes a weighted Gaussian filter and iterative processes to predict known station values (sequentially removed from the analysis) and subsequent extrapolated values. A DEM is required to account for the vertical variation in the landscape and to calculate the horizontal relationships of temperature in complex terrain in a spatially explicit manner. DAYMET has been shown to be a robust method of generating meteorological surfaces in complex terrain, with reported mean absolute errors (MAE) in yearly temperature prediction between 0.7 °C and 1.2 °C for maximum temperature ( $T_{max}$ ) and minimum temperature ( $T_{min}$ ), respectively. Daily errors in temperature are reported at 1.8 °C and 2.0 °C, respectively, for  $T_{max}$  and  $T_{min}$ . The spatial resolution of DAYMET grids used in this study was 1 km, matching composite AVHRR resolution.

#### Synchronization of DAYMET and $T_a$ estimate

For this study,  $T_{max}$  and  $T_{min}$  for each pixel in the study area were extracted from the DAYMET record. Temperature at the time of the satellite overpass, discussed later in the paper, was calculated and served as “ground truth” for comparison with satellite estimates of near-surface air temperature.

Time of satellite overpass was estimated following the method of Thornton (1998), which uses solar zenith angle and the equation defining Earth-Sun geometry used in the DAYMET model for incoming radiation calculations. Hour angle ( $h$ , where 0° is local solar noon) is then converted to local solar time with the following equations:

$$h = \text{acos}\{[\cos(z) - \sin(l)\sin(\text{decl})]/[\cos(l)\cos(\text{decl})]\} \quad (10)$$

$$\text{decl} = -23.45 \cos[(\text{year}-\text{day} + 11.25)0.9863] \quad (11)$$

$$T_{\text{so}} = 12.0 + h/15 \quad (12)$$

where  $z$  is the solar zenith angle,  $l$  is the latitude of each pixel,  $T_{\text{so}}$  is the time of the satellite overpass, and  $\text{decl}$  is the earth's rotational axis declination relative to the principal plane.

With the time of the satellite overpass known,  $T_{\text{max}}$  and  $T_{\text{min}}$  from DAYMET and the equations of Campbell and Norman (1998) were used to calculate a DAYMET-derived temperature coincident with the satellite overpass. DAYMET temperature was calculated in the following manner:

$$\Gamma_{(t)} = 0.44 - 0.46 \sin(\omega t + 0.9) + 0.11 \sin(2\omega t + 0.9) \quad (13)$$

where  $\omega = \pi/12$  and  $t$  is time of day (in hours, 12 at solar noon). This function can then be used to calculate temperature at any time of day ( $t$ ):

$$\begin{aligned} T_{(t)} &= T_{x,i-1}\Gamma_{(t)} + T_{n,i}[1 - \Gamma_{(t)}] & 0 < t \leq 5 \\ T_{(t)} &= T_{x,i}\Gamma_{(t)} + T_{n,i}[1 - \Gamma_{(t)}] & 5 < t \leq 14 \\ T_{(t)} &= T_{x,i}\Gamma_{(t)} + T_{n,i+1}[1 - \Gamma_{(t)}] & 14 < t \leq 24 \end{aligned} \quad (14)$$

where  $T_x$  and  $T_n$  are the daily maximum and minimum temperature, respectively;  $i$  is the present day;  $i - 1$  is the previous day; and  $i + 1$  is the next day.

### Validation strategy

The validation strategy relies on comparisons of co-located grids of DAYMET temperature at the time of satellite overpass and the AVHRR-based  $T_a$  estimate. As mentioned previously, the AVHRR-based  $T_a$  estimate is day-specific. The DAYMET surface was developed to match that day at the time of the satellite overpass and is treated as the ground truth. For example, the composite period date raster was used as a mask in the DAYMET surfaces so that the individual DAYMET pixels match the day of satellite acquisition. The result is day-specific comparisons between pixels in the two rasters. Results are reported primarily as pooled results (entire record) except when specific interests warrant otherwise. Reported statistics include mean absolute error (MAE) defined as the average of the absolute difference between DAYMET  $T_a$  and satellite  $T_a$  estimates, Pearson's correlation coefficient, and regression coefficients.

## Results

The results are presented in three parts. The first reports the original implementation with no relative azimuth constraint and a satellite zenith maximum of  $40^\circ$  (labeled version 1). The second set of results was a consequence of increased scrutiny of

the satellite zenith angles, where a  $30^\circ$  limit was imposed (version 2: SATZ30). Lastly, the third set of results shows the impact of relative azimuth constraints added to the estimation of  $T_a$  (version 3: SATZ30 & REL AZ). The addition of relative azimuth constraints was due to the observed impact that westerly look directions had on the error structure of  $T_a$  estimates ( $T_{\text{est}}$ ).

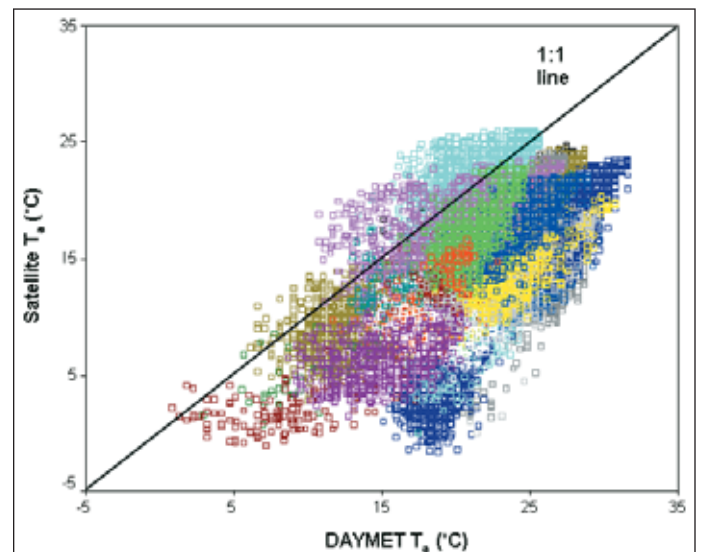
### Version 1

With the initial implementation, an attempt was made to use all composite periods (P10–P19) coupled with cloud screening and satellite zenith angle constraints ( $40^\circ$  as per Prihodko and Goward, 1997). Due to these constraints and the regression requirements listed previously, the algorithm was unable to calculate any air temperatures for composite periods P10–P13 (9 May – 3 July). The resulting period of assessment is 4 July – 25 September, composed of composite periods P14–P19.

Plots of estimated  $T_a$  versus DAYMET  $T_a$  are shown in **Figure 3**, and statistical summaries of regression analysis and correlation assessment are shown in **Table 1**. Additionally, mean absolute error (average absolute difference between  $T_{\text{est}}$  and DAYMET  $T_a$ ) was derived. Air temperature estimates and DAYMET actual air temperatures showed a Pearson's correlation coefficient of 0.576 (**Table 2**) and an MAE of  $6.96^\circ\text{C}$  for all days and a total of 28 unique day-based predictions ( $n = 28$ ).

### Version 2: SATZ30

When comparing satellite zenith angles with predicted temperature, poor predictive ability was observed when zenith angles were greater than  $30^\circ$  (data not shown). The code was modified to change the satellite zenith angle constraint from  $40^\circ$  to  $30^\circ$ , and the model was rerun for all composite periods.



**Figure 3.** Scatterplot of satellite air temperature ( $T_a$ ) estimates versus DAYMET  $T_a$  estimates for the original implementation (satellite zenith angle constraint  $\leq 40^\circ$ ). MAE =  $6.96$ ;  $R = 0.576$ . Color indicates individual day.

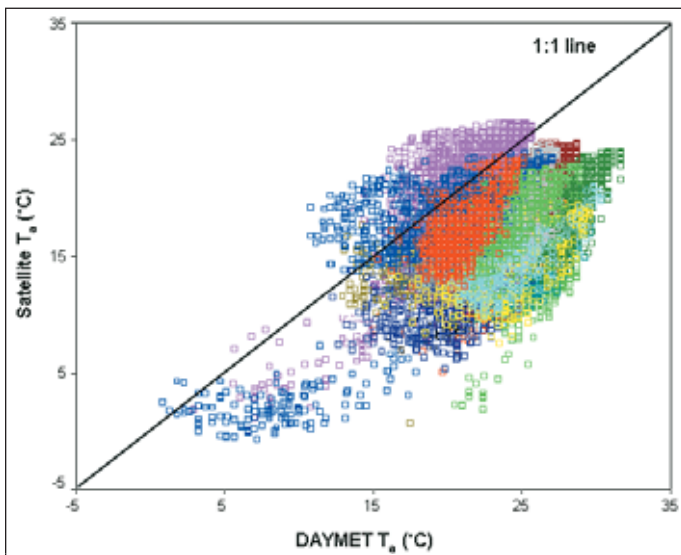
**Table 1.** Regression coefficients for the three implementations of satellite  $T_a$ .

	Regression parameters			Correlation coefficient ( $R$ )	SE
	$n$	$m$	$B$		
Original implementation	33 065	0.634	1.403	0.576	3.69
Satellite zenith 30 (SATZ30)	28 498	0.449	6.109	0.410	3.38
Satellite zenith 30 and relative azimuth (SATZ30 & REL AZ)	7 460	0.818	-2.149	0.742	2.73

**Note:** The dependant variable ( $Y$ ) is satellite  $T_a$  and the independent variable ( $X$ ) is DAYMET-derived  $T_a$ , with equations taking the form  $Y = mX + B$ . SE, standard error of the estimate.

**Table 2.** Mean absolute error (MAE), correlation coefficients, number of days, and membership in each MAE difference class ( $0 \leq$  class  $1 \leq 2.0$  °C;  $2.0$  °C < class  $2 \leq 5.0$  °C;  $5.0$  °C < class  $3 \leq 8.0$  °C).

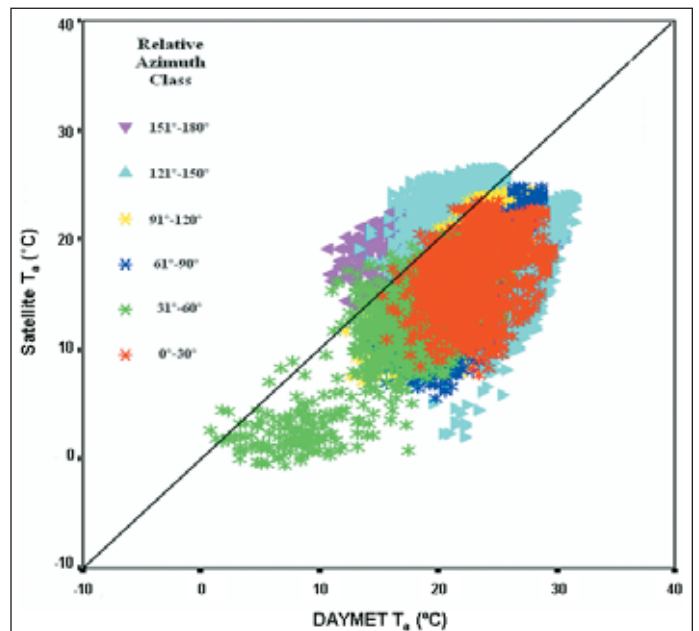
	MAE	Correlation, $R$	No. of days	Class 1		Class 2		Class 3		Total no. of pixels
				$n$	% of total	$n$	% of total	$n$	% of total	
Original implementation	6.96	0.576	28	3336	0.10	7261	0.22	9423	0.28	33 065
Satellite zenith 30 (SATZ30)	6.47	0.410	20	3430	0.12	6671	0.23	8270	0.29	28 498
Satellite zenith 30 and relative azimuth (SATZ30 & REL AZ)	6.09	0.742	12	463	0.06	2226	0.30	3024	0.41	7 460



**Figure 4.** Satellite zenith angle  $\leq 30^\circ$ . MAE = 6.47;  $R = 0.413$ . Color indicates individual day.

The enhanced zenith angle constraint had the effect of removing even more days from the total analysis ( $n = 20$ ). The MAE improved to 6.47 °C, but the correlation coefficient dropped to 0.410 (see **Table 1** and **Figure 4**).

With the satellite zenith angle information, look directions were calculated and the analysis was split accordingly. A satellite zenith angle of  $90^\circ$  denotes a nadir view, anything greater than  $90^\circ$  is a westerly look, and anything less than  $90^\circ$  is an easterly look. As shown in **Figure 5**, when the satellite is recording information with a westerly look, the predictive ability appears to suffer. Westerly looks give an MAE of 6.75 °C and an  $R$  of 0.163, whereas easterly looks have an MAE of 5.74 °C and an  $R$  of 0.766, a remarkable difference.



**Figure 5.** DAYMET  $T_a$  versus satellite  $T_a$  ( $30^\circ$  satellite zenith angle constraint). Markers denote relative azimuth classes. Classes are  $0^\circ$ – $30^\circ$  for class 1,  $31^\circ$ – $60^\circ$  for class 2, etc. up to  $180^\circ$ .

It is worth noting that the previous look-direction assessments were made on the final air temperature prediction product. There was no masking based on look direction in the original estimation of valley air temperature or subsequent temperatures based on lapse rates. Rather, they were used to stratify pixels from the analysis after the lapse rates were applied. Although it appears an incorporation of look angle will dramatically improve the correlation in the final product, there was no attempt to constrain pixels used in the previous analysis by look direction. The result is a likely decrease in the

predictability of the overall air temperature estimate due to the undesirable look directions. Based on this finding, the relative azimuth information was included as an initial masking constraint in the third version discussed in the following section.

### Version 3: SATZ30 & REL AZ

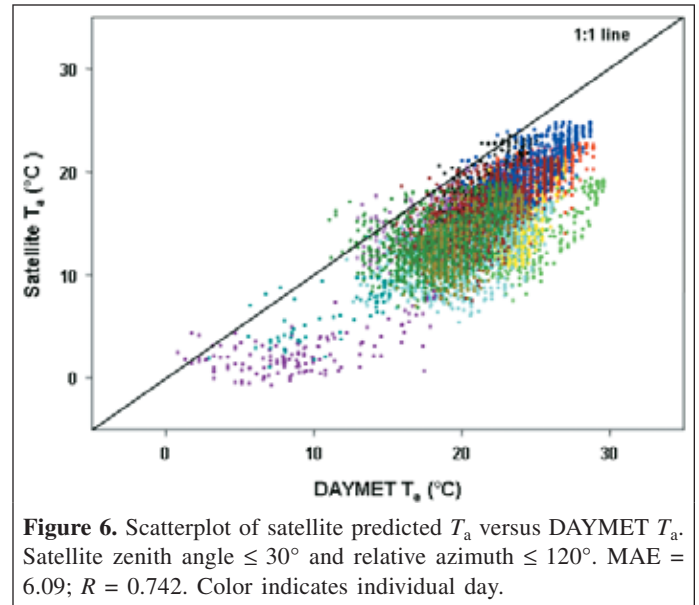
Relative azimuth is typically provided in the AVHRR dataset for atmospheric correction algorithms. It is defined as the absolute difference between the solar azimuth and the satellite azimuth angles and it ranges from  $0^\circ$  to  $180^\circ$ . Satellite and solar azimuth are the actual true azimuths from a given pixel to both the satellite and the sun, respectively. Therefore, relative azimuths of  $180^\circ$  define a sun–surface–sensor geometry where the satellite is looking into the solar plane.

Due to the directional impacts discussed previously, the relative azimuth data were categorized to provide an initial assessment of the impact that look direction has on air temperature estimates. Classes were defined at  $30^\circ$  increments of relative azimuth, such that class 1 contained  $0^\circ$ – $30^\circ$ , class 2 contained  $31^\circ$ – $60^\circ$ , and so on up to  $180^\circ$ . The result is six categories of relative azimuth. **Figure 5** shows the previous predicted versus observed air temperatures using the  $30^\circ$  satellite zenith constraint, and they are indicated by relative azimuth class.

Based on **Figure 5**, relative azimuth classes 5 and 6 were included in the mask and the algorithm was rerun. By incorporating this relative azimuth constraint, the extreme westerly looks in the initial valley-bottom air temperature prediction were removed. Subsequent lapse-rate calculations were made as before. Adding the relative azimuth mask resulted in further reduction of the dataset. For the composite periods P14–P19, there were 12 individual days of data. Interestingly, the MAE dropped to  $6.09^\circ\text{C}$  and the correlation coefficient rose to 0.742 (**Table 1**) when using the satellite zenith constraint of  $30^\circ$  and the relative azimuth mask. **Figure 6** shows estimated air temperature versus DAYMET-derived temperature stratified by day.

### Summary

**Table 2** lists the three methods of predicting air temperature and the landscape membership of difference classes. Class 1 was defined as the absolute difference in  $T_{\text{est}}$  and DAYMET  $T_a$  of  $0$ – $2.0^\circ\text{C}$ , class 2 of  $2.1$ – $5.0^\circ\text{C}$ , and class 3 of  $5.1$ – $8.0^\circ\text{C}$ . As discussed earlier, the addition of relative azimuth and  $30^\circ$  satellite zenith constraints reduced the total number of available pixels while improving the overall performance of the technique. The greatest improvement came in difference class 2, and a reduction in class 1 was observed. Overall, approximately 60% of all  $T_{\text{est}}$  pixels fell within  $8^\circ\text{C}$  of DAYMET surfaces. However, the more stringent analysis provided increased correlation and decreased MAE for the entire record. Regression parameters for all implementations are shown in **Table 1**.



**Figure 6.** Scatterplot of satellite predicted  $T_a$  versus DAYMET  $T_a$ . Satellite zenith angle  $\leq 30^\circ$  and relative azimuth  $\leq 120^\circ$ . MAE =  $6.09$ ;  $R = 0.742$ . Color indicates individual day.

## Discussion

### Sources of error

There are numerous sources of potential error in this study, both in air temperature estimation and in the DAYMET-derived reference source. Surface temperature errors, including water vapor impacts and emissivity errors, residual cloud contamination, and choice of static environmental lapse rates, all impact the estimated air temperature. Conversely, there is a known error structure associated with DAYMET-derived surfaces that were used as “ground truth” and some error in interpolated air temperature at the time of the satellite overpass.

Perhaps the largest potential error in this study comes from surface temperature calculations. The method of Ulivieri et al. (1994) was chosen because of its applicability to various NOAA AVHRR sensors and its reported resistance to errors in land surface emissivity calculations (Ulivieri et al., 1994; Vazquez et al., 1997). However, there is still a large potential error in the  $T_s$  calculation when atmospheric water vapor is not considered in the model. Ouaidrari et al. (2002) report possible errors in  $T_s$  due to water vapor ranging from a minimum of  $-7.73\text{ K}$  to a maximum of  $3.08\text{ K}$  when using a low water vapor formulation of the  $T_s$  equation in the presence of high actual atmospheric water vapor content. These authors state that analysis of data from the NASA Data Assimilation Office (DAO) shows that the base assumption of low atmospheric water content while formulating many split window techniques may not be valid. The presence of higher water vapor content than originally assumed is fairly common. Although the confounding effect of atmospheric water vapor was not specifically studied, it is reasonable to assume that there is a potential influence in the study site, particularly in 1997, which was an atypically rainy year in western Montana.

Emissivity is a critical component in calculating  $T_s$ . While the method of Ulivieri et al. (1994) has shown great resistance to errors in emissivity, the potential for miscalculation of  $T_s$  does exist. Our choice in estimating emissivity following the techniques of van de Griend and Owe (1994) and Thornton (1998) was based on the desire to develop a robust yet efficient means of estimating  $T_a$  at landscape scales. Ouaidrari et al. (2002) calculated potential errors in  $T_s$  from faulty emissivity calculations and showed that great variation may occur when emissivity is not adequately estimated. The reported potential root mean square (RMS) errors with the  $T_s$  method of Ulivieri et al. (1994) range from 1.10 K to 3.91 K depending on the magnitude of emissivity error.

Lastly, the gridded DAYMET-derived  $T_{\max}$  and  $T_{\min}$  surfaces were used to interpolate temperatures at the time of the satellite overpass. DAYMET, however, has its own set of possible associated errors. Thornton et al. (1997) report a daily MAE for  $T_{\max}$  and  $T_{\min}$  of 1.8 °C and 2.0 °C, respectively. Since an absolute value of the differences between predicted and observed is what the MAE assessment gives, one does not know the direction of this reported error. However, adding and subtracting 1.9 °C to the DAYMET temperatures and recalculating the MAEs and correlations show interesting results. Correlations stay constant, but subtracting results in a new MAE in predicted air temperature and DAYMET air temperature of 4.3 °C for the final implementation of the model (data not shown).

The numerous constraints imposed on the estimation of  $T_a$  warrant some discussion, primarily of the NDVI limits and the minimum regression parameters used in the initial valley-bottom estimation of air temperature. There is ample opportunity to adjust many of the constraints made in this project, and future work should examine the impact these choices have on the performance of the air temperature estimation.

The minimum NDVI value of 0.24 was empirically derived and is similar to values reported by other authors of uncontaminated bare ground pixels (Carlson et al., 1994; Riddering et al., 1999). The inclusion of low NDVI values (less than 0.24) results in a decreased ability to derive a strong  $T_s$ -NDVI relationship. Similarly, the choice of 0.86 as the extrapolation value was based on other published values and the fact that this number corresponds to known full-canopy vegetation in western Montana, as shown in the long-term NDVI records (Burgan and Chase, 1998). Though other studies have used different values of NDVI, the absolute difference in air temperature estimates as a function of NDVI score is minor. Prihodko and Goward (1997) report changes in NDVI from 0.860 to 0.873 resulted in air temperature changes of only 0.15 °C.

Since this technique uses composite data, neighborhood functions could not be used to derive an estimate of air temperature. Therefore, constraints were placed on the regression component of the central air temperature estimation loop. Since the site mask removes a great number of valley-bottom pixels, regression parameters were kept fairly

nonrestrictive ( $n \geq 5$  and  $r^2 \geq 0.5$ ) to allow for a larger number of air temperatures to be calculated. Additionally, there were many occasions in the data where a significant number of pixels occurred on a given date, but only a few pixels from that date occurred in the valley bottom (for the initial  $T_a$  estimation). By keeping the regression requirements nonrestrictive ( $n \geq 5$ ) we were able to perform estimates of  $T_a$  in this situation. Interestingly, although there were few dates that had a low number of valley-bottom pixels ( $n = 5-10$ ), when this occurred the correlations between the satellite estimate and DAYMET temperature were similar to those for dates with a larger  $n$  (data not shown).

### Efficiency

Developing a simple, efficient, and robust method of estimating air temperature was the primary goal of this study. This goal was achieved through the use of composite satellite data, a robust surface temperature algorithm, and basic assumptions on temperature variations with elevation. These results are similar to those reported in previous studies. With continued investigation, these techniques can be improved even further. To improve future performance of the algorithm, some suggestions are made. First, through the use of weekly composites, it is likely the total number of  $T_a$  estimates throughout the growing season can be improved, thereby giving a more complete assessment of air temperature. Though biweekly composites encompass 14 days of data, it is more typical to see only 4-5 days of retained data within an area defined by a fourth HUC. The increased temporal resolution of weekly composites would likely provide a more distributed and complete coverage of air temperature estimates throughout the growing season.

Second, seasonal variations in lapse rate need to be studied further. These data exist in the Selway-Bitterroot Wilderness of western Montana and northern Idaho (Finklin, 1983). Further studies should investigate the utility of this record for application to the Bitterroot Valley and the applicability to typical complex terrain in this region. It is highly unlikely that a single environmental lapse rate is sufficient to describe the complex atmospheric relationships between elevation and temperature throughout the year. Such simple components as increased atmospheric water vapor or inversions can seriously impact the assessment of air temperature as a function of elevation.

Lastly, a more complete assessment of sun-surface-sensor geometry needs to be made as evidenced by the results in the relative azimuth mask. The precise reasons why large relative azimuths tend to result in poor agreement in estimated versus observed air temperatures are not known, as that was beyond the scope of this study. However, it is hypothesized that by looking into the solar plane, the sensor is suffering from solar contamination, through either increased path radiance or other means. A look-direction constraint based on the differentiation of forward scatter and (or) backscatter would likely benefit future research in satellite-based air temperature estimates.

## Conclusions

Considering the list of potentially confounding issues discussed in this paper, the algorithm described in the paper performed remarkably well, particularly when compared to previous studies. For example, Goward et al. (1994) report a root mean square error (RMSE) of  $\pm 5.4$  °C in their air temperature prediction. Prince et al. (1998), from diverse landscapes, express RMS errors for estimates of  $T_a$  as high as 5.4 °C in certain study areas while reporting a total error for all study sites of 3.9 °C. In the Prihodko and Goward (1997) study from Kansas, USA, clouds were shown to be confounding factors in the estimation of air temperature. Before incorporating a cloud-detection routine, mean absolute error (MAE) was reported at 4.0 °C and 2.92 °C after cloud removal. Lastly, Lakshmi et al. (2001) used a combination of AVHRR and TIROS operational vertical sounder (TOVS) in Arkansas, USA, and show an RMSE of 5.3 °C. The final implementation of our algorithm resulted in an MAE of 6.09 °C and a standard error of the estimate of 2.73 °C (**Tables 1 and 2**), which is of a magnitude similar those from previous studies.

It is expected that improved capability may be achieved with additional site-specific corrections to the algorithm presented in this paper. Examples include improved lapse rate corrections (dry-wet lapses), a better understanding of the Earth-Sun-sensor geometric issues presented, and a more explicit treatment of the potential slope and aspect impacts in estimating air temperature in complex terrain. Additionally, the methods presented here would likely benefit from the use of advanced sensors and data currently available, such as the moderate-resolution imaging spectroradiometer (MODIS). With the improved calibration and radiometry of MODIS and the efficiency of this  $T_a$  technique, relatively simple estimates of  $T_a$  may be computed with high degrees of confidence.

## References

- Aston, A.R., and van Bavel, C.H.M. 1972. Soil surface water depletion and leaf temperature. *Agronomy Journal*, Vol. 64, pp. 368–373.
- Barry, R.G., and Chorley, R.J. 1998. *Atmosphere, weather, and climate*. Routledge, New York. 409 pp.
- Burgan, R.E., and Chase, C.H. 1998. *NDVI and derived image data*. Rocky Mountain Research Station, Missoula Fire Sciences Laboratory, Missoula, Mont. General Technical Report RMRS-GTR-9-CD. CD-ROM.
- Campbell, G.S., and Norman, J.M. 1998. *An introduction to environmental biophysics*. Springer-Verlag, New York. 286 pp.
- Carlson, T.N., Gillies, R.R., and Perry, E.M. 1994. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. *Remote Sensing Reviews*, Vol. 9, pp. 161–173.
- Carlson, T.N., Gillies, R.R., and Schmugge, T.J. 1995. An interpretation of methodologies for indirect measurement of soil water content. *Agricultural and Forest Meteorology*, Vol. 77, pp. 191–205.
- Czajkowski, K.P., Goward, S.N., Stadler, S.J., and Walz, A. 2000. Thermal remote sensing of near surface environmental variables: application over the Oklahoma mesonet. *Professional Geographer*, Vol. 52, No. 2, pp. 345–357.
- Finklin, A.I. 1983. *Weather and climate of the Selway-Bitterroot Wilderness*. University Press of Idaho, Moscow, Idaho. 144 pp.
- Flasse, S.P., and Ceccato, P. 1996. A contextual algorithm for AVHRR fire detection. *International Journal of Remote Sensing*, Vol. 17, No. 2, pp. 419–424.
- Florio, E.N., Lele, S.R., Chang, Y.C., Sterner, R., and Glass, G.E. 2004. Integrating AVHRR satellite data and NOAA ground observations to predict surface air temperature: a statistical approach. *International Journal of Remote Sensing*, Vol. 25, No. 15, pp. 2979–2994.
- Gates, D.M. 1980. *Biophysical ecology*. Springer-Verlag, New York. 611 pp.
- Geiger, R. 1965. *Climate near the ground*. Harvard University Press, Cambridge, Mass. 611 pp.
- Goward, S.N., Waring, R.H., Dye, D.G., and Yang, J. 1994. Ecological remote sensing at OTTER: satellite macroscale observations. *Ecological Applications*, Vol. 4, No. 2, pp. 322–343.
- Holben, B.N. 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, Vol. 7, No. 11, pp. 1417–1434.
- Lakshmi, V., Czajkowski, K., Dubayah, R., and Susskind, J. 2001. Land surface air temperature mapping using TOVS and AVHRR. *International Journal of Remote Sensing*, Vol. 22, No. 4, pp. 643–662.
- Nemani, R.R., and Running, S.W. 1989. Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *Journal of Applied Meteorology*, Vol. 28, No. 4, pp. 276–284.
- Nemani, R., Pierce, L., Running, S., and Goward, S. 1993. Developing satellite-derived estimates of surface moisture status. *Journal of Applied Meteorology*, Vol. 32, No. 3, pp. 548–557.
- Ouaidrari, H., Goward, S.N., Czajkowski, K.P., Sobrino, J.A., and Vermote, E. 2002. Land surface temperature estimation from AVHRR thermal infrared measurements: an assessment for the AVHRR Land Pathfinder II data set. *Remote Sensing of Environment*, Vol. 81, pp. 114–128.
- Prihodko, L., and Goward, S.N. 1997. Estimation of air temperature from remotely sensed surface observations. *Remote Sensing of Environment*, Vol. 60, pp. 335–346.
- Prince, S.D., Goetz, S.J., Dubayah, R.O., Czajkowski, K.P., and Thawley, M. 1998. Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using advanced very high-resolution radiometer satellite observations: comparison with field observations. *Journal of Hydrology*, Vol. 212, No. 213, pp. 230–249.
- Riddering, J.P., Seielstad, C.A., and Queen, L.P. 1999. Developing a computationally efficient fire potential index from satellite derived estimates of surface moisture status. In *From Image to Information: Proceedings of the American Society of Photogrammetry and Remote Sensing Annual Conference*, 17–21 May 1999, Portland, Oreg. American Society of Photogrammetry and Remote Sensing (ASPRS), Bethesda, Md. pp. 296–307.
- Seielstad, C.A., Riddering, J.P., Brown, S.R., Queen, L.P., and Hao, W.M. 2002. Testing the sensitivity of a MODIS-like daytime active fire detection model in Alaska using NOAA/AVHRR infrared data. *Photogrammetric Engineering and Remote Sensing*, Vol. 68, No. 8, pp. 831–838.

- Stoms, D.M., Bueno, M.J., and Davis, F.W. 1997. Viewing geometry of AVHRR image composites derived using multiple criteria. *Photogrammetric Engineering and Remote Sensing*, Vol. 63, No. 6, pp. 681–689.
- Thornton, P.E. 1998. *Regional ecosystem simulation: combining surface- and satellite-based observations to study linkages between terrestrial energy and mass budgets*. Ph.D. dissertation, University of Montana, Missoula, Mont. 280 pp.
- Thornton, P.E., Running, S.W., and White, M.A. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, Vol. 190, pp. 214–251.
- Ulivieri, C., Castronuovo, M.M., Francioni, R., and Cardillo, A. 1994. A split window algorithm for estimating land surface temperature from satellites. *Advances in Space Research*, Vol. 14, No. 3, pp. 59–65.
- US Geological Survey. 2000. *North American land cover characteristics data base*. US Geological Survey, Washington, D.C. Available from [http://edcdaac.usgs.gov/glcc/na\\_int.asp](http://edcdaac.usgs.gov/glcc/na_int.asp) [accessed October 2000].
- van de Griend, A.A., and Owe, M. 1994. Bare soil surface resistance to evaporation by vapor diffusion under semiarid conditions. *Water Resources Research*, Vol. 30, No. 2, pp. 181–188.
- Vazquez, D.P., Reyes, F.J.O., and Arboledas, L.A. 1997. A comparative study of algorithms for estimating land surface temperature from AVHRR data. *Remote Sensing of Environment*, Vol. 62, pp. 215–222.
- Vidal, A., and Devaux-Ros, C. 1995. Evaluating forest fire hazard with a LANDSAT TM derived water stress index. *Agricultural and Forest Meteorology*, Vol. 77, pp. 207–224.
- $T_x, T_n$  daily maximum and minimum air temperatures, respectively
- $T_4$  AVHRR channel 4 brightness temperature
- $T_5$  AVHRR channel 5 brightness temperature
- $\Delta T_{45}$  brightness temperature difference between channels 4 and 5
- $\Delta T_{34}$  brightness temperature difference between channels 3 and 4
- $z$  solar zenith angle
- $\epsilon$  mean emissivity of  $\epsilon_4$  and  $\epsilon_5$
- $\epsilon_4$  estimated emissivity of AVHRR channel 4
- $\epsilon_5$  estimated emissivity of AVHRR channel 5
- $\Delta\epsilon$  emissivity difference from channels 4 and 5
- $\omega$   $\pi/12$
- $\Gamma_{(t)}$  dimensionless diurnal temperature function at time  $t$

## List of symbols

- $B$   $Y$  intercept
- decl earth's rotational axis declination relative to the principal plane
- $h$  hour angle ( $0^\circ$  is local solar noon)
- $i$  present day
- $l$  latitude of pixel
- $m$  slope of the regression line
- $n$  number of observations
- $Q$  ratio of near-infrared (NIR) reflectance (channel 2) to red reflectance (channel 1)
- $r^2$  coefficient of determination
- $R$  correlation coefficient
- $t$  time of day (hours, 12 at solar noon)
- $T_a$  air temperature
- $T_{\text{est}}$  estimated air temperature
- $T_s$  surface temperature
- $T_{\text{so}}$  time of satellite overpass
- $T_{(t)}$  DAYMET temperature at time of satellite overpass