Evaluation of Disaggregation Methods for Downscaling MODIS Land Surface Temperature to Landsat Spatial Resolution in Barrax Test Site

Mar Bisquert, Juan Manuel Sánchez, and Vicente Caselles

***Abstract*—Thermal infrared (TIR) data are usually acquired at a coarser spatial resolution (CR) than visible and near infrared (VNIR). Several disaggregation methods have been recently developed to enhance the TIR spatial resolution using VNIR data. These approaches are based on the retrieval of a relation between TIR and VNIR data at CR, or training of a neural network, to be applied at the fine resolution afterward. In this work, different disaggregation methods are applied to the combination of two different sensors in the experimental test site of Barrax, Spain. The main objective is to test the feasibility of these techniques when applied to satellites provided with no TIR bands. Landsat and moderate imaging spectroradiometer (MODIS) images were used for this work. Land surface temperature (LST) from MODIS images was disaggregated to the Landsat spatial resolution using Landsat VNIR data. Landsat LST was used for the validation and comparison of the different techniques. Best results were obtained by the method based on a linear regression between normalized difference vegetation index (NDVI) and LST. An average $\mathrm{RMSE} = \pm 1.9$ K was observed between disaggregated and Landsat LST from four different dates in a study area of 120 km^2.**

***Index Terms*—Image enhancement, image resolution, remote sensing, temperature.**

I. Introduction

TIME series of fine spatial and temporal resolution images are key inputs in numerous studies, e.g., water resources management [1], [2]. However, there is a limitation in the existing satellites since revisit time for fine spatial resolution sensors is typically poor, while those with a high revisit frequency are characterized by a coarse spatial resolution. This is especially true when focusing on the thermal infrared (TIR) since spatial resolution for the TIR bands is always coarser than that for the visible and near infrared (VNIR) bands onboard the same sensor [2].

Disaggregation methods allow downscaling the TIR coarse resolution (CR) to finer resolutions. In [3], a review of land surface temperature (LST) disaggregation methods is performed. Zhan *et al.* [3] classified the disaggregation methods in two main groups: thermal sharpening (TSP) and temperature unmixing (TUM). The main difference is that TSP is used to obtain the LST of smaller resolution cells, while TUM aims at obtaining the LSTs of the existing elements within CR cells [3]. A bibliographical review showed that TSP is more frequently used than TUM. The TUM approaches need sufficient information of the component temperatures and become more useful when focused on obtaining temperatures of different surfaces (e.g., soil vs. vegetation, and asphalt vs. bare soil). In [4], a physics-based unmixing method was presented to estimate the relative proportion and the temperature of each material composing the mixed pixel. Emissivity and temperature over pure pixels of the different components are required in this method. The number of components has to be previously identified and a classification image is needed. Note this method is constrained to the existence of pure CR pixels of each element present in the fine resolution (FR) images.

Several TSP techniques have been proposed in the recent literature to enhance the spatial resolution of the TIR domain over vegetated areas by linking TIR and reflectance information [5]–[8]. Most of these techniques are based on the assumption that there exists a relation between the vegetation cover and the LST. According to these approaches, a relation between TIR and VNIR bands is first obtained at CR and then applied at the finer resolution of the VNIR bands, assuming that this relation is scale invariant. Kustas *et al.* [5] developed the disTrad (disaggregation procedure for radiometric surface temperature) based on a quadratic relationship between normalized difference vegetation index (NDVI) and LST. This technique was used in [9] to downscale TIR data to the VNIR resolution with Landsat and moderate imaging spectroradiometer (MODIS) images. Agam *et al.* [6] tested three new variants of the disTrad and the results from the four methods were compared. The disTrad variants were: a linear relationship between NDVI and LST, a linear relationship between fractional vegetation cover (fc) and LST, and a simplified version of the fc-LST variant (hereafter called TsHARP). Since results might depend on the land surface cover, TsHARP was applied separately to crop and natural vegetation [10]. The relation between NDVI and LST was too poor in the natural vegetation. In the crop area, the disaggregation led to better results compared to a nondisaggregation method uniTr (using the LST value of the CR image). Lower errors were found when applying the

Manuscript received May 06, 2015; revised January 08, 2016; accepted January 12, 2016. This work was supported in part by the Spanish Ministry of Economy and Competitiveness (projects CGL2013-46862-C2-1/2-P, cofinanced with European Union FEDER funds), and in part by the Generalitat Valenciana (project PROMETEOII/2014/086).

M. Bisquert and J. M. Sánchez are with the Department of Applied Physics, Polytechnic School of Cuenca, University of Castilla-La Mancha, Cuenca 16071, Spain (e-mail: mar.bisquert@uclm.es; juanmanuel.sanchez@uclm.es).

V. Caselles are with the Department of Earth Physics and Thermodynamics, University of Valencia, Valencia 46100, Spain (e-mail: vicente.caselles@uv.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2016.2519099

TsHARP to the entire scene than when focusing on the crop areas. Neural networks (NNs) were applied in [11] and [12] using land cover and several land surface parameters obtained from VNIR images as inputs (different combinations for each land cover). Jeganathan *et al.* [7] tested four different versions of the TsHARP: resolution-adjusted TsHARP, piecewise (a regression per NDVI intervals [0–0.2], [0.2–0.5], [0.5–1]), stratified (applied per land cover class), and local (applied within a moving window). Better results were obtained with the local disaggregation procedures. Gao *et al.* [13] developed a data mining (DM) approach between TIR and spectral reflectances of homogeneous pixels applied separately at global and local (within a moving window) scales and subsequently combined. Bindhu *et al.* [8] developed a nonlinear method (NL-disTrad) based on the estimation of a relationship between NDVI and LST from the pixels belonging to the hot edge (pixels forming the upper envelope of the NDVI–LST distribution). The relationship obtained is applied to CR pixels to calculate the difference between the original and the predicted LST (residuals). These residuals are trained in an NN using the NDVI values of a 3×3 window as inputs. Then, the trained NN is applied to the FR pixels to obtain the residuals at this FR. A random forest approach was introduced in [14] using VNIR bands, topography, and land cover classes as inputs. Chen *et al.* [15] applied a combination of TsHARP and Thin Plate Spline interpolation. Both methods were applied separately and weights were then calculated for each one and further combined. Mukherjee *et al.* [16] compared three of the previous approaches (disTrad, TsHARP, and TsHARP with local variant) to two new methods based on the adjustment of the linear regression between NDVI and LST using a least median square (LMS) regression and a pace regression. In [17], disaggregated microwave brightness temperatures were obtained at 1-km MODIS resolution. These authors established a relationship between NDVI and the microwave polarization difference index from the Advance Microwave Scanning Radiometer–Earth Observing System (AMSR-E). Disaggregated temperatures were then used as inputs in a soil moisture algorithm.

Performance of the different methods above is not comparable since results may be site dependent. Different conclusions and root-mean-square error (RMSE) values are obtained in different works. Most of these studies compare the results of the new proposed methods to the TsHARP, but even with TsHARP results are quite different depending on the study area and the spatial resolution of the sensors used. The majority of papers focus on Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), MODIS or Landsat images, with the objective of disaggregating from their TIR to VNIR resolution, e.g., from MODIS 1 km to MODIS 250 m. Few papers plan to apply these methods to data from various satellites, e.g., obtaining a regression from one sensor and applying it to another one provided with no TIR bands. Up to now most of these works simulated CR images by aggregating FR scenes to the coarser spatial resolution of a different satellite, using afterward these FR images to assess the disaggregation procedure. A typical example is aggregating Landsat bands to simulate MODIS, then disaggregation and validation using the Landsat original bands as a basis. This is the traditional procedure for the assessment of the disaggregation methods. However, if these methods are to be applied from two different sensors, the validation should also be based on the original data from both sensors. Very few papers have dealt with the real application from two satellites [8]. Mukherjee *et al.* [16] used both Landsat and MODIS images as the basis for the disaggregation but with no mixing between them. Better performance was observed using Landsat [16]. Some works in the literature analyze the effect of the spatial resolution on the disaggregation. Better results are obtained for smaller differences in the spatial resolution between the input and output images. For example, in [6], Landsat images were aggregated to 960 m and then disaggregated to 240, 120, and 60 m with RMSE values ranging from ± 0.7 K (240 m) to ± 2.2 K (60 m).

In this research, we are interested in providing LST at fine spatial and temporal resolutions to fulfill the requirements in the estimation of surface energy fluxes and evapotranspiration. Several study fields are situated within the experimental test site of Barrax, Spain. At least 30-m spatial resolution is desired. This can be provided by the Landsat VNIR bands. However, revisit cycle of Landsat is poor and a higher frequency is required. The recently launched Sentinel-2a has a 10-day repeat cycle, and 10–20 m spatial resolution in the VNIR bands, whereas no TIR information is available. With the coming Sentinel-2b, the combination of both satellites will offer a 5-day repeat cycle. Furthermore, the combination of Sentinel-2 and Sentinel-3 could offer the desired solution of spatial and temporal resolutions. The relationship between TIR and VNIR bands could be extracted from Sentinel-3 and then applied to Sentinel-2. The aim of this paper is to test the application of disaggregation techniques in the Barrax area from two different satellites, so that these approaches can be applied to sensors without thermal bands (e.g., Sentinel-2). Landsat and MODIS imagery were selected for this work due to the similar characteristics of these sensors and Sentinel-2 and Sentinel-3. We analyzed both, classical methods based on the VNIR-LST regression and newer methods using NN and DM. Three main experiments were developed to evaluate the classical approaches. A first experiment focused on the methods developed in [5], [6], and [10] based on the least square (LS) adjustment of the VNIR-LST regression. In a second experiment, new adjustments were tested (LMS) as well as the application of the method at local scale (using a moving window). In the last experiment, the disaggregation was applied separately for the different land covers present in the study area. The classical method leading to the best results was then compared to NN and DM approaches.

II. Materials and Methods

A. Study Site

The study area is located in Barrax, Central Spain, including "Las Tiesas" experimental farm (39°03' 35" N, 2°06' W). This is a very flat area with an average altitude of 700 m a.s.l close to Albacete. Barrax is one of the traditional ESA test sites for different international campaigns, such as SEN2FLEX

[18], SPARC [19], ImagineS [20], and DAISEX [21]. Several field campaigns have been carried out in recent years including ground measurements of LST and energy fluxes, as well as biophysical parameters, for different periods. These experimental campaigns are used with different objectives related to agricultural water management [22]–[26].

B. Satellite Images

The selection of satellite images was constrained to the coincidence of Landsat and MODIS overpasses with low MODIS viewing angle and cloud coverage for the growing season of 2014. We focused on Landsat 7 since no Landsat 8 images matched our requirements for the selected period. Four images were finally selected in 2014 for DOYs 67, 163, 195, and 202.

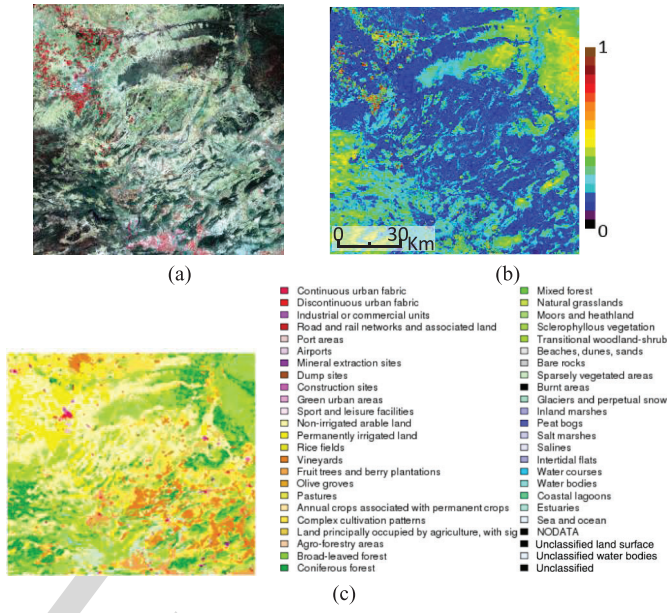
VNIR data were extracted from the Landsat 7 ETM+ CDR (http://earthexplorer.usgs.gov/) and the MODIS MOD09GA and MOD09GQ products (glovis.usgs.gov). Landsat CDR provides 30-m surface reflectances (atmospherically corrected). MOD09GQ offers red and NIR bands at 250-m spatial resolution, and MOD09GA contains seven spectral bands (VNIR), together with quality information, at a resolution of 500 m. Radiances in the TIR domain were obtained from band 6 of Landsat 7 ETM+, while the MOD11_L2 product offers LST directly. ETM+ TIR data are acquired at 60-m spatial resolution; however, these data are resampled and provided at 30-m resolution. For validation purposes, we used an aggregated LST at 60 m that corresponds to the original resolution in the TIR domain. MOD11_L2 is provided at 1 km. MODIS images were resampled to 240, 480, and 960 m in order to have pixel dimensions that are multiple of the Landsat spatial resolutions (30 and 60 m). All scenes were reprojected to UTM WGS 1984 zone 30 N.

Fig. 1 shows a false color composite of a Landsat image used in this work (RGB: bands 4, 3, 2), an NDVI image, and the Corine Land Cover 2006 (http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster) covering the entire study area ($\sim$ 120km^2). The images illustrate a sparsely vegetated area, with NDVI < 0.4 for most pixels. Largest NDVI values correspond to crop fields [red areas in Fig. 1(a) and (b)].

Table I lists a summary of the NDVI and LST values from the MODIS and Landsat images used at both, their original and aggregated 960-m spatial resolution. Note that wider value ranges are present in the original resolution compared to the aggregated 960-m resolution.

C. Image Processing

Quality bands from each product were used to mask pixels containing clouds and shadows. Landsat CDR product includes a mask obtained using the Fmask code [27]. Based on this mask only, pixels assigned "clear land pixel" were kept in this study, discarding those with clouds, snow, shadows, and water. TIR Landsat data were atmospherically corrected to retrieve LST following the procedure described in [28] and using the Atmospheric Correction Tool of Barsi *et al.* [29] and [30] (http://atm-corr.gsfc.nasa.gov/).



F1:1
Fig. 1. (a) Landsat false color image (RGB: bands 4, 3, 2). (b) NDVI corresponding to DOY 195. (c) Corine Land Cover of the study area. F1:2

The data aggregation is a key step in the disaggregation procedures. The aggregation of the VNIR bands was carried out by averaging the reflectance values in the red and NIR bands of all the FR pixels within an equivalent CR pixel. Following Gao *et al.* [11], the aggregation of the TIR band was done through the Stefan-Boltzmann law with the assumption of similar emissivity values for adjacent pixels Q1

$$T_{agg} = \sqrt[4]{(1/n)\sum_{i=1}^{n} T_i^4}. \quad (1)$$

The application of the disaggregation techniques to different sensors needs equivalent spectral data from both sensors. Differences between both sensors VNIR data may exist due to several effects such as spectral resolution, atmospheric correction, viewing angle, and pixel footprint. In the TIR domain, the main difference between both sensors may be due to the different acquisition time. A normalization procedure can be applied to minimize these discrepancies between MODIS and Landsat images [31], [32]. A linear regression between the aggregated Landsat and MODIS images at 960 m was obtained, and later applied to the desired spatial resolution (e.g., 60 m). This normalization step should be applied to each pair of FR and CR images to be used. Note that in this paper, we used Landsat ETM+ and MODIS images coincident in date with the aim of using the Landsat TIR band to validate the disaggregated temperatures. However, the disaggregation method could be applied to images from different dates if they are close enough in time to consider that no significant changes in the vegetation cover have occurred. In case of using images from different dates, the same normalization procedure should be applied to each pair of FR and CR images, and the disaggregated temperature obtained will correspond to the date of the CR image (from which the VNIR-TIR relationship is obtained).

T1:1 TABLE I
T1:2 NDVI AND LST AVERAGE, MAXIMUM AND MINIMUM VALUES FOR THE MODIS AND LANDSAT IMAGES AT THEIR ORIGINAL RESOLUTION (30
T1:3 AND 60 M) AND THE AGGREGATED 960-M RESOLUTION, FOR TWO DATES (DOYS 67 AND 195)

Doy	Image	NDVI					LST				
		Median	Max	Min	Q25	Q75	Median	Max	Min	Q25	Q75
67	Landsat 30 m / 60 m (NDVI / LST)	0.27	0.97	0.0001	0.19	0.4	294	312	280	291	296
	Landsat 960 m	0.25	0.7	0.09	0.2	0.34	294	303	285	292	296
	MODIS 240 m	0.29	0.91	0.002	0.22	0.39					
	MODIS 960 m	0.29	0.69	0.07	0.22	0.38	296	305	287	294	298
195	Landsat 30m / 60m (NDVI / LST)	0.21	0.9	0.0001	0.15	0.31	316	331	288	313	319
	Landsat 960 m	0.21	0.6	0.1	0.17	0.29	316	325	303	312	318
	MODIS 240 m	0.22	0.87	0.001	0.17	0.37					
	MODIS 960 m	0.22	0.72	0.02	0.18	0.3	319	327	303	317	321

Percentiles Q25 and Q75 indicate that 25% or 75%, respectively, of the total data are lower than these values.

D. Disaggregation Methods

Several disaggregation methods extracted from the recent literature were applied in this work with images from two different satellites. These methods are traditionally applied and tested with images from a single satellite. First, the FR images are aggregated to a CR and then the disaggregation method is applied, finally the original TIR images are used for validation. In this work, we tested the disaggregation from two different sensors. The further objective is assessing the performance of the disaggregation when applied to FR images from satellites with no thermal bands. To compare the application with one or two sensors, the two versions were tested here as follows. 1) The relationship between NDVI and LST (training) was extracted from the Landsat ETM+ aggregated images (960 m) and later applied to Landsat 60 m (traditional procedure). 2) The relationship was obtained from the original MODIS images (960 m) and then applied to Landsat 60 m images. Both procedures provided disaggregated 60 m LST using 60 m Landsat NDVI as input. The original 60 m Landsat LST was reserved for the validation of the disaggregated LST outputs.

1) NDVI–LST Regressions: Fig. 2 shows a flowchart of the methodology.

Experiment 1 consisted on applying the most used disaggregation methods developed in [5], [6], and [10]. In [5], a quadratic relationship between LST and NDVI was proposed (disTrad, 2). The relationship was obtained from CR pixels and applied to the FR pixels. The relationship obtained was also applied at CR to obtain the difference between the original and predicted LST (residuals). The CR residuals were then added to the estimated LST at FR; this means that the same residual was used in all the FR pixels belonging to the same CR pixel. With this action actual CR LST information was included in the disaggregated results. In [6], the quadratic relationship and other relationships between NDVI and LST were tested, including a linear approach between LST and NDVI (3), a linear approach between the fraction of vegetation cover (fc, expressed in terms of NDVI) and LST (4), and a simplified version of the fc approach (5). As suggested by [5], all regressions in this Experiment 1 were adjusted from the most homogeneous pixels in three NDVI intervals: [0–0.2], [0.2–0.5], and [0.5–1]. The residual correction was included in all the cases considered in this Experiment 1

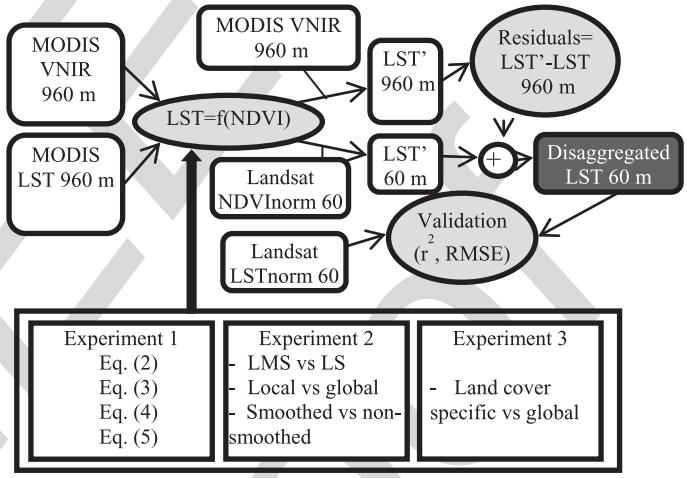


Fig. 2. Flowchart of the methodology applied, including the different processing steps and the disaggregation methods used. F2:1 F2:2

$$LST = a_0 + a_1 NDVI + a_2 NDVI^2 \quad (2)$$

$$LST = a_0 + a_1 NDVI \quad (3)$$

$$LST = a_0 + a_1 \left(1 - \left(\frac{NDVI_{\max} - NDVI}{NDVI_{\max} - NDVI_{\min}}\right)^{0.625}\right) \quad (4)$$

$$LST = a_0 + a_1 (1 - NDVI)^{0.625} \quad (5)$$

where a_0, a_1, and a_2 are the adjusted parameters.

More recent studies introduce some modifications to the disaggregation procedures described above. These were analyzed in Experiment 2. Note that only the method yielding the

best results in Experiment 1 was considered for Experiment 2. In [16], better results were obtained using LMS than LS regression. Thus, the LMS adjustment was tested and compared to the LS. The disaggregation methods using the residual correction present some boxy effect linked to these residuals themselves [10], [13]. To reduce this effect, the CR residuals can be smoothed using a Gaussian filter [33]. We adopted this technique and analyzed its effect visually and in terms of RMSE. All these regressions evaluated in Experiment 2 were obtained from the most homogeneous pixels and including the residual correction, as done in Experiment 1. In [7], (3) was applied at a local scale using a moving window. This local application showed better results than obtaining a global regression for the entire image. This approach consisted in obtaining a relationship between NDVI and LST at CR for each window and then applying it to the FR pixels belonging to that window. Each pixel is only considered in one window. The result for each pixel is then the LST obtained from the NDVI–LST relationship adjusted at the CR window to which it belongs. We also tested this moving window approach with different window sizes (from 3×3 to 13×13) and compared it to the global approach. In contrast to the previous methods, since few pixels are available in the local application all of them were used for adjusting the regression and no residual correction was applied due to the local scale.

Some authors have used the land cover information to derive different regressions for each particular land cover class [7], [10]. Several tests were also conducted in this work for the different land cover classes (Experiment 3) present in the images. Corine Land Cover data at 100-m resolution was used. Land cover information was needed at 60 and 960 m. For the 60-m resolution, a resampling was performed using nearest neighbor resampling. For 960 m, the predominant class (the most abundant) was assigned. The disaggregation method leading to the best results in the previous analyses (Experiments 1 and 2) was applied here specifically per land cover classes. The NDVI–LST relation was obtained considering all the pixels included in each land cover class at CR and then applied to the pixels of the same class at FR. The residuals correction was also accounted here. An analysis of the performance of the disaggregation technique was conducted using the land cover class distinction as a basis. Results were compared to those obtained not accounting for this land cover distinction.

2) NN and DM Approaches: NNs using different inputs (spectral bands, land cover, topography, etc.) have also been applied in several papers. The method developed by Bindhu *et al.* [8] was applied in this work. These authors used NN for training the residuals. The relationship between NDVI and LST was obtained from the pixels forming the hot edge. These pixels were selected as those showing the highest temperature values for different NDVI ranges. This relationship was then applied to the CR pixels to estimate the residuals. The CR residuals were trained in a NN which inputs were the NDVI values in a 3×3 pixel window. The trained NN was then applied at FR. Other tests performed with NN consisted in using different combinations of spectral bands, and in some cases land cover, as inputs in the NN, and the LST as output.

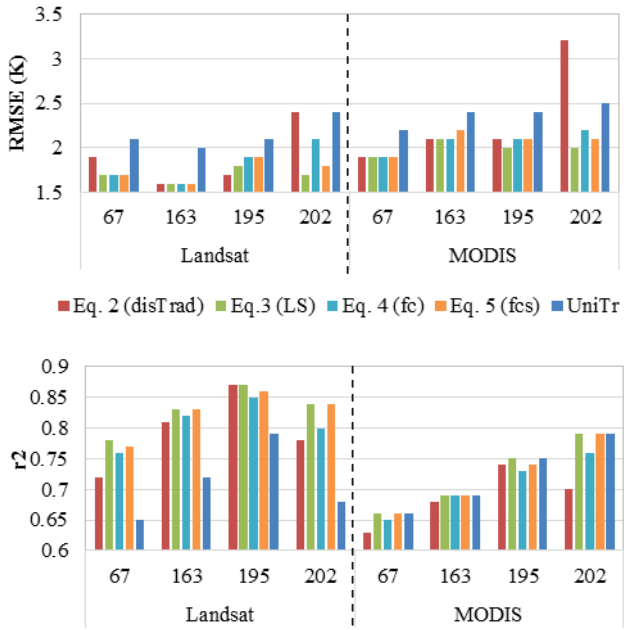


Fig. 3. Single- and dual-sensor analysis of the validation results (RMSE and r2) of the different methods (1)–(4). On the left, Landsat images were first degraded to the MODIS resolution and then disaggregated to the Landsat TIR original resolution. On the right, the relationship was obtained from the MODIS images and then applied to the Landsat images at 60-m spatial resolution. The residual correction was applied in both cases.

The DM approach developed in [13] was also tested here. This method combines a local and global application of regression trees in a DM approach. The Cubist package in R [34] was used in this work. This method uses the reflectance from all the bands. Thus, when applied to two different sensors, their bands have to be the same or similar, and the normalization of each band between both sensors has to be done. The regression tree method generates rule-based linear multivariate regressions. In this work, the 6 Landsat ETM+ VNIR bands (1–5 and 7), and the equivalent MODIS bands (bands 1–6) were used as inputs.

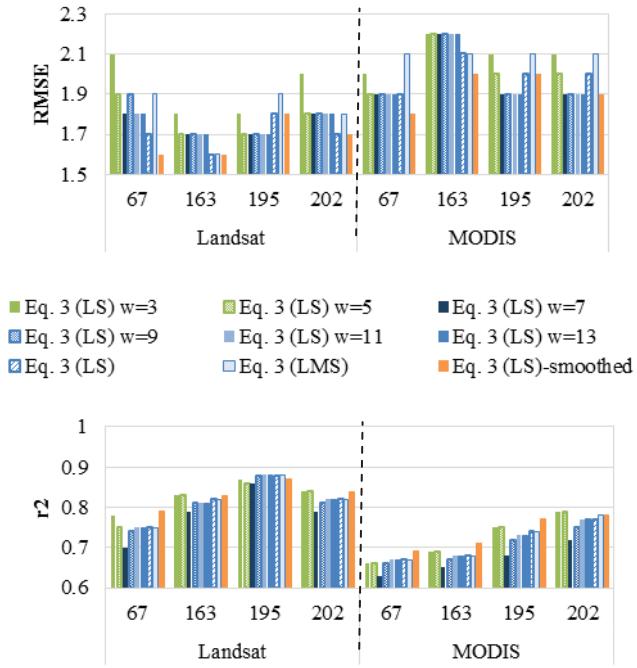
E. Performance Assessment

The comparison of the disaggregation methods was carried at a regional scale. All methods were analyzed in terms of the determination coefficient (r2) and the RMSE, and compared to a nondisaggregation method, uniTr [6]. According to this approach, each FR pixel is assigned at the value of the corresponding CR pixel.

III. Results

A. NDVI–LST Regressions

In the first experiment (Fig. 2), the methods presented in [6] were applied separately to a single sensor (Landsat ETM+) and to the combination of two sensors (Landsat and MODIS) as described above. Fig. 3 shows the validation results (RMSE and r2), for the 120-km^2 study area, corresponding to the regression obtained from both, Landsat aggregated images (960 m) on one hand and original MODIS images on the other hand, then



F4:1 Fig. 4. Validation results (RMSE and r2), of the local disaggregation method
F4:2 with different windows sizes (w) compared to the global application and to the
F4:3 global application with smoothed residuals.

applied to the original 60-m Landsat images. Previous works have tested the disaggregation methods with images belonging to the same sensor (by aggregation to coarser resolution) with the aim of being applied with two different sensors, but they rarely showed the real dual-sensor application. Results presented in Fig. 3 point out the larger errors obtained when these techniques are applied to two different sensors instead of a single one, despite the normalization adjustment applied. Better results were obtained using any of the disaggregation methods compared to the uniTr approach (using the LST of the CR image) in both the single and dual-sensor applications. Contrary to previous studies, the simplest method [linear regression between NDVI and LST, (3)] led to similar or better results than using other regressions in most of the cases analyzed. Equation (2) did not perform well for day 202 since the quadratic regression applied to NDVI outliers not present in the primitive retrieval of this equation yielded extreme LST values. Also, the NDVI–LST adjustment in this image was better fitted to a linear regression than to the quadratic regression. The same conclusions can be drawn from the method applied with one or two sensors despite the different errors obtained.

In the second experiment, the local application [7] was tested with a window size ranging from 3×3 to 13×13 pixels, and results were compared to those obtained from the global application using (3), (3) adjusted by LMS [16], and (3) with smoothed residuals (Fig. 4). Fig. 4 shows that lower errors were always obtained when applying the disaggregation with a single satellite. Regarding the local application, only the RMSE was slightly reduced in one of the images (DOY 195) when using one sensor (Landsat ETM+) and in two dates (DOYs 195 and 202) when using two sensors (ETM+ and MODIS), while

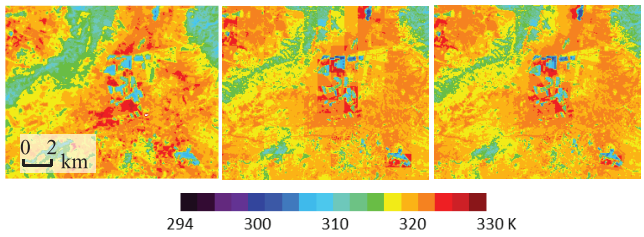
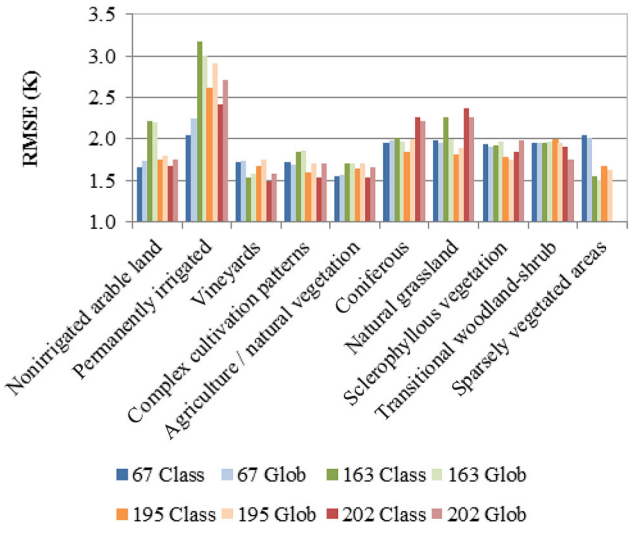


Fig. 5. Landsat LST (left), disaggregated LST with (3) LS (center), disaggregated with smoothed residuals (right) for DOY 195. F5:1 F5:2

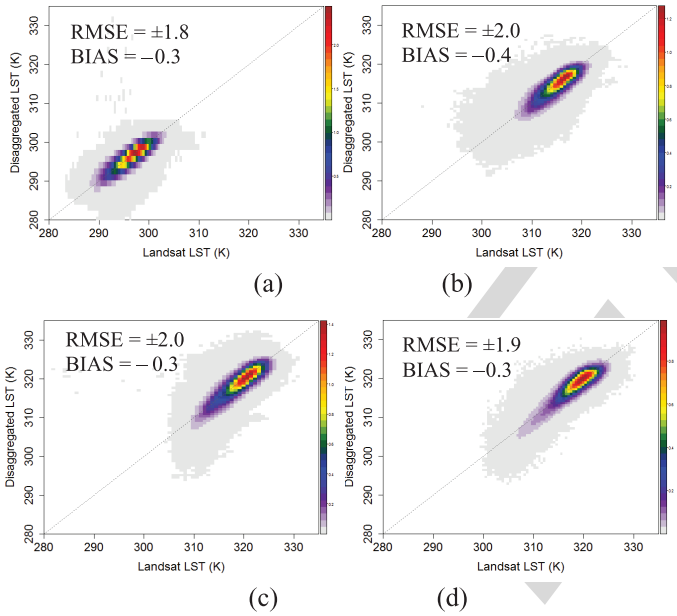
no significant differences were observed for the rest. The LMS adjustment did not improve the results from the LS adjustment. The smoothing of the residuals generally improved the results from the LS method when using two sensors. When applied to a single sensor, it improved the results in one date (67) and led to the same results for the other dates. Fig. 5 shows an example of the original Landsat LST and disaggregation results from two sensors, with and without residuals smoothing. Note the residuals introduced an important boxy effect that was reduced after the smoothing. The ratio between the CR and the FR is responsible for this boxy effect, lower differences between the spatial resolutions would mitigate this boxy effect.

Finally, specific disaggregation was conducted in Experiment 3 for different land cover classes. Only the dual-sensor application is shown here due to the similar conclusions obtained from the single- and dual-sensor applications in the previous experiments. Results considering separate land covers were compared to those obtained from applying the global equation to the same set of pixels. We focus on the linear equation LST–NDVI (3) and smoothed residuals, based on the results above. Following Agam *et al.* [10], two main classes were first considered: crop and natural vegetation. No significant differences, compared to the global results, were observed accounting for this distinction. The specific disaggregation was expanded to all vegetation classes present in the image with a significant presence. Classes with less than 100 CR pixels were excluded from the analysis. Detailed class-dependent results are shown in Fig. 6. Performance depended on the particular land cover class and date, and no firm conclusion could be extracted about the value of separating in classes. Average RMSE values of 1.8, 2.2, 1.9, and 2.0 K were obtained with the class-specific approach and 1.8, 2.0, 2.0, and 1.9 K with the global approach for DOYs 67, 163, 195, and 202, respectively. Only for DOY 195 was a slight enhancement observed with the class-specific approach. Note that larger errors were observed overall for "permanently irrigated" areas. This is probably due to the small size of these fields and also to irrigation effects. After an irrigation event, the LST may decrease dramatically, whereas the disaggregation methods do not account for this variation since the NDVI is not significantly affected. This effect is most obvious in bare soil areas such as croplands in the first stage of the crop growing season. This should be mitigated by the residuals obtained from the CR image but if the fields are small this is not well captured either.

Since no significant enhancement was observed from the land cover class-specific application, we decided to use (3)

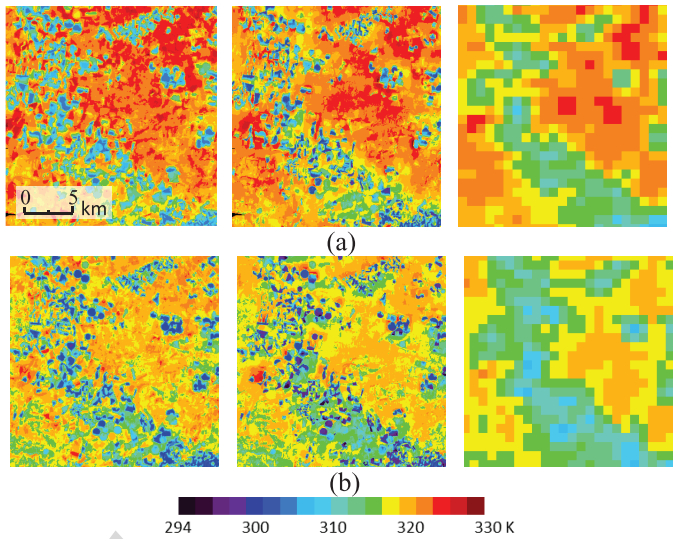


F6:1 Fig. 6. Validation results (RMSE) of class-specific disaggregation method
F6:2 (Class) compared to the global (Glob) application. Residuals have been
F6:3 smoothed in both.



F7:1 Fig. 7. Disaggregated LST with (3) LS and smoothed residuals versus Landsat
F7:2 LST for images corresponding to DOY. (a) 67. (b) 163. (c) 195. (d) 202.

485 globally together with smoothed residuals. Fig. 7 shows the
486 density scatter plots of the disaggregated LST with (3) LS and
487 smoothed residuals from two sensors versus the reference LST
488 (original Landsat normalized). The main percentage of pixels
489 (in red) concentrates around the 1:1 line although the large
490 scatter of a minority of them leads to important errors.
491 Fig. 8 shows a subset example of 24 km × 22 km of the
492 reference (left), disaggregated from two sensors (central) and
493 nondisaggregated uniTr (right) LST images. This subset is
494 dominated by permanently irrigated lands. Overall patterns
495 of the disaggregated LST are similar to reference LST. Note
496 the evident improvement of the disaggregation faced to the
497 nondisaggregated LSTs.



F8:1 Fig. 8. Subset (22 × 24 km^2) of the Landsat LST (left column) and disag-
F8:2 gregated LST with (3) LS and smoothed residuals (central column) and the
F8:3 nondisaggregated uniTr (right column) for images corresponding to DOYs.
F8:4 (a) 195. (b) 202. The gaps due to the SLC error have been filled for visual
F8:5 inspection of results using a triangulation method.

B. NN and DM Approaches

498

499 The methods presented in [8] and [13] were also tested
500 and compared to the results obtained using the linear regres-
501 sion with smoothed residuals. In the previous section, we have
502 already analyzed the difference between the single- and the
503 dual-sensor disaggregation, and similar conclusions can be
504 extracted when comparing different approaches. Therefore, in
505 this section only, the results of the dual-sensor application are
506 shown, which is the main interest of this work.
507 Results from the NN and DM approaches for the entire study
508 area are summarized in Table II as well as the results from the
509 LS method. Larger errors were observed when applying the NN
510 and DM approaches. Bindhu *et al.* [6] had already observed Q2
511 high errors for low NDVI values, dominant in our study site [see
512 Fig. 2(b)]. In addition to these methods, a battery of tests was
513 conducted with NN using different spectral bands as inputs with
514 no improvement in the results obtained by the linear regres-
515 sion in any case. NN and DM (trained at 960 m and applied
516 at 60 m), as well as other nonlinear regressions, have the risk
517 of overfitting and are very sensitive to noise in the samples. In
518 this work, the input data did not cover the full range of values
519 present at 60 m (see Table I), and the overfitting of the methods
520 may lead to important errors in the outliers not present in the
521 input data. Therefore, the simpler NDVI–LST linear regression
522 led to better results than the more sophisticated NN and DM
523 approaches.

IV. Discussion

524

525 Findings in this work show that there is a difference between
526 applying disaggregation methods from one or two sensors
527 although a normalization procedure is applied. The normaliza-
528 tion step is very important when dealing with two sensors in

TABLE II
COMPARISON, IN TERMS OF RMSE (K), OF THE VALIDATION RESULTS OF THE NN DISAGGREGATION METHOD, AND THE LINEAR DISAGGREGATION METHOD (3) WITH SMOOTHED RESIDUALS

DOY	67	163	195	202
DM (Gao *et al*. 2012)	2.6	2.5	2.3	2.5
NN (Bindhu *et al*. 2014)	2.2	2.5	2.3	3.1
LST -NDVI regression	1.8	2.1	1.9	2.0

order to account for the possible spectral and temporal differences between them. A good coregistration of the images from both sensors is also critical. These are the two main points that make the difference when applying the disaggregation with one or two sensors. The normalization procedure was addressed in this work at 960 m; however, at this spatial resolution, we observed a shorter range of values than those present at the FR (Table I). Consequently, the relationships obtained did not include values in the full FR range. This not only affects the normalization process but also affects the NDVI–LST relationship used in the disaggregation procedure. This difference in the spectral range at the FR and CR is site dependent and is related to the heterogeneity of the area and to the ratio between the FR and the CR being used. Thus, this constraint is common to the single- and dual-sensor applications and it may contribute to the different errors obtained by other authors when disaggregating to different spatial resolutions. For example, in [6], the RMSE errors are very different depending on the spatial resolution and dates. These authors disaggregated two images, affected by 1 week difference, from 960 to 60 m in July 2002. The linear NDVI–LST approach yielded RMSE values of ± 2.2 and ± 1.2 K. Disaggregation to 120 m led to an enhancement in RMSE of ± 0.3 K in both images, and when disaggregating to 240 m, the enhancement was of ± 0.8 K in one image and ± 0.5 K in the other.

In our study area, RMSE values ranged between ± 1.6 and ± 1.7 K when using the Landsat images, and ± 1.8 and ± 2.1 K using the MODIS ones. These results are in agreement with those obtained by Agam *et al.* [6]. In [7], the local disaggregation of Aster images, previously aggregated to 990-m spatial resolution, resulted in RMSE values ranging between ± 0.7 and ± 3 K for disaggregation to 810–90 m. Large errors were obtained when disaggregating to spatial resolutions finer than 270 m, consequently MODIS images were disaggregated only to 250 m with RMSE values between ± 2 and ± 3 K depending on the approach. Better results were obtained by Bindhu *et al.* [8] in a smaller area with high vegetation cover, with the highest errors corresponding to low NDVI values.

The application of a land cover-specific disaggregation showed the highest errors for the class "permanently irrigated land." This is probably due to the small size of the fields and the errors caused by irrigation. This is reinforced by the fact that errors in this class increase significantly in the summer images, both with the class-specific and the global approaches. Most likely, this is a consequence of the higher differences in LST between the crop fields and the surrounding bare soil during summer season.

Future work will deal with a more comprehensive validation at a field scale using ground temperature measurements concurrent to satellite overpasses. Also, we will check the application to other combinations of platforms, including fine spatial resolution sensors with no TIR band (e.g., Sentinel 2).

V. CONCLUSION

Several disaggregation methods were compared when applied with two different sensors, MODIS and Landsat ETM+, in a heterogeneous and sparsely vegetated area of central Spain. This dual-sensor application needed an additional step of normalization between the spectral bands of the two sensors and special attention to the coregistration between them. Although the errors with the dual-sensor application were higher than those from the single-sensor application, results obtained prove the potential and usefulness of the disaggregation techniques applied to two different sensors.

The disaggregation was applied to downscale images from 960-m spatial resolution (MODIS) to 60 m (Landsat). Best results were obtained with the procedure based on the linear regression between NDVI and LST, even when compared to NN and DM approaches. RMSE values around ± 2.0 K were obtained for four different images. Focusing on the experimental agricultural area of Barrax, classed as "permanently irrigated land" in Corine Land Cover, RMSE values were the highest (± 2 to ± 3 K). Since this area showed the largest heterogeneity within the image, we may conclude that the disaggregation procedures perform better in homogeneous areas. These results are encouraging, and reinforce the application of disaggregation procedures to sensors provided with no thermal bands. The application to Sentinel-2, with a high revisit cycle, could provide the time series of disaggregated LSTs at high temporal and spatial resolutions required for studies related to hydrology, climatology or agriculture.

REFERENCES

[1] M. C. Anderson *et al.*, "Mapping daily evapotranspiration at landsat spatial scales during the BEAREX'08 field campaign," *Bushland Evapotranspiration Agric. Remote Sens. Exp.*, vol. 50, pp. 162–177, Dec. 2012.

[2] W. Ha, P. Gowda, and T. Howell, "A review of downscaling methods for remote sensing-based irrigation management: Part I," *Irrig. Sci.*, vol. 31, no. 4, pp. 831–850, Jul. 2013.

[3] W. Zhan *et al.*, "Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats," *Remote Sens. Environ.*, vol. 131, pp. 119–139, Apr. 2013.

[4] M. Cubero-Castan, J. Chanussot, V. Achard, X. Briottet, and M. Shimoni, "A physics-based unmixing method to estimate subpixel temperatures on mixed pixels," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1894–1906, Apr. 2015.

[5] W. P. Kustas, J. M. Norman, M. C. Anderson, and A. N. French, "Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship," *Remote Sens. Environ.*, vol. 85, no. 4, pp. 429–440, Jun. 2003.

[6] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and C. M. U. Neale, "A vegetation index based technique for spatial sharpening of thermal imagery," *Remote Sens. Environ.*, vol. 107, no. 4, pp. 545–558, Apr. 2007.

[7] C. Jeganathan, N. A. S. Hamm, S. Mukherjee, P. M. Atkinson, P. L. N. Raju, and V. K. Dadhwal, "Evaluating a thermal image sharpening model over a mixed agricultural landscape in India," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 13, no. 2, pp. 178–191, Apr. 2011.

[8] V. M. Bindhu, B. Narasimhan, and K. P. Sudheer, "Development and verification of a non-linear disaggregation method (NL-DisTrad) to downscale MODIS land surface temperature to the spatial scale of Landsat thermal data to estimate evapotranspiration," *Remote Sens. Environ.*, vol. 135, pp. 118–129, Aug. 2013.
[9] J. M. Sánchez *et al.*, "Monitoring surface energy fluxes at different spatial resolutions. Effects on fluxes variability in the Basilicata Italian region," in *Proc. Recent Adv. Quant. Remote Sens. Conf.*, Torrent, Spain, 2006.
[10] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and P. D. Colaizzi, "Utility of thermal sharpening over Texas high plains irrigated agricultural fields," *J. Geophys. Res.*, vol. 112, no. D19, Oct. 2007.
[11] G. Yang, R. Pu, W. Huang, J. Wang, and C. Zhao, "A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 2170–2178, Apr. 2010.
[12] G. Yang, R. Pu, C. Zhao, W. Huang, and J. Wang, "Estimation of subpixel land surface temperature using an endmember index based technique: A case examination on ASTER and MODIS temperature products over a heterogeneous area," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1202–1219, May 2011.
[13] F. Gao, W. Kustas, and M. Anderson, "A data mining approach for sharpening thermal satellite imagery over land," *Remote Sens.*, vol. 4, no. 12, pp. 3287–3319, Oct. 2012.
[14] C. Hutengs and M. Vohland, "Downscaling land surface temperatures from MODIS data to mesoscale resolution with Random Forest regression," Gemeinsame Tag. 2014 DGfK DGPF GfGI GiN DGPF Tagungsband 23/2014, 2014.
[15] X. Chen, W. Li, J. Chen, Y. Rao, and Y. Yamaguchi, "A combination of TsHARP and thin plate spline interpolation for spatial sharpening of thermal imagery," *Remote Sens.*, vol. 6, no. 4, pp. 2845–2863, Mar. 2014.
[16] S. Mukherjee, P. K. Joshi, and R. D. Garg, "A comparison of different regression models for downscaling Landsat and MODIS land surface temperature images over heterogeneous landscape," *Adv. Space Res.*, vol. 54, no. 4, pp. 655–669, Aug. 2014.
[17] C. Song, L. Jia, and M. Menenti, "Retrieving high-resolution surface soil moisture by downscaling AMSR-E brightness temperature using MODIS LST and NDVI data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 3, pp. 935–942, Mar. 2014.
[18] J. A. Sobrino *et al.*, "Thermal remote sensing in the framework of the SEN2FLEX project: field measurements, airborne data and applications," *Int. J. Remote Sens.*, vol. 29, nos. 17–18, pp. 4961–4991, Sep. 2008.
[19] J. F. Moreno *et al.*, "The SPECTRA Barrax campaign (SPARC): Overview and first results from CHRIS data," *presented at the Eur. Space Agency (Special Publication) ESA SP*, 2004, pp. 30–39.
[20] C. Latorre, F. Camacho, F. de la Cruz, R. Lacaze, M. Weiss, and F. Baret, "Seasonal monitoring of FAPAR over the Barrax cropland site in Spain, in support of the validation of PROBA-V products at 333 m," in *Proc. Recent Adv. Quant. Remote Sens. Conf.*, Torrent, Spain, 2014.
[21] M. Berger *et al.*, "The DAISEX campaigns in support of a future land-surface-processes mission," *ESA Bull.*, vol. 105, pp. 101–111, 2001.
[22] R. López-Urrea, F. M. de S. Olalla, C. Fabeiro, and A. Moratalla, "An evaluation of two hourly reference evapotranspiration equations for semi-arid conditions," *Agric. Water Manage.*, vol. 86, no. 3, pp. 277–282, Dec. 2006.
[23] R. López-Urrea, A. Montoro, J. González-Piqueras, P. López-Fuster, and E. Fereres, "Water use of spring wheat to raise water productivity," *Agric. Water Manage.*, vol. 96, no. 9, pp. 1305–1310, Sep. 2009.
[24] J. M. Sánchez, R. López-Urrea, E. Rubio, and V. Caselles, "Determining water use of sorghum from two-source energy balance and radiometric temperatures," *Hydrol. Earth Syst. Sci.*, vol. 15, no. 10, pp. 3061–3070, 2011.
[25] R. López-Urrea, A. Montoro, F. Mañas, P. López-Fuster, and E. Fereres, "Evapotranspiration and crop coefficients from lysimeter measurements of mature 'Tempranillo' wine grapes," *Agric. Water Manage.*, vol. 112, pp. 13–20, Sep. 2012.
[26] J. M. Sánchez, R. López-Urrea, E. Rubio, J. González-Piqueras, and V. Caselles, "Assessing crop coefficients of sunflower and canola using two-source energy balance and thermal radiometry," *Agric. Water Manage.*, vol. 137, pp. 23–29, May 2014.
[27] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, Mar. 2012.
[28] J. M. Sánchez, G. Scavone, V. Caselles, E. Valor, V. A. Copertino, and V. Telesca, "Monitoring daily evapotranspiration at a regional scale from Landsat-TM and ETM+ data: Application to the Basilicata region," *J. Hydrol.*, vol. 351, nos. 1–2, pp. 58–70, Mar. 2008.
[29] J. A. Barsi, J. L. Barker, and J. R. Schott, "An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS'03)*, 2003, vol. 5, pp. 3014–3016.
[30] J. A. Barsi, J. R. Schott, F. D. Palluconi, and S. J. Hook, "Validation of a web-based atmospheric correction tool for single thermal band instruments," 2005, pp. 58820E–1–58820E–7.
[31] G. le Maire, S. Dupuy, Y. Nouvellon, R. A. Loos, and R. Hakamada, "Mapping short-rotation plantations at regional scale using MODIS time series: Case of eucalypt plantations in Brazil," *Remote Sens. Environ.*, vol. 152, pp. 136–149, Sep. 2014.
[32] M. Bisquert, G. Bordogna, A. Bégué, G. Candiani, M. Teisseire, and P. Poncelet, "A simple fusion method for image time series based on the estimation of image temporal validity," *Remote Sens.*, vol. 7, no. 1, pp. 704–724, 2015.
[33] M. C. Anderson, J. M. Norman, J. R. Mecikalski, R. D. Torn, W. P. Kustas, and J. B. Basara, "A multiscale remote sensing model for disaggregating regional fluxes to micrometeorological scales," *J. Hydrometeorol.*, vol. 5, no. 2, pp. 343–363, Apr. 2004.
[34] J. R. Quinlan, "Combining instance-based and model-based learning," in *Proc. 10th Int. Conf. Mach. Learn.*, 1993, pp. 236–243.

Mar Bisquert received the B.Sc. degree in physics and the M.Sc. and Ph.D. degrees in remote sensing from the Universtiy of Valencia, Valencia, Spain, in 2007, 2009, and 2011, respectively.

She is currently working with the Department of Applied Physics, University of Castilla-La Mancha, Cuenca, Spain. She is working on the enhancement of fine resolution image time series with the futher objective of monitoring surface energy fluxes from remote sensing satellites. Her research interests include the environmental applications of remote sensing techniques.

Juan Manuel Sánchez received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 2003, 2005, and 2008, respectively.

He is currently a Lecturer with the Department of Applied Physics, University of Castilla-La Mancha, Cuenca, Spain. He has authored 25 papers in international journals, 5 book chapters, and more than 60 conference papers. He has participated in 20 national and international projects. His research interest include the thermal infrared remote sensing in general and the surface-energy-flux retrieval in particular.

Dr. Sánchez is a Referee in 17 international journals. He was the recipient of the Nobert Gerbier-MUMM International Award 2010 by the World Meteorological Organization.

Vicente Caselles received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 1979, 1980, and 1983, respectively.

He is currently a Full Professor of Earth Physics and the Head of the Thermal Remote Sensing Group, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. He has 38 years expertise in the physical processes involved in both temperature measurement and evapotranspiration using remote sensing techniques, which has been documented through 15 books, 25 doctoral theses, 110 papers in international journals, 60 conference papers, and 30 reports. He was collaborating with the European Space Agency as a Member of the Advisory Group for the Land-Surface Processes and Interactions Mission. He was the Chairman of the Spanish Remote Sensing Society and the Manager of Human Resources and Researchers Mobility General Direction, Spanish Ministry of Science.

Dr. Caselles was the recipient of the Norbert Gerbier–MUMM International Award for 2010, conferred by the Executive Council of the World Meteorological Organization.

QUERIES

Q1: The authors "Gao et al." does not match with Ref. [11]. Please check.
Q2: The authors "Bindhu *et al.*" does not match with Ref. [6]. Please check.
Q3: Please check if journal title for Ref. [1] is correct.
Q4: Please provide page range for Refs. [9], [10], and [20].
Q5: Please provide institution location for Ref. [19].
Q6: Please provide complete details for Ref. [30].
Q7: The authors "Bindhu *et al.* 2014" does not match with the reference list. Please check.

Evaluation of Disaggregation Methods for Downscaling MODIS Land Surface Temperature to Landsat Spatial Resolution in Barrax Test Site

Mar Bisquert, Juan Manuel Sánchez, and Vicente Caselles

***Abstract*—Thermal infrared (TIR) data are usually acquired at a coarser spatial resolution (CR) than visible and near infrared (VNIR). Several disaggregation methods have been recently developed to enhance the TIR spatial resolution using VNIR data. These approaches are based on the retrieval of a relation between TIR and VNIR data at CR, or training of a neural network, to be applied at the fine resolution afterward. In this work, different disaggregation methods are applied to the combination of two different sensors in the experimental test site of Barrax, Spain. The main objective is to test the feasibility of these techniques when applied to satellites provided with no TIR bands. Landsat and moderate imaging spectroradiometer (MODIS) images were used for this work. Land surface temperature (LST) from MODIS images was disaggregated to the Landsat spatial resolution using Landsat VNIR data. Landsat LST was used for the validation and comparison of the different techniques. Best results were obtained by the method based on a linear regression between normalized difference vegetation index (NDVI) and LST. An average $\mathrm{RMSE} = \pm 1.9$ K was observed between disaggregated and Landsat LST from four different dates in a study area of 120 km^2.**

***Index Terms*—Image enhancement, image resolution, remote sensing, temperature.**

I. Introduction

TIME series of fine spatial and temporal resolution images are key inputs in numerous studies, e.g., water resources management [1], [2]. However, there is a limitation in the existing satellites since revisit time for fine spatial resolution sensors is typically poor, while those with a high revisit frequency are characterized by a coarse spatial resolution. This is especially true when focusing on the thermal infrared (TIR) since spatial resolution for the TIR bands is always coarser than that for the visible and near infrared (VNIR) bands onboard the same sensor [2].

Disaggregation methods allow downscaling the TIR coarse resolution (CR) to finer resolutions. In [3], a review of land surface temperature (LST) disaggregation methods is performed. Zhan *et al.* [3] classified the disaggregation methods in two main groups: thermal sharpening (TSP) and temperature unmixing (TUM). The main difference is that TSP is used to obtain the LST of smaller resolution cells, while TUM aims at obtaining the LSTs of the existing elements within CR cells [3]. A bibliographical review showed that TSP is more frequently used than TUM. The TUM approaches need sufficient information of the component temperatures and become more useful when focused on obtaining temperatures of different surfaces (e.g., soil vs. vegetation, and asphalt vs. bare soil). In [4], a physics-based unmixing method was presented to estimate the relative proportion and the temperature of each material composing the mixed pixel. Emissivity and temperature over pure pixels of the different components are required in this method. The number of components has to be previously identified and a classification image is needed. Note this method is constrained to the existence of pure CR pixels of each element present in the fine resolution (FR) images.

Several TSP techniques have been proposed in the recent literature to enhance the spatial resolution of the TIR domain over vegetated areas by linking TIR and reflectance information [5]–[8]. Most of these techniques are based on the assumption that there exists a relation between the vegetation cover and the LST. According to these approaches, a relation between TIR and VNIR bands is first obtained at CR and then applied at the finer resolution of the VNIR bands, assuming that this relation is scale invariant. Kustas *et al.* [5] developed the disTrad (disaggregation procedure for radiometric surface temperature) based on a quadratic relationship between normalized difference vegetation index (NDVI) and LST. This technique was used in [9] to downscale TIR data to the VNIR resolution with Landsat and moderate imaging spectroradiometer (MODIS) images. Agam *et al.* [6] tested three new variants of the disTrad and the results from the four methods were compared. The disTrad variants were: a linear relationship between NDVI and LST, a linear relationship between fractional vegetation cover (fc) and LST, and a simplified version of the fc-LST variant (hereafter called TsHARP). Since results might depend on the land surface cover, TsHARP was applied separately to crop and natural vegetation [10]. The relation between NDVI and LST was too poor in the natural vegetation. In the crop area, the disaggregation led to better results compared to a nondisaggregation method uniTr (using the LST value of the CR image). Lower errors were found when applying the

Manuscript received May 06, 2015; revised January 08, 2016; accepted January 12, 2016. This work was supported in part by the Spanish Ministry of Economy and Competitiveness (projects CGL2013-46862-C2-1/2-P, cofinanced with European Union FEDER funds), and in part by the Generalitat Valenciana (project PROMETEOII/2014/086).

M. Bisquert and J. M. Sánchez are with the Department of Applied Physics, Polytechnic School of Cuenca, University of Castilla-La Mancha, Cuenca 16071, Spain (e-mail: mar.bisquert@uclm.es; juanmanuel.sanchez@uclm.es).

V. Caselles are with the Department of Earth Physics and Thermodynamics, University of Valencia, Valencia 46100, Spain (e-mail: vicente.caselles@uv.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2016.2519099

TsHARP to the entire scene than when focusing on the crop areas. Neural networks (NNs) were applied in [11] and [12] using land cover and several land surface parameters obtained from VNIR images as inputs (different combinations for each land cover). Jeganathan *et al.* [7] tested four different versions of the TsHARP: resolution-adjusted TsHARP, piecewise (a regression per NDVI intervals [0–0.2], [0.2–0.5], [0.5–1]), stratified (applied per land cover class), and local (applied within a moving window). Better results were obtained with the local disaggregation procedures. Gao *et al.* [13] developed a data mining (DM) approach between TIR and spectral reflectances of homogeneous pixels applied separately at global and local (within a moving window) scales and subsequently combined. Bindhu *et al.* [8] developed a nonlinear method (NL-disTrad) based on the estimation of a relationship between NDVI and LST from the pixels belonging to the hot edge (pixels forming the upper envelope of the NDVI–LST distribution). The relationship obtained is applied to CR pixels to calculate the difference between the original and the predicted LST (residuals). These residuals are trained in an NN using the NDVI values of a 3×3 window as inputs. Then, the trained NN is applied to the FR pixels to obtain the residuals at this FR. A random forest approach was introduced in [14] using VNIR bands, topography, and land cover classes as inputs. Chen *et al.* [15] applied a combination of TsHARP and Thin Plate Spline interpolation. Both methods were applied separately and weights were then calculated for each one and further combined. Mukherjee *et al.* [16] compared three of the previous approaches (disTrad, TsHARP, and TsHARP with local variant) to two new methods based on the adjustment of the linear regression between NDVI and LST using a least median square (LMS) regression and a pace regression. In [17], disaggregated microwave brightness temperatures were obtained at 1-km MODIS resolution. These authors established a relationship between NDVI and the microwave polarization difference index from the Advance Microwave Scanning Radiometer–Earth Observing System (AMSR-E). Disaggregated temperatures were then used as inputs in a soil moisture algorithm.

Performance of the different methods above is not comparable since results may be site dependent. Different conclusions and root-mean-square error (RMSE) values are obtained in different works. Most of these studies compare the results of the new proposed methods to the TsHARP, but even with TsHARP results are quite different depending on the study area and the spatial resolution of the sensors used. The majority of papers focus on Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), MODIS or Landsat images, with the objective of disaggregating from their TIR to VNIR resolution, e.g., from MODIS 1 km to MODIS 250 m. Few papers plan to apply these methods to data from various satellites, e.g., obtaining a regression from one sensor and applying it to another one provided with no TIR bands. Up to now most of these works simulated CR images by aggregating FR scenes to the coarser spatial resolution of a different satellite, using afterward these FR images to assess the disaggregation procedure. A typical example is aggregating Landsat bands to simulate MODIS, then disaggregation and validation using the Landsat original bands as a basis. This is the traditional procedure for the assessment of the disaggregation methods. However, if these methods are to be applied from two different sensors, the validation should also be based on the original data from both sensors. Very few papers have dealt with the real application from two satellites [8]. Mukherjee *et al.* [16] used both Landsat and MODIS images as the basis for the disaggregation but with no mixing between them. Better performance was observed using Landsat [16]. Some works in the literature analyze the effect of the spatial resolution on the disaggregation. Better results are obtained for smaller differences in the spatial resolution between the input and output images. For example, in [6], Landsat images were aggregated to 960 m and then disaggregated to 240, 120, and 60 m with RMSE values ranging from ± 0.7 K (240 m) to ± 2.2 K (60 m).

In this research, we are interested in providing LST at fine spatial and temporal resolutions to fulfill the requirements in the estimation of surface energy fluxes and evapotranspiration. Several study fields are situated within the experimental test site of Barrax, Spain. At least 30-m spatial resolution is desired. This can be provided by the Landsat VNIR bands. However, revisit cycle of Landsat is poor and a higher frequency is required. The recently launched Sentinel-2a has a 10-day repeat cycle, and 10–20 m spatial resolution in the VNIR bands, whereas no TIR information is available. With the coming Sentinel-2b, the combination of both satellites will offer a 5-day repeat cycle. Furthermore, the combination of Sentinel-2 and Sentinel-3 could offer the desired solution of spatial and temporal resolutions. The relationship between TIR and VNIR bands could be extracted from Sentinel-3 and then applied to Sentinel-2. The aim of this paper is to test the application of disaggregation techniques in the Barrax area from two different satellites, so that these approaches can be applied to sensors without thermal bands (e.g., Sentinel-2). Landsat and MODIS imagery were selected for this work due to the similar characteristics of these sensors and Sentinel-2 and Sentinel-3. We analyzed both, classical methods based on the VNIR-LST regression and newer methods using NN and DM. Three main experiments were developed to evaluate the classical approaches. A first experiment focused on the methods developed in [5], [6], and [10] based on the least square (LS) adjustment of the VNIR-LST regression. In a second experiment, new adjustments were tested (LMS) as well as the application of the method at local scale (using a moving window). In the last experiment, the disaggregation was applied separately for the different land covers present in the study area. The classical method leading to the best results was then compared to NN and DM approaches.

II. Materials and Methods

A. Study Site

The study area is located in Barrax, Central Spain, including "Las Tiesas" experimental farm (39°03' 35" N, 2°06' W). This is a very flat area with an average altitude of 700 m a.s.l close to Albacete. Barrax is one of the traditional ESA test sites for different international campaigns, such as SEN2FLEX

[18], SPARC [19], ImagineS [20], and DAISEX [21]. Several field campaigns have been carried out in recent years including ground measurements of LST and energy fluxes, as well as biophysical parameters, for different periods. These experimental campaigns are used with different objectives related to agricultural water management [22]–[26].

B. Satellite Images

The selection of satellite images was constrained to the coincidence of Landsat and MODIS overpasses with low MODIS viewing angle and cloud coverage for the growing season of 2014. We focused on Landsat 7 since no Landsat 8 images matched our requirements for the selected period. Four images were finally selected in 2014 for DOYs 67, 163, 195, and 202.

VNIR data were extracted from the Landsat 7 ETM+ CDR (http://earthexplorer.usgs.gov/) and the MODIS MOD09GA and MOD09GQ products (glovis.usgs.gov). Landsat CDR provides 30-m surface reflectances (atmospherically corrected). MOD09GQ offers red and NIR bands at 250-m spatial resolution, and MOD09GA contains seven spectral bands (VNIR), together with quality information, at a resolution of 500 m. Radiances in the TIR domain were obtained from band 6 of Landsat 7 ETM+, while the MOD11_L2 product offers LST directly. ETM+ TIR data are acquired at 60-m spatial resolution; however, these data are resampled and provided at 30-m resolution. For validation purposes, we used an aggregated LST at 60 m that corresponds to the original resolution in the TIR domain. MOD11_L2 is provided at 1 km. MODIS images were resampled to 240, 480, and 960 m in order to have pixel dimensions that are multiple of the Landsat spatial resolutions (30 and 60 m). All scenes were reprojected to UTM WGS 1984 zone 30 N.

Fig. 1 shows a false color composite of a Landsat image used in this work (RGB: bands 4, 3, 2), an NDVI image, and the Corine Land Cover 2006 (http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster) covering the entire study area ($\sim$ 120km^2). The images illustrate a sparsely vegetated area, with $\mathrm{NDVI} < 0.4$ for most pixels. Largest NDVI values correspond to crop fields [red areas in Fig. 1(a) and (b)].

Table I lists a summary of the NDVI and LST values from the MODIS and Landsat images used at both, their original and aggregated 960-m spatial resolution. Note that wider value ranges are present in the original resolution compared to the aggregated 960-m resolution.

C. Image Processing

Quality bands from each product were used to mask pixels containing clouds and shadows. Landsat CDR product includes a mask obtained using the Fmask code [27]. Based on this mask only, pixels assigned "clear land pixel" were kept in this study, discarding those with clouds, snow, shadows, and water. TIR Landsat data were atmospherically corrected to retrieve LST following the procedure described in [28] and using the Atmospheric Correction Tool of Barsi *et al.* [29] and [30] (http://atm-corr.gsfc.nasa.gov/).

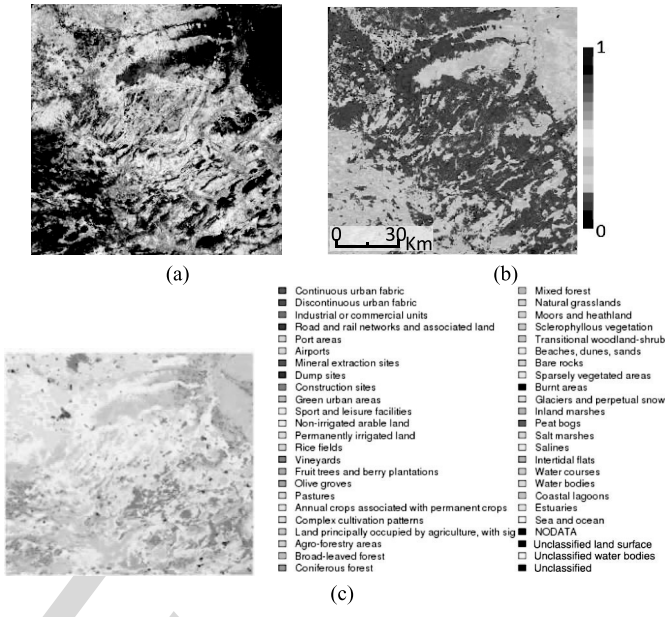


Fig. 1. (a) Landsat false color image (RGB: bands 4, 3, 2). (b) NDVI corresponding to DOY 195. (c) Corine Land Cover of the study area.

The data aggregation is a key step in the disaggregation procedures. The aggregation of the VNIR bands was carried out by averaging the reflectance values in the red and NIR bands of all the FR pixels within an equivalent CR pixel. Following Gao *et al.* [11], the aggregation of the TIR band was done through the Stefan-Boltzmann law with the assumption of similar emissivity values for adjacent pixels

$$T_{agg} = \sqrt[4]{(1/n)\sum_{i=1}^{n} T_i^4}. \quad (1)$$

The application of the disaggregation techniques to different sensors needs equivalent spectral data from both sensors. Differences between both sensors VNIR data may exist due to several effects such as spectral resolution, atmospheric correction, viewing angle, and pixel footprint. In the TIR domain, the main difference between both sensors may be due to the different acquisition time. A normalization procedure can be applied to minimize these discrepancies between MODIS and Landsat images [31], [32]. A linear regression between the aggregated Landsat and MODIS images at 960 m was obtained, and later applied to the desired spatial resolution (e.g., 60 m). This normalization step should be applied to each pair of FR and CR images to be used. Note that in this paper, we used Landsat ETM+ and MODIS images coincident in date with the aim of using the Landsat TIR band to validate the disaggregated temperatures. However, the disaggregation method could be applied to images from different dates if they are close enough in time to consider that no significant changes in the vegetation cover have occurred. In case of using images from different dates, the same normalization procedure should be applied to each pair of FR and CR images, and the disaggregated temperature obtained will correspond to the date of the CR image (from which the VNIR-TIR relationship is obtained).

T1:1 TABLE I
T1:2 NDVI AND LST AVERAGE, MAXIMUM AND MINIMUM VALUES FOR THE MODIS AND LANDSAT IMAGES AT THEIR ORIGINAL RESOLUTION (30
T1:3 AND 60 M) AND THE AGGREGATED 960-M RESOLUTION, FOR TWO DATES (DOYS 67 AND 195)

Doy	Image	NDVI					LST				
		Median	Max	Min	Q25	Q75	Median	Max	Min	Q25	Q75
67	Landsat 30 m / 60 m (NDVI / LST)	0.27	0.97	0.0001	0.19	0.4	294	312	280	291	296
	Landsat 960 m	0.25	0.7	0.09	0.2	0.34	294	303	285	292	296
	MODIS 240 m	0.29	0.91	0.002	0.22	0.39					
	MODIS 960 m	0.29	0.69	0.07	0.22	0.38	296	305	287	294	298
195	Landsat 30m / 60m (NDVI / LST)	0.21	0.9	0.0001	0.15	0.31	316	331	288	313	319
	Landsat 960 m	0.21	0.6	0.1	0.17	0.29	316	325	303	312	318
	MODIS 240 m	0.22	0.87	0.001	0.17	0.37					
	MODIS 960 m	0.22	0.72	0.02	0.18	0.3	319	327	303	317	321

Percentiles Q25 and Q75 indicate that 25% or 75%, respectively, of the total data are lower than these values.

D. Disaggregation Methods

Several disaggregation methods extracted from the recent literature were applied in this work with images from two different satellites. These methods are traditionally applied and tested with images from a single satellite. First, the FR images are aggregated to a CR and then the disaggregation method is applied, finally the original TIR images are used for validation. In this work, we tested the disaggregation from two different sensors. The further objective is assessing the performance of the disaggregation when applied to FR images from satellites with no thermal bands. To compare the application with one or two sensors, the two versions were tested here as follows. 1) The relationship between NDVI and LST (training) was extracted from the Landsat ETM+ aggregated images (960 m) and later applied to Landsat 60 m (traditional procedure). 2) The relationship was obtained from the original MODIS images (960 m) and then applied to Landsat 60 m images. Both procedures provided disaggregated 60 m LST using 60 m Landsat NDVI as input. The original 60 m Landsat LST was reserved for the validation of the disaggregated LST outputs.

1) NDVI–LST Regressions: Fig. 2 shows a flowchart of the methodology.

Experiment 1 consisted on applying the most used disaggregation methods developed in [5], [6], and [10]. In [5], a quadratic relationship between LST and NDVI was proposed (disTrad, 2). The relationship was obtained from CR pixels and applied to the FR pixels. The relationship obtained was also applied at CR to obtain the difference between the original and predicted LST (residuals). The CR residuals were then added to the estimated LST at FR; this means that the same residual was used in all the FR pixels belonging to the same CR pixel. With this action actual CR LST information was included in the disaggregated results. In [6], the quadratic relationship and other relationships between NDVI and LST were tested, including a linear approach between LST and NDVI (3), a linear approach between the fraction of vegetation cover (fc, expressed in terms of NDVI) and LST (4), and a simplified version of the fc approach (5). As suggested by [5], all regressions in this Experiment 1 were adjusted from the most homogeneous pixels in three NDVI intervals: [0–0.2], [0.2–0.5], and [0.5–1]. The residual correction was included in all the cases considered in this Experiment 1

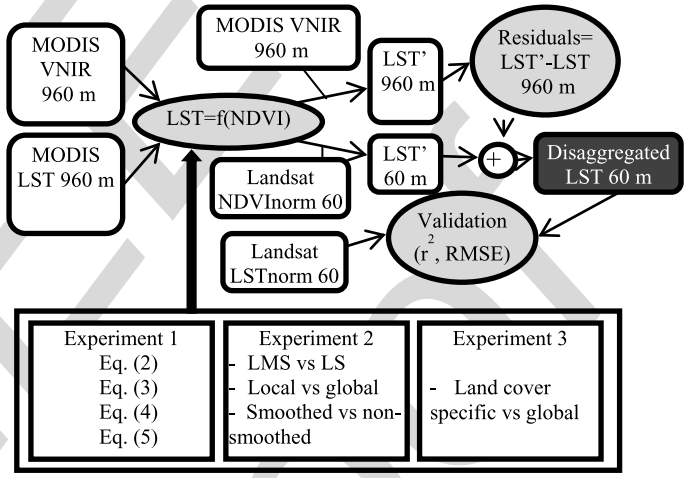


Fig. 2. Flowchart of the methodology applied, including the different processing steps and the disaggregation methods used. F2:1 F2:2

$$LST = a_0 + a_1 NDVI + a_2 NDVI^2 \quad (2)$$

$$LST = a_0 + a_1 NDVI \quad (3)$$

$$LST = a_0 + a_1 \left(1 - \left(\frac{NDVI_{\max} - NDVI}{NDVI_{\max} - NDVI_{\min}}\right)^{0.625}\right) \quad (4)$$

$$LST = a_0 + a_1 (1 - NDVI)^{0.625} \quad (5)$$

where a_0, a_1, and a_2 are the adjusted parameters.

More recent studies introduce some modifications to the disaggregation procedures described above. These were analyzed in Experiment 2. Note that only the method yielding the

best results in Experiment 1 was considered for Experiment 2. In [16], better results were obtained using LMS than LS regression. Thus, the LMS adjustment was tested and compared to the LS. The disaggregation methods using the residual correction present some boxy effect linked to these residuals themselves [10], [13]. To reduce this effect, the CR residuals can be smoothed using a Gaussian filter [33]. We adopted this technique and analyzed its effect visually and in terms of RMSE. All these regressions evaluated in Experiment 2 were obtained from the most homogeneous pixels and including the residual correction, as done in Experiment 1. In [7], (3) was applied at a local scale using a moving window. This local application showed better results than obtaining a global regression for the entire image. This approach consisted in obtaining a relationship between NDVI and LST at CR for each window and then applying it to the FR pixels belonging to that window. Each pixel is only considered in one window. The result for each pixel is then the LST obtained from the NDVI–LST relationship adjusted at the CR window to which it belongs. We also tested this moving window approach with different window sizes (from 3×3 to 13×13) and compared it to the global approach. In contrast to the previous methods, since few pixels are available in the local application all of them were used for adjusting the regression and no residual correction was applied due to the local scale.

Some authors have used the land cover information to derive different regressions for each particular land cover class [7], [10]. Several tests were also conducted in this work for the different land cover classes (Experiment 3) present in the images. Corine Land Cover data at 100-m resolution was used. Land cover information was needed at 60 and 960 m. For the 60-m resolution, a resampling was performed using nearest neighbor resampling. For 960 m, the predominant class (the most abundant) was assigned. The disaggregation method leading to the best results in the previous analyses (Experiments 1 and 2) was applied here specifically per land cover classes. The NDVI–LST relation was obtained considering all the pixels included in each land cover class at CR and then applied to the pixels of the same class at FR. The residuals correction was also accounted here. An analysis of the performance of the disaggregation technique was conducted using the land cover class distinction as a basis. Results were compared to those obtained not accounting for this land cover distinction.

2) NN and DM Approaches: NNs using different inputs (spectral bands, land cover, topography, etc.) have also been applied in several papers. The method developed by Bindhu *et al.* [8] was applied in this work. These authors used NN for training the residuals. The relationship between NDVI and LST was obtained from the pixels forming the hot edge. These pixels were selected as those showing the highest temperature values for different NDVI ranges. This relationship was then applied to the CR pixels to estimate the residuals. The CR residuals were trained in a NN which inputs were the NDVI values in a 3×3 pixel window. The trained NN was then applied at FR. Other tests performed with NN consisted in using different combinations of spectral bands, and in some cases land cover, as inputs in the NN, and the LST as output.

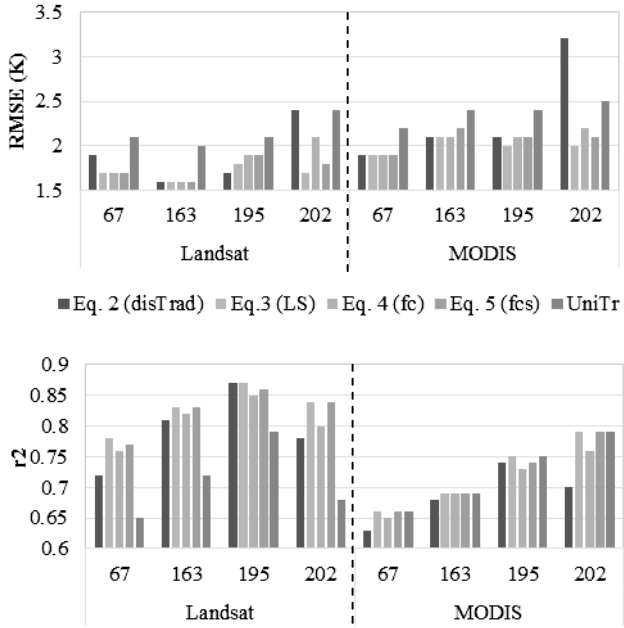


Fig. 3. Single- and dual-sensor analysis of the validation results (RMSE and r2) of the different methods (1)–(4). On the left, Landsat images were first degraded to the MODIS resolution and then disaggregated to the Landsat TIR original resolution. On the right, the relationship was obtained from the MODIS images and then applied to the Landsat images at 60-m spatial resolution. The residual correction was applied in both cases.

The DM approach developed in [13] was also tested here. This method combines a local and global application of regression trees in a DM approach. The Cubist package in R [34] was used in this work. This method uses the reflectance from all the bands. Thus, when applied to two different sensors, their bands have to be the same or similar, and the normalization of each band between both sensors has to be done. The regression tree method generates rule-based linear multivariate regressions. In this work, the 6 Landsat ETM+ VNIR bands (1–5 and 7), and the equivalent MODIS bands (bands 1–6) were used as inputs.

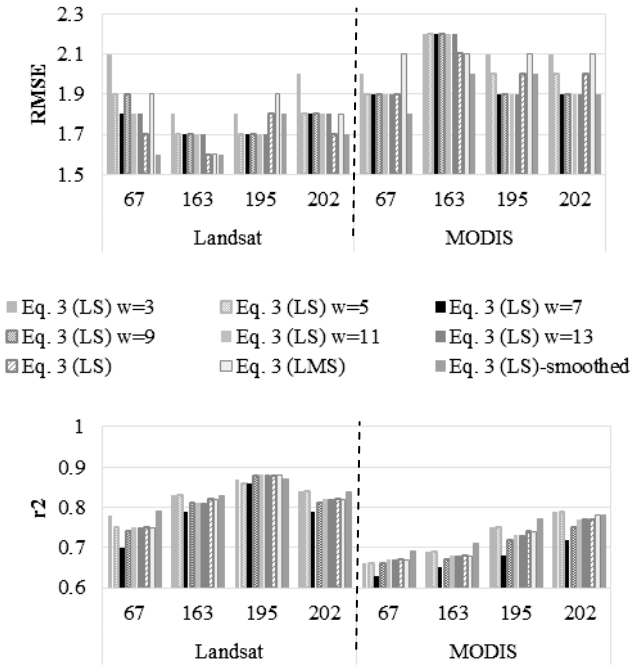
E. Performance Assessment

The comparison of the disaggregation methods was carried at a regional scale. All methods were analyzed in terms of the determination coefficient (r2) and the RMSE, and compared to a nondisaggregation method, uniTr [6]. According to this approach, each FR pixel is assigned at the value of the corresponding CR pixel.

III. Results

A. NDVI–LST Regressions

In the first experiment (Fig. 2), the methods presented in [6] were applied separately to a single sensor (Landsat ETM+) and to the combination of two sensors (Landsat and MODIS) as described above. Fig. 3 shows the validation results (RMSE and r2), for the 120-km^2 study area, corresponding to the regression obtained from both, Landsat aggregated images (960 m) on one hand and original MODIS images on the other hand, then



F4:1 Fig. 4. Validation results (RMSE and r2), of the local disaggregation method
F4:2 with different windows sizes (w) compared to the global application and to the
F4:3 global application with smoothed residuals.

applied to the original 60-m Landsat images. Previous works have tested the disaggregation methods with images belonging to the same sensor (by aggregation to coarser resolution) with the aim of being applied with two different sensors, but they rarely showed the real dual-sensor application. Results presented in Fig. 3 point out the larger errors obtained when these techniques are applied to two different sensors instead of a single one, despite the normalization adjustment applied. Better results were obtained using any of the disaggregation methods compared to the uniTr approach (using the LST of the CR image) in both the single and dual-sensor applications. Contrary to previous studies, the simplest method [linear regression between NDVI and LST, (3)] led to similar or better results than using other regressions in most of the cases analyzed. Equation (2) did not perform well for day 202 since the quadratic regression applied to NDVI outliers not present in the primitive retrieval of this equation yielded extreme LST values. Also, the NDVI–LST adjustment in this image was better fitted to a linear regression than to the quadratic regression. The same conclusions can be drawn from the method applied with one or two sensors despite the different errors obtained.

In the second experiment, the local application [7] was tested with a window size ranging from 3×3 to 13×13 pixels, and results were compared to those obtained from the global application using (3), (3) adjusted by LMS [16], and (3) with smoothed residuals (Fig. 4). Fig. 4 shows that lower errors were always obtained when applying the disaggregation with a single satellite. Regarding the local application, only the RMSE was slightly reduced in one of the images (DOY 195) when using one sensor (Landsat ETM+) and in two dates (DOYs 195 and 202) when using two sensors (ETM+ and MODIS), while

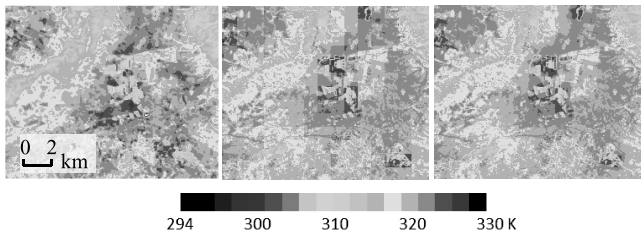
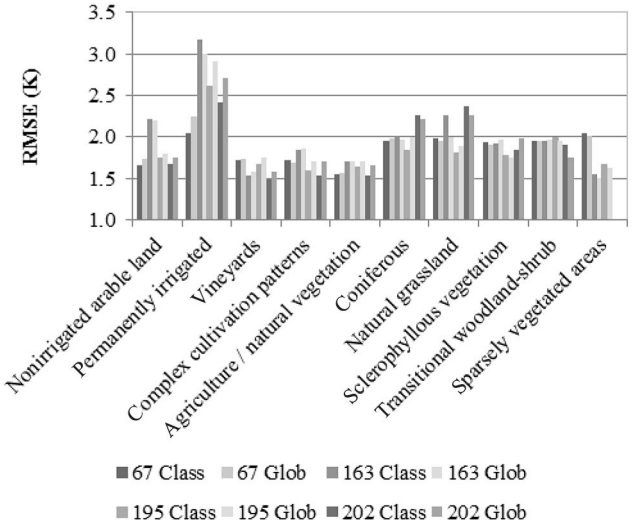


Fig. 5. Landsat LST (left), disaggregated LST with (3) LS (center), disaggregated with smoothed residuals (right) for DOY 195. F5:1 F5:2

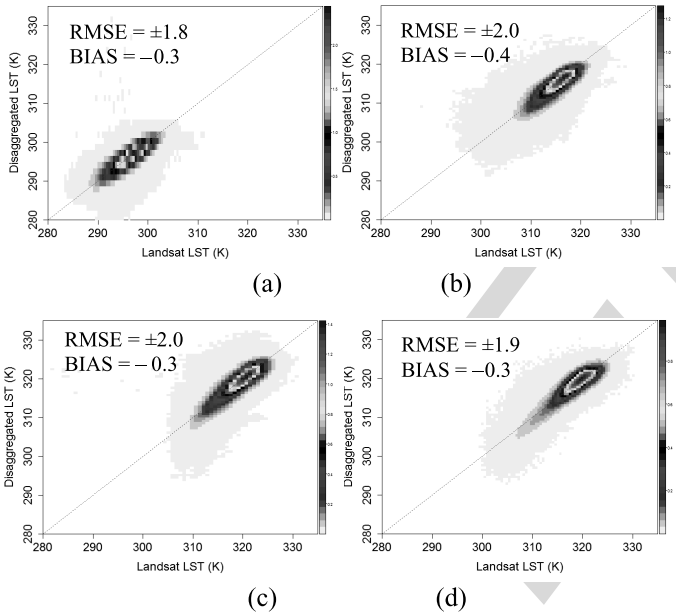
no significant differences were observed for the rest. The LMS adjustment did not improve the results from the LS adjustment. The smoothing of the residuals generally improved the results from the LS method when using two sensors. When applied to a single sensor, it improved the results in one date (67) and led to the same results for the other dates. Fig. 5 shows an example of the original Landsat LST and disaggregation results from two sensors, with and without residuals smoothing. Note the residuals introduced an important boxy effect that was reduced after the smoothing. The ratio between the CR and the FR is responsible for this boxy effect, lower differences between the spatial resolutions would mitigate this boxy effect.

Finally, specific disaggregation was conducted in Experiment 3 for different land cover classes. Only the dual-sensor application is shown here due to the similar conclusions obtained from the single- and dual-sensor applications in the previous experiments. Results considering separate land covers were compared to those obtained from applying the global equation to the same set of pixels. We focus on the linear equation LST–NDVI (3) and smoothed residuals, based on the results above. Following Agam *et al.* [10], two main classes were first considered: crop and natural vegetation. No significant differences, compared to the global results, were observed accounting for this distinction. The specific disaggregation was expanded to all vegetation classes present in the image with a significant presence. Classes with less than 100 CR pixels were excluded from the analysis. Detailed class-dependent results are shown in Fig. 6. Performance depended on the particular land cover class and date, and no firm conclusion could be extracted about the value of separating in classes. Average RMSE values of 1.8, 2.2, 1.9, and 2.0 K were obtained with the class-specific approach and 1.8, 2.0, 2.0, and 1.9 K with the global approach for DOYs 67, 163, 195, and 202, respectively. Only for DOY 195 was a slight enhancement observed with the class-specific approach. Note that larger errors were observed overall for "permanently irrigated" areas. This is probably due to the small size of these fields and also to irrigation effects. After an irrigation event, the LST may decrease dramatically, whereas the disaggregation methods do not account for this variation since the NDVI is not significantly affected. This effect is most obvious in bare soil areas such as croplands in the first stage of the crop growing season. This should be mitigated by the residuals obtained from the CR image but if the fields are small this is not well captured either.

Since no significant enhancement was observed from the land cover class-specific application, we decided to use (3)



F6:1 Fig. 6. Validation results (RMSE) of class-specific disaggregation method
F6:2 (Class) compared to the global (Glob) application. Residuals have been
F6:3 smoothed in both.



F7:1 Fig. 7. Disaggregated LST with (3) LS and smoothed residuals versus Landsat
F7:2 LST for images corresponding to DOY. (a) 67. (b) 163. (c) 195. (d) 202.

globally together with smoothed residuals. Fig. 7 shows the density scatter plots of the disaggregated LST with (3) LS and smoothed residuals from two sensors versus the reference LST (original Landsat normalized). The main percentage of pixels (in red) concentrates around the 1:1 line although the large scatter of a minority of them leads to important errors.

Fig. 8 shows a subset example of 24 km × 22 km of the reference (left), disaggregated from two sensors (central) and nondisaggregated uniTr (right) LST images. This subset is dominated by permanently irrigated lands. Overall patterns of the disaggregated LST are similar to reference LST. Note the evident improvement of the disaggregation faced to the nondisaggregated LSTs.

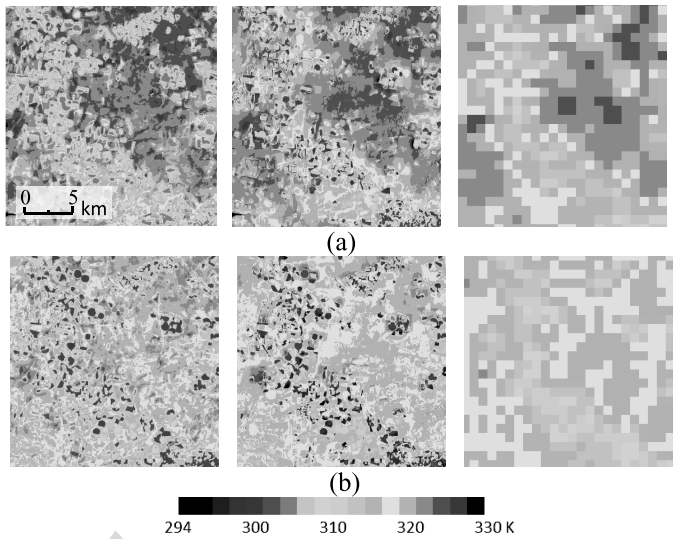


Fig. 8. Subset (22×24 km^2) of the Landsat LST (left column) and disaggregated LST with (3) LS and smoothed residuals (central column) and the nondisaggregated uniTr (right column) for images corresponding to DOYs. (a) 195. (b) 202. The gaps due to the SLC error have been filled for visual inspection of results using a triangulation method.

B. NN and DM Approaches

The methods presented in [8] and [13] were also tested and compared to the results obtained using the linear regression with smoothed residuals. In the previous section, we have already analyzed the difference between the single- and the dual-sensor disaggregation, and similar conclusions can be extracted when comparing different approaches. Therefore, in this section only, the results of the dual-sensor application are shown, which is the main interest of this work.

Results from the NN and DM approaches for the entire study area are summarized in Table II as well as the results from the LS method. Larger errors were observed when applying the NN and DM approaches. Bindhu *et al.* [6] had already observed high errors for low NDVI values, dominant in our study site [see Fig. 2(b)]. In addition to these methods, a battery of tests was conducted with NN using different spectral bands as inputs with no improvement in the results obtained by the linear regression in any case. NN and DM (trained at 960 m and applied at 60 m), as well as other nonlinear regressions, have the risk of overfitting and are very sensitive to noise in the samples. In this work, the input data did not cover the full range of values present at 60 m (see Table I), and the overfitting of the methods may lead to important errors in the outliers not present in the input data. Therefore, the simpler NDVI–LST linear regression led to better results than the more sophisticated NN and DM approaches.

IV. Discussion

Findings in this work show that there is a difference between applying disaggregation methods from one or two sensors although a normalization procedure is applied. The normalization step is very important when dealing with two sensors in

T2:1 TABLE II
T2:2 COMPARISON, IN TERMS OF RMSE (K), OF THE VALIDATION RESULTS
T2:3 OF THE NN DISAGGREGATION METHOD, AND THE LINEAR
T2:4 DISAGGREGATION METHOD (3) WITH SMOOTHED RESIDUALS

DOY	67	163	195	202
DM (Gao *et al.* 2012)	2.6	2.5	2.3	2.5
NN (Bindhu *et al.* 2014)	2.2	2.5	2.3	3.1
LST-NDVI regression	1.8	2.1	1.9	2.0

order to account for the possible spectral and temporal differences between them. A good coregistration of the images from both sensors is also critical. These are the two main points that make the difference when applying the disaggregation with one or two sensors. The normalization procedure was addressed in this work at 960 m; however, at this spatial resolution, we observed a shorter range of values than those present at the FR (Table I). Consequently, the relationships obtained did not include values in the full FR range. This not only affects the normalization process but also affects the NDVI–LST relationship used in the disaggregation procedure. This difference in the spectral range at the FR and CR is site dependent and is related to the heterogeneity of the area and to the ratio between the FR and the CR being used. Thus, this constraint is common to the single- and dual-sensor applications and it may contribute to the different errors obtained by other authors when disaggregating to different spatial resolutions. For example, in [6], the RMSE errors are very different depending on the spatial resolution and dates. These authors disaggregated two images, affected by 1 week difference, from 960 to 60 m in July 2002. The linear NDVI–LST approach yielded RMSE values of ± 2.2 and ± 1.2 K. Disaggregation to 120 m led to an enhancement in RMSE of ± 0.3 K in both images, and when disaggregating to 240 m, the enhancement was of ± 0.8 K in one image and ± 0.5 K in the other.

In our study area, RMSE values ranged between ± 1.6 and ± 1.7 K when using the Landsat images, and ± 1.8 and ± 2.1 K using the MODIS ones. These results are in agreement with those obtained by Agam *et al.* [6]. In [7], the local disaggregation of Aster images, previously aggregated to 990-m spatial resolution, resulted in RMSE values ranging between ± 0.7 and ± 3 K for disaggregation to 810–90 m. Large errors were obtained when disaggregating to spatial resolutions finer than 270 m, consequently MODIS images were disaggregated only to 250 m with RMSE values between ± 2 and ± 3 K depending on the approach. Better results were obtained by Bindhu *et al.* [8] in a smaller area with high vegetation cover, with the highest errors corresponding to low NDVI values.

The application of a land cover-specific disaggregation showed the highest errors for the class "permanently irrigated land." This is probably due to the small size of the fields and the errors caused by irrigation. This is reinforced by the fact that errors in this class increase significantly in the summer images, both with the class-specific and the global approaches. Most likely, this is a consequence of the higher differences in LST between the crop fields and the surrounding bare soil during summer season.

Future work will deal with a more comprehensive validation at a field scale using ground temperature measurements concurrent to satellite overpasses. Also, we will check the application to other combinations of platforms, including fine spatial resolution sensors with no TIR band (e.g., Sentinel 2).

V. CONCLUSION

Several disaggregation methods were compared when applied with two different sensors, MODIS and Landsat ETM+, in a heterogeneous and sparsely vegetated area of central Spain. This dual-sensor application needed an additional step of normalization between the spectral bands of the two sensors and special attention to the coregistration between them. Although the errors with the dual-sensor application were higher than those from the single-sensor application, results obtained prove the potential and usefulness of the disaggregation techniques applied to two different sensors.

The disaggregation was applied to downscale images from 960-m spatial resolution (MODIS) to 60 m (Landsat). Best results were obtained with the procedure based on the linear regression between NDVI and LST, even when compared to NN and DM approaches. RMSE values around ± 2.0 K were obtained for four different images. Focusing on the experimental agricultural area of Barrax, classed as "permanently irrigated land" in Corine Land Cover, RMSE values were the highest (± 2 to ± 3 K). Since this area showed the largest heterogeneity within the image, we may conclude that the disaggregation procedures perform better in homogeneous areas. These results are encouraging, and reinforce the application of disaggregation procedures to sensors provided with no thermal bands. The application to Sentinel-2, with a high revisit cycle, could provide the time series of disaggregated LSTs at high temporal and spatial resolutions required for studies related to hydrology, climatology or agriculture.

REFERENCES

Q3

[1] M. C. Anderson *et al.*, "Mapping daily evapotranspiration at landsat spatial scales during the BEAREX'08 field campaign," *Bushland Evapotranspiration Agric. Remote Sens. Exp.*, vol. 50, pp. 162–177, Dec. 2012.

[2] W. Ha, P. Gowda, and T. Howell, "A review of downscaling methods for remote sensing-based irrigation management: Part I," *Irrig. Sci.*, vol. 31, no. 4, pp. 831–850, Jul. 2013.

[3] W. Zhan *et al.*, "Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats," *Remote Sens. Environ.*, vol. 131, pp. 119–139, Apr. 2013.

[4] M. Cubero-Castan, J. Chanussot, V. Achard, X. Briottet, and M. Shimoni, "A physics-based unmixing method to estimate subpixel temperatures on mixed pixels," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1894–1906, Apr. 2015.

[5] W. P. Kustas, J. M. Norman, M. C. Anderson, and A. N. French, "Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship," *Remote Sens. Environ.*, vol. 85, no. 4, pp. 429–440, Jun. 2003.

[6] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and C. M. U. Neale, "A vegetation index based technique for spatial sharpening of thermal imagery," *Remote Sens. Environ.*, vol. 107, no. 4, pp. 545–558, Apr. 2007.

[7] C. Jeganathan, N. A. S. Hamm, S. Mukherjee, P. M. Atkinson, P. L. N. Raju, and V. K. Dadhwal, "Evaluating a thermal image sharpening model over a mixed agricultural landscape in India," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 13, no. 2, pp. 178–191, Apr. 2011.

[8] V. M. Bindhu, B. Narasimhan, and K. P. Sudheer, "Development and verification of a non-linear disaggregation method (NL-DisTrad) to downscale MODIS land surface temperature to the spatial scale of Landsat thermal data to estimate evapotranspiration," *Remote Sens. Environ.*, vol. 135, pp. 118–129, Aug. 2013.
[9] J. M. Sánchez *et al.*, "Monitoring surface energy fluxes at different spatial resolutions. Effects on fluxes variability in the Basilicata Italian region," in *Proc. Recent Adv. Quant. Remote Sens. Conf.*, Torrent, Spain, 2006.
[10] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and P. D. Colaizzi, "Utility of thermal sharpening over Texas high plains irrigated agricultural fields," *J. Geophys. Res.*, vol. 112, no. D19, Oct. 2007.
[11] G. Yang, R. Pu, W. Huang, J. Wang, and C. Zhao, "A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 2170–2178, Apr. 2010.
[12] G. Yang, R. Pu, C. Zhao, W. Huang, and J. Wang, "Estimation of subpixel land surface temperature using an endmember index based technique: A case examination on ASTER and MODIS temperature products over a heterogeneous area," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1202–1219, May 2011.
[13] F. Gao, W. Kustas, and M. Anderson, "A data mining approach for sharpening thermal satellite imagery over land," *Remote Sens.*, vol. 4, no. 12, pp. 3287–3319, Oct. 2012.
[14] C. Hutengs and M. Vohland, "Downscaling land surface temperatures from MODIS data to mesoscale resolution with Random Forest regression," Gemeinsame Tag. 2014 DGfK DGPF GfGI GiN DGPF Tagungsband 23/2014, 2014.
[15] X. Chen, W. Li, J. Chen, Y. Rao, and Y. Yamaguchi, "A combination of TsHARP and thin plate spline interpolation for spatial sharpening of thermal imagery," *Remote Sens.*, vol. 6, no. 4, pp. 2845–2863, Mar. 2014.
[16] S. Mukherjee, P. K. Joshi, and R. D. Garg, "A comparison of different regression models for downscaling Landsat and MODIS land surface temperature images over heterogeneous landscape," *Adv. Space Res.*, vol. 54, no. 4, pp. 655–669, Aug. 2014.
[17] C. Song, L. Jia, and M. Menenti, "Retrieving high-resolution surface soil moisture by downscaling AMSR-E brightness temperature using MODIS LST and NDVI data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 3, pp. 935–942, Mar. 2014.
[18] J. A. Sobrino *et al.*, "Thermal remote sensing in the framework of the SEN2FLEX project: field measurements, airborne data and applications," *Int. J. Remote Sens.*, vol. 29, nos. 17–18, pp. 4961–4991, Sep. 2008.
[19] J. F. Moreno *et al.*, "The SPECTRA Barrax campaign (SPARC): Overview and first results from CHRIS data," *presented at the Eur. Space Agency (Special Publication) ESA SP*, 2004, pp. 30–39.
[20] C. Latorre, F. Camacho, F. de la Cruz, R. Lacaze, M. Weiss, and F. Baret, "Seasonal monitoring of FAPAR over the Barrax cropland site in Spain, in support of the validation of PROBA-V products at 333 m," in *Proc. Recent Adv. Quant. Remote Sens. Conf.*, Torrent, Spain, 2014.
[21] M. Berger *et al.*, "The DAISEX campaigns in support of a future land-surface-processes mission," *ESA Bull.*, vol. 105, pp. 101–111, 2001.
[22] R. López-Urrea, F. M. de S. Olalla, C. Fabeiro, and A. Moratalla, "An evaluation of two hourly reference evapotranspiration equations for semi-arid conditions," *Agric. Water Manage.*, vol. 86, no. 3, pp. 277–282, Dec. 2006.
[23] R. López-Urrea, A. Montoro, J. González-Piqueras, P. López-Fuster, and E. Fereres, "Water use of spring wheat to raise water productivity," *Agric. Water Manage.*, vol. 96, no. 9, pp. 1305–1310, Sep. 2009.
[24] J. M. Sánchez, R. López-Urrea, E. Rubio, and V. Caselles, "Determining water use of sorghum from two-source energy balance and radiometric temperatures," *Hydrol. Earth Syst. Sci.*, vol. 15, no. 10, pp. 3061–3070, 2011.
[25] R. López-Urrea, A. Montoro, F. Mañas, P. López-Fuster, and E. Fereres, "Evapotranspiration and crop coefficients from lysimeter measurements of mature 'Tempranillo' wine grapes," *Agric. Water Manage.*, vol. 112, pp. 13–20, Sep. 2012.
[26] J. M. Sánchez, R. López-Urrea, E. Rubio, J. González-Piqueras, and V. Caselles, "Assessing crop coefficients of sunflower and canola using two-source energy balance and thermal radiometry," *Agric. Water Manage.*, vol. 137, pp. 23–29, May 2014.
[27] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, Mar. 2012.
[28] J. M. Sánchez, G. Scavone, V. Caselles, E. Valor, V. A. Copertino, and V. Telesca, "Monitoring daily evapotranspiration at a regional scale from Landsat-TM and ETM+ data: Application to the Basilicata region," *J. Hydrol.*, vol. 351, nos. 1–2, pp. 58–70, Mar. 2008.
[29] J. A. Barsi, J. L. Barker, and J. R. Schott, "An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS'03)*, 2003, vol. 5, pp. 3014–3016.
[30] J. A. Barsi, J. R. Schott, F. D. Palluconi, and S. J. Hook, "Validation of a web-based atmospheric correction tool for single thermal band instruments," 2005, pp. 58820E–1–58820E–7.
[31] G. le Maire, S. Dupuy, Y. Nouvellon, R. A. Loos, and R. Hakamada, "Mapping short-rotation plantations at regional scale using MODIS time series: Case of eucalypt plantations in Brazil," *Remote Sens. Environ.*, vol. 152, pp. 136–149, Sep. 2014.
[32] M. Bisquert, G. Bordogna, A. Bégué, G. Candiani, M. Teisseire, and P. Poncelet, "A simple fusion method for image time series based on the estimation of image temporal validity," *Remote Sens.*, vol. 7, no. 1, pp. 704–724, 2015.
[33] M. C. Anderson, J. M. Norman, J. R. Mecikalski, R. D. Torn, W. P. Kustas, and J. B. Basara, "A multiscale remote sensing model for disaggregating regional fluxes to micrometeorological scales," *J. Hydrometeorol.*, vol. 5, no. 2, pp. 343–363, Apr. 2004.
[34] J. R. Quinlan, "Combining instance-based and model-based learning," in *Proc. 10th Int. Conf. Mach. Learn.*, 1993, pp. 236–243.

Mar Bisquert received the B.Sc. degree in physics and the M.Sc. and Ph.D. degrees in remote sensing from the Universtiy of Valencia, Valencia, Spain, in 2007, 2009, and 2011, respectively.

She is currently working with the Department of Applied Physics, University of Castilla-La Mancha, Cuenca, Spain. She is working on the enhancement of fine resolution image time series with the futher objective of monitoring surface energy fluxes from remote sensing satellites. Her research interests include the environmental applications of remote sensing techniques.

Juan Manuel Sánchez received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 2003, 2005, and 2008, respectively.

He is currently a Lecturer with the Department of Applied Physics, University of Castilla-La Mancha, Cuenca, Spain. He has authored 25 papers in international journals, 5 book chapters, and more than 60 conference papers. He has participated in 20 national and international projects. His research interest include the thermal infrared remote sensing in general and the surface-energy-flux retrieval in particular.

Dr. Sánchez is a Referee in 17 international journals. He was the recipient of the Nobert Gerbier-MUMM International Award 2010 by the World Meteorological Organization.

Vicente Caselles received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 1979, 1980, and 1983, respectively.

He is currently a Full Professor of Earth Physics and the Head of the Thermal Remote Sensing Group, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. He has 38 years expertise in the physical processes involved in both temperature measurement and evapotranspiration using remote sensing techniques, which has been documented through 15 books, 25 doctoral theses, 110 papers in international journals, 60 conference papers, and 30 reports. He was collaborating with the European Space Agency as a Member of the Advisory Group for the Land-Surface Processes and Interactions Mission. He was the Chairman of the Spanish Remote Sensing Society and the Manager of Human Resources and Researchers Mobility General Direction, Spanish Ministry of Science.

Dr. Caselles was the recipient of the Norbert Gerbier–MUMM International Award for 2010, conferred by the Executive Council of the World Meteorological Organization.

QUERIES

Q1: The authors "Gao et al." does not match with Ref. [11]. Please check.
Q2: The authors "Bindhu *et al.*" does not match with Ref. [6]. Please check.
Q3: Please check if journal title for Ref. [1] is correct.
Q4: Please provide page range for Refs. [9], [10], and [20].
Q5: Please provide institution location for Ref. [19].
Q6: Please provide complete details for Ref. [30].
Q7: The authors "Bindhu *et al.* 2014" does not match with the reference list. Please check.