# Downscaling land surface temperatures from MODIS data to mesoscale resolution with Random Forest regression

CHRISTOPHER HUTENGS[1] & MICHAEL VOHLAND[1]

*Zusammenfassung: Für diese Studie wurde ein ca. 32.000 km² großes Testgebiet im östlichen Mittelmeerraum ausgewählt, das sich durch wechselhafte Verhältnisse hinsichtlich Topographie und Landnutzung auszeichnet. Untersucht wurde, mit welcher Genauigkeit aus MODIS-Tagesprodukten der Landoberflächentemperatur LST (mit nomineller Sensorauflösung von ~1 km) räumlich besser aufgelöste LST-Daten (240 m) abgeleitet werden können. Der für das „downscaling" gewählte Random Forest-Regressionsansatz erwies sich unter den genannten Randbedingungen als sehr praktikabel, da nur wenige und standardmäßig verfügbare Datenfelder zur zuverlässigen Modellierung der 240 m-LST ausreichten. Wie bei jedem „downscaling"-Ansatz erwies sich die Abbildung lokal auftretender Werteausprägungen im unteren bzw. oberen Randbereich der Werteverteilung als problematisch.*

## 1 Introduction

Land surface temperature (LST) derived from thermal-infrared (TIR) satellite imagery is a key parameter in environmental modeling (KUSTAS & ANDERSON, 2009). Biophysical processes from local to global scales are governed by the spatial distribution of LST through the modulation of surface energy fluxes. Among the numerous applications, estimation of evapotranspiration for water resources management is of particular interest in times of increasingly limited freshwater supplies around the world.

Routine monitoring of ET requires satellite imagery with fine spatial and temporal resolutions at the scale of human influence (KUSTAS ET AL., 2003). Currently available satellite imagery reflects a tradeoff in resolution requirements. Landsat TM/ETM+ thermal imagery provides spatial resolutions sufficient to map LST at field scales (~100 m) but has a long repeat cycle of 16 days. Continuous monitoring efforts are further limited by cloud cover during times of image acquisition. MODIS LST products from NASA's Terra/Aqua platforms are available daily, greatly enhancing temporal sampling of temperature distributions, but are restricted to coarse spatial resolutions (~1 km). Downscaling (also referred to as thermal sharpening or disaggregation) refers to methods enhancing the spatial resolution of remotely-sensed imagery commonly by regression-type approaches (ATKINSON, 2013). Visible (VIS) and near-infrared (NIR) reflectance data provide information about vegetation cover and surface albedo which are physically linked to LST through the surface energy budget (SANDHOLT ET AL., 2002) and are generally available at higher spatial resolution than TIR imagery. Accordingly, earlier downscaling efforts utilized the VI – LST relationship to estimate subpixel variations in surface temperature.

Most of these methods exploit the well-known negative correlation between LST and vegetation indices (VI) to fit linear regression models at the coarser resolution of TIR imagery and apply

1) Institut für Geographie, Universität Leipzig, Geoinformatik und Fernerkundung, Johannisallee 19a, 04103 Leipzig; christopher.hutengs@studserv.uni-leipzig.de, michael.vohland@uni-leipzig.de

this functional relationship to finer resolution VI data fields to generate LST maps with increased spatial resolution (KUSTAS ET AL., 2003; AGAM ET AL., 2007). These approaches assume that one unique, scale invariant functional relationship between LST and VIs exists within the image. However, consistent VI – LST relationships may be limited to landscapes dominated by green vegetation with homogeneous air and soil conditions (GAO ET AL., 2012). This assumption is supported by INAMDAR AND FRENCH (2009) who found NDVI to be an inconsistent predictor of LST in large regions with heterogeneous environmental conditions and mixed landscapes. Recent studies explored non-linear downscaling methods to overcome existing limitations in complex landscapes. YANG ET AL. (2010, 2011) utilized artificial neural networks (ANN) to estimate subpixel surface temperatures based on land cover information. A regression tree approach which uses shortwave reflectance data to predict disaggregated LST was developed by GAO ET AL. (2012) and found to outperform VI-based linear models in irrigated agricultural areas and heterogeneous naturally vegetated landscapes. BINDHU ET AL. (2013) proposed a hybrid model which disaggregates LST by fitting a polynomial function to the hot edge of NDVI-LST feature space and subsequently models the resulting residuals with an ANN.
From the existing studies it may be concluded that recent non-linear models tend to outperform VI-based methods but are generally more difficult to implement. In this paper, we specifically address the issue of downscaling MODIS data in large, heterogeneous regions comprised of mixed landscapes. Our objective was to evaluate the usability of Random Forest regression to downscale MODIS 1-km LST products to 240 m spatial resolution.
Random Forest regression is adequate for modeling LST as it can handle continuous and categorical data, which allows, for example, the incorporation of land use information into the model. Additionally, no pre-specified functional relationship between dependent and independent variables is assumed. Target resolution and selection of predictor variables are among the most important issues to be considered in any downscaling approach. A modest target resolution of 240 m was chosen to align LST data fields with the maximum available resolution for MODIS surface reflectance bands. Predictor variables were restricted to readily available MODIS data and auxiliary data fields that can be derived from digital elevation models. Additionally, MODIS 1-km LSTs were corrected for an emissivity-related low temperature bias (WAN ET AL., 2002) before downscaling. The downscaling model was tested in a region in the eastern Mediterranean comprised of mixed landscapes with environmental conditions varying over short distances.
To assess the performance of the downscaling procedure, it was at first tested with synthetic low resolution data derived from Landsat imagery by applying an aggregation procedure that simulates the MODIS spatial response. The downscaling approach was subsequently applied to the bias-corrected MODIS imagery to evaluate model performance with real low resolution data and determine the influence of the bias correction on downscaling results.

# 2 Material and methods

## 2.1 Study area

The study area is located in the eastern Mediterranean and encompasses the Jordan River valley and its broader environs. Fig. 1 depicts a false color composite of the region that covers an area of about 32,000 km$^2$. Elevation ranges from -420 m to 2800 m above sea level. Climate is highly variable showing a steep temperature gradient from the Mediterranean Sea in the west to the Arabian Desert in the south east with most of the study area lying in the transition zone. Land cover is mixed including irrigated croplands, rainfed agricultural areas, forests, grass lands and barren surfaces. Land surface temperatures at the time of image acquisition ranged from 294.5 K to 325.4 K and 294.1 K to 327.1 K for Landsat and MODIS imagery, respectively, excluding water and clouds. Additionally, a subset covering the Hula valley, an irrigation agriculture site in the northern part of the study area with large small scale variations in surface temperature, was selected to evaluate the implemented downscaling model.

## 2.2 Data preprocessing and aggregation

This study used data from the Landsat-7 ETM+ and MODIS sensors acquired over the Jordan River region on 21 March 2001. MODIS surface reflectance (MOD09GA), land surface temperature (MOD11A1, MOD11B1) and land cover (MCD12Q1) products were obtained from the Land Processes Distributed Active Archive Center (LP DAAC) and registered to UTM WGS 84

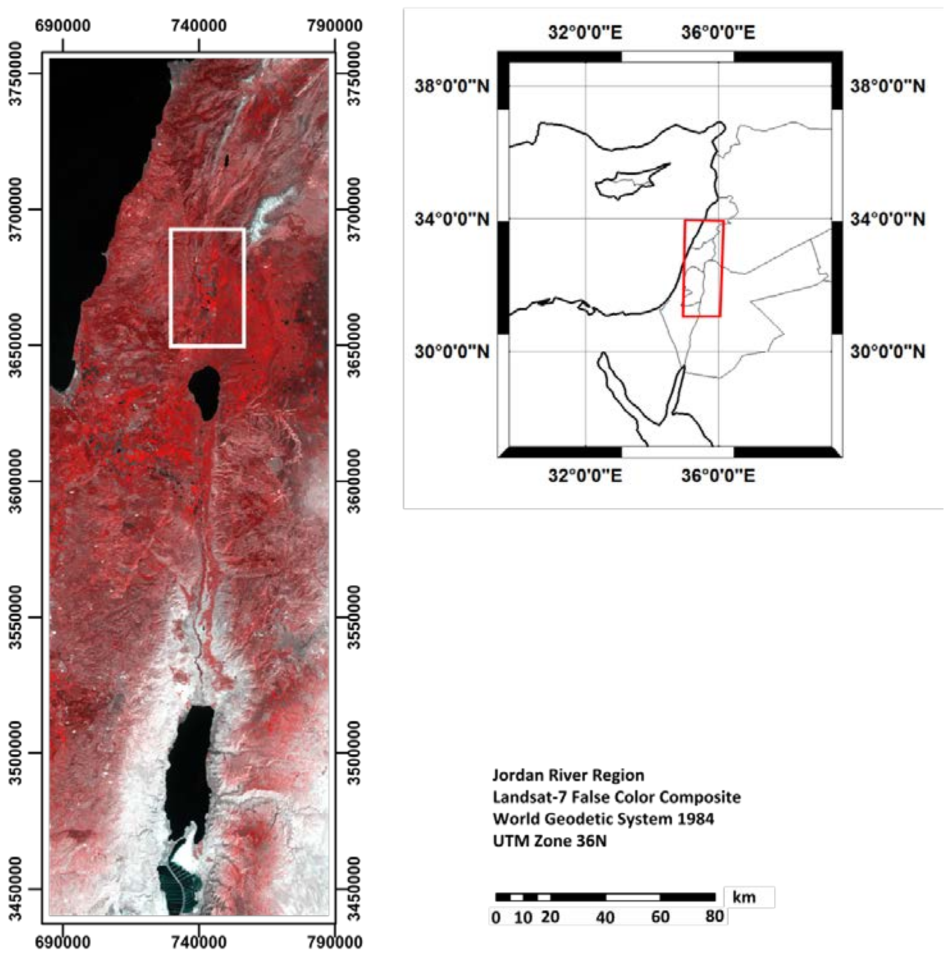


Fig. 1: Location of the study area from a Landsat-7 false color composite. The white rectangle corresponds to the Hula valley (used as spatial subset for evaluation purposes).

Zone 36 N. Reflectance products were resampled to 240 m resolution. Land surface temperature products were resampled to 960 m (MOD11A1) and 4800 m (MOD11B1) resolution, respectively. Land cover data were resampled to 480 m resolution. Two Landsat images covering the study area were mosaicked, radiometrically calibrated and geometrically registered to the MODIS data. Reflective Landsat channels were converted from radiance to reflectance by applying the empirical line method with MODIS surface reflectances as reference data fields. Surface temperatures were retrieved from Landsat TIR data using the single channel algorithm of JIMENEZ-MUNOZ ET AL. (2009). The MOD11B1 product was used to rectify MODIS LST data fields contained in MOD11A1 for errors in surface emissivity over bare and sparsely vegetated areas (WAN ET AL., 2002; LIU ET AL., 2006). SRTM elevation data were acquired from the Global Land Cover Facility (GLCF) at 3 arc second resolution, geometrically registered to MODIS imagery and resampled to 60 m nominal resolution with the cubic convolution algorithm. From the DEM slope and incidence angle maps were derived.
Landsat surface reflectances and LST were aggregated to resolutions required for model fitting and evaluation by applying a weighted average scheme that simulates the MODIS spatial response (NISHIMA ET AL., 1997). Scaling of thermal data was carried out by first converting LST to radiance using the Stefan-Boltzmann law. MODIS surface reflectances were aggregated for model fitting purposes from 240 m to 960 m by spatial averaging. Land cover data was scaled to 240 m resolution through nearest neighbor resampling. Coarse resolution land cover data was generated by assigning each 960 m resolution element to the class comprising the majority of 240 m pixels with ties being resolved by nearest neighbor resampling. DEM-derived data fields were generated by first upscaling the elevation data by spatial averaging and subsequently calculating slope, aspect and incidence angles.

## 2.3 Downscaling

Downscaling LST to higher spatial resolutions requires predictor variables that correlate with LST and are observable at higher spatial resolutions. Factors determining LST on local scales include the radiation budget, especially shortwave radiation input, and land surface characteristics as albedo, vegetation cover and soil moisture (SANDHOLT ET AL., 2002). On regional scales surface air temperature patterns become important as air temperatures and LST strive for thermal equilibrium through the transfer of sensible heat. Air temperatures and therefore LST generally decrease for example with increasing surface elevation. VIs are the most common predictors in LST downscaling as they generally show strong negative correlations with thermal imagery. For larger regions the predictive strength of global VI-based models diminishes as the VI – LST relationship is modified by additional factors (e.g. solar energy input, air temperature patterns, topography) and varies with different land cover types (ZAKŠEK & OŠTIR, 2012; INAMDAR & FRENCH, 2009).
Predictor variables in this study were limited to readily available and thus operationally usable datasets. Surface reflectances provide information about land surface albedo and vegetation cover. Solar energy input at the land surface is largely controlled by the terrain parameters slope and incidence angle that can be derived from digital elevation models. Land cover information was included to account for further dependencies between LST and land use.

Random Forests is a nonlinear statistical ensemble method that constructs a large collection of de-correlated regression trees to model the relationships between input and response variables (BREIMAN, 2001). Random Forest offers several advantages in LST prediction compared to simple linear regression. The relationship between predictor and response variables is entirely determined from the training sample. As no pre-specified functional relationship is assumed, the method is able to adapt to a wide range of environmental conditions. Additionally, the model can handle a large number of inputs, including continuous and categorical variables, and can be easily extended with additional predictors.

Most downscaling procedures follow a similar processing chain. Input data, available at higher spatial resolution, are aggregated to the coarser resolution of LST data fields to build a regression model that relates input variables to LST data. The regression model is applied to the fine resolution input variables to predict LST at finer spatial resolution. To ensure consistency with coarse resolution LST, downscaled fine resolution LST data fields are re-aggregated to the coarser resolution and the residuals between original and predicted values are calculated and added to the fine resolution predictions.

The Random Forest downscaling model was built using scikit-learn, a machine learning library implemented in Python (PEDREGOSA ET AL. 2011). Land surface temperature data fields at 960 m resolution were first converted to radiances. Surface reflectances from the near-infrared and visible red spectral bands, topography, and land cover data were used as independent variables to train the Random Forest regression model at 960 m, which was subsequently applied to the high resolution predictors to generate 240 m radiance data. These were re-aggregated to 960 m resolution and the differences between the re-aggregated model output and the original 960 m radiance data were computed. Residuals obtained at the coarse resolution were then added back to the downscaled data fields before final conversion to LST.

## 3 Results

Downscaling results for MODIS imagery and Landsat-derived coarse resolution LST were evaluated using the 240 m Landsat LST data fields as reference. Model quality was assessed by visual inspection of downscaled imagery and calculation of the statistical measures root mean squared error (RMSE) and coefficient of determination ($R^2$). Statistics were also computed for LST data fields that were uniformly disaggregated from 960 m to 240 m spatial resolution to provide a framework to further assess the downscaling model. The evaluation was carried out for the complete study area and a subset in the northern part of the region covering an irrigation agriculture site. Figure 2 depicts the Landsat reference surface temperatures and downscaled imagery from MODIS and Landsat at 240 m resolution for the complete region.

Major surface temperature patterns are similar in all images but the reference temperature field clearly conveys more high spatial frequency content than the downscaled LST maps. A distinct blurring effect is particularly visible in the downscaled MODIS LST data, as downscaling results based on degraded Landsat temperature fields show significantly more contrast.

Statistical results for the complete region are summarized in table 1. Downscaled imagery achieved RMSE of 1.47 K and 2.24 K for Landsat and MODIS data, respectively, while uniform

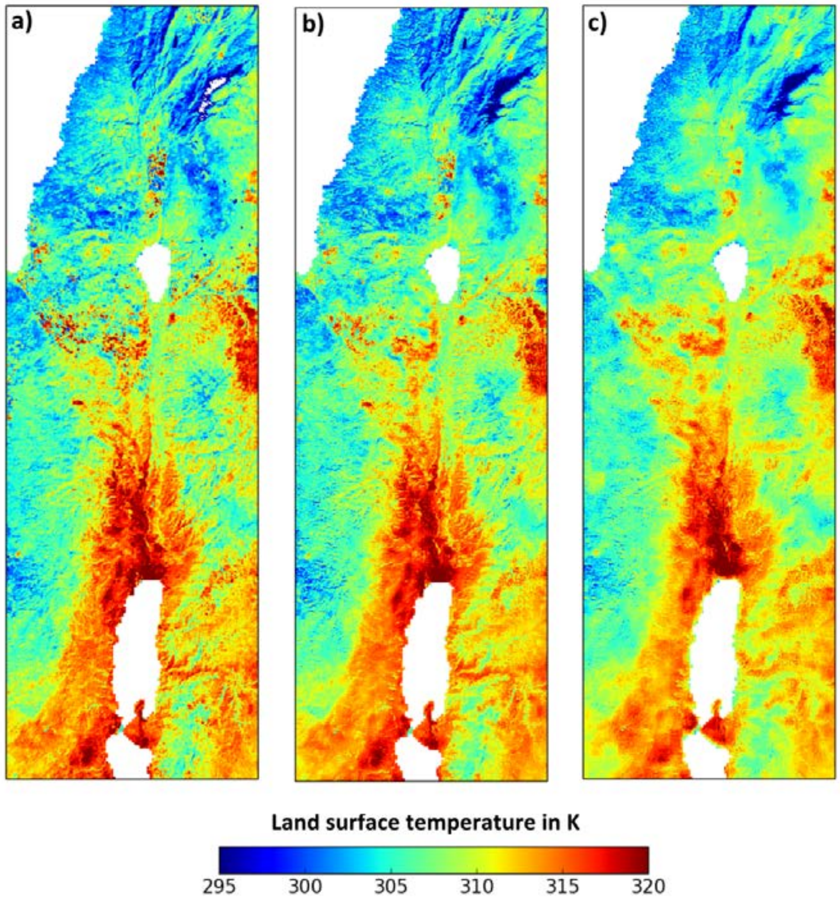


Fig. 2: Landsat reference land surface temperatures at 240 m resolution (a) and downscaled LST from Landsat (b) and Terra MODIS (c).

disaggregation yielded errors of 1.80 K and 2.51 K. Correlation with reference LST improved from $R^2$ = 0.84 to $R^2$ = 0.91 for downscaled Landsat LST and from $R^2$ = 0.72 to $R^2$ = 0.80 for MODIS surface temperatures.

Similar results were obtained for the Hula valley subset (table 2). Downscaling improved RMSE from 2.23 K to 1.61 K for Landsat data and from 2.84 K to 2.60 K for MODIS imagery compared to uniform disaggregation. Increases in correlation were comparatively larger than for the complete study area with $R^2$ values of 0.80 and 0.58 for downscaled Landsat and MODIS LST maps compared to the uniform case, which yielded $R^2$ values of 0.61 and 0.38, respectively.

Tab. 1: Downscaling statistics (uniform and Random Forest downscaling) for the complete region.

| Data | RMSE | $R^2$ |
|---|---|---|
| MODIS (uniform) | 2.51 | 0.72 |
| MODIS (Random Forest)) | 2.24 | 0.80 |
| ETM+ (uniform) | 1.80 | 0.84 |
| ETM+ (Random Forest) | 1.47 | 0.91 |

Tab. 2: Downscaling statistics (uniform and Random Forest downscaling) for the Hula valley subset.

| Data | RMSE | $R^2$ |
|---|---|---|
| MODIS (uniform) | 2.84 | 0.38 |
| MODIS (Random Forest) | 2.60 | 0.58 |
| ETM+ (uniform) | 2.23 | 0.61 |
| ETM+ (Random Forest) | 1.61 | 0.80 |

Land surface temperature distributions for original and downscaled data and both sensors are depicted for the subset area in figure 3 (next page), which displays the visual characteristics of downscaled LST maps in more detail. Downscaling of coarse resolution Landsat surface temperatures reproduced the spatial patterns visible in the reference data. Box-shaped artefacts commonly introduced in VI-based linear models through normalization with coarse resolution LST data (JEGANATHAN ET AL., 2011) are barely visible indicating improved prediction of fine resolution LSTs prior to the addition of model residuals derived at the coarse resolution. Spatial patterns are less well defined in the downscaled MODIS LST map, which shows a distinct blurring effect.

Figure 4 depicts scatterplots for downscaled and reference LSTs for the Hula valley subset. MODIS data shows more scatter around the 1:1-line compared to Landsat LSTs. Downscaled LSTs are biased in regions of high and low surface temperatures with a tendency to overrate LST for low and underrate LST for high reference temperatures. This bias is more pronounced for downscaled MODIS imagery and for high surface temperatures (in both cases).

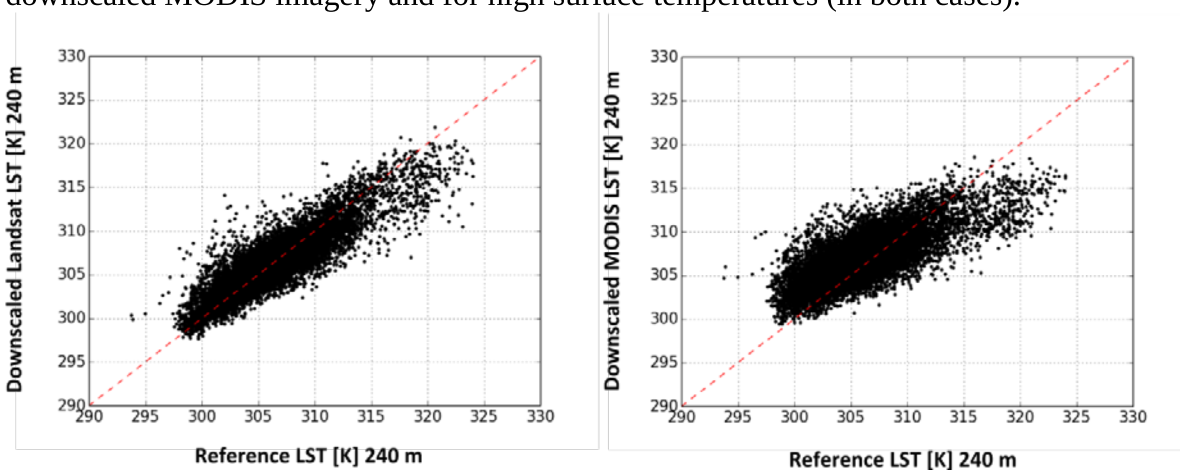


Fig. 4: Scatterplots for downscaled Landsat and MODIS LST (240 m) with respect to 240 m Landsat LST reference for the Hula valley subset.

## 4 Discussion

Evaluation results underline the usability of Random Forest regression for downscaling LST in mixed landscapes with large scale temperature variations, although the downscaling method performed significantly better for coarse resolution Landsat-derived surface temperatures than for MODIS LSTs. Landsat downscaling results are comparable to those reported in VI-based linear

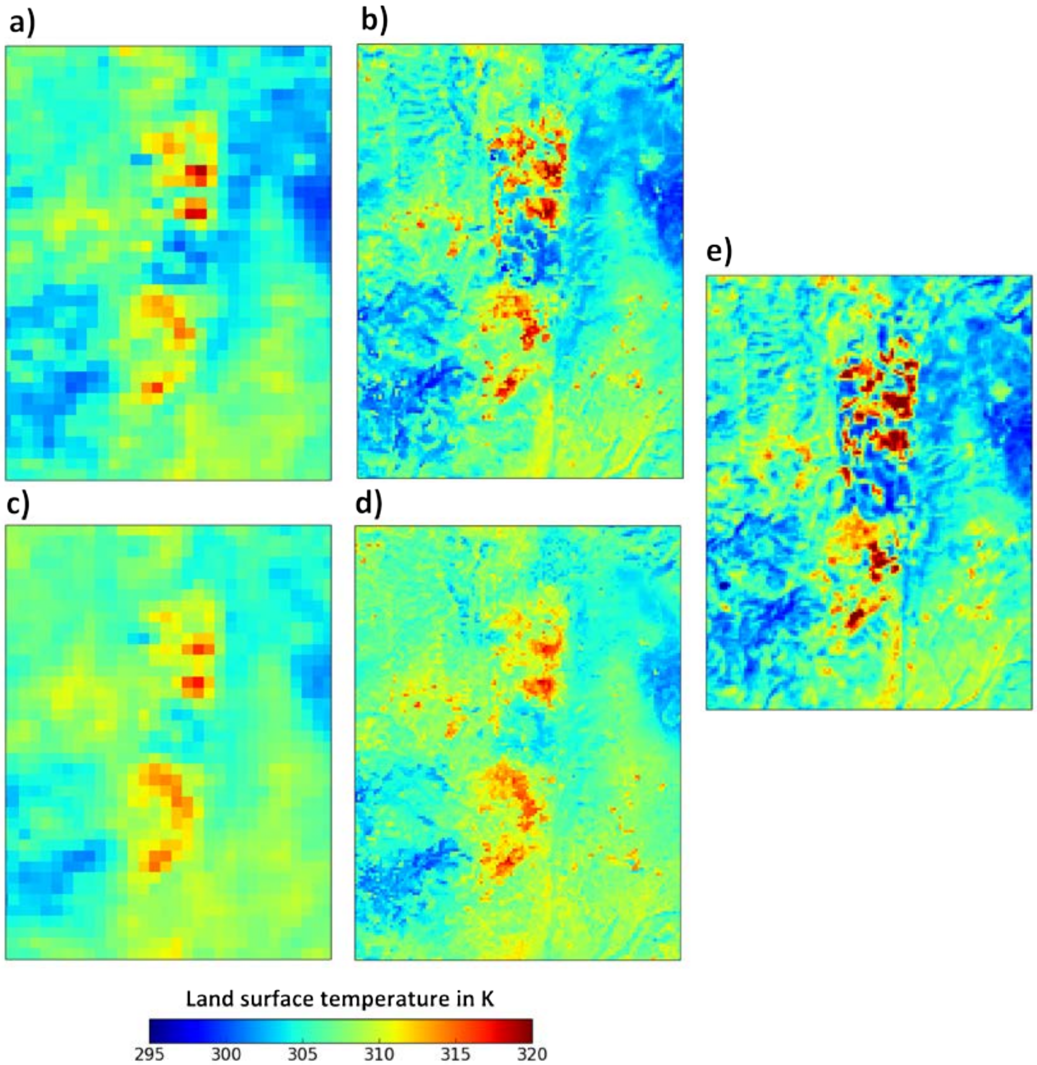


Fig. 3: LSTs for the Hula valley subset: a) Landsat LST 960 m, b) downscaled Landsat LST 240 m, c) MODIS LST 960 m, d) downscaled MODIS LST 240 m, e) Landsat reference LST 240 m.

regression downscaling studies (KUSTAS ET AL., 2003; AGAM ET AL., 2007; JEGANATHAN ET AL., 2011). In this case, Random Forest regression provides the advantage of adapting to temperature variation induced by variable topography and land cover type, while reducing box-shaped artefacts introduced by normalization of downscaled temperatures with coarse resolution data.

Discrepancies between MODIS and Landsat downscaling results can be partly attributed to sensor differences, since Landsat LSTs were used as reference at the target resolution. However, larger improvements of the match between downscaled and reference LST compared to simple uniform disaggregation for Landsat data and a sharper visual appearance indicate further complications regarding MODIS LST downscaling. The true resolution of the MODIS 1-km LST product is lower than its nominal resolution owing to the influence of the sensor's point spread function which smoothes high spatial frequency variation in the image. Coarse-resolution Landsat data were generated with the approximate MODIS spatial response to account for this effect; though downscaling results indicate that the degradation of real MODIS imagery is larger than anticipated. Additionally, MODIS 1-km LSTs were corrected with the corresponding 5-km

LST data fields to correct for the emissivity-related low temperature bias in the 1-km LST retrieval algorithm. While this improves the match between MODIS and Landsat-derived LST, it also further degrades the quality of the LST data fields used to fit the Random Forest model at the coarse resolution. The blurring effect exhibited by downscaled MODIS LST is therefore likely related to a lack of high spatial frequency information necessary to fit robust relationships between predictor and response variables.
The systematic bias associated with downscaled temperatures corresponding to pixels in the tails of the true temperature distribution at the fine scale is a common problem in all downscaling procedures (BINDHU ET AL., 2013), which can be linked to the characteristics of the data used to fit the model. Images acquired at coarse spatial resolution can be conceptualized as a low pass filtered representation of the landscape. With decreasing spatial resolution, intra-pixel variability also decreases and an increasing number of pixels will consist of mixed surfaces with different temperatures, so that at 1 km spatial resolution extreme temperatures are largely smoothed out. A lack of extreme temperature pixels in the training data thus prevents the correct prediction of temperature conditions for the fine target resolution.

# 5 Summary and conclusions

Random Forest regression was evaluated as a method to downscale LST from MODIS LST product resolution (~ 1 km) to a resolution of 240 m to align LST data fields with MODIS surface reflectance bands (in the visible red (620-670 nm) and near-infrared (841-876 nm)). The downscaling method is applicable at large spatial scales in regions with mixed landcover and requires only a small number of readily available input data sets. Downscaling results were comparatively better for Landsat-derived coarse resolution LST which can be attributed to the additional degradation of MODIS data caused by correcting the 1-km MODIS LST product for low temperature bias with the 5-km MODIS LST product. Failure of downscaling methods to account for surface temperature extremes at the fine spatial scale remains an unresolved problem. An approach to correct this systematic bias would improve downscaled LST estimates considerably. Furthermore, better downscaling accuracies could be achieved with additional model input data (e.g. additional surface reflectance bands, more detailed land use and land cover maps). Despite these unresolved issues, Random Forest regression represents an improvement over VI-based linear regression models for large scale applications and can easily be applied to LST data from different sensors or sensor combinations and adapted to available input datasets.

# 6 Literature

AGAM, N.; KUSTAS, W. P.; ANDERSON, M. C.; LI, F. & NEALE, C. M., 2007: A vegetation index based technique for spatial sharpening of thermal imagery. Remote Sensing of Environment, **107**, S. 545-558.

ATKINSON, P. M., 2013: Downscaling in remote sensing. International Journal of Applied Earth Observation and Geoinformation, **22**, S. 106-114.

BINDHU, V.; NARASIMHAN, B. & SUDHEER, K., 2013: 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 Sensing of Environment, **135**, S. 118-129.

BREIMAN, L., 2001: Random Forests. Machine Learning, **45**, S. 5-32.

GAO, F.; KUSTAS, W. P. & ANDERSON, M. C., 2012: A Data Mining Approach for Sharpening Thermal Satellite Imagery over Land. Remote Sensing, **4**, S. 3287-3319.

INAMDAR, A. K. & FRENCH, A., 2009: Disaggregation of GOES land surface temperatures using surface emissivity. Geophysical Research Letters, **36** (2), DOI: 10.1029/2008GL036544.

JIMENEZ-MUNOZ, J.; CRISTOBAL, J.; SOBRINO, J.; SORIA, G.; NINYEROLA, M. & PONS, X., 2009: Revision of the Single-Channel Algorithm for Land Surface Temperature Retrieval From Landsat Thermal-Infrared Data. IEEE Transactions on Geoscience and Remote Sensing, **47**, S. 339-349.

JEGANATHAN, C.; HAMM, N.; MUKHERJEE, S.; ATKINSON, P.; RAJU, P. & DADHWAL, V., 2011: Evaluating a thermal image sharpening model over a mixed agricultural landscape in India. International Journal of Applied Earth Observation and Geoinformation, **13**, S. 178-191.

KUSTAS, W. P.; NORMAN, J. M.; ANDERSON, M. C. & FRENCH, A. N., 2003: Estimating subpixel surface temperatures and energy fluxes from the vegetation index-radiometric temperature relationship. Remote Sensing of Environment, **85**, 429-440.

KUSTAS, W. & ANDERSON, M., 2009: Advances in thermal infrared remote sensing for land surface modelling. Agricultural and Forest Meteorology, **149**, S. 2071-2081.

LIU, Y.; HIYAMA, T. & YAMAGUCHI, Y., 2006: Scaling of land surface temperature using satellite data: A case examination on ASTER and MODIS products over a heterogeneous terrain area. Remote Sensing of Environment, **105**, S. 115-128.

NISHIMA, M.; WOLFE, R.; SOLOMON, D.; PATT, D.; BLANCHETTE, F.; FLEIG, J. & MASUOKA, E., 1997: MODIS Level 1A Earth Location. Algorithm Theroretical Basis Document. Version 3.0 NASA/Goddard Space Flight Center.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M. & DUCHESNAY, M. , 2011: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, **12**, S. 2825-2830.

SANDHOLT, I.; RASMUSSEN, K. & ANDERSEN, J., 2002: A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sensing of Environment, **79**, S. 213-224.

WAN, Z.; ZHANG, Y.; ZHANG, Q. & LIANG LI, Z., 2002: Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. Remote Sensing of Environment, **83**, S. 163-180.

YANG, G.; PU, R.; HUANG, W.; WANG, J. & ZHAO, C., 2010: A Novel Method to Estimate Subpixel Temperature by Fusing Solar-Reflective and Thermal-Infrared Remote-Sensing Data With an Artificial Neural Network. IEEE Transactions on Geoscience and Remote Sensing, **48**, S. 2170-2178.

YANG, G.; PU, R.; ZHAO, C.; HUANG, W. & WANG, J., 2011: 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 Sensing of Environment, **115**, S. 1202-1219.

ZAKŠEK, K. & K. OŠTIR, 2012: Downscaling land surface temperature for urban heat island diurnal cycle analysis. Remote Sensing of Environment, **117**, S. 114-124.