

Department of Earth and Environmental Sciences

PhD program in Chemical, Geological and Environmental Science

Cycle: XXXIV

Curriculum in Environmental Sciences

**DEVELOPMENT OF A NEW LAND SURFACE  
TEMPERATURE PRODUCT FOR  
IMPROVING SATELLITE-BASED  
EVAPOTRANSPIRATION MODELLING IN  
THE EUROPEAN ALPS**

Surname: Bartkowiak  
Registration number: 848463

Name: Paulina

Tutor: Prof. Roberto Colombo

Supervisor: Dr. Mariapina Castelli

Coordinator: Prof. Andrea Franzetti

**ACADEMIC YEAR 2021/2022**



---

## Table of Contents

---

Table of contents.....	i
Abstract.....	iii
Acknowledgments.....	v
Contributions.....	vii
Acronyms.....	ix
Symbols and variables .....	xi
1. Introduction.....	1
1.1. Background: thermal infrared remote sensing .....	1
1.2. Land surface temperature mapping from space .....	2
1.3. Energy balance modelling of evapotranspiration driven by satellite data.....	4
2. Scope of the thesis .....	7
2.1. Main aim of the thesis.....	7
2.2. Sub-objectives and structure .....	8
3. Downscaling Land Surface Temperature from MODIS Dataset with Random Forest Approach over Alpine Vegetated Areas ( <b>Paper 1</b> ) .....	11
4. Land Surface Temperature Reconstruction under Long-term Cloudy-sky Conditions at 250 m Spatial Resolution: Case Study of Vinschgau/Venosta Valley in the European Alps ( <b>Paper 2</b> ) .....	31
5. Two source energy balance modeling of evapotranspiration with enhanced thermal MODIS observations over Alpine region ( <b>Paper 3</b> ) .....	53
6. Conclusions.....	75
7. Outlook .....	77
Appendix.....	xiii
A1. Thermal variability of the reconstruction results over alpine vegetation.....	xiii
References.....	xv



The European Alps have been affected by intensification of meteorological droughts in recent years. Due to changing climatic conditions, the region is vulnerable to deviations in water cycling, which can be observed in the context of evapotranspiration (ET) anomalies. Land surface temperature (LST) is a key factor in regulating the exchange of water and energy between land and atmosphere, which directly relates it to ET. Development of two-source energy balance (TSEB) models driven by thermal remote sensing data has made a significant contribution to estimate ET at large scale. However, their coarse spatial resolution and sensitivity of TIR instruments to cloudy-sky conditions make them insufficient for complex ecosystems, such as mountain regions. To overcome these limitations, this thesis served to develop a new clear-sky land surface temperature product at 250 m spatial resolution, as an alternative to 1-km MODIS LST data, for estimating fine-resolution TSEB fluxes.

In the first part of the thesis, imbalance between spatial resolution of 1-km MODIS LST data was solved by applying a sharpening procedure to obtain daily LST at 250-m spatial resolution. Due to reduced capabilities of LST–VNIR statistical models in complex ecosystems, multi-source predictors, including normalized difference vegetation index (NDVI) and digital elevation model (DEM) were used. Inspired by superiority of machine learning for non-linear problems, relationships between coarse resolution LSTs and 250-m predictors with random forest (RF) algorithm were exploited. The obtained results indicate an improvement of 20% in the agreement between Landsat and the sharpened LST compared to statistics for the original MODIS dataset. The LST models determined averaged RMSE of 2.3°C and MAE of 1.8°C.

In order to reconstruct missing LSTs beneath the clouds, the author proposed a novel approach to predict invalid pixels by exploiting correlation between ground-based LST and air temperature in conjunction with auxiliary variables, e.g., downwelling solar radiation, albedo- and LAI-derived products under long-term cloudy-skies. Considering a high site dependency driven by different land-cover types, LST reconstruction was performed for aggregated stations that represent three vegetation groups: grassland, forest and permanent crops. The gap-filling was performed with two steps: site-based LST modelling from ground-derived variables under cloudy skies, and then applying the fitted models to gridded predictors at subpixel level corresponding to the downscaled output. The reconstruction achieved reliable performance with local data yielding  $R^2$  of 0.84 and RMSE of 2.12°C.

In the last part of the thesis, the resulting LST maps were incorporated into two-source energy balance model of Priestley-Taylor for estimating energy fluxes at 250-m spatial resolution. First, the performance of the model forced by local temperatures was evaluated with measured fluxes from eddy covariance towers. The benchmark simulations for latent (LE) and sensible heat (H) yielded an averaged RMSE of 57  $\text{Wm}^{-2}$  and mean absolute bias (MB) of 26  $\text{Wm}^{-2}$ . Next, the model estimates driven by satellite-based LSTs, i.e., original 1-km MODIS LST product and downscaled maps, were validated against in-situ data. Turbulent fluxes modelled with 250-m LSTs resulted in RMSE of 86  $\text{Wm}^{-2}$  and MB of 55  $\text{Wm}^{-2}$ , which translated to 8% and 15% decrease in the respective errors when compared to TSEB estimates combined with original MODIS LST.

**Keywords:** downscaling, random forest, cloudy-sky conditions, evapotranspiration, TSEB



---

## Acknowledgments

---

This work was partially supported by the Autonomous Province of Bozen-Bolzano in the framework of “Cycling of carbon and water in mountain ecosystems under changing climate and land use” (CYCLAMEN) project (Project Nr. 5/34).



### *Journal articles*

**Bartkowiak, P.**, Castelli, M. and Notarnicola, C., 2019. Downscaling Land Surface Temperature from MODIS Dataset with Random Forest Approach over Alpine Vegetated Areas. *Remote Sensing*, 11(11), p. 1319.

**Bartkowiak, P.**, Castelli, M., Crespi, A., Niedrist, G., Zanotelli, D., Colombo, R. and Notarnicola, C., 2022. Land Surface Temperature Reconstruction under Long-term Cloudy-sky Conditions at 250 m Spatial Resolution: Case Study of Vinschgau/Venosta Valley in the European Alps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

**Bartkowiak, P.**, Castelli, M., Colombo, R., Notarnicola, C. Two source energy balance modeling of evapotranspiration with enhanced thermal MODIS observations over Alpine region (in preparation).

### *Conferences*

**Bartkowiak, P.**, Castelli, M., Colombo, R. and Notarnicola, C., 2020, March. Thermal remote sensing data enhancement over Alpine Vegetated Areas for evapotranspiration modelling. In *EGU General Assembly Conference Abstracts* (p.3484).

Kitz, F., Wohlfahrt, G., Rotach, M.W., Tasser, E., Tscholl, S., **Bartkowiak, P.**, Castelli, M., Notarnicola, C., Dabhi, H. and Simon, T., 2020, May. Cycling of carbon and water in mountain ecosystems under changing climate and land use (CYCLAMEN). In *EGU General Assembly Conference Abstracts* (p. 13970).

Crespi, A., **Bartkowiak, P.**, Castelli, M., Greifeneder, F., Notarnicola, C., Petitta, M. and Zebisch, M., 2019. Drought characterization over Trentino–South Tyrol: assessment of trends in meteorological indices and impacts on vegetation from optical remote sensing imagery.

### *Accomplishments*

In December 2020 I was awarded the SpaceBrains Foundation prize for my PhD research in the field of Earth Observation, which European Space Agency grants to projects that are particularly ambitious and disruptive.



<b>AHI</b>	Advanced Himawari Imager
<b>AI</b>	Artificial Intelligence
<b>AIRS</b>	Atmospheric Infrared Sounder
<b>ASTER</b>	Advanced Spaceborne Thermal Emission and Reflection Radiometer
<b>CLDAS</b>	China Land Data Assimilation System
<b>DEM</b>	Digital Elevation Model
<b>DSSF</b>	Downward Surface Shortwave Flux
<b>EC</b>	Eddy-covariance
<b>ECMWF</b>	European Center for Medium-Range Weather Forecasting
<b>ECOSTRESS</b>	Ecosystem Spaceborne Thermal Radiometer Experiment on Space Station
<b>ESA</b>	European Space Agency
<b>ESRA</b>	European Solar Radiation Atlas
<b>ET</b>	Evapotranspiration
<b>EVI</b>	Enhanced Vegetation Index
<b>FR</b>	Fine spatial resolution
<b>GDAS</b>	Global Data Assimilation System
<b>GOES</b>	Geostationary Operational Environmental Satellite
<b>HypIRI</b>	Hyperspectral Infrared Radiometer
<b>ISS</b>	International Space Station
<b>LAI</b>	Leaf Area Index
<b>Landsat</b>	Land remote-sensing satellite
<b>LST</b>	Land Surface Temperature
<b>MERRA</b>	Modern Era Retrospective-analysis for Research and Applications
<b>MODIS</b>	Moderate Resolution Imaging Spectroradiometer
<b>MODTRAN</b>	Moderate Resolution Atmospheric Transmission
<b>MSG</b>	Meteosat Second Generation
<b>MV</b>	Microwave radiation
<b>NDVI</b>	Normalized Difference Vegetation Index
<b>NLDAS</b>	North American Land Data Assimilation System
<b>S-NPP</b>	Suomi National Polar-orbiting Partnership
<b>OSEB</b>	One-Source Energy Balance
<b>RF</b>	Random Forest
<b>RTM</b>	Radiative Transfer Model
<b>SC</b>	Single-Channel inversion algorithm for LST retrieval
<b>(S)EB</b>	(Surface) Energy Balance
<b>SEBAL</b>	Surface Energy Balance Algorithm for Land
<b>SEBS</b>	Surface Energy Balance System
<b>Sen-ET</b>	Sentinels for Evapotranspiration
<b>SEVIRI</b>	Spinning Enhanced Visible and InfraRed Imager
<b>SLSTR</b>	Sea and Land Surface Temperature Radiometer

<b>SW</b>	Split-Window LST algorithm
<b>TES</b>	Temperature Emissivity Separation
<b>TIR</b>	Thermal Infrared
<b>TIRS</b>	Thermal Infrared Sensor
<b>TSEB</b>	Two-Source Energy Balance
<b>VIIRS</b>	Visible Infrared Imaging Radiometer Suite
<b>VIS</b>	Visible light
<b>VNIR</b>	Visible and Near-InfraRed

---

## Symbols and variables

---

$B_\lambda$	Spectral radiance of a blackbody at a given wavelength [ $\text{W m}^{-2} \mu\text{m}^{-1} \text{sr}^{-1}$ ]
$f_C$	Fractional vegetation cover [-] / [%]
$H$	Sensible heat flux [ $\text{W m}^{-2}$ ]
$h_C$	Canopy height [m]
$H_C$	Sensible heat flux for canopy [ $\text{W m}^{-2}$ ]
$H_S$	Sensible heat flux for soil [ $\text{W m}^{-2}$ ]
$L_\lambda$	Spectral radiance of a non-blackbody material at a given wavelength [ $\text{W m}^{-2} \mu\text{m}^{-1} \text{sr}^{-1}$ ]
$L_\lambda^{sat}$	At-sensor radiance – top of the atmosphere (TOA) [ $\text{W m}^{-2} \mu\text{m}^{-1} \text{sr}^{-1}$ ]
$T_C$	Canopy temperature [K] / [°C]
$T_K$	Kinematic temperature [K] / [°C]
$T_R$	Radiant temperature measured by sensor [ $\text{W m}^{-2}$ ]
$T_S$	Soil temperature [K] / [°C]
$\alpha_{PT}$	potential transpiration coefficient of Priestley-Taylor (default = 1.26) [-]
$\alpha_\lambda$	Absorbance at a given wavelength [-]
$\epsilon_\lambda$	Spectral emissivity at a given wavelength [-]
$\lambda$	Wavelength [ $\mu\text{m}$ ]
$\lambda E$	Latent heat flux [ $\text{W m}^{-2}$ ]
$\lambda E_C$	Latent heat flux for canopy [ $\text{W m}^{-2}$ ]
$\lambda E_S$	Latent heat flux for soil [ $\text{W m}^{-2}$ ]
$\lambda E_{Ci}$	Initial canopy transpiration estimate [ $\text{W m}^{-2}$ ]
$\sigma$	Stefan-Boltzmann constant – $5.6697 \cdot 10^{-8}$ [ $\text{W m}^{-2} \text{K}^{-4}$ ]
$\tau_\lambda$	Atmospheric transmissivity [-]



### 1.1. Background: thermal infrared remote sensing

Thermal infrared (TIR) remote sensing is of crucial importance for better understanding land-atmosphere processes of the Earth system. It provides spatio-temporal information on thermal conditions of the surface for many environmental applications in the fields of meteorological and climatological studies (Czajkowski et al., 2004; Mildrexler et al., 2018), water use management (Sheffield et al., 2018), evapotranspiration modelling (Anderson et al., 2021), and drought assessment (West et al., 2019).

Thermal infrared radiation is an electromagnetic spectrum in the wavelength range between visible light (VIS) and microwave radiation (MW) (Budzier et al., 2011). Even though there is no strict definition of the TIR, many scientists estimate it within 3-14  $\mu\text{m}$  range, from which thermal conditions of the land surface can be estimated (Künzer & Dech, 2013). Remotely sensed TIR data are acquired in two infrared spectral regions: the 3-5  $\mu\text{m}$  and 8-14  $\mu\text{m}$ , known as “atmospheric windows”, where minimalized atmospheric absorption and scattering effect occur. While the first range is additionally contaminated by reflected solar irradiance that is difficult to separate from the daytime land surface radiance, for the latter one, only small part of its spectrum (9.6-10  $\mu\text{m}$ ) is disturbed by ozone. Therefore, for vegetation-oriented applications LST retrieval from the 8-14  $\mu\text{m}$  range is preferred rather than from the mid-wave infrared region.

Satellite-based measurements of radiant flux ( $L_{\lambda}^{sat}$ ) are affected by the atmosphere and land-driven impacts, including reflected incoming radiance from the ground, thermal path radiance between atmosphere and sensor, and Earth’s temperature emissions, which hampers reliable estimation of actual LSTs (Künzer & Dech, 2013). In order to derive the surface-emitted radiance ( $L_{\lambda}$ ), the atmospheric effect is mitigated by simulating its contributions using a radiative transfer models (RTM), such as MODTRAN (Berk, 2005) or by applying specially designed atmospheric functions (Jiménez-Muñoz & Sobrino, 2010). Furthermore, in contrast to the Planck radiation law, the land is not an ideal emitter with 100% efficiency. Therefore, the radiant energy released from the Earth needs to be additionally corrected. Most land-cover types, such as vegetation, urban-related materials, rocks, and soils emit less electromagnetic radiation than estimated from their real kinematic temperature ( $T_k$ ). This intrinsic property of the radiating objects is related to spectral emissivity ( $\varepsilon_{\lambda}$ ) that is expressed as a ratio of the actual radiance  $L_{\lambda}$  of the material to the corresponding radiance of blackbody ( $B_{\lambda}$ ) at the same temperature (Norman & Becker, 1995).  $\varepsilon_{\lambda}$  depends on object types with different thermal efficiencies that vary with the wavelength and physical characteristics of the Earth surface. Typical emissivities of terrestrial materials are from 0.7 to nearly 1.0 in the 8-14  $\mu\text{m}$  spectrum range (Baldrige et al., 2009). In remote sensing studies surface emissivity can be approximated using landcover maps and spectral indices, such as NDVI, to represent vegetation variability and its phenology. Additionally, scientific community estimates  $\varepsilon_{\lambda}$  from a priori emissivity measurements during field campaigns or from NASA ASTER Global Emissivity Database with an accuracy of 0.01 for all spectral bands (Sobrino et al., 2008; Hulley et al., 2015).

There have been proposed several approaches to obtain LST maps from satellite observations in recent years (Li et al., 2013). In general, they are related to sensor-specific characteristics and thus, can be aggregated into two groups. The methods of the first category are commonly used with satellite instruments with maximum two TIR bands using the single-channel (SC) and split-window (SW) algorithms (Jiménez-Muñoz et al., 2004; Jiménez-Muñoz & Sobrino, 2008). For the SC, used with Landsat 4-7 satellites, LST is retrieved by inverting the Planck function,  $B_\lambda$ , with the emitted surface radiance, given a priori emissivity measurements. The standard SW method exploits at-sensor brightness temperatures of two adjacent TIR bands at around 10  $\mu\text{m}$  and 12  $\mu\text{m}$ . The impact of the atmosphere is removed by the differential atmospheric absorption properties between the channels, and next LST is estimated. This concept has been used to derive skin temperatures from different instruments, including Terra/Aqua MODIS, Terra ASTER, Suomi NPP VIIRS, as well TIRS and SLSTR sensors onboard Landsat-8 and Sentinel-3 (Hulley et al., 2019). The second category of LST estimation, commonly applied to sensors with three or more TIR bands, rely on physical-based retrieval of both surface temperature and full emissivity spectrum. The temperature emissivity separation (TES) belongs to the most effective multichannel approaches and in contrast to the SC and SW, it exploits an improved water vapor scaling correction to increase the accuracy of the atmospheric parameters (Gillespie et al., 1999). The TES has been successfully applied to NASA's latest versions of Suomi NPP VIIRS, Terra/Aqua MODIS, and high-resolution ISS ECOSTRESS mission launched in 2018 (<https://ecostress.jpl.nasa.gov>).

## 1.2. Land surface temperature mapping from space

The main application of remotely sensed TIR measurements is understanding of the terrestrial processes through land surface temperature (LST) analysis (Hulley & Ghent, 2019). LST is an important input parameter in climatic, ecological, and hydrological models (Liang et al., 2019). Skin temperature drives the upwelling longwave radiation and turbulent heat fluxes and is therefore a key variable in forcing the SEB-based models to retrieve evapotranspiration (ET) estimates (Fisher et al., 2020; Semmens et al., 2016).

The advent of remotely sensed LST has made it possible to monitor land surface dynamics in more robust way, being one of the most commonly used techniques for environmental mapping at both local and global scales. This can be explained by accessibility and spatial capabilities of spaceborne data with much bigger extent when compared to traditional ground-based measurements. Currently, LST is acquired by a wide range of satellite sensors capturing spatial information with different pixel size (60 m-5 km) and frequency (10 min-16 days) as shown in Tab. 1.1. While geostationary satellites with sub-daily temporal resolution, like the MSG, GOES, and Himawari, allow observing diurnal temperatures of the surface, their coarse spatial resolutions ( $\geq 1$  km) limit their application to global-scale applications and/or relatively homogenous landscapes. The alternative to geostationary sensors is the increasing availability of high spatial resolution (hereafter named as *high-resolution*) platforms, such as Landsat (i.e., TM, ETM+, TIRS, TIRS-2) and Terra ASTER missions (60-120 m), and recently launched multispectral ECOSTRESS scanner capturing data with 70-m pixel size (Tab. 1.1). Although LST maps from these instruments fill the gap regarding spatial resolution, the measurements are rare, e.g., 16-day repeat cycle for Landsat or are irregular for Europe (ECOSTRESS). Moreover, the revisit time for full data use can be even longer due to TIR sensor vulnerability to clouds. Low temporal resolution is not adequate for routine environmental tasks,

such as drought assessment, plant seasonal phenology monitoring, and rapid land surface changes. It should be mentioned that apart from open-source satellite imagery of the government programs, high-resolution thermal data from commercial microsattelite constellations, like Hydrosat’s VanZyl-1 mission, are being designed to overcome the spatiotemporal limitations with the first planned launch in early 2022 (<https://spacenews.com/hydrosat-signs-with-loft-orbital>).

Satellite	Instrument	TIR bands [8-12.5 $\mu\text{m}$ ]	Pixel size	Revisit time	Launch year
MSG	SEVIRI	3	3000 m	15 min	2002
GOES-E	GOES-16	2	2000 m	15 min	2016
AHI	Himawari-8	3	2000 m	10 min	2015
Terra	MODIS	3	1000 m	twice daily	1999
Aqua					2002
S-NPP	VIIRS	4	750 m	daily	2011
Sentinel-3A	SLSTR	2	1000 m	daily	2016
Sentinel-3B					2018
Terra	ASTER	5	90 m	16 days	1999
Landsat-7	ETM+	1	60 m	16 days	1999
Landsat-8	TIRS	2	100 m	16 days	2013
Landsat-9	TIRS-2	2	100 m	16 days	2021
ISS	ECOSTRESS	5	70 m	3-5 days	2018
TBD	HyspIRI	7	60 m	5 days	> 2023

**Tab. 1.1.** Overview of current and future spaceborne thermal missions with their acquisition characteristics (<https://directory.eoportal.org/web/eoportal/satellite-missions/>).

So far, a trade-off between temporal and spatial resolution can be found in daily TIR data offered by MODIS sensor on board Terra and Aqua satellites. In contrast to the previous platforms, the MODIS LST instrument have spatial resolution of 1 km with over a 20-year image collection, which makes it an ideal solution for time-series analyses at regional scales. MODIS images are commonly used in water- and vegetation-oriented studies, including agriculture production and drought estimation (Wan et al., 2004; Swain et al. 2011; Zhang et al., 2015; Zhou et al., 2017). Nevertheless, these research cases and studies on MODIS-based ET modelling (Jia et al., 2009; Ruhoff et al., 2012; Guzinski et al., 2013; Rahimi et al., 2015) were typically conducted over wide-range regions characterized by non-heterogenous landscapes with relatively small topographical variations.

The spectral response from 1-km thermal MODIS data over complex areas, such as the Alps, may be not representative due to spatial details of the terrain and frequent cloud contamination effect resulting in non-valid imagery (Castelli et al., 2018). Until today there are few studies on LST enhancement for ET modelling considering both its coarse spatial resolution and overcast pixels simultaneously. In the framework of the ESA Sen-ET project (<https://www.esa-sen4et.org>) Guzinski & Nieto (2019) implemented thermal sharpening of Gao *et al.* (2012) to provide high-resolution ET estimates over vegetated sites in Denmark with a relatively homogenous and plain terrain. The disaggregation method was applied to the 1-km MODIS LST (Sentinel-3 SLSTR) to obtain downscaled temperature product using spectral information from 30-m Landsat (10-m Sentinel-2) imagery. Although pixel size of the thermal imagery was successfully improved, the potential cloud cover still remain an obstacle for full data use. In this regard, Liu *et al.* (2020)

enhanced 1-km MODIS LST over northern China by gap-filling cloud-covered areas from available clear-sky pixels, and then applied random forest downscaling to obtain LST at 250 m spatial resolution. Even though that LST reconstruction method is correct from the statistical point of view, LST values beneath the clouds physically differ from those under clear skies (Gallo et al., 2011; Good et al., 2016). Especially, if we downscale to regional scale, like mountain areas, the thermal deviations are much more intense, which hampers correct retrieval of actual thermal conditions of the surface from clear-sky pixels.

In the literature, there is little reported about gap-filling methods that deal with long-term cloudiness that modifies surface thermal conditions when compared to clear-sky LSTs. Many studies on passive MV-LST methods and surface energy balance (SEB) approaches provide actual LSTs under all-weather conditions, however, they are limited in terms of coarse spatial resolution and low retrieval precision of the input variables (Duan et al., 2017; Zhang et al., 2019; Martins et al., 2019; Yu et al., 2019). For instance, Xu *et al.* (2019) observed a substantial decrease in the accuracy metrics for recovered cloud-covered pixels from passive microwave instruments in the Tibet when compared to a relatively flatter area in the Heihe River Basin.

Low resolutions of SEB inputs and MV-TIR images has led to development of multi-source data fusion methods that integrate AI-driven modelling with physical-based assumptions to derive actual LST values under cloudy-sky conditions (Belgiu & Stein, 2019; Long et al., 2020; Wu et al., 2021). Zhao & Duan (2020) predicted overcast MODIS LST pixels from random forest algorithm by blending different predictor variables (e.g., elevation, NDVI, EVI, surface albedo) with coarse-resolution DSSF solar radiation maps from MSG/SEVIRI to depict the impact of cloudy-sky condition on the land surface temperature in Spain. Nonetheless, these approaches are still influenced by coarse-resolution data from geostationary satellites and land data assimilation systems, like GLDAS and NLDAS, which prevents their application from regions with high spatial heterogeneities, common over rugged terrains.

### **1.3. Energy balance modelling of evapotranspiration driven by satellite data**

Satellite remote sensing does not provide direct estimates of evapotranspiration (ET), but nevertheless energy balance (EB) modeling forced by spaceborne land surface temperature (LST) has been intensively applied as a valuable tool for ET retrieval from local to global scales (Allen et al., 2011; Kalma et al., 2008; Kustas et al., 2000). The TIR-ET approaches can minimize the need for meteorological input, since they do not require spatially distributed rainfall and soil moisture information, which is an advantage in mountainous areas due to the low spatial density of measurement instruments (Castelli et al., 2018; Kustas & Anderson, 2009). In general, the EB methods can be grouped into two categories: (1) one-source energy balance (OSEB) models that assume land surface as a single big leaf without distinction between soil and vegetation impacts (Allen et al., 2007; Bastiaanssen et al., 1998; Su et al., 2002), and (2) two-source energy balance (TSEB) approaches which estimate separate turbulent fluxes, i.e., sensible ( $H$ ) and latent ( $\lambda E$ ) heat fluxes, from bare soil and canopy sources (Anderson et al., 1997; Norman et al., 1995). In comparison, there are many other ground-based techniques that allow effectively monitoring evapotranspiration. For example, the eddy covariance (EC) method provides energy flux estimates at a field scale by measuring the exchanges of heat, mass, and momentum (Foken et al., 2012; Schotland, 1955). While on-site observations are meaningful for model validation and its adjustments (also for the purposes of this study), they are not

practical for estimating energy fluxes over large areas with highly complex landscapes. In fact, this also applies to the OSEB models, like SEBAL and SEBS, that generalize land surface details to one homogenous vegetation layer. Therefore, in this work, the TSEB model was considered for ET estimation.

The thermal-based two-source energy balance model was originally developed by Norman *et al.*, (1995) to obtain soil evaporation and vegetation transpiration forced by remotely sensed LST observations, vegetation structure parameters, like LAI and canopy height, and meteorological inputs. Given the low spatial resolution of satellite sensors and a related impurity of LST pixels, the model was modified by decomposing single temperatures with the fractional vegetation cover ( $f_C$ ) to separate surface temperature and surface energy fluxes from soil and canopy components (Kustas & Norman, 1999, 2000). Moreover, the partitioning skill of TSEB into two subcomponents gives ancillary information on the moisture of canopy and soil and provides insights into vertical distribution of moisture in the soil profile (Kustas *et al.*, 2013). Consequently, this two-layer approach has been successfully applied for estimating the convective fluxes over diversified ecosystems and was found to be practical in mountain areas with spatial heterogeneities of the land surface (Burchard-Levine *et al.*, 2019; Castelli, 2021; Elfarkh *et al.*, 2020; Kustas *et al.*, 2013). A more detailed information on the surface energy balance modelling is provided in Section 5 of the thesis.



### **2.1. Main aim of the thesis**

The PhD research interest is motivated by increasing demand for spatially continuous high-resolution land surface temperature imagery for water-related applications with main focus on the mountain areas being often underrepresented in global-oriented studies. Considering insufficient spatial resolution of daily remotely sensed LST products for both regional and local applications, the first part of the research was focused on the data-driven enhancement of those grids in respect to the landscape complexity of the Alpine region in Italy. In parallel, due to vulnerability of thermal infrared sensors to cloud contamination and resulting data gaps, the author concentrated on overcoming this limitation by synthesizing multi-source remote sensing datasets. More specifically, since intense overcast conditions modify the energy budget of the Earth, this work aimed to obtain physically meaningful skin temperature under long-term cloudy skies. Based on these two research components, next specific research topic was raised that is a practical application of the improved LST product to TSEB model for deriving evapotranspiration at finer spatial resolution and studying its performance at different levels with regard to various LST inputs and retrieval methods.

The main aim of the thesis is to provide a prototyping LST inputs for monitoring and sustainable management of water resources in the mountain areas by answering following research questions:

1. How can we downscale low spatial resolution land surface temperature data in the vegetated Alpine region?
2. How performant is the AI-based disaggregation approach compared to original coarse resolution data and high spatial resolution imagery offered by other satellite instruments?
3. What are the main controls of land surface temperature estimates considering landcover variability in the Alps?
4. How can we model and predict missing TIR-LST data under long-term cloudy sky conditions over vegetated ecosystems in the region?
5. What are the expected accuracy results of ET estimates from site-based flux measurements and existing remotely sensed LST inputs and approaches?
6. Is it possible to enhance the performance of TSEB modelling forced by the developed LST product?

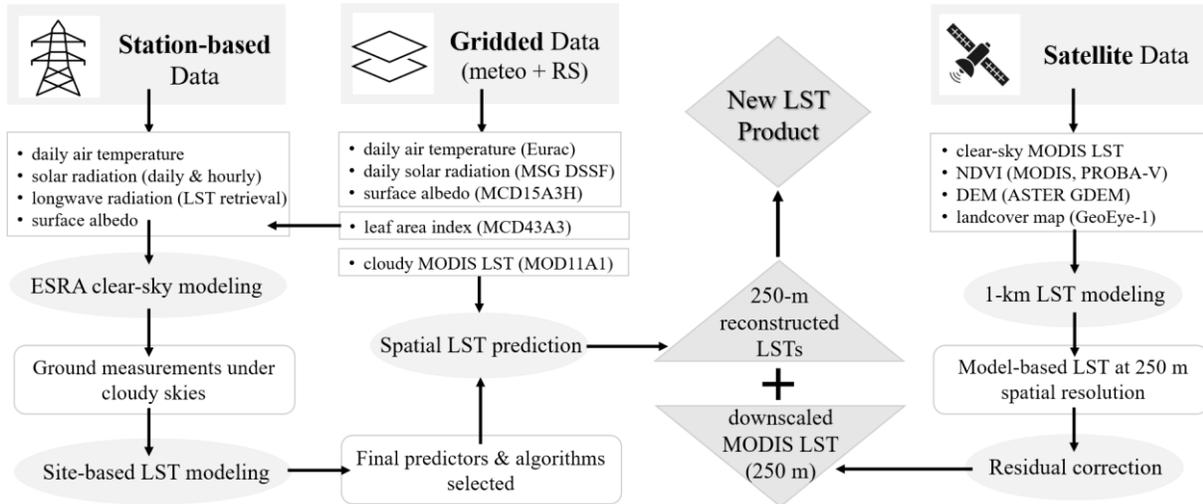
## 2.2. Sub-objectives and structure

This study is a synergy between data-driven analysis from remote sensing and physical-based modelling to assess water availability over vegetated areas in the European Alps. In this regard, an improved land surface temperature product from multi-source spatial data was generated as a key input for the TSEB modelling of evapotranspiration at regional scale.

The research includes the following three sub-objectives:

1. Downscaling 1-km LST data from satellite observations to account for complex terrain and landcover heterogeneities (Section 3).
2. Increasing the temporal coverage of MODIS LST by reconstructing invalid LST maps under cloudy-sky conditions at spatial resolution corresponding to the downscaled output (250 m) (Section 4).
3. Incorporating the downscaled and gap-filled LST in the TSEB model for retrieval of 250-m ET estimates (Section 5).

In Section 3 and Section 4, the development of enhanced land surface temperature product was proposed with main focus on the Alps, being often under-represented in global-scale studies. In this regard, this work aimed to develop methodology to produce 250-m Terra MODIS LSTs, specially tuned to highly heterogenous mountain ecosystems. The concept consists of two-step procedure that includes downscaling of daytime Terra MODIS LST followed by the reconstruction of missing pixels under long-term cloudy sky conditions. Since MODIS sensor after over 20 years in orbit comes to an end soon, the proposed procedures were designed for ongoing ESA Sentinel-3 mission. For this reason, Terra satellite was selected rather than Aqua due its closer observing time to Sentinel-3 overpass. An overview of the proposed methodology is shown in Fig. 1.1.

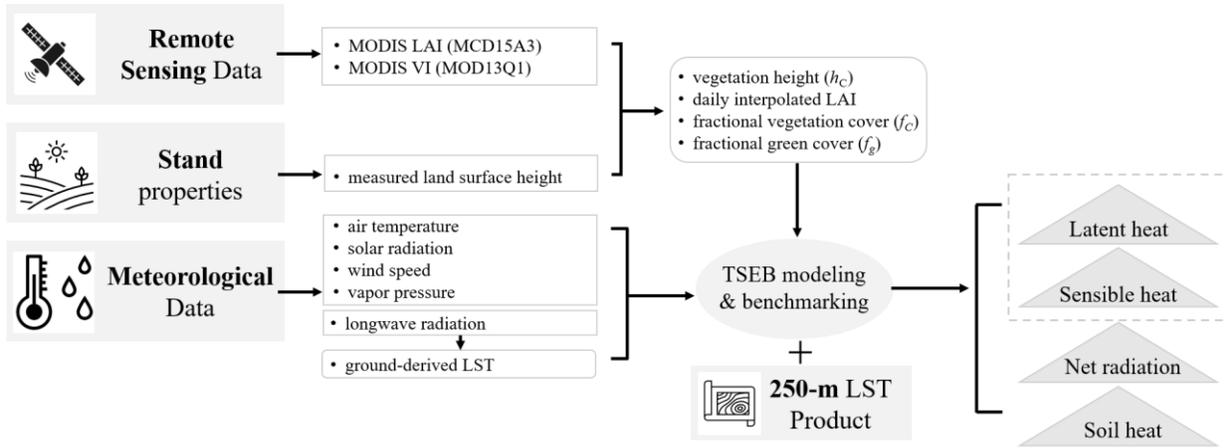


**Fig. 1.1.** Simplified scheme of the entire enhancement procedure (gap-filling and downscaling) for Terra MODIS Land surface temperature data (MOD11A1 Collection 6) developed in R. While rectangles and round rectangles indicate source parameters and intermediate results, processes and (sub)outputs are represented by ovals and (triangles)rhombus, respectively.

As mentioned previously, it is challenging to obtain remotely sensed TIR LST observations at both satisfactory spatial and temporal resolutions. Due to weaker sensitivity of the scanner to emitted radiance, TIR data have always smaller pixel size compared to spectral reflectance from the same satellite platform. For instance, MODIS (Landsat 8-9) VNIR bands register at 250 m (30 m) spatial resolution, whilst thermal data capture information with 1-km (100-m) grid cell (Mao et al., 2021). Considering a wide range of high-resolution images that may explain LST variability, kernel-driven approach was implemented by establishing regressions between 1-km MODIS LST and statistically correlated variables (kernels) at higher spatial resolution to obtain LST data with predictor-like pixel size. Due to reduced capabilities of simple linear LST-VNIR relationships proposed by Kustas *et al.* (2003) and Agam *et al.* (2007), multi-source predictors, including 250-m MODIS NDVI and ASTER DEM were exploited (Dong et al., 2020). Inspired by superiority of machine learning for non-linear problems and availability of AI-based regression tools, daily 250-m sharpened maps were obtained with random forest (RF) algorithm (Hutengs & Vohland, 2016; Li et al., 2019; Njuki et al., 2020; Wu et al., 2019; Yang et al., 2017). When compared to spatiotemporal data-fusion methods (Weng et al., 2014; Xia et al., 2019; Zhu et al., 2010), that combine high-resolution TIR images (e.g., ASTER, Landsat TIRS) with temporally frequent coarse resolution thermal data, like MODIS and VIIRS, the kernel-driven approach is more adequate for mountains, as it is not constrained by low-repeat cycle of high-resolution LST data with often invalid pixels due to clouds. In addition, there are few studies on thermal downscaling tuned to complex landscapes, thus, the novelty of the proposed method is to utilize RF-based sub-pixel mapping of LST to provide reliable spatial information adapted to mountainous ecosystems by considering spatial heterogeneities in terms of topography and landcover (Maeda, 2014; Neteler, 2010).

Next, a novel approach for reconstructing spaceborne TIR-LST influenced by intense overcast was proposed. Section 4 describes development of a station-based method for reconstructing cloud-contaminated LSTs at 250 m spatial resolution over different Alpine ecosystems. The method integrates data-driven analysis with physical-based modelling of sky conditions. For this reason, invalid pixels were predicted by exploiting AI-driven correlations between meteorological variables and surface properties under long-term cloudy-skies derived from ESRA clear-sky radiation model (Rigollier et al., 2000). Next, LST regressions were applied to 250-m gridded predictors to reconstruct cloud-covered MODIS images at subpixel spatial resolution. Areas of applicability were delineated considering spectral similarities and minimum distance between training data and target pixels.

In Section 5, the resulting LST maps within phenological cycle (April-October) were incorporated in the Two-Source Energy Balance (TSEB) model for retrieval of energy fluxes at subpixel level corresponding to the enhanced output. This work is based on the TSEB Priestley-Taylor (TSEB-PT) introduced by Norman *et al.* (1995) and further refined by Kustas & Norman (1999). The approach uses a single LST measurement for the both two layers (soil and vegetation) simultaneously, therefore, the model iteratively calculates soil and canopy sensible heat fluxes ( $H_s$ ,  $H_c$ ) along with their corresponding radiometric temperatures ( $T_s$ ,  $T_c$ ), with an initial estimate of potential canopy transpiration ( $\lambda E_{ci}$ ) according to the Priestley-Taylor equation (Priestley & Taylor, 1972). A general scheme of the model, including its parameters and outputs along with further evaluation is presented in Fig. 1.2. The spatial-based 250-m ET estimates were compared to time-coincident eddy-covariance (EC) measurements and gridded turbulent fluxes forced by original MOD11A1 within two-dimensional flux footprint area (Kljun et al., 2015).



**Fig. 1.2.** General workflow of TSEB modelling. Rectangles and round rectangles indicate source parameters and intermediate results. Processes and outputs are represented by ovals and triangles, respectively. The checked rectangle with TSEB outputs represents parameters of main interest. 250-m LST maps used in this study are time-series imagery produced from MOD11A1 data as described later.

The TSEB Priestley-Taylor model was newly implemented in an open-source Python code (<https://github.com/hectornieto/pyTSEB>). Even though the pyTSEB package was previously validated at different scales over a variety of agricultural sites and natural ecosystems (Burchard-Levine et al., 2021; Bellvert et al., 2020, Guzinski & Nieto, 2019; Nieto et al., 2019), to the best author’s knowledge it has not been tested in the mountainous areas so far.

The model implementation was designed to use both field observations and gridded data. To benefit from these advantages, in-situ measurements from the EC towers in the Matscher/Mazia Valley (Della Chiesa et al., 2014) and remote sensing inputs were applied for further estimation of turbulent fluxes. Since the pyTSEB enables to set user-customized parameters, it can be tuned to different vegetation types by configuring their structural and physiological properties, such as canopy height ( $h_c$ ), leaf spectral reflectance, and green fraction of vegetation ( $f_g$ ), which is beneficial in complex landscapes.

Section 6 and 7 summarizes the whole work conducted giving general conclusion of the outcomes and potential directions for future research on this subject. Further work, that has not been presented or published, is described in *Appendix* section.

All the work described above, including data pre-processing, implementation of the RF downscaling, development of LST reconstruction method, and running the TSEB model (known as pyTSEB), was performed in R and Python programming languages.

The outcomes from each research sub-objective presented above were summarized in the form of research article and published (or submitted) to highly ranked international journals indexed by *Scopus*. They are covered in Sections 3-5.

---

**Downscaling Land Surface Temperature from MODIS Dataset with Random Forest Approach over Alpine Vegetated Areas (Paper 1)**

---

This paper presents the utility of the kernel-driven downscaling with random forest algorithm for retrieval of 250-m MODIS land surface temperature over natural and agricultural ecosystems in the Alps. The LST output was predicted from NDVI and digital elevation model derived at the abovementioned spatial resolution. Three disaggregation models were established considering different thresholds for vegetation content. Time-coincident fine-resolution Landsat LST images were used to compare the downscaling results and original 1-km observations (MOD11A1 Collection 6), and then to test the performance of the disaggregation method with respect to the original MODIS dataset.

---

This section was published in *Remote Sensing* on 1 June 2019, ©MDPI, and it is available online: <https://www.mdpi.com/2072-4292/11/11/1319>. According to *Scopus* the journal received Q1 category in the field of Earth and Planetary Sciences in 2019.

---

Article

# Downscaling Land Surface Temperature from MODIS Dataset with Random Forest Approach over Alpine Vegetated Areas

Paulina Bartkowiak<sup>1,2,\*</sup> , Mariapina Castelli<sup>1</sup> and Claudia Notarnicola<sup>1</sup> 

<sup>1</sup> Eurac Research, Institute for Earth Observation, 39100 Bolzano, Italy; mariapina.castelli@eurac.edu (M.C.); claudia.notarnicola@eurac.edu (C.N.)

<sup>2</sup> Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1 e 4, I-20126 Milano, Italy

\* Correspondence: paulina.bartkowiak@eurac.edu; Tel.: +39-3272-203-972

Received: 24 April 2019; Accepted: 30 May 2019; Published: 1 June 2019



**Abstract:** In this study, we evaluated three different downscaling approaches to enhance spatial resolution of thermal imagery over Alpine vegetated areas. Due to the topographical and land-cover complexity and to the sparse distribution of meteorological stations in the region, the remotely-sensed land surface temperature (LST) at regional scale is of major area of interest for environmental applications. Even though the Moderate Resolution Imaging Spectroradiometer (MODIS) LST fills the gap regarding high temporal resolution and length of the time-series, its spatial resolution is not adequate for mountainous areas. Given this limitation, random forest algorithm for downscaling LST to 250 m spatial resolution was evaluated. This study exploits daily MODIS LST with a spatial resolution of 1 km to obtain sub-pixel information at 250 m spatial resolution. The nonlinear relationship between coarse resolution MODIS LST (CR) and fine resolution (FR) explanatory variables was performed by building three different models including: (i) all pixels (BM), (ii) only pixels with more than 90% of vegetation content (EM1) and (iii) only pixels with 75% threshold of homogeneity for vegetated land-cover classes (EM2). We considered normalized difference vegetation index (NDVI) and digital elevation model (DEM) as predictors. The performances of the thermal downscaling methods were evaluated by the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) between the downscaled dataset and Landsat LST. Validation indicated that the error values for vegetation fraction (EM1, EM2) were smaller than for basic modelling (BM). BM model determined averaged RMSE of 2.3 K and MAE of 1.8 K. Enhanced methods (EM1 and EM2) gave slightly better results yielding 2.2 K and 1.7 K for RMSE and MAE, respectively. In contrast to the EMs, BM showed a reduction of 22% and 18% of RMSE and MAE respectively with regard to Landsat and the original MODIS LST. Despite some limitations, mainly due to cloud contamination effect and coarse resolution pixel heterogeneity, random forest downscaling exhibits a large potential for producing improved LST maps.

**Keywords:** thermal downscaling; MODIS; land surface temperature; random forest; modelling

## 1. Introduction

In this study, we used MODIS datasets to increase spatial resolution of the land surface temperature (LST) images in the Alpine region. We performed modelling by using the digital elevation model (DEM) and normalized difference vegetation index (NDVI) as a set of predictors of 250 m spatial resolution (FR), and the 1 km resolution LST product (CR) as a dependent variable. This paper analyses the strength of random forest downscaling (RFD) using different variants for model creation. With the

use of different approaches, it was possible to examine in which way explanatory variables explain the spatio-temporal LST distribution within the study area.

Thermal remote sensing has significantly contributed to the enhancement of spatio-temporal information about temperature distribution on the surface of the Earth. One of the key possibilities provided by satellite instruments is data acquisition in the thermal infrared (TIR) domain, from which land surface thermal conditions are derived [1–3]. Land surface temperature (LST) retrieved from remote sensing data at different scales is an essential variable in environmental research studies, e.g., in agricultural management [4–7], in urban heat island assessment [8–12], for evapotranspiration (ET) modelling [13–18] and drought monitoring [5,6,19].

National Oceanic and Atmospheric Administration-Advanced Very High Resolution Radiometer (NOAA-AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS) images with daily temporal resolution have improved the quality of continuous Earth monitoring, however their moderate spatial resolution is not sufficient to perform analyses in areas characterized by high spatial heterogeneity in terms of topography and land-cover. Contrary to the aforementioned instruments, thermal images acquired by fine spatial resolution scanners installed on different satellites (e.g., Landsat 5, Landsat 8) are more effective in spatial pattern detection [20–22]. There is still a limitation in the existing satellite instruments since numerous satellites provide data at high spatial resolution but with non-satisfactory repeat cycle (e.g., 16-day revisit time for Landsat) [13,23]. In recent studies [14,24,25] associated with application of thermal remote sensing for continuous evapotranspiration modelling, the authors indicate the need to retrieve data at both high spatial resolution and short repeat cycle. Therefore, in the above-mentioned applications, TIR imagery with both small pixel size and high temporal resolution is highly desirable [26]. Among existing methods for LST pixel size enhancement, downscaling technique is one of the most commonly used approaches in many research studies [27,28].

Thermal downscaling is a technique to retrieve a new LST dataset at finer spatial resolution than original one based on independent variables that represent biophysical properties obtained by remotely-sensed data at higher spatial resolution. This technique exploits the correlation between co-registered fine resolution (FR) with long revisit time and coarse resolution (CR) with short revisit time data to obtain images with both small pixel size and short revisit time. In order to provide improved LST maps, regression modelling is performed based on aggregated independent variables and LST maps. For a better characterization of thermal variations on the Earth surface, it is common to use fine spatial resolution data in modelling, such as individual spectral reflectance bands in visible and infrared electromagnetic regions, spectral indices, digital elevation models or land-use/land-cover (LULC) information [29–33].

The robustness of downscaling methods for analyzing land surface thermal conditions in different landscapes has been well demonstrated in literature. Most of the studies related to LST downscaling concentrate on statistical thermal sharpening techniques exploiting vegetation-based spectral indices (VIs). VIs offer a satisfactory base for exploring relationships between LST and biophysical properties of different land-cover types because of the correlation between land surface temperature and spectral vegetation response [34–37]. The most widely-used method is Disaggregation Procedure for Radiometric Surface Temperature (DisTrad) algorithm which exploits performance of spectral indices to investigate their correlation with thermal bands [24]. In order to disaggregate land surface temperature pixels to the shortwave band resolution, Kustas et al. [24] assume that there is a least-square fitting between NDVI and TIR. Because of the original DisTrad limitations related to ill-defined NDVI distributions over complex regions, some modifications were introduced in the preliminary concept. Agam et al. [25] developed a technique for thermal sharpening (TsHARP) replacing NDVI by fractional vegetation cover (Fv), for which the correlation coefficients were higher compared to the ones obtained with simple spectral indices. Using high biomass area in the study, Qiu et al. [38] introduced a refinement evaluating a new spectral index in the DisTrad model and showed that performance of enhanced vegetation index (EVI) for LST sub-pixel mapping was a more robust approach for LST modelling.

Classical methods used successfully for topographically uniform areas and relatively homogenous land-cover are not efficient in complex regions [39,40]. Therefore, application of other explanatory variables instead of individual VI-based predictors has been performed in many studies [40–44]. As shown by Bechtel et al. [42], the introduction of additional parameters, such as averaged TIR images, land-cover based products or data obtained from image dimensionality reduction (e.g., Principle Component Analysis, PCA), improved the goodness-of-fit in the modelling. Many authors suggested that the application of simple multivariate modelling increases the accuracy of LST downscaling. Results obtained by Maeda [29] demonstrated that when applying DEM in combination with NDVI, higher coefficients of determination ( $R^2 \geq 0.95$ ) were observed. Furthermore, Duan & Li [43] proposed successfully to use LST, NDVI and altitude in geographically weighted regression (GWR) to eliminate stationary effect affected by simple univariate and multivariate regression approaches.

Although traditional statistical techniques seem to be promising in many case studies, some authors suggest the use of machine learning techniques in order to take into account the non-linearity between predictors and LST [41,45]. Even if we exploit more mathematically complex regressions, machine learning methods can incorporate many explanatory variables in the physical-deterministic modelling. Furthermore, due to the use of machine learning in thermal sharpening, automatic data production shows more robust results. Therefore, different techniques have been exploited, including Bayesian-based modelling [46], support vector regression [47], artificial neural networks [48] and random forests (RF) [39,49]. Pioneers in the use of RF technique for downscaling purposes were Hutengs & Vohland [39] who developed sharpening of simulated Landsat thermal maps with ground sampling distance (GSD) equals to 60 and 120 m to retrieve Landsat-like images of VNIR spatial resolution (GSD = 30 m). The authors applied multi-source data fusion in the model, including reflectance bands and DEM products over vegetated areas, yielding good model performance. No method was universally recognized as better than the others, though RF modelling is a promising technique over terrain influenced by multiple factors.

Since there are few studies related to thermal sub-pixel mapping for complex topography, the novelty of the paper is to implement RF image sharpening to overcome LST limitations for mountainous ecosystems by taking into account topographic features and land-cover heterogeneity [29,39,50]. Due to evolving meteorological conditions, the Alps are a hot spot of climate change and of drought extreme events. Thus, TIR remotely-sensed data at fine spatial resolution are highly desirable.

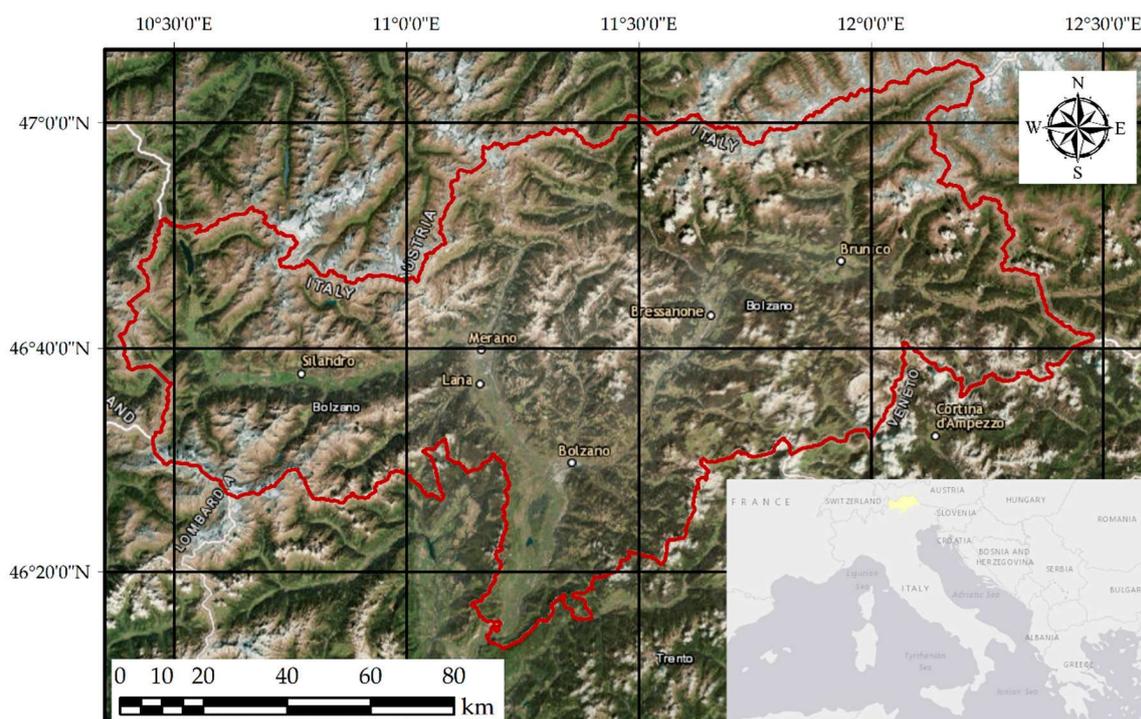
This paper consists of the following paragraphs: Section 2 shows information about data used and describes methodology as well as the applied RF downscaling algorithm, Section 3 presents evaluation of the results. The last two paragraphs deal with discussion related to the study findings and conclusion for the whole paper.

## 2. Materials and Methods

### 2.1. Study Area

The study area is located in the central-eastern Alps in the northern part of Italy, the Autonomous Province of Bolzano/Bozen, and it covers area of 7400 km<sup>2</sup> (Figure 1).

The region is predominantly mountainous, divided from north to south by the Adige/Etsch river valley. Elevation ranges from 110 m to about 3685 m a.s.l. The mountainous topography determines land-cover heterogeneity. The region is mainly covered by forests, grasslands, pastures and agricultural areas. Due to complexity of the environment and topographical location, the Province of Bolzano/Bozen is one of the driest areas in the Alps with diversified microclimate zones [51]. The climate type is differentiated depending on geographical location; sub-continental for the main valleys, continental in the mountain valleys and alpine within the regions above forest border [50].



**Figure 1.** Geographical location of the study area (source: map was created using ArcGIS® software by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under license. Copyright © Esri).

## 2.2. Input Data

In this study we used as coarse resolution input (CR) MODIS/Terra—the MOD11A1.006 Land Surface Temperature (LST) product [52], daily at 1 km spatial resolution, and MODIS/Terra and Aqua NDVI 4-Day composites (see Table 1 for details), at 250 m spatial resolution [53]. Since our main aim is to evaluate downscaling possibilities over vegetated areas, we selected imagery within vegetation phenological cycle, from April to November. MODIS data were acquired from April 2003 to June 2017. The data were produced and qualified in this time frame during the MONALISA project [53]. The grids were originally obtained in MODIS Sinusoidal projection (SR-ORG: 6974). The MOD11A1 products were subset and downloaded from Google Earth Engine, and consisted of daytime surface temperature distribution with an accuracy of approximately 1 K [54]. To get cloud uncontaminated image collection, before data downloading we applied the bitmask using JavaScript API available in Google Earth Engine platform. Only imagery with good quality flag (QA) was used in the study. The 4-Day NDVI MODIS data produced by Eurac research within the MONALISA project [55] included composites created with the use of Maximum Value Composite (MVC) algorithm for all vegetated areas (according to Corine Land Cover 2012) from MODIS/Aqua and Terra images (MOD/MYD09GQ and 09GA Version 6) [56–59]. The final dataset consisted in time-series of images from 2003 to 2017, from April to November, months with green vegetation cover.

**Table 1.** Overview of original and processed datasets in this study.

Dataset Name	Spatial Resolution	Short Description	Derivatives Products
MOD11A1	1000 m	MODIS Terra Land Surface Temperature	Resampled to NDVI spatial resolution (GSD = 250 m) for residual correction
NDVI 4-Day composites	250 m	MODIS Terra and MODIS Aqua Reflectance bands based on MOD09GA	NDVI aggregated to 1000 m pixel size by spatial averaging
ASTER GDEM	30 m	Global Digital Elevation Model acquired ASTER scanner	GDEM resampled to 250 m and 1000 m spatial resolution
LISS 2013 -Land Information System South Tyrol	-	Land Information System South Tyrol based on GeoEye-1 image collection from 2012	Vegetation vector masks for: - areas covered by minimum 90% of vegetation within 1km pixels - pure pixels for different types of vegetation (75% threshold of homogeneity)
Landsat 5	30 m 120 m (30) <sup>2</sup>	Atmospherically-corrected reflectance products (red, NIR), Thermal at-sensor radiance	Gaussian filtering and resampling to 250 m
Landsat 8	30 m 100 m (30) <sup>2</sup>	Atmospherically-corrected reflectance products (red, NIR), Thermal at-sensor radiance	Gaussian filtering and resampling 250 m
MOD05_L2	1000 m	MODIS Precipitable Water	Re-projected to MODIS Sinusoidal Projection

<sup>1</sup> GSD: Ground Sampling Distance. <sup>2</sup> resampled products by data provider.

In order to evaluate the performances of the proposed downscaling method in the area of interest, we validated our results with Landsat images (WRS-2 Path: 192, Row: 28) reprojected to MODIS Sinusoidal Projection (Table 2). Data applied for the algorithm were retrieved through the NASA Earthdata explorer [60] and the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC). The acquired multispectral images in VNIR domain are characterized by Surface Reflectance Level-2, while thermal bands are characterized by Level-1, therefore a pre-processing was needed. We selected for the validation Landsat images with the same acquisition dates as MODIS Terra LST. The time difference between Landsat (Landsat 5 and Landsat 8) and MODIS Terra LST composites did not exceed 1.5 hours. Due to the lack of good quality data caused by common cloud contamination and shadows effects, we selected seven time-coincident Landsat (Table 2) and seven MODIS LST (MOD11A1) images with the corresponding 4-day NDVI Terra/Aqua combined composites (Table 1) acquired in different seasons. To perform the atmospheric correction of Landsat thermal bands we used MODIS Water Vapor product (MOD05\_L2) [61,62]. To retrieve Landsat LST maps, we performed three main steps: (i) calculation of the at-sensor brightness temperature ( $T_B$ ) and of NDVI, (ii) calculation of emissivity ( $\epsilon$ ) based on vegetation fraction [62] and (iii) land surface temperature retrieval [61,62].

**Table 2.** Landsat dataset used in the validation process.

Date	Sensor	Granule ID	Overpass Time (GMT)
27.09.2004	Landsat 5 TM	LT05_L1TP_192028_20040927_20161129_01_T1	09:42
25.05.2005		LT05_L1TP_192028_20050525_20161126_01_T1	09:45
16.10.2005		LT05_L1TP_192028_20051016_20161124_01_T1	09:46
18.07.2007		LT05_L1TP_192028_20070718_20161112_01_T1	09:52
12.09.2010		LT05_L1TP_192028_20100912_20161013_01_T1	09:48
27.08.2016	Landsat 8 TIRS	LC08_L1TP_192028_20160827_20170321_01_T1	09:58
11.06.2017		LC08_L1TP_192028_20170611_20170627_01_T1	09:58

For the downscaling procedure, in addition to MODIS NDVI, we used information on elevation provided by the Global Digital Elevation Model (GDEM), acquired by the Advanced Spaceborne

Thermal Emission and Reflection Radiometer (ASTER), at approximately 30 m resolution [63]. For the original dataset we performed bilinear resampling to a resolution of 250 m and then we applied spatial averaging to 1000-m pixel using mean function within  $4 \times 4$  kernel size [39].

To select and characterize vegetated areas, in this study we exploited the LULC vector layers based on high spatial resolution RapidEye imagery (GSD = 6.5 m) of 2012, produced by the Institute for Earth Observation of Eurac research in the frame of the project “LISS-2013—Land-use information in South Tyrol: update, harmonization with European Standards and integration of research results” (© Autonome Provinz Bozen – Südtirol | 28.0.1 Landeskartographie und Koordination der Geodaten). This dataset includes minimum polygon at the level of 1600 m<sup>2</sup>, contrary to the more commonly used Corine Land Cover (CLC) database, for which the minimum area of individual object is 25 ha. An overview of the datasets used in the study is shown in Table 1.

### 2.3. Methodology

#### 2.3.1. The Random Forest Algorithm

RF regression is a machine learning technique exploiting statistical nonlinear relationship between variables. This dataset required for RF regression consists of observations ( $n$ ), including predictors ( $p$ ) and dependent variable. Basic component of random forest algorithm is regression tree, and the main concept is to build many different subsets. Randomness is introduced by the fact that each tree is created based on random sample with replacement of  $n$  from training data (bootstrap sample). The method determines the best split by increasing similarity between features in each sub-node. One method to achieve this is to minimize variance or to reduce standard deviation of variables on the basis of random sample. In practice,  $p/3$  of randomly sampled variables is used as split candidates. No pruning is applied when constructing the trees, which means any split is removed from the model. Random forest predictions are calculated for each regression tree separately and then arithmetic average of the trees as final forecast is performed. Basic equation describing final RF prediction for regression results based on created trees is presented as follows:

$$F(x) = \frac{\sum_{j=1}^N T_j(x)}{N} \quad (1)$$

where  $N$  indicates number of trees,  $T_j$  represents each tree and  $F$  is a prediction at a new point  $x$  as an averaged prediction based on created trees [64].

Since the random forest regression divides dataset into decorrelated trees and then predicts a target variable based on average value obtained from the subsets, it is robust for high dimensional grid data [65]. RF regression is considered easy in optimization by applying a relatively small number of parameters in modelling. Random forest procedure requires specifying a few user parameters, including tree quantity and minimum number of observations in each tree [64,65]. Many authors showed that RF regression enables to use continuous dataset as well as categorized predictors, such as land-cover information [39]. In theory, a RF model can explain LST values by multivariate relationship, considering different inputs, like remotely-sensed spectral indices, incoming solar radiation, digital elevation models, terrain slope and aspect, land-cover maps or soil moisture [39,49].

#### 2.3.2. Implementation of Random Forest for Thermal Sharpening

MODIS Land Surface Temperature, with its spatial resolution of 1 km, does not provide accurate spatial information for many environmental analyses performed within specific land-cover types at the level of individual regions and municipalities [13,14]. Therefore, in this study, the random forest algorithm was selected to model the relationship between the 1 km MODIS LST and the 250-m DEM and time-coincident 4-days composites of MODIS NDVI. Downscaling procedure was implemented in R software and GDAL Utilities using packages intended for that purpose [66–69].

Apart from the spectral indices, it is known that for mountainous areas topography variations have a large impact on vegetation types and phenology. Because of adiabatic temperature differences vegetation cover changes with elevation, which results in diversified flora zones. Furthermore, height gradients lead to big biodiversity of spatially distributed land-cover and therefore it influences LST distribution [70]. For this reason, in this work we modelled the relationship between low resolution daytime LST imagery and spatial information at 250 m pixel size, including NDVI and digital elevation model.

After resampling DEM to NDVI resolution, we performed mean aggregation on FR data to simulate parameters of CR dataset. Next, random forest regression was carried out. The number of trees was set to 1000 based on a “trial & error” approach as a compromise between processing burden and accuracy [4,71]. The whole downscaling procedure was based on the relationship between coarse resolution images and then applied to the FR pixels. Modelling was based on the following relationship:

$$LST_{CR} = f(NDVI_{FR_{deg}}, DEM_{FR_{deg}}) \quad (2)$$

where the subscript *CR* means variable at coarse resolution and the subscript *FR<sub>deg</sub>* indicates degraded fine resolution predictors to 1000 m pixel size, with *f* a nonlinear function of combined altitude and NDVI established by the RF method.

When finishing RF regressions for a single image from the image collection, models were applied to 250-m pixels to predict FR land surface temperature maps. Consequently, simulated FR LST values were obtained using the following Equation (3):

$$LST_{FR} = f(NDVI_{FR}, DEM_{FR}) \quad (3)$$

To approximate scale effect on the correlation between LST and predictors, many authors [39,41,72] have used residual correction originally implemented by Kustas et al. [24]. This operation involves the following steps: (i) aggregation by averaging of FR LST images to original MODIS LST pixel size, (ii) residuals calculation ( $\Delta$ ) by subtracting re-aggregated fine resolution LST pixels ( $LST_{FR_{deg}}$ ) from original LST images ( $LST_{CR}$ ), (iii) CR residuals resampling to  $LST_{FR}$ -like spatial resolution and adding these corrections to LST predictions from modelling ( $LST_{FR}$ ), as shown in Equation (5):

$$\Delta = LST_{CR} - LST_{FR_{deg}} \quad (4)$$

$$LST_{FR_{res}} = LST_{FR} + \Delta \quad (5)$$

where  $LST_{FR_{res}}$  indicates simulated LST pixels at 250 m spatial resolution after addition of residual correction.

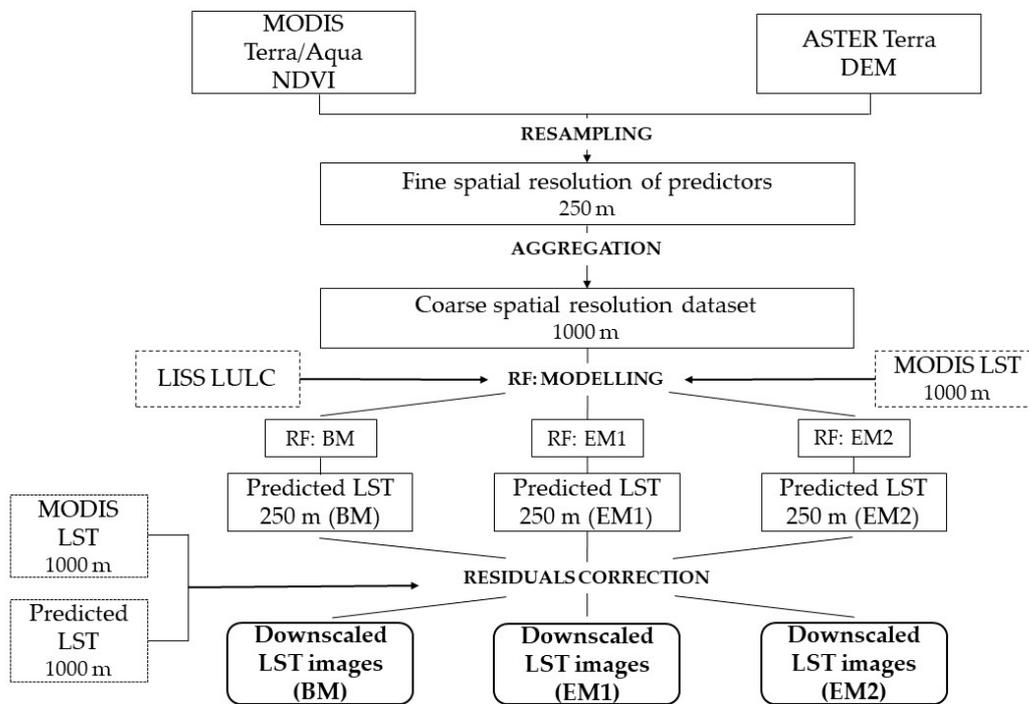
### 2.3.3. Random Forest Model Concepts

Due to topography complexity and land-cover heterogeneity in the Alpine region, prior to final selection of predictors we tested explanatory performance of different variables, including NDVI, DEM, aspect, sky view factor maps as well as MODIS near-infrared bands. NDVI and DEM were selected in the estimation of land surface temperature variations dependent on altitude as they showed to be the most relevant predictors. We tested the model as well based on all the predictors and we found that on average the difference among the two models was around 0.71 K.

Moreover, we evaluated three different approaches: (i) basic modelling (BM), (ii) enhanced modelling 1 (EM1) and (iii) enhanced modelling 2 (EM2). In the BM method all good quality pixels covering the study area were applied in RFD. It means that pixels including different land-cover types within squares of 1 km length were exploited for the image sharpening [39]. The two latter model concepts arise from the need to retrieve homogeneity characteristics at initial MODIS LST scale. In terms of homogenous pixels-based downscaling we evaluated regressions for every vegetated land cover class within the region, including forest, vineyards and orchards, annual crops, grassland and bushes.

As reported by Kustas et al. [24] and then by Essa et al. [73], the selection of homogenous low-resolution pixels improves correlation results between spectral indices and LST. Therefore, we conducted RFD using EM1-based approach as a piecewise modelling regarding fractional vegetation cover and then we applied stricter criteria to the third model, including only pure pixels (EM2). Contrary to BM, EM1 is a more advanced disaggregation method based only on pixels mostly covered by vegetation. The selection of pixels at low spatial resolution allowed extracting data that exhibit a certain degree of homogeneity. Since not all pixels contained vegetation, we applied a threshold of 90% for vegetation within 1 km pixel-based mask (EM1) and then using 75% homogeneity criteria based on vegetation classes obtained from the LISS 2013 (EM2). The last model assumed that pixels within the same vegetation class should be similar, considering spectral- and spatial-neighborhood patterns. In this way the model EM2 represents areas with full homogeneity characteristics.

Based on the above-mentioned thresholds we applied three LST prediction approaches within 1-km fishnet pixel-based mask. Due to high level of heterogeneity and common cloud contamination effect in the Alpine region, many pixels were excluded from further analysis. The performed image sub-pixel sharpening approach is summarized in Figure 2.



**Figure 2.** Overview of the procedure for downscaling land surface temperature from MODIS data.

#### 2.3.4. Data Preparation for the Validation Phase

To assess RF modelling performance, we calculated Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) between the downscaled MODIS LST and Landsat LST degraded to 250 m pixel size. We calculated RMSE and MAE as follows:

$$RMSE = \sqrt{\frac{\sum_{j=1}^N (LST_{FR_{res}} - LST_{Lt})^2}{N}}; MAE = \frac{\sum_{j=1}^N |LST_{FR_{res}} - LST_{Lt}|}{N} \quad (6)$$

where  $N$  is the number of observations, and subscript of  $Lt$  indicates land surface temperature for Landsat pixels degraded to downscaled MODIS-like spatial resolution.

To assess the model performances, we validated our results by using Landsat imagery acquired in different seasons. Downscaled LST images were validated with the use of Landsat 5 TM and Landsat 8

TIRS LST maps having a common extent of MODIS tiles. The Landsat dataset was resampled to 250 m pixel size to simulate the CR data.

After applying residuals calculated on the basis of Equations (4) and (5), we processed the Landsat dataset to retrieve LST at 250 m spatial resolution. Different approaches have been used to atmospherically correct satellite-based thermal data [74–76]. Among various techniques, the Single Channel (SC) algorithm has been widely applied as an efficient method for LST retrieval [61,77]. In this study we applied the SC of Jiménez Muñoz et al. [71]. This LST retrieval method is based on the following general formula developed by Jiménez-Muñoz et al. [62]:

$$LST = \gamma \left[ \frac{1}{\varepsilon} (\psi_1 L + \psi_2) + \psi_3 \right] + \delta \quad (7)$$

where  $\psi_1$ ,  $\psi_2$ ,  $\psi_3$  are derived from MODIS water vapor content for each sensor separately [77], and  $\gamma$ ,  $\delta$  are given by:

$$\gamma = \left\{ \frac{c_2 L}{T_B^2} \left[ \frac{\lambda^4}{c_1} L + \lambda^{-1} \right] \right\}^{-1}; \quad \delta = -\gamma L + T_B \quad (8)$$

where  $L$  is at-satellite radiance,  $c_1$  and  $c_2$  are sensor constants,  $\lambda$  indicates the effective wavelength of thermal bands.

Atmospheric water vapor content (WV) for the Landsat acquisition dates (Table 2) did not exceed  $3 \text{ g cm}^{-2}$ . Based on tests conducted by Sobrino & El Kharraz [78], MODIS water vapor product (MOD05\_L2) with WV concentration smaller than  $3 \text{ g cm}^{-2}$  is reliable input in Single Channel LST retrieval. According to authors, SC algorithm based on MOD05\_L2 at maximum level of  $3 \text{ g cm}^{-2}$ , provides RMSE values of about 1.5 K [61,77].

After applying smoothing Gaussian filtering and mean value cell area weighted resampling for upscaling, we calculated LST maps for the selected dates [39]. Landsat images used for the validation are summarized in Table 2.

### 3. Results

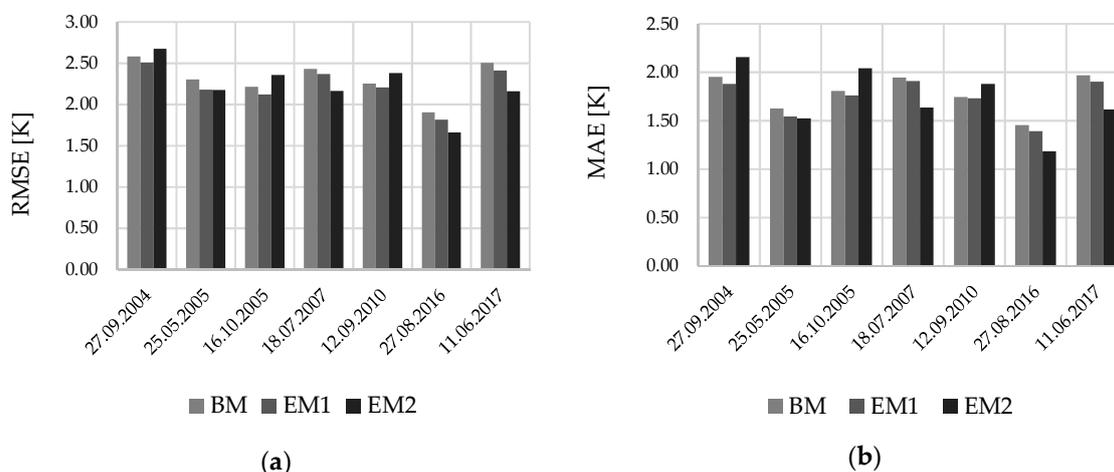
#### 3.1. Global Validation

As mentioned before, to evaluate the effectiveness of the BM, EM1 and EM2 random forest regression, we conducted spatial degradation of TIR bands from Landsat 5 and Landsat 8 to downscaled MODIS 250 m pixel size (Table 1). Moreover, a comparison with MODIS LST at 1000 m is carried out in order to understand what is the level of improvement obtained with the downscaling procedure. Since this paper is intended for vegetation analyses, RMSE and MAE were calculated for seven sharpened images acquired in different seasons (spring, summer, autumn). Figure 3 shows RMSE and MAE for the proposed model, averaged in the study area.

As can be observed, random forest algorithm, regardless of pixel selection method, was able to predict sub-pixel LST with similar accuracy. All disaggregation methods based on fractional vegetation mask slightly improved the accuracy of the downscaled 250 m LST layouts. Overall RMSEs (MAEs) ranged from 1.81 K to 2.51 K (1.39 K to 1.91 K) for EM1 and 1.66 K to 2.67 K (1.18 K to 2.16 K) for enhanced modelling 2. Uniform modelling (BM) yielded RMSE values from 1.90 K to 2.58 K and MAE from 1.45 K to 1.97 K, respectively. The largest RMSE was found for LST image dated on 27 September 2004 ( $RMSE_{BM} = 2.58 \text{ K}$ ,  $RMSE_{EM1} = 2.51 \text{ K}$ ). Similar situation applied to MAE evaluation index ( $MAE_{BM} = 1.95 \text{ K}$ ,  $MAE_{EM1} = 1.88 \text{ K}$ ), which corresponded to an accuracy improvement of 4% compared to BM.

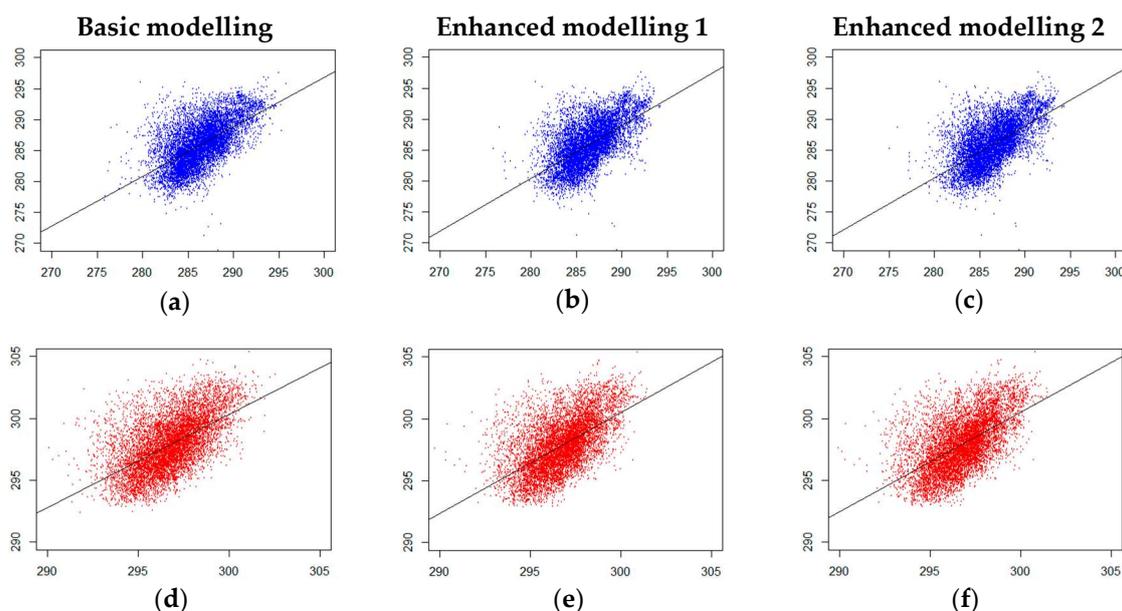
Regression performances differed depending on acquisition date. Apart from images acquired in September and October, we achieved a positive effect of enhanced modelling (EM1, EM2) on the accuracy results (Figure 3). The greatest errors appeared for the scenes from early autumn (27 September 2004) and from late spring (11 June 2017), with RMSE of about 2.6 K and 2.4 K, respectively.

Such discrepancies, documented in Figure 3, may have occurred because of different periods of growing season, the limited number of samples and their small variability incorporated in the RF modelling. This is likely due to weather conditions during image acquisitions, with dense cloud cover and shadowed surfaces over the Alps. These factors influenced negatively the determination of the impact of biomass content and elevation in the prediction of land surface temperature spatial patterns.



**Figure 3.** Validation results based on average Root Mean Square Error (RMSE) (a) and Mean Absolute Error (MAE) (b) between disaggregated images and reference images for Basic Modelling (BM), Enhanced Modelling 1 (EM1), and Enhanced Modelling 2 (EM2).

Similar behavior in the context of models' performances presents scatterplot comparison of the downscaled MODIS LST and Landsat references (Figures 3 and 4). The graphs created for dates with the highest and the lowest global RMSEs and MAEs (Figure 3) show almost the same spatial distributions for all the exploited approaches. The graphs indicate that different vegetation thresholds for applied models did not introduce noticeable improvements (Figure 4).



**Figure 4.** Scatterplots of the sharpened MODIS land surface temperature (LST) [K] [X-axis] versus Reference Image [K] [Y-axis] (degraded Landsat LST) for the applied RF models (BM, EM1, EM2). The top row shows results obtained on 27 September 2004 (a–c) and the lower line presents outcomes for 27 August 2016 (d–f).

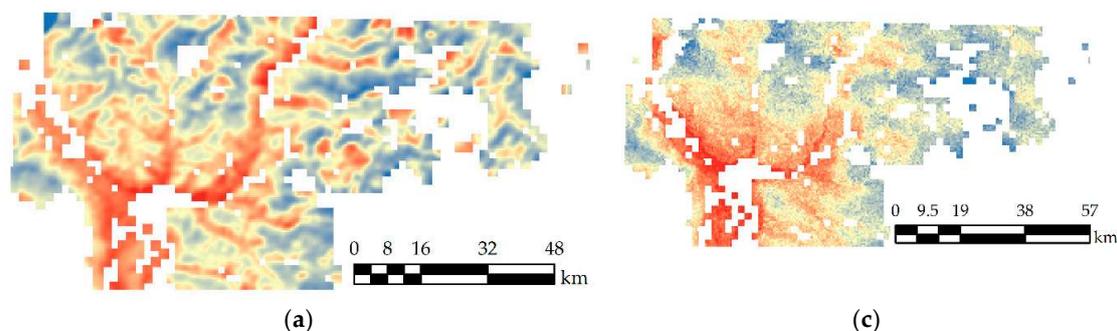
Considering the averaged global analysis, we achieved better results for BM with disaggregated MODIS LST than for the original datasets (Table 3). EM1 and EM2 yielded quite similar RMSE and MAE values, which indicates a reduced contribution of different variables incorporated in both LST sharpening methods. In contrast to the latter methods,  $RMSE_{mean}$  ( $MAE_{mean}$ ) for basic modelling was equal to 2.31 K (1.79 K), which corresponded to accuracy improvement of 22% (18%) compared to statistics obtained for Landsat and the original MODIS (Table 3). The above-mentioned outcomes (Figures 3 and 4) indicate that the proposed RF approaches gave quite similar LST spatial patterns.

**Table 3.** Overview of averaged Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) ( $RMSE_{mean}$ ,  $MAE_{mean}$ ) based on all Landsat validation data at global scale.

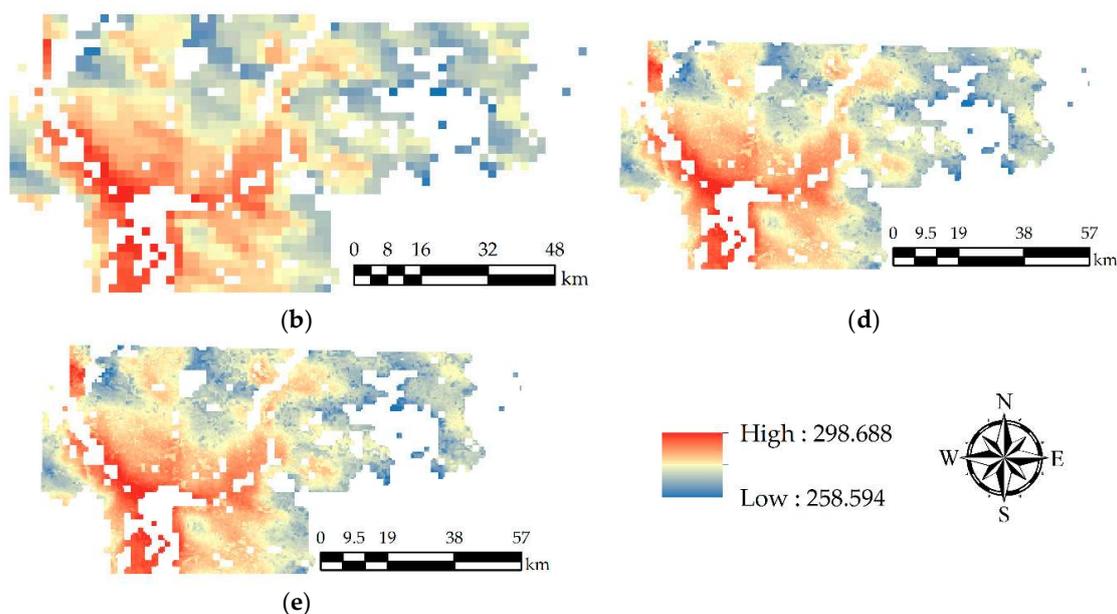
Landsat LST vs. original MODIS LST			Landsat LST vs. sharpened MODIS LST		
$RMSE_{mean}$					
BM	EM1	EM2	BM	EM1	EM2
2.97	2.38	2.44	2.31	2.23	2.22
$MAE_{mean}$					
BM	EM1	EM2	BM	EM1	EM2
2.18	1.83	1.85	1.79	1.73	1.72

In addition to the assessment of the model performances provided by RMSE and MAE, we conducted a qualitative assessment of the results based on visual analysis for the methods used, by comparing downscaled LST with original MODIS LST at 1 km resolution and with Landsat LST downgraded at 250 m resolution (Figure 5).

Differences in spatial patterns between the original MODIS and the downscaled imagery can be noticed (Figure 5b–e). This behavior is especially visible for the valleys within the study area. Notable visual improvements are noticeable in the north east valleys for which Landsat-like thermal variability with higher LST values was obtained (Figure 5c). Based on qualitative assessment, downscaling procedures allow identifying finer details. In contrast to the enhanced approaches (EM1, EM2), larger number of samples regarding BM modelling influenced bigger variability of downscaled pixels (Figure 5c–e). Comparing Landsat LST after applying vegetation mask (Figure 5a) with the downscaled MODIS Terra LST products (Figure 5c–e), we observed that sharpened pixels showed similar distribution triggered by the original CR imagery (Figure 5b). The above-mentioned outcomes (Figures 3–5) indicate that the proposed RF approaches gave similar LST spatial patterns as well. In most cases degraded Landsat LST maps had a broader temperature value distribution than the original MODIS LST at 1 km spatial resolution. This means that RFD affected by CR LST data could not be able to simulate thermal variations at sub-pixel level over the area (e.g., on 27 September 2004, Landsat LST values ranged from 258.6 K to 298.7 K and MODIS LST values ranged from 277.6 K to 294.6 K).



**Figure 5.** Cont.



**Figure 5.** Visual comparison between MODIS images and the time-coincident Landsat LST maps degraded to 250 m spatial resolution. The left column shows: (a) degraded Landsat reference image (250 m), and (b) the original MODIS LST (1000 m), and the right column corresponds to downscaled imagery sharpened by BM (c), EM1 (d), EM2 (e) on 27 September 2004.

### 3.2. Validation for the Different Land Cover Classes

In parallel with the global validation, to assess the influence of land-cover type on model results we performed the validation based on the RMSEs for each vegetated class obtained from LISS resources. Additionally, the original MODIS LST was compared against degraded Landsat LST products. Due to better performance obtained for basic modelling (Table 3), in Tables 4 and 5 we show BM prediction errors for five LULC, including forest, vineyards and orchards, annual crops, grassland and bushes.

**Table 4.** Local comparison of RMSE between reference images and sharpened MODIS data obtained by basic modelling (BM) for different land-cover types.

Landsat LST vs. Sharpened MODIS LST						
RMSE (K)						
Date	Forest	Vineyards & Orchards	Annual Crops	Grassland	Bushes	AVERAGE
27.09.2004	2.57	2.55	2.19	3.03	3.53	2.78
25.05.2005	1.84	2.67	1.24	3.30	5.09	2.83
16.10.2005	2.19	1.45	1.05	2.50	4.08	2.25
18.07.2007	2.10	4.33	3.60	2.95	3.26	3.25
12.09.2010	2.10	2.66	2.97	2.78	3.36	2.77
27.08.2016	1.58	2.39	2.47	2.66	3.22	2.46
11.06.2017	2.09	3.32	4.13	3.38	4.43	3.47
<b>AVERAGE</b>	2.07	2.77	2.52	2.94	3.85	-

**Table 5.** Local comparison of RMSE between reference images and original MODIS data obtained by BM approach.

Landsat LST vs. original MODIS LST						
RMSE (K)						
Date	Forest	Vineyards & Orchards	Annual Crops	Grassland	Bushes	AVERAGE
27.09.2004	2.73	2.75	2.30	3.53	3.85	3.03
25.05.2005	1.96	2.46	1.03	4.05	5.51	3.00
16.10.2005	2.24	1.62	2.71	3.06	4.64	2.86
18.07.2007	2.20	3.98	3.02	3.32	3.91	3.29
12.09.2010	2.21	2.56	3.41	3.30	4.57	3.21
27.08.2016	1.65	2.11	2.02	3.05	3.67	2.50
11.06.2017	2.19	2.97	3.23	3.85	4.48	3.34
<b>AVERAGE</b>	2.17	2.64	2.53	3.45	4.38	-

As can be seen in the above tables (Tables 4 and 5), at the level of different vegetation types, accuracies of thermal downscaling are diversified and therefore RMSE values increased in many cases. Considering global and local evaluation measures, we found that larger deviations between downscaled/original MODIS LST and Landsat LST were obtained at land-cover level. Regarding the size of RMSE values, the BM yielded less satisfactory results for bush class, for which standard deviation ranged 0.63–0.71 K and the  $RMSE_{mean}$  was 3.85 K for sharpened MODIS and 4.38 K for the original MODIS (Tables 4 and 5). Due to the complicated structure of vineyards and orchards, which influence sub-pixel variability, RMSE ranged from 1.45 K to 4.33 K. A similar behavior was observed on 11 July 2017 for annual crops, grassland and bushes classes yielding maximum RMSE errors of BM ranging from 3.23 K to 4.48 K (Tables 4 and 5). On average, for the forest class the prediction error (Table 4) was markedly lower ( $RMSE_{mean} = 2.17$  K), corresponding to an improvement of 2–6% in prediction accuracies (RMSE) against the original MODIS LST (Table 5). Despite bigger RMSE and MAE values than for other classes, the largest improvements after downscaling were observed for grasslands and bushes yielding 15% and 12% decreases in RMSE, respectively. As already noted for the global comparison, the BM at local scale did improve accuracy outcomes. On average, differences in  $RMSE_{BM}$  between the original MODIS and the downscaled outcomes ranged from  $-0.13$  K to 0.60 K. Only  $RMSE_{mean}$  ( $RMSE_{mean} = -0.13$  K) for 11 June 2017 indicated lack of improvements in the RFD procedure.

#### 4. Discussion

The proposed random forest algorithm allows predicting LST at the medium spatial resolution (250 m) yielding satisfactory results. Model performance assessment based on quantitative comparison between MODIS original and downscaled images showed BM as more robust method than other approaches. Considering the applied random forest models (BM, EM1, EM2) exploiting DEM and NDVI explanatory performance, findings from this study indicate that prediction error differences between downscaled images obtained by different implementation of RF regressions were insignificant (Figure 3, Table 3). Regardless of number of prediction variables and masks applied, differences in RMSEs between BM and EMs on average did not exceed 0.1 K. In this case the outcomes from the paper are not in agreement with results obtained by Kustas et al. [24] who showed that purity pixel selection caused improvement in downscaling procedure as a big potential for spatial ET modelling. When comparing the prediction errors (Figure 3) to the ones observed by Merlin et al. [79], Jeganathan et al. [41], Maeda [29] and Hutengs & Vohland [39], we noted that our evaluation measures had on average lower values. Hutengs & Vohland [39] applied random forest regression to MODIS LST data with RMSEs ranging from 1.41 K to 1.92 K. Similarly, good results were obtained by Maeda [29]. The author showed that by using simple multivariate regression (DEM and NDVI as predictors) it was possible to downscale daytime 1-km images with maximum standard error equals to 3.29 K and

coefficient of determination,  $R^2_{\min} = 0.78$ . Above-mentioned results are site-dependent and are highly influenced by topographic complexity as well as by vegetation heterogeneity over study area.

RFD exhibits a large potential for producing enhanced LST maps, despite some limitations, mainly due to cloud contamination and heterogeneity of 1-km pixels. In fact, within 2003–2017 image collection, the unavoidable presence of contaminated pixels may have impact on the final quality of the downscaling outputs. Furthermore, since the downscaled images maintained the spatial characteristics of the original 1 km thermal data, the possibility of extracting homogenous pixels based on vegetation cover fraction is hampered.

Vegetation content in the Alpine region is differentiated, meaning that within 1 km fishnet based on MODIS LST many pixels were not dominated by vegetation. Since in today's research on downscaling techniques many authors have shown that the introduction of a big number of predictors improves model performance [39,49], our preliminary model concept included diversified explanatory variables, like topography-derived variables (terrain slope and aspect) and reflectance bands. Nonetheless, by introducing these additional data we achieved similar resampled Landsat-downscaled MODIS relationship and, as a result, we obtained the same spatial information as with the predictors which we have chosen. Moreover, the exploitation of masks derived from the LISS land-cover, based on vegetation content thresholds, led to similar mean prediction errors both generally and for single land-cover classes. Based on preliminary tests, the high precision of the vector layer did not improve the 250 m outputs with a difference ranging from 0.5 to 1.1 K. Therefore, for the application of the proposed method to larger areas, Corine Land Cover maps with a Minimum Mapping Unit (MMU) of 25 ha could be an alternative to LISS 2013.

Statistical comparison between 250 m spatial resolution Landsat LST and the sharpened as well as the original MODIS LST maps provided an overview for the assessment of the downscaling accuracy. By applying Gaussian filtering and then upscaling resampling to Landsat LST it was possible to partially smooth spatial details registered by high resolution sensors. However, there were still some LST discrepancies caused by complex terrain affected by thermal anisotropy, changeable meteorological conditions between time acquisitions or different viewing angle of sensors. Another issue was related to uniformly disaggregated MODIS residuals which kept spatial information at CR pixel base. Thus, future work will focus on the development of residual correction based on its normalization prior to applying it to predicted pixels [30,32].

Unfortunately, due to the lack of ground data for the chosen Landsat acquisitions (Table 2), comparison of the results with field measurements was not conducted. Future work should include multiple ground-based radiometric measurements, distributed spatially in order to incorporate LST values for different land cover classes within test sites.

As shown in Table 3, since there is a high level of agreement between the three applied model concepts, further evaluation of random forest should be considered. The spectral response for heterogeneous areas of the Alps represents various biomes, which are spatially and temporarily limited. Due to the common cloud contamination over the study area, it was not possible to combine these methods with data acquired from different sensors, such as Landsat 8 TIRS-1. On the other hand, by using high spatial resolution data for the downscaling procedure and then applying FR mask, it would be more effective to retrieve thermal contrast within smaller pixels corresponding to different materials. Given this approach, heterogeneity of LULC could be overcome by extraction of spectrally pure grid cells representing unique vegetated land-cover classes. This could bring to significant accuracy improvements for random forest modelling over the complex area of the Autonomous Province of Bolzano/Bozen.

## 5. Conclusions

This study presents an evaluation of the random forest algorithm for downscaling MODIS LST, based on the relationship between land surface temperature and static and dynamic variables, including digital elevation model and NDVI, in the heterogeneous ecosystems of the Alps. The

application of three RF model concepts was needed to investigate whether the selection of specific pixels based on land-cover criteria contributed to explain LST distribution over the Province of Bolzano/Bozen. This paper demonstrates that RF machine learning regression, regardless of the method used, was capable of modelling non-linear relationships between variables in a very robust way.

The performance of the proposed regressions (BM, EM1, EM2) against co-registered Landsat images yielded quite similar results. A comparison based on statistical measures indicated that, on average, RMSE (MAE) ranged from 1.66 K to 2.67 K (from 1.18 K to 2.16 K). At level of single land cover classes, temperature deviations were also observed with significant LST differences for grasslands and bushes (Tables 4 and 5).

Considering these difficulties, further studies should firstly focus on the implementation of approaches to predict missing pixels, in order to perform LST downscaling based on full image coverage and enhance modelling robustness. Future work will also concentrate on the development of improved LST sharpening methods to reduce the gap between FR data with low temporal resolution and coarse resolution imagery acquired at daily basis. Considering the recent progress in the availability of satellites acquiring FR imagery, such as Landsat and Sentinel missions; further modifications of the LST downscaling will exploit these datasets, which is an urgent need for studies related to climatology, drought monitoring, and water management. These activities require high spatial resolution datasets to retrieve detailed information about spatial variability of LST. Considering this issue, downscaling procedures exploiting both MODIS LST and new high spatial resolution space-borne instruments have a large potential for regional water availability assessments.

**Author Contributions:** Conceptualization, P.B., M.C. and C.N.; Methodology, P.B.; Formal analysis, P.B.; Investigation, P.B.; Writing—original draft preparation, P.B.; Writing—review and editing, P.B., M.C. and C.N.; Supervision, M.C., C.N.

**Funding:** This study has been carried out within the project “Cyclamen” financed by Provincia Autonoma di Bolzano, Alto Adige/Süd Tirol, Ripartizione Diritto allo studio, Università e ricerca scientifica, within the scope of Provincial Law 14 from 13. December 2006 (Project Nr. 5/34).

**Acknowledgments:** The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Czajkowski, K.P.; Goward, S.N.; Mulhern, T.; Goetz, S.J.; Walz, A.; Shirey, D.; Dubayah, R.O. Estimating environmental variables using thermal remote sensing. In *Thermal Remote Sensing in Land Surface Processes*; Quattrochi, D.A., Luvall, J.C., Eds.; CRC Press LLC: Boca Raton, FL, USA, 2004; pp. 11–32.
2. Weng, Q. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS J. Photogramm.* **2009**, *64*, 335–344. [[CrossRef](#)]
3. Niclòs, R.; Tomás, S.; Juan, M.; Valiente, J.A.; Barberà, M.J.; Caselles, D.; Caselles, V. *Evaluation of Landsat-8 Thermal Bands to Monitor Land Surface Temperature*; Serie A; Publicaciones de la Asociación Española de Climatología: Sevilla, Spain, 2014; Volume 9.
4. Park, S.; Im, J.; Jang, E.; Rhee, J. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agric. For. Meteorol.* **2016**, *216*, 157–169. [[CrossRef](#)]
5. Swain, S.; Wardlow, B.D.; Narumalani, S.; Tadesse, T.; Callahan, K. Assessment of vegetation response to drought in Nebraska using Terra-MODIS land surface temperature and normalized difference vegetation index. *GISci. Remote Sens.* **2011**, *48*, 432–455. [[CrossRef](#)]
6. Anderson, M.C.; Zolin, C.A.; Sentelhas, P.C.; Hain, C.R.; Semmens, K.; Yilmaz, M.T.; Tetrault, R. The Evaporative Stress Index as an indicator of agricultural drought in Brazil: An assessment based on crop yield impacts. *Remote Sens. Environ.* **2016**, *174*, 82–99. [[CrossRef](#)]

7. Anderson, M.C.; Hain, C.; Wardlow, B.; Pimstein, A.; Mecikalski, J.R.; Kustas, W.P. Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States. *J. Clim.* **2011**, *24*, 2025–2044. [[CrossRef](#)]
8. Voogt, J.A.; Oke, T.R. Thermal remote sensing of urban climates. *Remote Sens. Environ.* **2003**, *86*, 370–384. [[CrossRef](#)]
9. Lu, D.; Song, K.; Zang, S.; Jia, M.; Du, J.; Ren, C. The effect of urban expansion on urban surface temperature in Shenyang, China: An analysis with landsat imagery. *Environ. Model. Assess.* **2015**, *20*, 197–210. [[CrossRef](#)]
10. Balçık, F.B. Determining the impact of urban components on land surface temperature of Istanbul by using remote sensing indices. *Environ. Monit. Assess.* **2014**, *186*, 859–872. [[CrossRef](#)]
11. Hardy, C.H.; Nel, A.L. Data and techniques for studying the urban heat island effect in Johannesburg. In Proceedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Berlin, Germany, 11–15 May 2015; Volume XL-7/W3, p. 203.
12. Deilami, K.; Kamruzzaman, M. Modelling the urban heat island effect of smart growth policy scenarios in Brisbane. *Land Use Policy* **2017**, *64*, 38–55. [[CrossRef](#)]
13. Anderson, M.C.; Allen, R.G.; Morse, A.; Kustas, W.P. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sens. Environ.* **2012**, *122*, 50–65. [[CrossRef](#)]
14. Castelli, M.; Anderson, M.C.; Yang, Y.; Wohlfahrt, G.; Bertoldi, G.; Niedrist, G.; Notarnicola, C. Two-source energy balance modeling of evapotranspiration in Alpine grasslands. *Remote Sens. Environ.* **2018**, *209*, 327–342. [[CrossRef](#)]
15. Tang, R.; Li, Z.L. Evaluation of two end-member-based models for regional land surface evapotranspiration estimation from MODIS data. *Agric. For. Meteorol.* **2015**, *202*, 69–82. [[CrossRef](#)]
16. Rahimi, S.; Gholami Sefidkouhi, M.A.; Raeini-Sarjaz, M.; Valipour, M. Estimation of actual evapotranspiration by using MODIS images (a case study: Tajan catchment). *Arch. Agron. Soil Sci.* **2015**, *61*, 695–709. [[CrossRef](#)]
17. Anderson, M.C.; Norman, J.M.; Mecikalski, J.R.; Otkin, J.A.; Kustas, W.P. A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *J. Geophys. Res. Atmos.* **2007**, *112*. [[CrossRef](#)]
18. Carpintero, E.; Dugo, M.G.; Hain, C.; Nieto, H.; Gao, F.; Andreu, A.; Anderson, M.C. Continuous evapotranspiration monitoring and water stress at watershed scale in a Mediterranean oak savanna. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*; International Society for Optics and Photonics: Leiden, The Netherlands, 2016; Volume 9998, p. 99980N.
19. Park, S.; Feddema, J.J.; Egbert, S.L. Impacts of hydrologic soil properties on drought detection with MODIS thermal data. *Remote Sens. Environ.* **2004**, *89*, 53–62. [[CrossRef](#)]
20. Xing, Q.; Li, L.; Lou, M.; Bing, L.; Zhao, Z.; Li, Z. Observation of oil spills through Landsat thermal infrared imagery: A case of deepwater horizon. *Aquat. Procedia* **2015**, *3*, 151–156. [[CrossRef](#)]
21. Semmens, K.A.; Anderson, M.C.; Kustas, W.P.; Gao, F.; Alfieri, J.G.; McKee, L.; Xia, T. Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach. *Remote Sens. Environ.* **2016**, *185*, 155–170. [[CrossRef](#)]
22. Allan, M.G.; Hamilton, D.P.; Trolle, D.; Muraoka, K.; McBride, C. Spatial heterogeneity in geothermally-influenced lakes derived from atmospherically corrected Landsat thermal imagery and three-dimensional hydrodynamic modelling. *Int. J. Appl. Earth Obs.* **2016**, *50*, 106–116. [[CrossRef](#)]
23. Wulder, M.A.; White, J.C.; Loveland, T.R.; Woodcock, C.E.; Belward, A.S.; Cohen, W.B.; Roy, D.P. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* **2016**, *185*, 271–283. [[CrossRef](#)]
24. Kustas, W.P.; Norman, J.M.; Anderson, M.C.; French, A.N. Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship. *Remote Sens. Environ.* **2003**, *85*, 429–440. [[CrossRef](#)]
25. Agam, N.; Kustas, W.P.; Anderson, M.C.; Li, F.; Neale, C.M. A vegetation index based technique for spatial sharpening of thermal imagery. *Remote Sens. Environ.* **2007**, *107*, 545–558. [[CrossRef](#)]
26. Weng, Q.; Fu, P.; Gao, F. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sens. Environ.* **2014**, *145*, 55–67. [[CrossRef](#)]
27. Zurita-Milla, R.; Kaiser, G.; Clevers, J.G.; Schneider, W.; Schaepman, M.E. Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sens. Environ.* **2009**, *113*, 1874–1885. [[CrossRef](#)]
28. Atkinson, P.M. Downscaling in remote sensing. *Int. J. Appl. Earth Obs.* **2013**, *22*, 106–114. [[CrossRef](#)]

29. Maeda, E.E. Downscaling MODIS LST in the East African mountains using elevation gradient and land-cover information. *Int. J. Remote Sens.* **2014**, *35*, 3094–3108. [[CrossRef](#)]
30. Bindhu, V.M.; Narasimhan, B.; Sudheer, K.P. 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.* **2013**, *135*, 118–129. [[CrossRef](#)]
31. Bisquert, M.; Sánchez, J.M.; Caselles, V. Evaluation of disaggregation methods for downscaling MODIS land surface temperature to Landsat spatial resolution in Barrax test site. *IEEE J. Sel. Top. Appl.* **2016**, *9*, 1430–1438. [[CrossRef](#)]
32. Essa, W.; Verbeiren, B.; van der Kwast, J.; Batelaan, O. Improved DisTrad for Downscaling Thermal MODIS Imagery over Urban Areas. *Remote Sens.* **2017**, *9*, 1243. [[CrossRef](#)]
33. Lillo-Saavedra, M.; García-Pedrero, A.; Merino, G.; Gonzalo-Martín, C. 2018 TS2uRF: A New Method for Sharpening Thermal Infrared Satellite Imagery. *Remote Sens.* **2018**, *10*, 249. [[CrossRef](#)]
34. Zhou, X.; Wang, Y.C. Dynamics of Land Surface Temperature in Response to Land-Use/Cover Change. *Geogr. Res.* **2011**, *49*, 23–36. [[CrossRef](#)]
35. Wan, Z.; Wang, P.; Li, X. Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA. *Int. J. Remote Sens.* **2004**, *25*, 61–72. [[CrossRef](#)]
36. Yue, W.; Xu, J.; Tan, W.; Xu, L. The relationship between land surface temperature and NDVI with remote sensing: Application to Shanghai Landsat 7 ETM+ data. *Int. J. Remote Sens.* **2007**, *28*, 3205–3226. [[CrossRef](#)]
37. Karnieli, A.; Agam, N.; Pinker, R.T.; Anderson, M.; Imhoff, M.L.; Gutman, G.G.; Goldberg, A. Use of NDVI and land surface temperature for drought assessment: Merits and limitations. *J. Clim.* **2010**, *23*, 618–633. [[CrossRef](#)]
38. Qiu, J.; Yang, J.; Wang, Y.; Su, H. A comparison of NDVI and EVI in the DisTrad model for thermal sub-pixel mapping in densely vegetated areas: A case study in Southern China. *Int. J. Remote Sens.* **2018**, *39*, 2105–2118. [[CrossRef](#)]
39. Hutengs, C.; Vohland, M. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sens. Environ.* **2016**, *178*, 127–141. [[CrossRef](#)]
40. Mukherjee, S.; Joshi, P.K.; Garg, R.D. A comparison of different regression models for downscaling Landsat and MODIS land surface temperature images over heterogeneous landscape. *Adv. Space Res.* **2014**, *54*, 655–669. [[CrossRef](#)]
41. Jeganathan, C.; Hamm, N.A.S.; Mukherjee, S.; Atkinson, P.M.; Raju, P.L.N.; Dadhwal, V.K. Evaluating a thermal image sharpening model over a mixed agricultural landscape in India. *Int. J. Appl. Earth Obs.* **2011**, *13*, 178–191. [[CrossRef](#)]
42. Bechtel, B.; Zakšek, K.; Hoshyaripour, G. Downscaling land surface temperature in an urban area: A case study for Hamburg, Germany. *Remote Sens.* **2012**, *4*, 3184–3200. [[CrossRef](#)]
43. Duan, S.B.; Li, Z.L. Spatial downscaling of MODIS land surface temperatures using geographically weighted regression: Case study in northern China. *IEEE Trans. Geosci. Remote* **2016**, *54*, 6458–6469. [[CrossRef](#)]
44. Zhang, X.; Zhao, H.; Yang, J. Spatial downscaling of land surface temperature in combination with TVDI and elevation. *Int. J. Remote Sens.* **2019**, *40*, 1875–1886. [[CrossRef](#)]
45. Chen, Y.; Zhan, W.; Quan, J.; Zhou, J.; Zhu, X.; Sun, H. Disaggregation of Remotely sensed land surface temperature: A generalized paradigm. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5952–5965. [[CrossRef](#)]
46. Fasbender, D.; Tuia, D.; Bogaert, P.; Kanevski, M.F. Support-Based Implementation of Bayesian Data Fusion for Spatial Enhancement: Applications to ASTER Thermal Images. *IEEE Geosci. Remote Sens.* **2008**, *5*, 598–602. [[CrossRef](#)]
47. Keramitsoglou, I.; Kiranoudis, C.T.; Weng, Q. Downscaling geostationary land surface temperature imagery for urban analysis. *IEEE Geosci. Remote Sens.* **2013**, *10*, 1253–1257. [[CrossRef](#)]
48. Bai, Y.; Wong, M.S.; Shi, W.Z.; Wu, L.X.; Qin, K. Advancing of land surface temperature retrieval using extreme learning machine and spatio-temporal adaptive data fusion algorithm. *Remote Sens.* **2015**, *7*, 4424–4441. [[CrossRef](#)]
49. Yang, Y.; Cao, C.; Pan, X.; Li, X.; Zhu, X. Downscaling Land Surface Temperature in an Arid Area by Using Multiple Remote Sensing Indices with Random Forest Regression. *Remote Sens.* **2017**, *9*, 789. [[CrossRef](#)]
50. Neteler, M. Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data. *Remote Sens.* **2010**, *2*, 333–351. [[CrossRef](#)]

51. Isotta, F.A.; Frei, C.; Weilguni, V.; Perčec Tadić, M.; Lassegues, P.; Rudolf, B.; Munari, M. The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int. J. Climatol.* **2014**, *34*, 1657–1675. [[CrossRef](#)]
52. Wan, Z.; Hook, S.; Hulley, G. *MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006*; NASA EOSDIS LP DAAC; NASA: Washington, DC, USA, 2015.
53. Asam, S.; Callegari, M.; Matiu, M.; Fiore, G.; De Gregorio, L.; Jacob, A.; Menzel, A.; Zebisch, M.; Notarnicola, C. Relationship between Spatiotemporal Variations of Climate, Snow Cover and Plant Phenology over the Alps—An Earth Observation-Based Analysis. *Remote Sens.* **2018**, *10*, 1757. [[CrossRef](#)]
54. Wan, Z. *MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD)*; Institute for Computational Earth System Science: Santa Barbara, CA, USA, 1999; p. 75.
55. MONALISA. Available online: <http://www.monalisa-project.eu> (accessed on 20 December 2018).
56. Vermote, E.; Wolfe, R. *MOD09GQ MODIS/Terra Surface Reflectance Daily L2G Global 250 m SIN Grid V006*; NASA EOSDIS LP DAAC; NASA: Washington, DC, USA, 2015.
57. Vermote, E.; Wolfe, R. *MYD09GQ MODIS/Aqua Surface Reflectance Daily L2G Global 250 m SIN Grid V006*; NASA EOSDIS LP DAAC; NASA: Washington, DC, USA, 2015.
58. Vermote, E.; Wolfe, R. *MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1 km and 500 m SIN Grid V006*; NASA EOSDIS LP DAAC; NASA: Washington, DC, USA, 2015.
59. Vermote, E.; Wolfe, R. *MYD09GA MODIS/Aqua Surface Reflectance Daily L2G Global 1 km and 500 m SIN Grid V006*; NASA EOSDIS LP DAAC; NASA: Washington, DC, USA, 2015.
60. EARTHDATA Search. Available online: <https://search.earthdata.nasa.gov> (accessed on 10 November 2018).
61. Jiménez-Muñoz, J.C.; Sobrino, J.A.; Skoković, D.; Mattar, C.; Cristóbal, J. Land surface temperature retrieval methods from Landsat-8 thermal infrared sensor data. *IEEE Geosci. Remote Sens.* **2014**, *11*, 1840–1843. [[CrossRef](#)]
62. Jiménez-Muñoz, J.C.; Cristóbal, J.; Sobrino, J.A.; Sòria, G.; Ninyerola, M.; Pons, X. Revision of the single-channel algorithm for land surface temperature retrieval from Landsat thermal-infrared data. *IEEE Trans. Geosci. Remote* **2009**, *47*, 339–349. [[CrossRef](#)]
63. *NASA/METI/AIST/Japan Space Systems, and U.S./Japan ASTER Science Team ASTER Global Digital Elevation Model [Version 2, 49oN, 6oE]*; NASA EOSDIS Land Processes DAAC; NASA: Washington, DC, USA, 2009.
64. Hastie, T.; Tibshirani, R.; Friedman, J. Random forests. In *The Elements of Statistical Learning*; Series in Statistics; Springer: New York, NY, USA, 2001; pp. 587–602.
65. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
66. GDAL/OGR Contributors. *GDAL/OGR Geospatial Data Abstraction Software Library*; Open Source Geospatial Foundation: Chicago, IL, USA, 2019; Available online: <https://gdal.org> (accessed on 20 December 2018).
67. Hijmans, R.J. Raster: Geographic Data Analysis and Modeling. R Package Version 2.8-19. 2019. Available online: <https://CRAN.R-project.org/package=raster> (accessed on 20 December 2018).
68. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.
69. Bivand, R.; Keitt, T.; Rowlingson, B. Rgdal: Bindings for the ‘Geospatial’ Data Abstraction Library; R Package Version 0.8-16. Available online: <http://CRAN.R-project.org/package=rgdal> (accessed on 20 December 2018).
70. Barry, R.G. *Mountain Weather and Climate*; Psychology Press: Hove, UK, 1992.
71. Li, M.; Im, J.; Beier, C. Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington Wildlife Forest. *GISci. Remote Sens.* **2013**, *50*, 361–384. [[CrossRef](#)]
72. Chen, X.; Li, W.; Chen, J.; Rao, Y.; Yamaguchi, Y. A combination of TsHARP and thin plate spline interpolation for spatial sharpening of thermal imagery. *Remote Sens.* **2014**, *6*, 2845–2863. [[CrossRef](#)]
73. Essa, W.; van der Kwast, J.; Verbeiren, B.; Batelaan, O. Downscaling of thermal images over urban areas using the land surface temperature–impervious percentage relationship. *Int. J. Appl. Earth Obs.* **2013**, *23*, 95–108. [[CrossRef](#)]
74. Yu, X.; Guo, X.; Wu, Z. Land surface temperature retrieval from Landsat 8 TIRS—Comparison between radiative transfer equation-based method, split window algorithm and single channel method. *Remote Sens.* **2014**, *6*, 9829–9852. [[CrossRef](#)]
75. Sattari, F.; Hashim, M. A brief review of land surface temperature retrieval methods from thermal satellite sensors. *Middle-East J. Sci. Res.* **2014**, *22*, 757–768.

76. Dissanayake, D.; Morimoto, T.; Murayama, Y.; Ranagalage, M. Impact of Landscape Structure on the Variation of Land Surface Temperature in Sub-Saharan Region: A Case Study of Addis Ababa using Landsat Data (1986–2016). *Sustainability* **2019**, *11*, 2257. [[CrossRef](#)]
77. Sobrino, J.A.; Jimenez-Munoz, J.C.; Paolini, L. Land surface temperature retrieval from LANDSAT TM 5. *Remote Sens. Environ.* **2004**, *90*, 434–440. [[CrossRef](#)]
78. Sobrino, J.A.; El Kharraz, J.; Li, Z.L. Surface temperature and water vapour retrieval from MODIS data. *Int. J. Remote Sens.* **2003**, *24*, 5161–5182. [[CrossRef](#)]
79. Merlin, O.; Duchemin, B.; Hagolle, O.; Jacob, F.; Coudert, B.; Chehbouni, G.; Kerr, Y. Disaggregation of MODIS surface temperature over an agricultural area using a time series of Formosat-2 images. *Remote Sens. Environ.* **2010**, *114*, 2500–2512. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

---

**Land Surface Temperature Reconstruction under Long-term Cloudy-sky Conditions at 250 m Spatial Resolution: Case Study of Vinschgau/Venosta Valley in the European Alps (Paper 2)**

---

In the second part of the thesis, the main goal was to improve MODIS data with regard to cloud contamination effect translating into blank LST grids. In this context, a novel reconstruction method was proposed that integrates data-driven modelling with physical-based assumptions to represent long-term overcast conditions. The reconstruction was performed with two steps: year-round LST modelling from station-based variables under cloudy skies, and then applying the fitted models to invalid grids at subpixel corresponding to the downscaled output in Section 3.

---

This chapter was published in the *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* on 1 February 2022, and it is available online: <https://ieeexplore.ieee.org/document/9699062>. According to *Scopus* the journal received Q1 category in the field of Earth and Planetary Sciences in 2020.

---

# Land Surface Temperature Reconstruction Under Long-Term Cloudy-Sky Conditions at 250 m Spatial Resolution: Case Study of Vinschgau/Venosta Valley in the European Alps

Paulina Bartkowiak <sup>1</sup>, *Student Member, IEEE*, Mariapina Castelli, Alice Crespi, Georg Niedrist, Damiano Zanotelli, Roberto Colombo, and Claudia Notarnicola <sup>2</sup>, *Member, IEEE*

**Abstract**—In this article, we present a new concept for predicting satellite-derived land surface temperature (LST) under cloudy skies over vegetated areas in the Alps. Although many different reconstruction methods have been developed, they require rarely available inputs, or they restore missing pixels from clear-sky observations with low spatial resolution (1–5 km), which makes them unreliable in heterogeneous ecosystems. Given these limitations, we propose a station-based procedure to predict cloud-covered grids from 1-km Terra MODIS LST at 250 m spatial resolution. First, we explored correlations between ground-measured LST and air temperature in conjunction with other geo-biophysical variables under cloudy-sky conditions derived from ESRA clear-sky radiation model. Considering a high site dependency driven by different landcovers, in-situ data were aggregated into three groups (forest, permanent crops, grassland) and then, models were established. Next, the regressions were applied to 250-m gridded predictors to estimate cloud-covered LST pixels for six Terra MODIS LST images in 2014. While for permanent crops and forest group linear modelling was the most efficient, neural networks achieved the best performance for grasslands. The reconstructions showed reasonable LST distribution considering landscape heterogeneity of the region. The results were validated against timeseries of ground-measured LST in 2014. The models achieved reliable performance with an average  $R^2$  of 0.84 and root-mean-square error of

2.12 °C. Despite some limitations, mainly due to diversified character of cloudy-sky conditions and high heterogeneity of gridded predictors, the method can effectively reconstruct overcast MODIS data at subpixel level, which shows great potential for producing cloud-free LSTs in complex ecosystems.

**Index Terms**—Cloudy-sky conditions, land surface temperature, machine learning, reconstruction.

## I. INTRODUCTION

**E**ARTH'S skin temperature is a fundamental property regulating the exchange of water and energy between land and the atmosphere. Thus, it influences water and surface energy budget that is needed to estimate the impacts of climate change on water cycling, landcover, and to examine water anomalies in vegetation through evapotranspiration modeling [1]–[3]. Moreover, it allows monitoring vegetation conditions and studying climate change and impacts of extreme events on vegetation. As a result, land surface temperature (LST) is required as a baseline information for many environmental applications, such as management of water resources, climate change studies, sustainable agricultural production, drought predictions, and also land degradation monitoring [4]–[6].

Since the global network of meteorological stations is sparse, especially with regards to radiometers monitoring thermal infrared radiation, and LSTs vary over short distances, thermal remote sensing has shown large potential due to its spatial coverage and accessibility [7]–[9]. The rapid development of spaceborne thermal infrared (TIR) instruments followed by robust LST retrieval methods has allowed monitoring spatially and temporally continuous LSTs at different scales [10]. In particular, MODIS instrument has been frequently used due to its short repeat cycle (four times per day), global coverage and long-term image collection (since 2000) [11]. MODIS LST product has been applied in multiple research fields, including urban heat island assessment [12]–[15], drought detection [16]–[19], agricultural management [7], [20], [21], and energy and water balance modelling [22]–[25].

Although MODIS LST maps have been commonly used in many studies, TIR sensors are prone to overcast conditions. TIR instruments cannot acquire spatial information beneath clouds

Manuscript received September 16, 2021; revised December 1, 2021 and December 16, 2021; accepted January 22, 2022. Date of publication February 1, 2022; date of current version March 7, 2022. This work was supported in part by the Project “CYCLAMEN,” financed by Provincia Autonoma di Bolzano, Alto Adige/Südtirol, Ripartizione Diritto allo studio, Università e ricerca Scientifica, within the scope of Provincial Law December 13–14, 2006, under Project Nr. 5/34 and in part by The Mazia Station, funded by the Autonomous Province of Bolzano/Bozen within the framework of the LTSER site Matsch/Mazia. (*Corresponding author: Paulina Bartkowiak.*)

Paulina Bartkowiak is with the Institute for Earth Observation, Eurac Research, 39100 Bozen-Bolzano, Italy, and also with the Department of Earth and Environmental Sciences, University of Milano-Bicocca, I-20126 Milan, Italy (e-mail: paulina.bartkowiak@eurac.edu).

Mariapina Castelli, Alice Crespi, and Claudia Notarnicola are with the Institute for Earth Observation, Eurac Research, 39100 Bozen-Bolzano, Italy (e-mail: mariapina.castelli@eurac.edu; alice.crespi@eurac.edu; claudia.notarnicola@eurac.edu).

Georg Niedrist is with the Institute for Alpine Environment, Eurac Research, 39100 Bozen-Bolzano, Italy (e-mail: georg.niedrist@eurac.edu).

Damiano Zanotelli is with the Faculty of Science and Technology, Free University of Bozen-Bolzano, 39100 Bozen-Bolzano, Italy (e-mail: damiano.zanotelli@unibz.it).

Roberto Colombo is with the Department of Earth and Environmental Sciences, University of Milano-Bicocca, I-20126 Milan, Italy (e-mail: roberto.colombo@unimib.it).

Digital Object Identifier 10.1109/JSTARS.2022.3147356

that translates to invalid LST images with a strong impact on high-frequency cloud-contaminated areas [26]. According to Jin [27], cloudy skies “represent more than half of the actual day-to-day weather conditions” resulting in minimum 50% blank LST scenes from TIR data. If we focus on mountain regions, cloud cover is much more common and intense, which causes substantial gaps in LST images in space and time. Therefore, spatially, and temporally continuous LST information is of great importance and its complete creation is an urgent issue among the scientific community.

Methods for reconstructing land surface temperature at cloud-covered pixels have been well demonstrated in many research studies. In general, they can be aggregated into empirical and physical-based categories. The methods of the physical-driven group rely on the assumption that cloud-covered pixels differ from LST values under clear-sky conditions. Prediction of these temperatures requires knowledge about physical relationship between targets and their adjacent cloud-free pixels. Jin [27] developed a “neighboring-pixel” (NP) method that uses surface energy balance (SEB) to recover invalid values from their spatially (100–300 km) or temporally ( $\leq$  two days) neighboring cloud-free pixels. Due to temporal limitations of the NP method, Lu *et al.* [28] proposed an enhanced solution based on geostationary satellite data from the MSG SEVIRI radiometer with shorter repeat cycle with respect to polar-orbiting satellites. Later, Yu *et al.* [29] adapted the NP by exploiting the spatio-temporal domain offered by MODIS LST product. Because of the physical complexity of this method and scarcity of the station-based inputs, such as wind speed and latent heat flux, some new approaches were proposed. Using MSG satellite data, Zhang *et al.* [30] predicted cloud-covered LST from a simplified heat transfer formula with the reduced number of station-based parameters. Next, Zeng *et al.* [31] developed a multisource approach that estimates invalid pixels from cloud-free MODIS LST in reference to normalized difference vegetation index (NDVI) and then applies a SEB-based factor to obtain values under cloudy skies. After that, Yang *et al.* [32] simplified the NP and showed that applying only solar radiation as the auxiliary feature can well estimate missing MODIS LST for areas significantly covered by clouds. Meanwhile, Martins *et al.* [33] developed an operational “all-weather land surface temperature” product based on the clear-sky 3-km MSG SEVIRI scenes and LST data under cloudy conditions derived from the surface energy balance model through the LSA-SAF ET v2 algorithm. In addition to the SEB-based strategies, Fu *et al.* [34] proposed a novel reconstruction method for urban areas by exploiting relationship between LSTs from the physical-based WRF/UCM system and clear-sky MODIS data using random forest algorithm. However, the performance of these approaches depends on the complexity of the study area or the availability of clear-sky LST, which makes it difficult to implement in ecosystems with extremely high variability with regard to weather conditions, altitude, soil, and landcover. Apart from the physical-based methods, cloud-covered LSTs can be recovered by data fusion approach by combining data from different resources. Many studies showed that TIR-based land surface temperature can be predicted by its integration with temporally adjacent

images from passive microwave instruments that are capable of penetrating clouds [35]–[39]. Furthermore, Long *et al.* [40] combined clear-sky MODIS data with 7-km LST dataset from China Land Data Assimilation System using ESTARFM algorithm and obtained all-weather 1-km MODIS-like scenes with root-mean-square error (RMSE) yielding from 2.77 K to 3.96 K. Similarly, the performances of these methods have limitations in terms of low spatial resolution of the cloud-free inputs.

The second category for LST reconstruction belongs to empirical methods and treats pixels obscured by clouds as those under clear skies. A widely used technique of this group is geostatistical interpolation that exploits similarities from neighboring clear-sky pixels in spatial, temporal, and spatiotemporal domain [41]–[45]. Furthermore, Crosson *et al.* [46] implemented multisensor approach to predict Aqua MODIS LST from clear-sky Terra MODIS LST and increased daytime and nighttime data availability of Aqua product by 24% and 30%, respectively. In addition, Wang *et al.* [47] approximated cloudy LSTs for Terra MODIS and Landsat-8 TIRS by spatiotemporal fusion of clear-sky multitemporal MODIS LST composites (MOD11A2) and synthetic surface temperatures with “solar-cloud-satellite geometry” derived from MODIS cloud and geolocation products (MOD06, MOD03) and Landsat-8 data. The empirical approaches, similarly to physical-based methods, rely on the availability of time-coincident cloud-free LSTs and input accuracies, which decreases their spatial prediction performance. To overcome this limitation, Ke *et al.* [48] exploited the relationship between MODIS LST and other environmental variables using regression kriging technique. As shown by Fan *et al.* [49], incorporation of NDVI, soil moisture and landcover information in conjunction with artificial neural networks and regression tree modeling allowed predicting missing LST with RMSEs ranging from 1.32 K to 1.66 K. Given the high capabilities of artificial intelligence (AI), new machine learning-based approaches have been recently developed. For example, Wu *et al.* [50] combined a convolutional neural network (CNN) with spatiotemporal information offered by geostationary instruments and obtained a prediction error of approximately 1 K for images with 70% cloud-contaminated pixels. Meanwhile, Zhao and Duan [51] proposed a random forest approach to predict MODIS LST by combining multisource remote sensing predictors with solar radiation to represent a cloud cover impact on missing LST pixels. However, the accuracies of these approaches are still highly influenced by the availability of clear pixels and low spatial resolution of satellite data that limits their application to a relatively homogenous terrain.

To address these limitations, in this article, we concentrate on the development of a new method to reconstruct cloud-covered 1-km MODIS LST at 250 m spatial resolution over vegetated areas in the European Alps. Considering high spatio-temporal dynamics of surface temperatures, we propose a new approach that integrates data-driven modeling with physical-based assumptions to detect long-term cloudy-sky conditions at sub-pixel level with respect to the 1-km original MODIS LST. The proposed method exploits relationship between ground-based LST and commonly accessible input parameters under

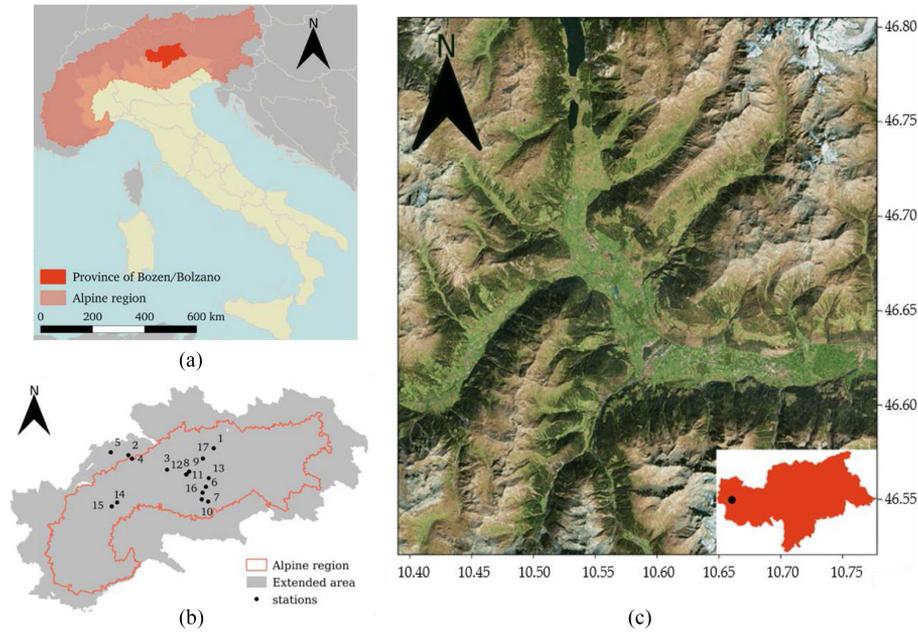


Fig. 1. Overview of the study case. (a) Location of Province of Bozen/Bolzano. (b) Positions of the stations utilized in the ground-based LST modeling. Some stations were situated outside the official border of the Alps (in red).<sup>1</sup> Since they were located in areas with similar climatic conditions, they were included in the modeling. (c) Satellite image of the experimental area for the LST reconstruction in Vinschgau/Venosta Valley obtained from the Express Kosmosnimki service.<sup>2</sup>

cloudy-sky conditions, like air temperature (TA), downwelling solar radiation ( $SW_{in}$ ), surface albedo ( $\alpha$ ), and leaf area index (LAI), for spatial reconstructions of cloud-covered MODIS LST data. To our best knowledge, gap-filling from station-based models in conjunction with a physical-based approach to obtain *in situ* observations under cloudy skies has never been performed before. We analyze the prediction performance of the developed method based on different machine learning algorithms to reconstruct missing MODIS LST values. Through the application of different modelling scenarios, it is feasible to investigate in which way algorithms and predictors can explain LST variability within diversified mountain ecosystems. Due to limitations of MODIS LST pixel size, our concept predicts invalid grids at 250 m spatial resolution in order to minimize the impact of topography and landscape heterogeneity of the study area. Additionally, the reconstructed maps implemented in this article will be combined with 250-m downscaled Terra MODIS LST [52] for further energy balance modeling of evapotranspiration with relevant implications on water assessment in the Alps.

## II. DATA AND METHODOLOGY

### A. Study Site

Our area of interest is an approximately 520-km<sup>2</sup> region of Vinschgau/Venosta Valley in the northwestern part of the Province of Bozen/Bolzano, located in the Eastern Italian Alps (see Fig. 1(a)–(c)). The area is predominantly mountainous with elevation ranging from about 700 m to 3740 m a.s.l. Alpine orography has a significant impact on the structure of the landscape with highly patched landcovers over the region.

The area is mainly dominated by forests, grasslands, and apple orchards. Because of the insular location of the valley and the sheltering effect of the neighboring mountain ridges, the climate is warmer and drier than in other parts of the Alps, which translates to higher evaporation [53]. This results in unfavorable effects for vegetation, especially for grasslands and agricultural crops with higher demand for water. Considering the particular climatic conditions and agriculture-oriented land management, this region would particularly benefit of accurate reconstructed LST maps for monitoring vegetation conditions and water availability.

Land surface temperature reconstruction was performed based on year-round station records from the Fluxnet network and other stations located over different vegetated ecosystems in the Alps [see Fig. 1(a) and (b)] [54]. Eddy covariance data outside the Fluxnet network were processed in the framework of the project CYCLAMEN (<https://www.eurac.edu/>).

### B. Input Data

1) *Ground Measurements*: Since this article is intended for vegetation analyses over the Alps, *in situ* measurements were collected at 17 meteorological sites distributed over different ecosystems, during the phenological cycle, between April and October (see Table I). The station loggers had unique time-series measurements recorded between 2002 and 2019 with 15- and 30-min intervals. As Table I shows, 10 sites are located above 1000 m and covered by grassland or forest, while the remaining seven stations lie in agricultural areas at altitudes below 1000 m.

In this article, we exploited ground-based upwelling and downwelling longwave radiation ( $LW_{in}$ ,  $LW_{out}$ ) to retrieve land surface temperature for each station record (see Table II) [55], [56]. In order to keep time consistency with MODIS LST data,

<sup>1</sup>[Online]. Available: <http://www.eurac.edu/>

<sup>2</sup>[Online]. Available: <http://kosmosnimki.ru/>

TABLE I  
METEOROLOGICAL STATIONS UTILIZED IN THE LST MODELING

Model	LULC	Group	Station	Fluxnet site	Altitude [m a.s.l.]	Time span	
(1)	Grassland	Grassland	Rotholz (1)	AT-Rtz	523	2008	2012
			Chamau (2)	CH-Cha	393	2006	2012
			Früebüel (4)	CH-Fru	982	2006	2012
			Neustift (17)	AT-Neu	970	2002	2012
			Oensingen (5)	CH-Oe1	452	2003	2008
			Monte Bondone (10)	IT-MBo	1553	2003	2013
			Mazia 1 (11)	-	1450	2014	2017
			Mazia 2 (12)	-	1550	2014	2017
			Mazia 3 (8)	-	1909	2019	2019
			Mazia 4 (9)	-	2688	2016	2019
			Torgnon 1 (14)	(IT-Tor)	2160	2008	2017
(2)	Apple Orchard	Permanent	Caldaro (6)	-	240	2014	2015
	Vineyard	crops	Valle dell'Adige (16)	IT-VdA	206	2008	2010
(3)	Evergreen needleleaf forest	Forest	Lavarone (7)	IT-Lav	1349	2003	2014
	Deciduous needleleaf forest		Davos (3)	CH-Dav	1639	2006	2011
			Renon (13)	IT-Ren	1730	2004	2013
			Torgnon 2 (15)	IT-TrF	2091	2010	2016

\*Numbers in parentheses in the forth column refer to station locations presented in Fig. 1(b).

TABLE II  
STATION-BASED PARAMETERS FROM YEAR-ROUND OBSERVATIONS AND THEIR CORRESPONDING VARIABLES USED FOR LST MODELING

Source data	Variable	Short description
Outgoing and incoming longwave radiation	Land Surface Temperature [ $^{\circ}\text{C}$ ] $\left(\frac{LW_{\text{out}} - (1 - \varepsilon)LW_{\text{in}}}{\sigma\varepsilon}\right)^{1/4}$	LST formula based on Stefan–Boltzmann law [55], where $\sigma$ is the Stefan–Boltzmann constant, $\varepsilon$ is the surface emissivity from M*D21A1D [57], and $LW_{\text{out}}/LW_{\text{in}}$ is the outgoing/incoming longwave radiation retrieved from ground observations
Air temperature	Mean air temperature ( $TA_{\text{mean}}$ ) [ $^{\circ}\text{C}$ ] Maximum air temperature ( $TA_{\text{max}}$ ) [ $^{\circ}\text{C}$ ]	daily mean and maximum air temperature derived from in-situ data
Solar radiation	Daily incoming shortwave radiation ( $SW_{\text{in}}$ ) [ $\text{MJ m}^{-2}\text{day}^{-1}$ ]	daily cumulative $SW_{\text{in}}$ retrieved from ground observations
Surface albedo	Vegetation parameter [-] $\frac{\log(z_0)}{\alpha}$	aerodynamic roughness length ( $z_0$ ) and albedo ( $\alpha$ ) to describe canopy structure

longwave radiation was extracted that corresponded to the local MODIS observing time. To compute surface temperature, we calculated broadband surface emissivity from daily MODIS land surface emissivity (LSE) product (M\*D21A1D Collection 6) [57], as described by Wang *et al.* [58]. M\*D21A1D was selected rather than daily M\*D11A1, as it was created using physical-based temperature emissivity separation algorithm instead of simplified landcover-derived emissivity retrieval [59]. Even though 90-m ASTER LSE product offers finer spatial resolution, its small spatial extent, and 16-day repeat cycle in conjunction with frequent cloudiness over the study area hampered its practical applicability.

Due to strong correlation of LST with air temperature (TA) and availability of daily TA maps (see Section II-B2), observations of daily mean and maximum air temperature ( $TA_{\text{mean}}$ ,  $TA_{\text{max}}$ ) were computed from the *in situ* measurements and considered as the baseline input for modeling [60], [61]. In mountain regions with heterogenous landscape, such as the Alps, LST is influenced by complex interactions between land and atmosphere, solar energy, topography, soil moisture, and landcover. In this context, we exploited auxiliary ground-based biophysical parameters that may explain spatial variation in LST

over the study area (see Table II). In addition to  $TA_{\text{mean}}$  and  $TA_{\text{max}}$ , we incorporated diurnal incoming shortwave radiation ( $SW_{\text{in}}$ ) that regulates ground heating process with an important impact on land-atmosphere energetics [53]. Owing to the energy fluxes between atmosphere and various vegetated landcovers, we combined surface albedo ( $\alpha$ ) with aerodynamic surface roughness ( $z_0$ , as in Table II) assigned for each landcover group separately as a complementary biophysical predictor for LST modeling [61]–[63].

All collected measurements were averaged to hourly resolution corresponding to MODIS acquisition time and they were inspected for the presence of outliers, including detection of inconsistent minimum and maximum station records and unusual temporal variations. Additionally, we performed a specific quality control procedure for solar radiation data by applying physical thresholds and step tests, specially adapted to Alpine conditions [64]. After these checks, the ground-based predictors for the LST modelling were computed (see Table II).

2) *Gridded Data*: Gridded variables were generated for reconstructing cloud-contaminated 1-km MODIS LST pixels at subpixel (250 m) spatial resolution from the fitted models (see Table III).

TABLE III  
GRIDDED DATASETS USED IN THIS ARTICLE

Source data	Variable	Pixel size	Short description
<i>Prediction of missing MODIS LST pixels</i>			
In-situ weather station records	Daily grids of: Mean air temperature ( $TA_{\text{mean}}$ ) [ $^{\circ}\text{C}$ ] Maximum air temperature ( $TA_{\text{max}}$ ) [ $^{\circ}\text{C}$ ]	250 m	Maps derived from daily station observations through a spatial interpolation scheme
MSG/SEVIRI DSSF	Downscaled daily incoming shortwave radiation maps [ $\text{MJ m}^{-2}\text{day}^{-1}$ ]	250 m	Downscaled MSG downwelling surface shortwave flux (DSSF) through RK interpolation
MCD43A3	Vegetation parameter [-] $\frac{\log(z_0)}{\alpha}$	500 m	Relationship between aerodynamic roughness length ( $z_0$ ) and albedo ( $\alpha$ ) from MCD43A3 composites
MCD15A3H <sup>1</sup>	$f_1(z_0, \text{LAI}): z_0 \cdot (h)^{-1}_{\text{CM88}}$ [-] $f_2(z_0, \text{LAI}): z_0 \cdot (h)^{-1}_{\text{SD00}}$ [-]	500 m	Vegetation structure parametrization ( $f_1, f_2$ ) with roughness length ( $z_0$ ) and leaf area index (LAI) obtained from MCD15A3H. Parameter $h$ refers to vegetation height, and CM88 and SD00 subscripts indicate adopted formulas of Choudhury & Monteith [68] and Schaudt & Dickinson [69], respectively
MOD21A1D <sup>2</sup> MYD21A1D <sup>2</sup>	Emissivity bands: 29,31,32 [-]	1000 m	Daily emissivity maps from M*21A1D composites for station-based LST retrieval (see Table II for more details)
<i>Determination of spatial applicability of the reconstructed LST maps</i>			
MOD11A1	Cloud cover [-]	1000 m	Cloud cover from daily Terra MODIS LST product to identify cloudy pixels
Snow Cover Area (SCA)	Snow cover [-]	250 m	Daily snow maps based on Terra and Aqua MODIS Reflectance for masking areas covered by snow
LISS 2013 - Land Information System South Tyrol	Vegetation mask [-]	-	Vegetation mask for identification of homogeneous landcovers within 250-m pixels (80% threshold of homogeneity)
MOD13Q1	Enhanced Vegetation Index (EVI) [-]	250 m	EVI grids from Terra MODIS Vegetation Indices product to delineate pixel-wise areas of applicability for the models

<sup>1</sup>Data used for both the station-based modelling and the reconstruction of missing MODIS LST pixels; <sup>2</sup>Data used for station-based LST retrieval.

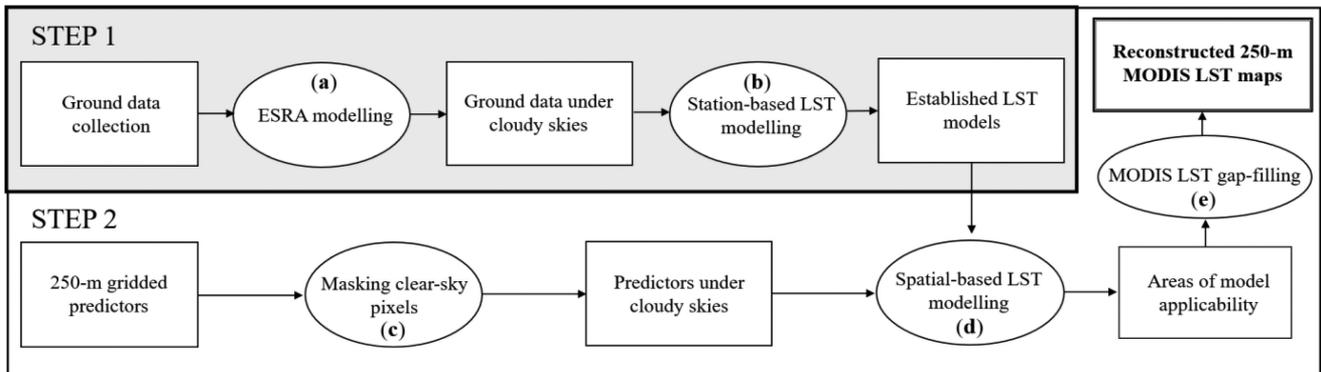


Fig. 2. Workflow for LST gap-filling procedure under cloudy-sky conditions for MODIS data. While rectangular boxes refer to data used, ellipse shapes represent processing steps.

The 250-m air temperature maps for Trentino-South Tyrol region were obtained by interpolating the daily meteorological observations provided by the regional weather station networks of Meteotrentino for the Province of Trento and the Hydrological Department of the Province of Bozen/Bolzano. All observations were checked for quality and homogeneity and harmonized in a dense archive of more than 200 station series covering the region. The daily grids of  $TA_{\text{mean}}$  and  $TA_{\text{max}}$  were then derived for the period 1980–2018 by applying an interpolation method combining the reference mean climate, i.e., the climatology, and the daily temperature anomalies. Due to the mountainous terrain of the area, the interpolation applied a regression-based procedure

modelling the local relationships between temperature spatial gradients and orographic features, including elevation and slope characteristics, which were derived from the Digital Elevation Model (DEM) Copernicus EU-DEM v1.1<sup>3</sup> and then aggregated to the target 250 m resolution [65], [66]. Cross-validated average RMSE values in a spatial cross-validation approach for  $TA_{\text{mean}}$  and  $TA_{\text{max}}$  outputs were around 1.9  $^{\circ}\text{C}$  and 2.4  $^{\circ}\text{C}$ , respectively.

Daily solar radiation grids at 250 m resolution were derived by applying a geostatistical downscaling to the 2004–2018 daily

<sup>3</sup>[Online]. Available: <https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1>

downwelling surface shortwave flux (DSSF) product derived from MSG/SEVIRI, available on the LSA-SAF system.<sup>4</sup> In particular, the sharpening of daily solar radiation ( $SW_{in}$ ) was performed by means of a regression kriging (RK) in conjunction with the main topographic drivers, e.g., elevation, slope steepness and its orientation. In this scheme, the linear regression model was first estimated, the resulting residuals were interpolated onto the target grid through Ordinary Kriging with automatic fit of the variogram, and the final daily fields were obtained as sum of the spatialized residuals and regression predictions at each target grid cell. The average RMSE (bias) error was  $2.64 \text{ MJ m}^{-2}\text{day}^{-1}$  ( $0.11 \text{ MJ m}^{-2}\text{day}^{-1}$ ) translating into the mean absolute percentage difference of 0.15 when compared to the ground-derived datasets.

In this article, we exploited remotely sensed surface albedo from a 16-day Terra/Aqua MODIS Albedo product (MCD43A3 Version 6) with 500-m pixel size [67]. The time-coincident MCD43A3 with good quality assurance (QA) flags were used for gap-filling of invalid MODIS LST. Additionally, an impact of landcover on LST was examined by parametrization of vegetation properties using aerodynamic roughness ( $z_0$ ) and LAI (see Table III for details) [68], [69]. Spatially continuous LAI maps were obtained from 4-day Terra/Aqua MODIS LAI composites (MCD15A3H Version 6) at 500 m spatial resolution [70]. To increase spatial availability of LAI pixels, the 4-day MODIS LAI was upscaled to 14-day composite considering the highest quality of the QA flags as well as the closest acquisition time between station records and MODIS time overpass. While gridded albedo was only applied to the fitted models, LAI was used for both LST modeling and reconstruction of missing MODIS LSTs under cloudy-sky conditions. To keep spatial consistency with the reconstruction outputs, both MCD43A3 and MCD15A3H were disaggregated to 250-m pixel size using nearest neighbor resampling approach.

In order to identify cloud-covered areas for the LST reconstruction, we used invalid values that were assigned to the QA flags in the 1-km MODIS product. Since the focus of the article was on the LST reconstruction for vegetation analyses, daytime Terra MODIS LST (MOD11A1 Version 6) was utilized [71]. To examine the performance of the proposed method, we reconstructed missing pixels for six MOD11A1 images acquired in different seasons in 2014 (May 2nd, June 29th, July 8th, September 19th, October 11th, and 26th).

Areas of applicability for the reconstructed LST maps were determined by exploiting relationship between 250-m pixels within model vegetation groups (see Table I) and EVI from Terra MODIS Vegetation Indices (MOD13Q1 Version 6) granules [72]. In this article, we used the detailed land use landcover (LULC) data with minimum mapping area equal to  $1600 \text{ m}^2$ , as shown in Appendix [73]. The pixelwise selection procedure will be explained in Section II-CD. In addition, we applied daily 250-m MODIS snow cover mask to gridded predictors in order to exclude nonvegetated pixels from LST reconstructions [74].

### C. Methodology

1) *Model Concept*: The Alpine region benefits from dense network of ground stations, which translates to time-series of climate data and generation of spatially continuous meteorological grids.<sup>5</sup> To take the full advantage of these 250-m inputs, we propose station-based modelling to reconstruct daytime Terra MODIS LST (MOD11A1) under cloudy-sky conditions at 250 m pixel size, that is a sufficient scale for capturing spatial details at regional level [75]–[78]. Subpixel mapping is considered as a reasonable solution in mountain regions characterized by complex structure of the landscape, where 1-km MODIS LST images cannot fully represent spatial heterogeneities of the terrain. The land surface temperature reconstruction was performed with two steps as follows:

- 1) Year-round LST modelling from station-based environmental variables under long-term cloudy skies; and
- 2) Applying the fitted models to cloud-contaminated MODIS LST pixels.

The conceptual scheme of the proposed method is summarized in Fig. 2.

The diurnal cycle of LST is highly affected by the surface energy balance and the surface thermal inertia that depend on ground characteristics, such as landcover, soil type, and its moisture [79]. Since clouds affect the energy budget of the ground, the LST modeling was defined under the assumption that LSTs beneath clouds are different than those under clear skies. In this context, we used the European Solar Radiation Atlas (ESRA) to define sky conditions for each hourly observation from station records [see Fig. 2(a)] [64], [80]–[82]. Determination of cloudy- and clear-sky observations was based on the hourly incident solar radiation ( $SW_{in}$ ) from the meteorological stations and its corresponding maximum theoretical value computed from the ESRA model [60]. A ground-derived  $SW_{in}$  record was considered as a clear-sky observation provided that it exceeded 90% of maximum theoretical ESRA-based estimation, while the measured solar radiation with less than 50% of the maximum theoretical value was assigned to the cloudy-sky group. Additionally, the measurements registered two hours before and two hours after a respective observation had to meet the cloudy-sky criteria in order to extract long-term overcast observations.

Given a strong relationship between instantaneous air TA and ground-derived LST, daily mean air temperature ( $TA_{mean}$ ) was considered as a baseline predictor in the LST modeling [7], [8], [60], [83], [84]. Fig. 3 presents the daytime and yearly cycle of the mean difference between hourly LST and daily  $TA_{mean}$  considering cloud-free and overcast conditions for all available observations (see Table I).

In general, the hourly [see Fig. 3(a)] and daily [see Fig. 3(b)] differences between LST and  $TA_{mean}$  under clear skies were noticeably bigger when compared to cloudy-sky variations. Overcast conditions alter the energy budget of land, resulting in smaller variations between LST and  $TA_{mean}$  [see Fig. 3(a) and (b)]. As shown in Fig. 3(a), cloud-free  $LST - TA_{mean}$  values varied with hours, with the biggest deviation reaching  $9.2 \text{ }^\circ\text{C}$  close to

<sup>4</sup>[Online]. Available: <https://landsaf.ipma.pt>

<sup>5</sup>[Online]. Available: <https://doi.pangaea.de>

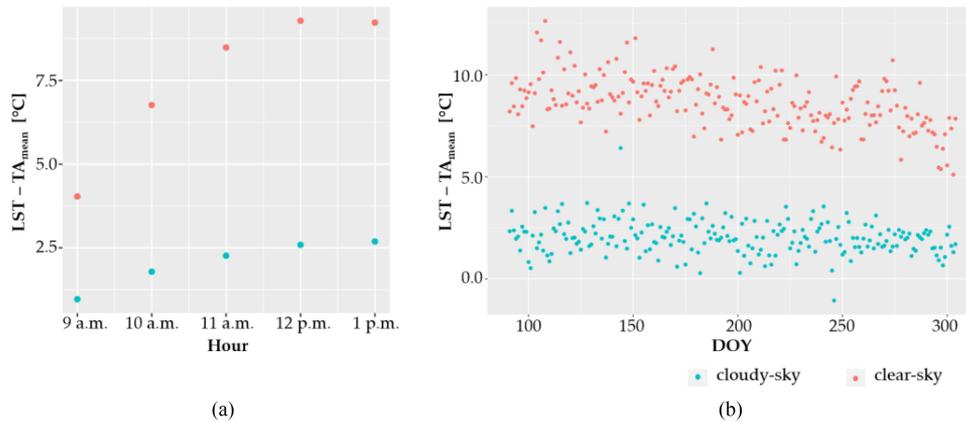


Fig. 3. Mean differences between hourly LST and daily  $TA_{\text{mean}}$  for all stations used in the study grouped by: (a) hour corresponding to approximate daytime Terra MODIS acquisition time and (b) day of year (DOY) between April and October.

TABLE IV  
MEAN DIFFERENCE BETWEEN HOURLY LST AND DAILY  $TA_{\text{MEAN}}$  UNDER CLEAR AND LONG-TERM CLOUDY SKIES FOR EACH STATION

Station	LST- $TA_{\text{mean}}$ [°C]	
	Cloudy-sky	Clear-sky
Rotholz	2.60	7.66
Chamau	1.29	6.45
Früebüel	1.06	6.69
Neustift	2.76	8.51
Oensingen	3.70	10.16
Monte Bondone	4.07	12.86
Mazia 1	2.76	9.63
Mazia 2	3.84	13.31
Mazia 3	4.08	13.04
Mazia 4	3.74	15.23
Torgnon 1	4.17	13.18
Caldaro	1.43	4.13
Valle dell'Adige	1.53	7.83
Lavarone	0.07	2.82
Davos	-1.04	2.07
Renon	0.97	5.01
Torgnon 2	1.20	5.60

solar noon. For overcast conditions, however, the discrepancies between LST and  $TA_{\text{mean}}$  were smaller ( $\leq 2.7$  °C), and they were more consistent within a day. This resulted in hourly-based standard deviation equal to 0.71 °C that was approximately 1.5 °C smaller than under clear-sky conditions [see Fig. 3(a)]. Similarly, as for the hourly observations, the day-of-year (DOY) based differences under cloudy-sky conditions were fairly stable throughout the year when compared to clear-sky observations [see Fig. 3(b)].

An additional analysis performed for the individual stations showed that LST tends to be greater than  $TA_{\text{mean}}$  with increased values for clear-sky observations (see Table IV). While there was one station (Davos) with  $TA_{\text{mean}}$  greater than LST under cloudy skies, for the rest of the sites the mean differences were positive and ranged from 0.07 °C (Lavarone) to 4.17 °C in Torgnon 1. The mean variations between LST and  $TA_{\text{mean}}$  for cloud-free conditions were larger than for overcast conditions for all stations yielding values from 2.07 °C in Davos to nearly 15.2 °C in grass-covered Mazia 4 (see Table IV). The paired t-test verified that the clear-sky LST- $TA_{\text{mean}}$  observations for the individual stations were significantly greater ( $p < 0.001$ ) than the corresponding differences under cloudy-sky conditions.

Based on the abovementioned analyses (see Fig. 3, Table IV), daily mean air temperature was examined as a principal variable to explain LST deviations as described in further sections of the article.

2) *Algorithms*: LST observations under long-term cloudy-sky conditions were modeled by three different regression algorithms [see Fig. 2(b)]. A multivariate linear regression (MLM) model was used as a baseline estimator to explain the relationship between ground-based LST and independent variables. The MLM is considered as an intuitive tool with lower complexity for data interpretation, which makes it commonly used in modeling and preprocessing tasks [7], [9], [49]. As an alternative to the standard MLM, we exploited artificial neural network (ANN) and random forest (RF) [85]–[87]. Selection of these approaches was dictated by two main reasons. First, they belong to two algorithm families with different assumptions, which make them useful for comparison analysis [88]. Second, the ANN and RF have ability to account for both linear problems and nonlinearities between predictors and dependent variable. The algorithms have been successfully used in many studies, including spatial mapping and remote sensing enhancement tasks [9], [52], [89]. More detailed information about the algorithms is provided by Kuhn *et al.* [86].

3) *Model Calibration and Evaluation*: Spatio-temporal predictive tasks entail an appropriate data management and modelling strategies to retrieve reliable estimations for new locations. One of the most common problems regarding modelling from geographical features is the existence of dependence between predictors and time-neighboring observations resulting in model overfit that is revealed by well-fitted regression for training data and poor predictions beyond known points [9], [90].

Prior model training process, dependencies between predictors were investigated by exploiting variance inflation factor (VIF) to solve multicollinearity issue among features [91]. The VIF is defined as follows:

$$VIF_k = \frac{1}{(1 - R_k^2)} \quad (1)$$

TABLE V  
RESULTS FROM THE HYPERPARAMETER OPTIMIZATION FOR THE MODELS WITH THEIR FINAL VALUES

Algorithm	Short description	Hyperparameter search grid		Final value
eNet	alpha: elastic mixing parameter	alpha: 0.0-1.0	alpha <sup>forest</sup> : 1.0	lambda <sup>forest</sup> : 0.05
	lambda: regularization parameter	lambda: 0.001-0.6	alpha <sup>agri</sup> : 0.8, 1.0	lambda <sup>agri</sup> : 0.0, 0.05, 0.7
			alpha <sup>grass</sup> : 1.0	lambda <sup>grass</sup> : 0.0
ANN	size: units per hidden layer	size: 1-10	size <sup>forest</sup> : 1, 9	decay <sup>forest</sup> : 0.0, 0.4
	decay: weight penalty parameter	decay: 0.0-0.5	size <sup>agri</sup> : 1, 2, 3	decay <sup>agri</sup> : 0.0
			size <sup>grass</sup> : 1, 2, 3	decay <sup>grass</sup> : 0.0, 0.1, 0.2, 0.5
RF	mtry: number of prediction variables			mtry <sup>forest</sup> : 2, 3
	randomly sampled as candidates at each split	mtry: 1-4		mtry <sup>agri</sup> : 2, 3, 4 mtry <sup>grass</sup> : 2

ANN = feed forward network with one hidden layer, RF = random forest based on 1000 trees, eNet = elastic net model used as an alternative to the MLM. The abbreviation *agri* refers to permanent crops.

where the  $R_k^2$  indicates the unadjusted coefficient of determination calculated by regressing the  $k$ th independent variable on the remaining predictors.

For regression tasks in spatio-temporal domain the overfitting problem is not only related to the redundant features but also to an incorrect model calibration or inappropriate predictor selection [9]. In this context, a model regularization needs to include these two aspects. Spatial cross validation (SCV) with a “leave-one-station-out” approach that excludes all observations of one station iteratively using station-fold splits (as  $k$ -fold subsets), and then compares all fold-driven simulations, could allow evaluating prediction power beyond training data, while simultaneously reducing autocorrelation between observations. On the other hand, forward feature selection (FFS) estimates the significance of predictors by inspecting possible model combinations and selects only those features, which improve model performance in terms of accuracy metrics [92]. In this article, we combined the FFS and SCV to determine the final features considering the highest model prediction performance in terms of its accuracy [9], [93], [94]. Average RMSE and  $R^2$  together with their corresponding standard deviations ( $SD_{RMSE}$  and  $SD_{R^2}$ ) were calculated as evaluation metrics to choose the best modeling approach.

As mentioned previously, the regularization was performed by applying forward feature selection to all predictor variables in conjunction with the leave-one-station-out SCV [9]. The tuning was additionally conducted to obtain optimal algorithm hyperparameters for the models [95]. In the case of the ANN and RF, parameters for each model were evaluated within the iterative SCV process using tuning search grids (see Table V) [96], [97]. In addition, we applied elastic net (eNet) model to test regularization impact on the MLM. We performed multiple modeling for each vegetation group (see Table I) considering different combinations of hyperparameters in the search grids. The optimal values were chosen according to SCV accuracy scores.

Since the main aim of the calibrated models was to apply them to unseen gridded data, the test subsets were created during the spatial cross-validation approach to estimate model robustness

beyond training points. The spatial predictive performance of all models (MLM, eNet, ANN, RF) was compared by conducting multiple paired t-tests.

In this article, for model creation we used Caret and CAST packages available in the R statistical software that contains the MLM, eNet, ANN, and RF model implementations from other R libraries [96]–[99].

4) *Extension to Gridded Data:* After the definition of the models, based on ground station data, they were applied to gridded predictors to reconstruct cloudy LSTs from MODIS LST at 250 m spatial resolution [see Fig. 2(c)–(e)]. Since year-round modeling was split into separate LULC-driven estimators, we defined potential areas for each model by considering similarities in predictor variable space.

In order to estimate model transferability, we made a quantitative comparison between data used for model training and gridded features representing new locations for the LST reconstruction. Similarity measure between a target pixel (i.e., a new point to predict) and training data was assessed by minimum Euclidean distance in the multidimensional predictor space with respect to an average distance between points used during training process [90]

$$d_{\text{target}}^i = \min [d(i, j)] \quad (2)$$

where  $\min [\cdot]$  represents minimum function,  $d(i, j)$  indicates distance between a new point  $i$  and  $j$ th observation from training data, and  $d_{\text{target}}^i$  is a minimum distance between a new location  $i$ th and a point used in the modeling [96].

Based on this, a standardized distance ( $d_{\text{target}}^\sigma$ ) for each new location was derived, as shown in

$$d_{\text{target}}^\sigma = \frac{d_{\text{target}}^i}{d_{\text{mean}}} \quad (3)$$

where  $d_{\text{mean}}$  indicates an average Euclidean distance from a target  $i$  to all training points.

Areas of applicability (AOA) for the considered models were defined based on the concept developed by Meyer and Pebesma [90]. We selected this approach because it deals with a problem

TABLE VI  
SUMMARY OF THE MODEL EVALUATION MEASURES FROM MULTIVARIATE LINEAR REGRESSION, ELASTIC NET, NEURAL NETWORKS, AND RANDOM FOREST

Group	Algorithm	RMSE [°C]	SD <sub>RMSE</sub> [°C]	R <sup>2</sup>	SD <sub>R<sup>2</sup></sub>	Final predictors
Permanent crops	<u>MLM</u>	<u>2.67</u>	<u>1.62</u>	<u>0.74</u>	<u>0.15</u>	TA <sub>mean</sub> , TA <sub>max</sub>
	eNet	2.61	1.74	0.76	0.16	TA <sub>max</sub> - TA <sub>mean</sub>
	ANN	2.74	1.51	0.72	0.13	SW <sub>in</sub>
	RF	3.05	1.49	0.69	0.11	log(z <sub>0</sub> )α <sup>-1</sup>
Forest	<u>MLM</u>	<u>1.84</u>	<u>0.42</u>	<u>0.88</u>	<u>0.04</u>	TA <sub>mean</sub> , TA <sub>max</sub>
	eNet	1.84	0.45	0.88	0.04	SW <sub>in</sub>
	ANN	1.84	0.42	0.88	0.04	TA <sub>max</sub> - TA <sub>mean</sub>
	RF	1.91	0.41	0.86	0.04	
Grassland	MLM	2.07	0.32	0.77	0.07	TA <sub>mean</sub> , TA <sub>max</sub>
	eNet	2.08	0.32	0.77	0.07	TA <sub>max</sub> - TA <sub>mean</sub>
	<u>ANN</u>	<u>2.05</u>	<u>0.33</u>	<u>0.78</u>	<u>0.07</u>	SW <sub>in</sub>
	RF	2.12	0.32	0.77	0.07	z <sub>0</sub> · (h) <sup>-1</sup> <sub>SD00</sub>

Underlined records represent final algorithms for reconstructing cloud-covered MODIS LST and their corresponding predictors (seventh column) selected in the forward feature selection approach (see Sections III-B and III-C for more details).

of model transfer into unknown environments that have never been seen by fitted regressions. This approach is relevant to reduce uncertainties in spatial predictions beyond the training data, especially in mountainous regions with fragmented landscape for which cross-validation provides only global accuracy metrics limited to feature variability covered by training observations [100], [101]. Therefore, AOA were determined considering the range of the predictor values used for establishing the models. Specifically, the AOA was derived from standardized distances based on training data records ( $d_{\text{training}}^{\sigma}$ ) with respect to the spatial cross-validation folds as follows:

$$d_{\text{training}}^{\sigma} = \frac{d_{\text{training}}^j}{d_{\text{mean}}^{\text{training}}} \quad (4)$$

where  $d_{\text{training}}^j$  represents minimum Euclidean distance between a  $j$ th training observation and a point from remaining station-fold subsets, and  $d_{\text{mean}}^{\text{training}}$  indicates a mean distance between a  $j$ th location and other points included in other CV splits.

It means that  $d_{\text{training}}^{\sigma}$  was derived based on training points that did not appear in the same station-fold subset in the CV (see Section II-C3) as we assumed that the model performance estimates (see Table VI) apply to  $d_{\text{target}}^{\sigma}$ , which is comparable to  $d_{\text{training}}^{\sigma}$  values.

Model transfer to new geographic locations was realized by applying the 0.95 quantile of the  $d_{\text{training}}^{\sigma}$  as a threshold for the target standardized distances ( $d_{\text{target}}^{\sigma}$ ) [90]. Gridded data beyond that range were flagged as outliers (outside AOA), and thus, those areas were excluded from further LST reconstruction. Estimation of the AOA was performed in the CAST package implemented in the R software [98].

In addition to AOA, we determined potential locations for LST prediction by investigating spectral similarities between pure pixels represented by LULC-driven groups (see Section III-A) and mixed grid cells. Previous studies showed that land surface temperature is highly correlated with vegetation indices that capture spectral differences between plant species [31], [102], [103]. Thus, MODIS EVI (MOD13Q1 Version 6) was used for similar pixel extraction for vegetation groups defined in Table I. First, we extracted homogenous EVI with a minimum

threshold of 80% within 250-m mask obtained from the LISS-2013 landcover (Appendix, Fig. 9). We exploited time-series of EVI images acquired from 2014 to 2017 within months with phenological vegetation cycle. Target pixels were classified to the closest vegetation group using the following condition:

$$|\beta_i^T - \beta_{\text{target}}^T| < \sqrt{\frac{\sum_{n=1}^n (\beta_n^T - \beta_{\text{mean}}^T)^2}{n(n-1)}} \quad (5)$$

where  $\beta_i^T$  indicates the closest homogenous pixel  $i$  to a target pixel  $\beta_{\text{target}}^T$ ,  $T$  is a DOY corresponding to MODIS acquisition time,  $\beta_{\text{mean}}^T$  represents an average EVI value of available pure pixels  $\beta_n^T$  for a given DOY,  $n$  is number of homogenous pixels for each model group.

Pixels were assigned to a biome group if the majority of the individual  $\beta_{\text{target}}^T$  from the multiyear classification fulfilled the above constraint. Considering high landcover heterogeneity over the study area, some EVI pixels may contain fractional landcover, which hampers delineation of areas of applicability for the models. Many authors showed that LST of a nonpure pixel can be a linear mixture of subpixel components [104], [105]. In this context, when target pixels were classified to more than one group, fractional vegetation mask from the LULC map was generated, and weighted average values of surface temperature from respective biome-based models were computed (Appendix, Fig. 10).

### III. RESULTS

#### A. $LST_{\text{mean}} - TA_{\text{mean}}$ Comparison Under Cloudy Skies

Considering the complex character of the surface-atmosphere processes in the Alps, we first examined the overall relationships between ground-based  $LST_{\text{mean}}$  and  $TA_{\text{mean}}$  under-cloudy-sky conditions for each station independently. Fig. 4 illustrates the  $LST_{\text{mean}} - TA_{\text{mean}}$  scatterplots for sites located in forests and grasslands.

The regressions show a site-dependency driven by landcover of the stations (see Table I). The scatterplots represented by forest, however, approximated a 1:1 relationship more closely than sites covered by grasslands (see Fig. 4). For forest sites [see

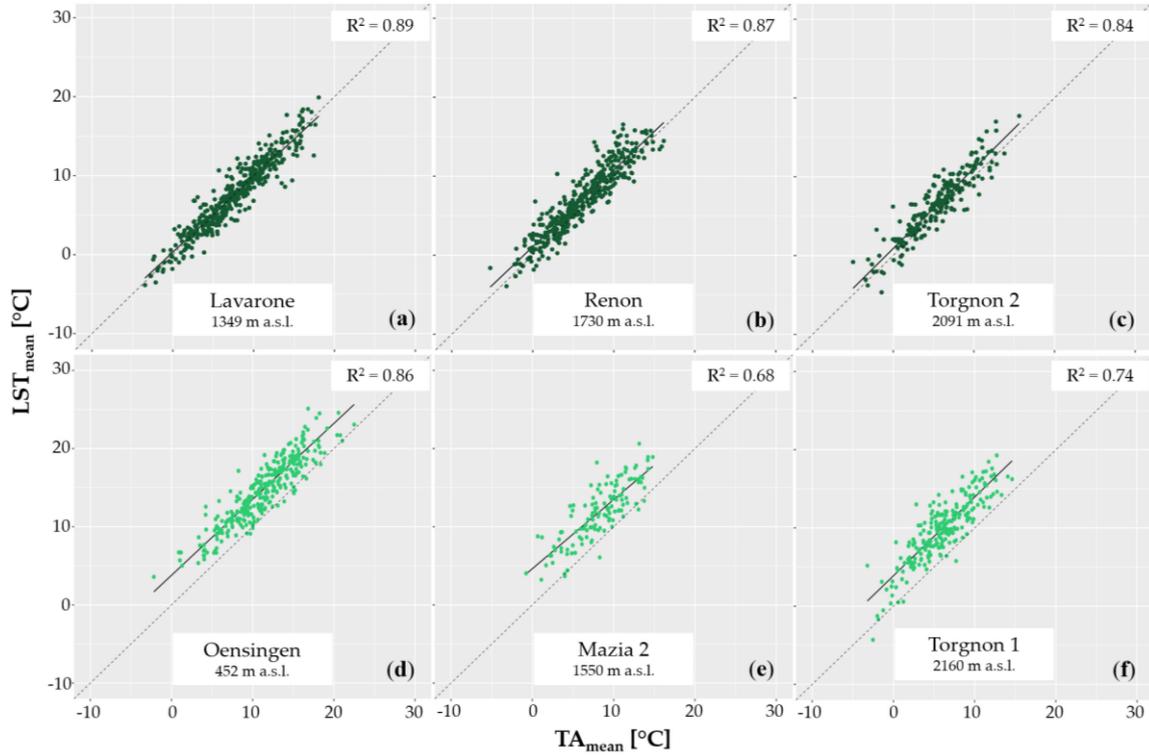


Fig. 4. Scatterplots between  $LST_{mean}$  and  $TA_{mean}$  under cloudy skies for six example weather stations covered by forest (in dark green) and grassland (in light green).  $LST_{mean}$  was calculated as an average from hourly LSTs recorded between 9 A.M. and 1 P.M. as an approximate range for daytime Terra MODIS acquisition time over the study area. Dashed lines in the scatterplots depict divergence between  $LST_{mean}$  and  $TA_{mean}$ .

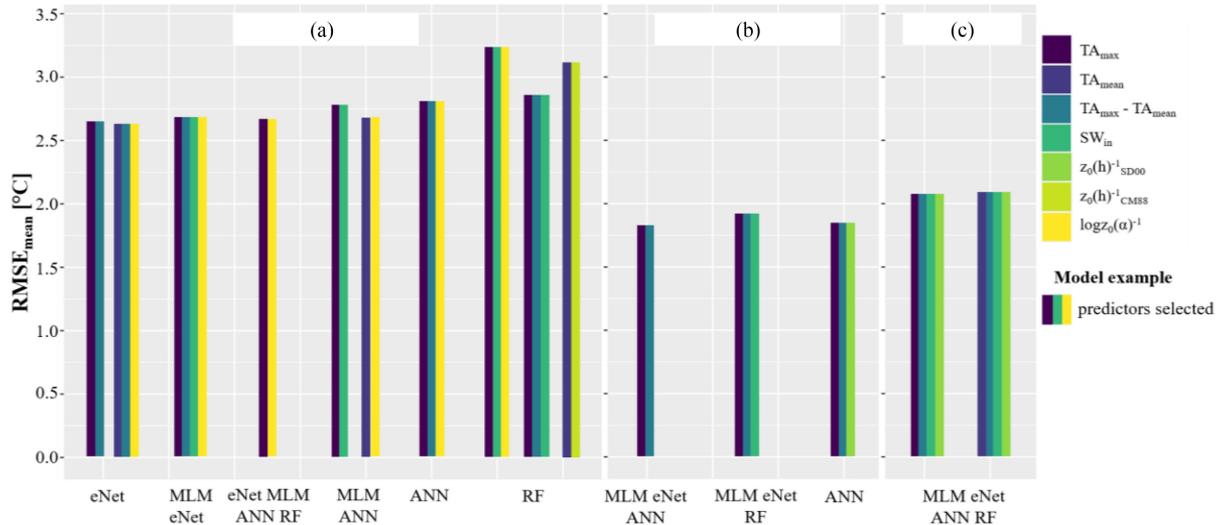


Fig. 5. Selected combinations of predictors evaluated by the averaged RMSE ( $RMSE_{mean}$ ) from all algorithms for: (a) permanent crops, (b) forest, and (c) grassland during the FFS procedure.  $RMSE_{mean}$  was computed considering hour-round LST models corresponding to MODIS acquisition time.

Fig. 4(a)–(c)  $LST_{mean}$  and  $TA_{mean}$  under cloudy skies ranged roughly the same values regardless their elevation, while  $TA_{mean}$  for grass-covered areas may underestimate surface temperatures [see Fig. 4(d)–(f)]. In particular,  $LST_{mean}$  over grasslands tends to be higher than  $TA_{mean}$  and the difference grows with temperature, especially at higher altitudes [see Fig. 4(e)–(f)]. Considering different  $LST_{mean} - TA_{mean}$  behavior among the landcover types, the LST models were built based on aggregated stations that represent similar environmental conditions

(see Table I). Forest, grasslands, and permanent crops were considered as three separate model groups for the final LST reconstruction.

### B. Selected Predictors

Although LST can be explained from  $TA_{mean}$ , the biophysical impact of landcover on LST, as shown in Fig. 4, suggested incorporation of additional variables that can describe complex

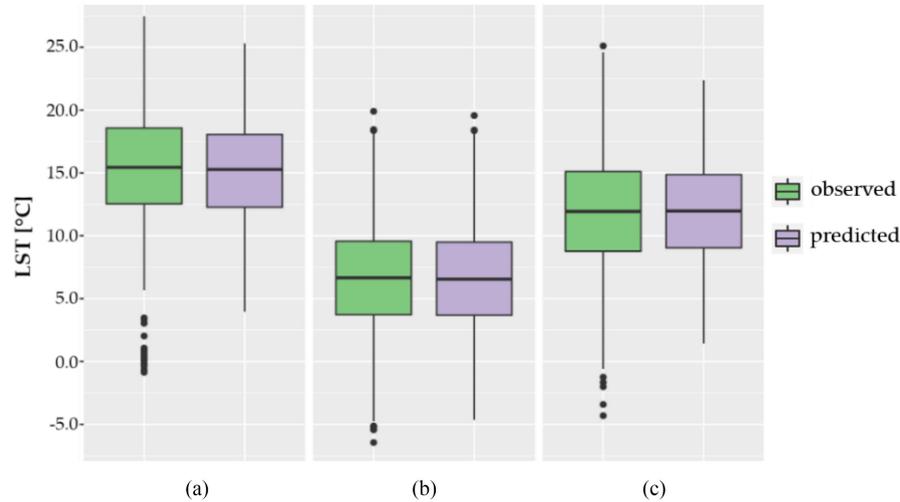


Fig. 6. Comparison between observed and predicted LSTs from the final models for: (a) permanent crops, (b) forest, and (c) grassland, considering combined MODIS-like time splits.

interplay between ground and atmosphere. The final features for each model group, as explained in Section II-C3, were selected using FFS procedure in the SCV approach. Fig. 5 displays selected explanatory variables for each model group based on all algorithms applied. Their relevance was evaluated by RMSE, averaged with respect to MODIS-like time splits ( $RMSE_{\text{mean}}$ ).

As can be observed in Fig. 5, LSTs over grassland and forest ecosystems, regardless of the algorithm used, were explained by one and three unique sets of predictors, respectively. During the FFS procedure for forest group, parameters were mainly reduced to air temperature ( $TA_{\text{max}}$ ,  $TA_{\text{mean}}$ ) and daily incoming solar radiation yielding  $RMSE_{\text{mean}}$  ranging from 1.82 °C to 1.92 °C, while for grass-covered areas vegetation structure parameter ( $z_0 \cdot h^{-1} SD_{00}$ ) was additionally selected as significant predictor in all MODIS-like time models [see Fig. 5(b) and (c)]. Due to complex structure of the permanent crops, which influences temporal variability in LST, eight different combinations of features were chosen, and they varied between algorithms [see Fig. 5(a)]. Considering all predictor sets of permanent crops, albedo-based product  $-\log(z_0) \cdot (\alpha)^{-1}$  and air temperature reported the highest frequencies yielding 69% and 100% of times they occurred in the models, respectively.

### C. Model Comparison and Assessment

Along with the predictor selection procedure, performance of the proposed LST reconstruction method was compared for each vegetation group separately (see Table I). As mentioned previously in Section II-C3, we checked regression overfit on test data in the SCV approach to evaluate predictive strength of the fitted models. Table VI shows averaged accuracy scores obtained from each algorithm with combined MODIS-like time splits.

As can be observed in Table VI, the proposed LST reconstruction concept gave quite similar accuracy statistics for each vegetation classes with small differences between the tested algorithms. Considerably uniform RMSE ( $R^2$ ) metrics, regardless

of the algorithm used, were present for forest and grassland ranging from 1.84 °C to 1.91 °C (0.86 to 0.88) and 2.05 °C to 2.12 °C (0.77 to 0.78), respectively (see Table VI). The greatest errors appeared over permanent crops yielding overall cross-validated RMSEs from 2.61 °C (MLM) to 3.05 °C using RF algorithm. Similar situation applied to standard deviation of RMSE ( $SD_{RMSE}$ ) and  $R^2$  scores ( $SD_{R^2}$ ), resulting in higher values for all tested algorithms when compared to forest and grass-covered model groups (see Table VI).

The multivariate linear regression was found as the best LST estimator with SCV RMSEs ( $SD_{RMSE}$ ) of 2.67 °C (1.62 °C) and 1.84 °C (0.42 °C) for permanent crops and forest, respectively. Although slightly lower RMSE over permanent crops was identified for the eNet, we selected MLM due to its smaller  $SD_{RMSE}$  score (see Table VI). For forest group eNet and ANN gave the same results ( $RMSE = 1.84$  °C) when compared to the linear model, however, we excluded these algorithms because of their longer computation time to tune hyperparameters (see Tables V and VI). On average, for the grassland models, we obtained the highest predictive performance from ANN yielding 0.78 and 2.05 °C for  $R^2$  and RMSE, respectively. Slightly lower accuracy statistics were noted for MLM and eNet with a 1.5% increase in RMSE. Random forest depicted the poorest predictive performance among all LULC model groups (see Table VI).

The performance of the regressions differed at the level of landcover classes. For forest we found smaller errors, while larger deviations in LSTs appeared over grasslands and permanent crops with 12% and 48% increases in RMSE when compared to forest accuracy metrics (see Table VI). Due to the unique atmospheric coupling of forest, this class maintained the strongest LST-TA relationship ( $R^2 \approx 90\%$ ) among all model groups (see Fig. 5, Table VI) [61], [106]. LSTs over grassland and permanent crops with a relatively smaller impact of turbulent mixing were additionally explained by other biophysical parameters, e.g., surface albedo, leaf area index, and aerodynamic roughness. Instability of crops model group revealed by higher  $SD_{RMSE}$  and  $SD_{R^2}$  values documented in Table VI, can be

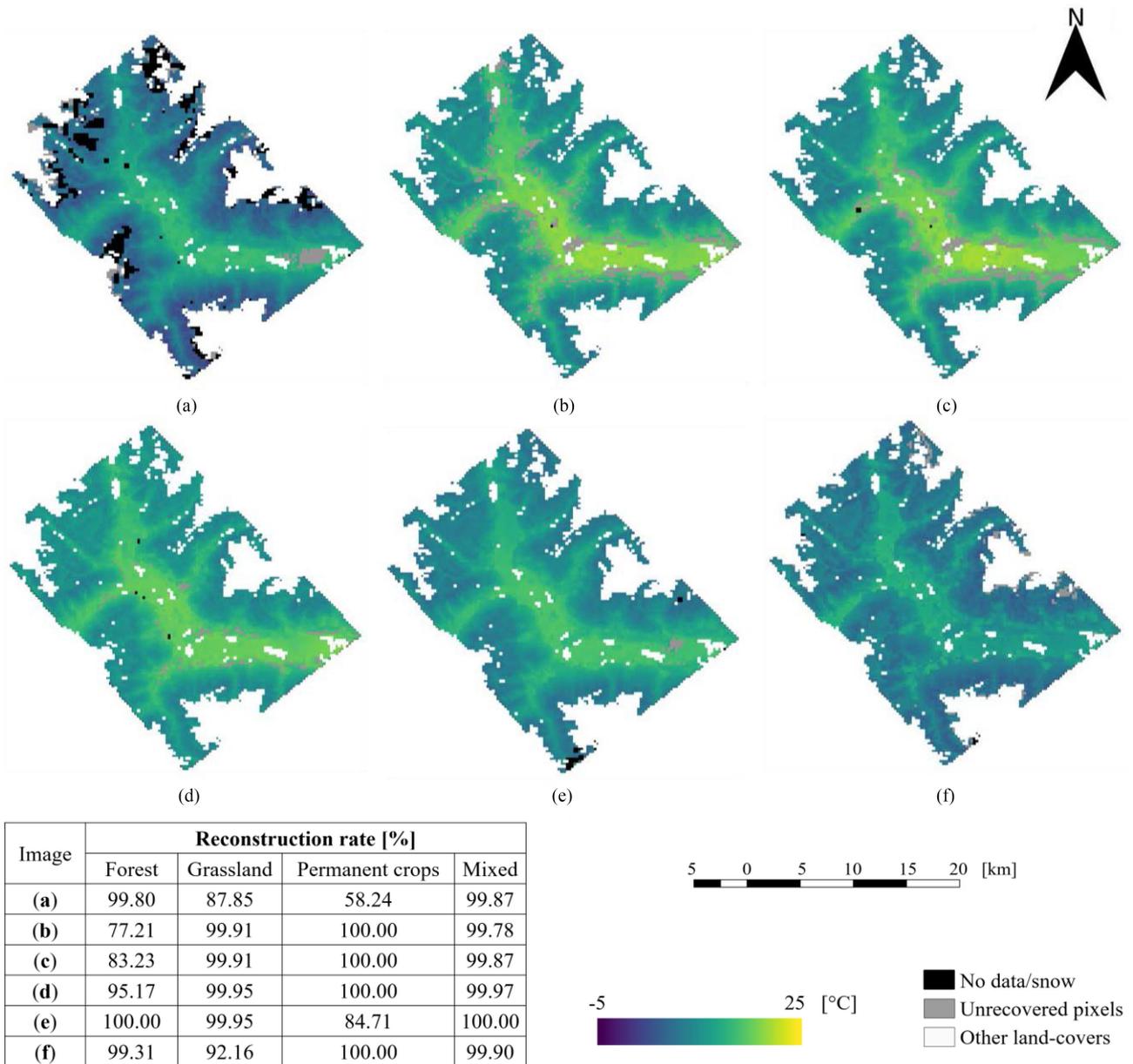


Fig. 7. Results of the proposed subpixel LST reconstruction of 1-km daytime Terra MODIS LST for long-term cloudy skies acquired on (a) May 2, 2014, (b) June 29, 2014, (c) July 8, 2014, (d) September 19, 2014, (e) October 11, 2014, and (f) October 26, 2014 in Vinschgau/Venosta. The attached table presents rate success of reconstructions for each vegetation group (see Appendix). According to the QA layer all original MODIS LST matrices were 100% covered by clouds.

related to the limited number of samples and heterogeneities of the combined land-use types (orchards and vineyards) incorporated in the modeling.

Considering the averaged accuracy statistics in Table VI, we achieved the strongest performance from ANN for grassland, while linear model gave the best results for permanent crops and forest groups. Therefore, only these algorithms were considered in the further analyses to reconstruct MODIS LST pixels obscured by clouds.

To compare prediction performance of the vegetation groups we investigated distribution of ground-based LST with their corresponding values predicted by the selected algorithms (see Fig. 6). On average, the interquartile ranges (IQR) of observed LSTs for all vegetation groups were in accordance with predicted

data. We found that the models were incapable to reconstruct observations beyond the IQRs, especially over grasslands and permanent crops [see Fig. 6(a)–(c)]. These patterns were mainly present for very low and high values (shown as circles), which appeared rarely in the models. Generally, this demonstrates that the chosen algorithms were able to capture LST variability with respect to different landcovers.

#### D. LST Reconstruction Under Cloudy-Sky Conditions

In the second part of the LST reconstruction the regressions were applied to gridded variables to estimate missing 250-m MODIS LST under cloudy-sky conditions. The outcomes of restoring invalid MODIS LSTs are presented in Fig. 7.

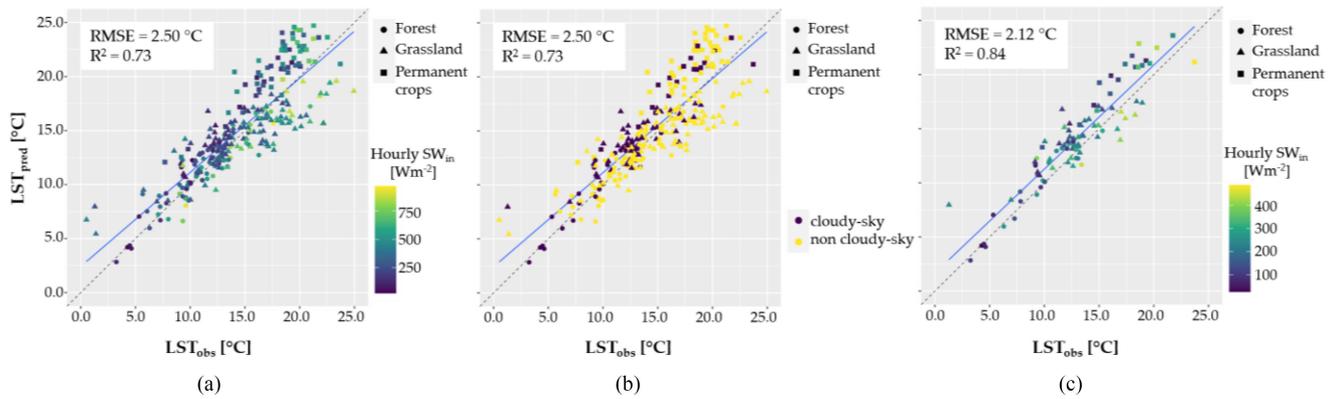


Fig. 8. Scatterplots between LST from ground measurements ( $LST_{obs}$ ) and recovered values ( $LST_{pred}$ ) based on the MODIS acquisition dates for the year 2014 with respect to: (a) instantaneous solar radiation (hourly  $SW_{in}$ ), (b) sky conditions, (c) available stations with reduced number of observations. Noncloudy-sky conditions in Fig. 8(b) refer to station records that were neither classified as long-term cloudy-sky nor clear-sky observations. The validated sites are located in relatively uniform areas with minimum 85% of homogeneity for related land-cover types within 250 m pixel [73], [107].

As demonstrated in Fig. 7, the approach proposed in this article provided a satisfactory data recovery over the study area. Invalid values were effectively predicted and LST downscaling to 250 m spatial resolution allowed simulating LST variability at subpixel level. For all images, regardless their acquisition date and observing time, more than 92% of blank pixels were filled properly. The highest reconstruction rate was observed for scenes recorded in late summer and autumn [see Fig. 7(d) and (e)] and ranged from 97.59% to 99.66% on October 26th and 11th, respectively. For other three dates [see Fig. 7(a)–(c)], however, we noticed larger number of unrecovered pixels ( $\leq 7.16\%$ ) that were classified as outliers by the models. On June 29th and July 8th areas of nonapplicability were identified for forests located in the center of the main valley and on its edges [see Fig. 7(b)–(c)]. Slightly less unpredicted pixels (3.75%) were found on May 2nd, which corresponded to high-mountain grasslands and permanent crops in the main valley [see Fig. 7(a)]. We found that the increased proportion of invalid LSTs, as seen in Fig. 7, may have occurred because of the smaller number of samples with cloudy-sky conditions and limited variability of observations to fit the models. This can be explained by different meteorological conditions in summer, with less intense and shorter cloud cover periods when compared to spring and autumn seasons. In addition, some pixels were not predicted due to coarse resolution of gridded parameters, such as MODIS LAI or albedo products. These factors had a negative impact on the LST reconstruction over the entire study area.

The reconstructed maps generally follow thermal patterns of the study area with higher temperatures in the valleys and colder conditions at high altitudes. On average, pixels of permanent crops located in the southwestern part of the region had relatively higher temperatures when compared to other biome groups situated at higher altitudes. We found that lower LST values were obtained over grasslands and forests. In contrast to the permanent crops, these biomes exhibited higher spatial heterogeneities with regard to terrain and thermal variability. This was confirmed in grass-covered areas with high deviations in elevation and LST equal to 546 m and 3.65 °C, respectively. Although for forest and grassland we noted similar thermal variability, standard

deviation of elevation for the latter vegetation group was over 200 m greater than for forest.

In addition, the reconstructed LST maps were compared with clear-sky MODIS LST grids acquired on temporally adjacent days (May 4th, July 1st, September 22nd, October 27th). Table VII shows an overview of statistical measures obtained from LSTs under cloudy skies (see Fig. 7) and their corresponding original MODIS images.

As shown in Table VII, the averaged differences in LST between clear-sky MODIS and the reconstructed grids revealed a cooling role of thick cloud cover that reduces amount of incoming shortwave radiation. In general, mean surface temperatures for clear-sky MODIS maps were higher than for the reconstructed grid cells, although the differences varied between time acquisitions. The largest deviations could be observed during spring and summer, with a maximum difference of 4.54 °C between June 29th and July 1st (see Table VII). Although higher values of the clear-sky MODIS LSTs generally indicate a stronger impact of solar radiation on the surface heating process,  $LST_{mean}$  on October 26th was very close to the average from MODIS LST acquired one day later (see Table VII).

Due to the complex land surface interactions in highly heterogeneous mountain ecosystems, the restored LST pixels could have been additionally explained by other predictors that influence surface thermal properties, such as landcover classes, biomass content, or surface albedo (see Fig. 5, Table VI). On the other hand, 1-km MODIS data may not represent LST variability at the subpixel scale of the reconstructed maps resulting in lower LST values over the study area.

#### E. Validation With In Situ Data

To assess the effectiveness of our LST reconstruction, we performed a quantitative assessment by comparing station-based LSTs with their corresponding gap-filled pixels at 250 m spatial resolution for the entire year of 2014. Validation results are illustrated in Fig. 8.

The evaluation results indicate a close agreement between ground measurements ( $LST_{obs}$ ) and the LST reconstructions

TABLE VII  
ANALYTICAL COMPARISON BETWEEN THE RECOVERED LST AND THE TIME-COINCIDENT ORIGINAL MODIS DATA, AVERAGED FOR THE STUDY AREA

Reconstructed cloudy LST		Clear-sky MODIS LST	
date	LST <sub>mean</sub> [°C]	date	LST <sub>mean</sub> [°C]
2 May 2014	8.50	4 May 2014	10.41
29 June 2014	12.95	1 July 2014	17.49
19 September 2014	12.56	22 September 2014	13.60
26 October 2014	8.69	27 October 2014	8.72

Clear-sky MODIS LST maps for the remaining recovered days (July 8, 2014, October 11, 2014) were not available.

TABLE VIII  
LOCAL COMPARISON OF RMSE AND BIAS VALUES BETWEEN *IN SITU* DATA AND RECONSTRUCTED GRIDS FOR CLOUD-COVERED MODIS LST PIXELS (DERIVED FROM QA FLAGS) FOR THE FOUR STATIONS BASED ON TIMESERIES FROM 2014

Station	Reconstructed LST RMSE [°C] (BIAS [°C])	
	all cloudy-sky conditions (long- & short-term)	long-term cloudy-sky conditions
Caldaro	3.26	2.81
	(-2.85)	(-2.38)
Lavarone	1.37	1.24
	(0.12)	(-0.44)
Mazia 1	2.40	2.42
	(-0.29)	(-1.47)
Mazia 2	2.88	1.57
	(1.10)	(-0.71)

The reconstruction results are shown for all cloudy-sky conditions and for observations under long-term cloudy-sky conditions [see Fig. 4 and Fig. 8(c)].

(LST<sub>pred</sub>). Considering all available *in situ* observations for the time acquisitions in 2014, the proposed approach was able to predict missing values with average  $R^2$  of 0.73 and RMSE equal to 2.50 °C and [see Fig. 8(a) and (b)]. As can be observed in the scatterplots [see Fig. 8(a)–(c)], the estimated grids have a relatively strong coherence with the stations, however, some observations depicted divergence from a 1:1 relationship. The largest deviations occurred for the agricultural (Caldaro) and grass-covered sites (Mazia 1, Mazia 2) with higher values of the RMSE and instantaneous  $SW_{in}$  (over 500  $Wm^{-2}$ ), which reflects no presence of long-term cloudy skies during MODIS acquisition times [see Fig. 8(b)]. From this aspect, we eliminated those points from further analysis and achieved better evaluation results yielding 2.12 °C and 0.84 for RMSE and  $R^2$ , respectively [see Fig. 8(c)]. Considering each station separately, accuracies of the LST reconstruction under long-term cloudy-sky conditions differed between stations with the lowest errors for Lavarone and Mazia 2 (RMSE = 1.24–1.57 °C) and higher ones for Mazia 1 and Caldaro (RMSE = 2.42–2.81 °C) (see Table VIII). Similarly, biases between observed and predicted LSTs were dependent on the site location. In general, for Caldaro and Mazia 1, regardless of sky conditions, the model led to overestimation ranging from –2.85 °C to –2.38 °C and –0.29 °C to –1.47 °C, respectively (see Table VIII). While LST under all cloudy-sky conditions (long-term and short-term cloudy sky) was underestimated at some sites (specifically at Lavarone and Mazia 2), on average LST predictions under long-term cloudy skies were higher than the observed values. Overall, the reconstructions perform better for long-term cloudy-sky observations with lower RMSE values when compared to all cloudy-sky conditions (see Table VIII).

Such discrepancies, as shown in Fig. 8 and Table VIII, may have appeared because of differences in cloud cover conditions defined by the models and MODIS QA layer. A possible explanation is that although MODIS cells were classified as overcast during satellite observing time, those LSTs might not have been influenced by the long-term cloudy-sky conditions of our approach. This led to lower values of the reconstructed temperatures when compared to the ground measurements (see Table VIII, Fig. 8). In addition, spatial heterogeneities of land surface over the Alps could not have been captured by coarse resolution of gridded predictors, such as solar radiation, LAI, and albedo products (see Table III). These factors would have impacts on the accuracy of the gap-filling models leading to biases in the computation of the actual LSTs (see Table VIII).

To test the potential impact of the meteorological input uncertainties on the model outputs under cloudy skies, we additionally predicted LSTs using the gridded datasets that were perturbed with input uncertainties of Type B by including their  $\pm 100\%$  values to the input grids for the year 2014 [108]. The reconstructed pixels were assessed in terms of RMSE and percentage LST change with respect to *in situ* observations and original LST predictions, respectively. In general, among all predictors analyzed, we observed the biggest model deviations influenced by  $TA_{max}$  ranging from  $\pm 16\%$  for Lavarone (forest) to  $\pm 8\%$  for grassland (Mazia 1, Mazia 2) and permanent crops in Caldaro (see Appendix, Fig. 11). For other variables we observed smaller impact on the LST estimations. As shown in Fig. 12, RMSEs obtained between ground-derived and modelled LSTs (using both source and modified parameters) depended

on model biome with the largest absolute differences ( $\sim 1$  °C) for permanent crops (Caldaro) when  $TA_{\text{mean}}$  and  $TA_{\text{max}}$  were perturbed.

#### IV. DISCUSSION

##### A. Advantages of the Proposed Reconstruction

Thermal remote sensing images are prone to overcast conditions resulting in spatial gaps in LST data. Although many LST reconstruction methods were developed, their implementation is hampered either due to rarely available biophysical parameters or coarse resolution outputs, which makes them irrelevant in heterogeneous environments. To overcome these limitations, the main objective of this article was to develop a robust procedure for restoring invalid coarse-resolution MODIS LSTs at 250 m spatial resolution by combining data-driven modeling from meteorological stations with physical-based approach to retrieve variables under long-term cloudy-sky conditions. The restored LST maps for six selected MOD11A1 images demonstrated the effectiveness of the AI-based reconstructions among all vegetation groups. The fitted models were able to identify parameters playing a key role in explaining LST variability driven by different landcovers. SCV with high coefficients of determination and relatively small errors confirmed the strong model performance, yielding on average an  $R^2$  of 0.80 and a RMSE of 2.19 °C. From a visual assessment of the restored maps, regardless of different seasons, no notable irregularities in LST patterns were observed.

Due to lack of *in situ* measurements in 2014 for all sites, the gap-filled LST images were validated with observations from four available stations (see Fig. 8, Table VIII). Results demonstrated that predicted LST data were in accordance with ground-based records obtaining  $R^2$  of 0.84 and RMSE of 2.12 °C. Compared with other reconstructions dealing with mountainous areas, our approach showed satisfactory evaluation scores. On average, Ke *et al.* [48] and Sun *et al.* [109] obtained lower accuracy with RMSEs ranging from 1.42 K to 3.16 K. Similarly, the SEB-based methods implemented by Yu *et al.* [29] and Yang *et al.* [32] led to weaker correlation between observed and predicted LSTs resulting in increased RMSEs ranging from 3.16 °C to 4.78 °C. However, we should keep in mind that these outcomes depend on site locations and are affected by accuracy of input data and spatial heterogeneity of study area.

This article demonstrated that surface temperatures under cloudy-sky conditions differ significantly from those under clear skies leading to greater differences between LST and daily  $TA_{\text{mean}}$  (see Fig. 3, Table IV) [60], [110]. Although they are correct from geostatistical point of view, they should not be used for retrieval of actual thermal conditions of the surface. On average, quantitative comparison between reconstructed maps and temporally adjacent clear-sky MODIS pixels showed that LSTs under cloudy-sky conditions were smaller when compared to valid datasets (see Table VII). This indicates an impact of clouds on amount of incident shortwave radiation, which regulates the land heating process. In this case findings from this article are not in agreement with spatio-temporal gap-filling proposed by Weiss *et al.* [44], Sun *et al.* [109], Li *et al.* [111], and Sarafanov *et al.*

[43] who predicted overcast surface temperatures from adjacent cloud-free pixels.

The developed AI-based approach provides more accurate understanding of additional controls on land surface temperature at local scale. Although strong correlations between LST and  $TA_{\text{mean}}$  under overcast conditions were observed (see Fig. 4), applying auxiliary variables to the models helped to explain LST variations among different vegetation groups. Based on preliminary tests, a split into separate landcover classes, which was also found favorable by Huang *et al.* [7], improved the accuracy of the models resulting in unique selection of final algorithms and predictors for each analyzed biome (see Fig. 5, Table VI). Permanent crops and forest models achieved the most satisfactory results from multivariate linear regression, while for grasslands the highest prediction performance was obtained using neural network algorithm. Forward feature selection revealed impacts of different factors controlling LSTs in heterogeneous ecosystems. Similar to the results of Bertoldi *et al.* [53] and Mildrexler *et al.* [61], forest maintained the strongest relationship with daily air temperature ( $TA_{\text{mean}}/TA_{\text{max}}$ ), while agricultural and grass-covered areas were additionally influenced by incident shortwave radiation and surface properties, such as albedo, aerodynamic roughness, and biomass content (see Figs. 4 and 5). The method developed in this article works well for the areas with clouds, which makes it comparable to physical-based LST reconstructions [30], [112]. Additionally, it does not require rarely available input parameters to describe complex physical mechanisms between ground and atmosphere.

Although our reconstruction approach showed a strong performance, about 7% of pixels of the study area were not predicted due to intentionally imposed constraints on the models (see Fig. 7). Prediction power of the reconstruction is limited by SCV strategy and similarity factors (Euclidean distance and spectral similarities in predictor variable space) that compare coherence between station-based training dataset and gridded predictors. These factors reduce areas for LST estimation to known environments where reconstruction errors apply. In this context, this approach seems to be objective in terms of spatial predictive tasks into new areas by avoiding locations with nonstandard observations identified by models [9], [90], [109].

##### B. Limitations

The proposed gap-filling approach exhibits large potential for producing high quality LST maps. Nevertheless, reliable reconstructions of LST in the Alps still pose challenges, mainly due to landscape heterogeneities and thermal variability in complex mountain environments.

Despite strong correlation between gridded data and ground measurements, the RMSE is 2.12 °C, which reflects that for some points prediction performance was still poor. Such discrepancies are likely related to input parameters leading to overestimation or underestimation of the restored maps. Differences in spatial scales between station records and gridded predictors, e.g., surface albedo, solar radiation, and air temperature can explain errors in the reconstructed data. Furthermore, the black antihail net used to protect the orchards (Caldaro) may introduce some

bias in the final LST output. In addition, some remotely sensed variables, including leaf area index and surface albedo were obtained from composite products (4-day MCD15A3H and daily 16-day MCD43A3 products), which introduces uncertainties associated with neglected temporal variations. In this context, more advanced daily interpolation methods could be explored [113]–[115].

When comparing the cross-validation results (see Table VI) with other data-driven methods [50]–[51], we found that our evaluation measures had less satisfactory scores. Wu *et al.* [50] obtained considerably stronger model performance with the CNN algorithm, attaining RMSE below 1 °C. Similarly, good results were reported by Zhao and Duan [51]. The authors applied random forest model from clear-sky pixels to cloud-covered MODIS LST with RMSE of 1.14 °C and  $R^2$  equal to 0.94. Such discrepancies between our method and the abovementioned [50], [51] could be explained by a relatively small number of observations for the agricultural landcover class, which led to instability of the model reflected by greater deviations in RMSE between cross-validated ensembles. In fact, diversified character of the permanent crops (orchards and vineyards) and the resulting data randomness had an additional impact on the model performance. Therefore, further study will concentrate on the extension of ground data by augmenting the timeseries of underrepresented training data for modelling prior to LST reconstructions [116], [117]. Since recurrent neural network can learn patterns and associations between sequential data over time, deep learning models, such as the long short-term memory (LSTM) algorithm exhibit a great potential in generating high-quality observable variables for better model training effects [118]. In addition, combining ancillary ground observations outside the study area along with postprocessed outputs from physical models would be beneficiary to learn new spatio-temporal input parametrizations of predictors [119]. On the other hand, the performance of the discussed methods [50], [51] relied on low-resolution satellite-based inputs and random cross-validation strategy, which makes them less strict and more prone to autocorrelation than our “leave-one-station-out” SCV method [9]. Additionally, our approach exploits data that capture spatial detail at station level, which translates to stronger thermal variability when compared to heterogeneous information from coarse resolution MODIS and MSG/SEVIRI LST pixels. Therefore, the prediction errors for our method can be generally acceptable.

While the predictions were robust in areas with high-frequency cloudy-sky conditions, in some cases cloud-covered pixels from MODIS QA layer could lead to disagreement between reconstructions and their actual temperatures. When comparing our estimations to ground data, we found that increased proportion of solar radiation caused underestimation of restored values (see Fig. 8, Table VIII). While LST reconstruction was resilient in areas with high-frequency cloudy-sky conditions, the method was limited for short-term cloudy skies. The proposed method assumes that overcast conditions are present when minimum 5-h constant cloudiness is recorded. Thus, the approach is suitable for long-term overcast conditions with thick clouds. Otherwise, it can lead to underestimation of LST values,

which is related to less intense insolation blocked by clouds. Thus, developing an extended approach for diversified overcast conditions could bring to significant accuracy improvements to the LST reconstruction over complex Alpine ecosystems.

Finally, the developed relationships are tuned to the ecosystems that characterize the study area. However, they could be easily extended to different regions if new observations (e.g., representing unique climatic conditions and surface properties) were provided. From this point of view, the method is less general than other data-driven approaches utilizing spatial information from remote sensing measurements [50], [51], [120], [121], because it specifically aims to gap-fill LST over Alpine regions, which are often under-represented in wide-ranging models. Nonetheless, LST reconstruction over different areas is still feasible, provided that required variables are available.

## V. CONCLUSION

Cloud cover has a significant impact on quality of remotely sensed LST observations, especially in high-frequency overcast areas, such as complex mountain regions. Therefore, a robust reconstruction of actual LST is a major research priority.

To address this problem, we presented a new method to reconstruct MODIS LST values under cloudy skies from station-based models at subpixel (250 m) spatial resolution. The proposed approach reconstructs actual thermal conditions of land surface under cloudy skies as an alternative way to the SEB-based methods that require complex parameters and are computationally expensive. Results demonstrate that the proposed data-driven analysis is capable of restoring invalid MODIS maps in a very robust way.

The reconstructed maps have reasonable LST values when compared to temporally neighboring clear-sky MODIS days. The validation against *in situ* data confirmed the strong agreement between estimations and ground observations. The performance of the models had satisfactory results. Prediction power showed landcover dependency resulting in individual predictors for each biome. Overall differences in RMSEs between machine learning algorithms and linear regression were minor for forest and permanent crops, while for grassland neural networks slightly improved the performance. At level of single vegetation groups, RMSE values ranged from 1.84 °C in forest to 2.67 °C for permanent crops.

Notably, it should be mentioned that the restored maps represent LSTs under long-term cloudy skies and may underestimate LST cells affected by brief overcast conditions. Therefore, future research should include an implementation of a hybrid approach for recovering missing pixels affected by short-term and long-term cloud contamination. Moreover, additional work will focus on the development of the enhanced reconstructions under long-term cloudy-sky conditions to increase spatial prediction performance of the established models. In parallel to the gap-filling procedure, data enhancement will be also applied to clear-sky observations of 1-km MODIS LST data to produce the full collection of 250-m resolution images considering different sky conditions. Furthermore, the proposed reconstruction

method by applying the AI-models can be extended to other low-resolution TIR sensors, such as Sentinel-3 SLSTR to provide data continuity for the study area. The availability of a spatially and temporally continuous set of thermal data would allow monitoring multitemporal trends of thermal conditions of the surface. These outputs could be value-added products for studies related to climatology, drought detection, and sustainable agriculture production, where land surface temperature is a baseline information for monitoring ecosystem dynamics over high frequency cloud-covered areas, such as mountain regions.

APPENDIX

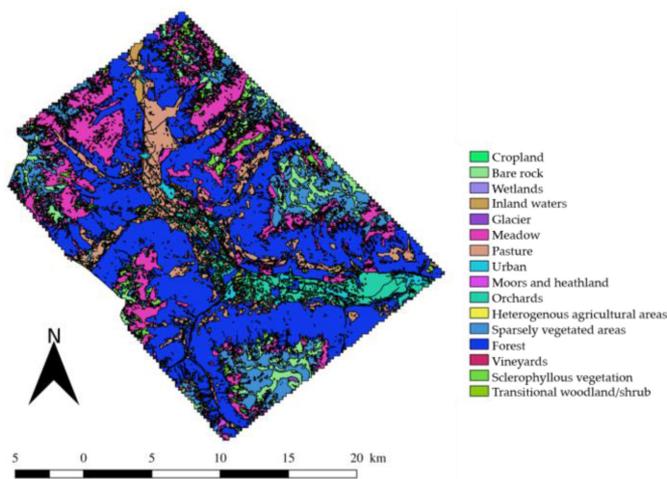


Fig. 9. Detailed land use landcover (LULC) map from the LISS-2013 cropped to the reconstruction area.

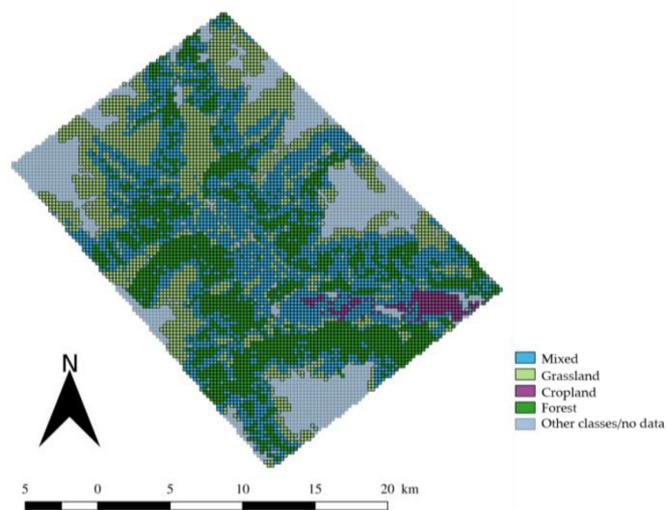


Fig. 10. 250-m fractional vegetation map from the LISS-2013 for the landcover-based reconstruction models. Group “mixed” indicates 250-m pixels with more than one vegetation biome.

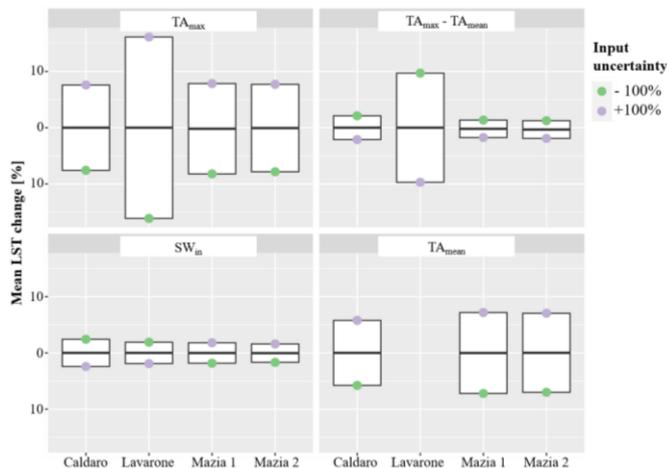


Fig. 11. Mean change in modeled LSTs with respect to surface temperature estimations from source meteorological input and the selected models (see Table VI) for the year 2014. Each climate predictor ( $TA_{max}$ ,  $TA_{mean}$ ,  $TA_{max}-TA_{mean}$ ,  $SW_{in}$ ) was perturbed by its uncertainty and included in the model together with remaining (unchanged) variables.

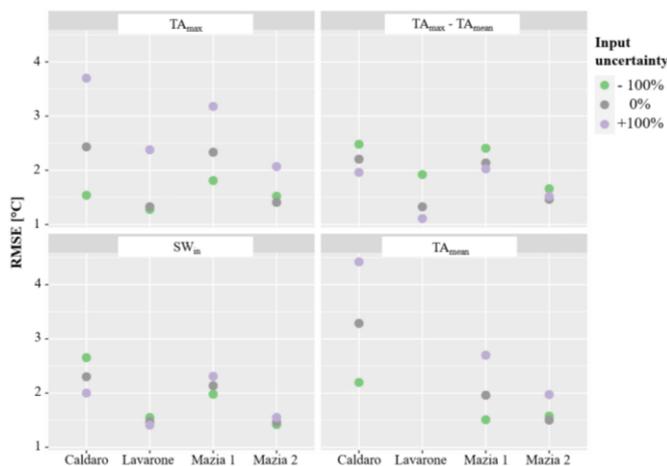


Fig. 12. RMSE values based on station-derived LST and predicted observations for two input scenarios: original variables (0%), and perturbed variables by  $\pm 100\%$  uncertainty values. For each model run only one predictor was changed.

ACKNOWLEDGMENT

The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs. The authors would like to thank G. Wohlfahrt for sharing data from Fluxnet sites and to H. Meyer for her helpful insights into spatial mapping.

REFERENCES

[1] M. C. Anderson, R. G. Allen, A. Morse, and W. P. Kustas, “Use of landsat thermal imagery in monitoring evapotranspiration and managing water resources,” *Remote Sens. Environ.*, vol. 122, pp. 50–65, Jul. 2012.

[2] M. Castelli *et al.*, “Two-source energy balance modeling of evapotranspiration in alpine grasslands,” *Remote Sens. Environ.*, vol. 209, pp. 327–342, May 2018.

- [3] G. R. Diak *et al.*, "Estimating land surface energy budgets from space: Review and current efforts at the University of Wisconsin—Madison and USDA—ARS," *Bull. Amer. Meteorological Soc.*, vol. 85, no. 1, pp. 65–78, Jan. 2004.
- [4] S. Leuzinger, R. Vogt, and C. Körner, "Tree surface temperature in an urban environment," *Agricultural Forest Meteorol.*, vol. 150, no. 1, pp. 56–62, Jan. 2010.
- [5] I. Sandholt, K. Rasmussen, and J. Andersen, "A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status," *Remote Sens. Environ.*, vol. 79, no. 2/3, pp. 213–224, Feb. 2002.
- [6] D. A. Sims *et al.*, "A new model of gross primary productivity for north american ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS," *Remote Sens. Environ.*, vol. 112, no. 4, pp. 1633–1646, Apr. 2008.
- [7] R. Huang, C. Zhang, J. Huang, D. Zhu, L. Wang, and J. Liu, "Mapping of daily mean air temperature in agricultural regions using daytime and nighttime land surface temperatures derived from TERRA and AQUA MODIS data," *Remote Sens.*, vol. 7, no. 7, pp. 8728–8756, Jul. 2015.
- [8] N. Janatian *et al.*, "A statistical framework for estimating air temperature using MODIS land surface temperature data," *Int. J. Climatol.*, vol. 37, no. 3, pp. 1181–1194, May 2016.
- [9] H. Meyer *et al.*, "Mapping daily air temperature for Antarctica based on MODIS LST," *Remote Sens.*, vol. 8, no. 9, Sep. 2016, Art. no. 732.
- [10] Z. L. Li *et al.*, "Satellite-derived land surface temperature: Current status and perspectives," *Remote Sens. Environ.*, vol. 131, pp. 14–37, Apr. 2013.
- [11] Z. Wan and J. Dozier, "A generalized split-window algorithm for retrieving land-surface temperature from space," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 892–905, Jul. 1996.
- [12] F. B. Balçık, "Determining the impact of urban components on land surface temperature of Istanbul by using remote sensing indices," *Environ. Monit. Assessment*, vol. 186, pp. 859–872, Sep. 2013.
- [13] K. Deilami and M. Kamruzzaman, "Modelling the urban heat island effect of smart growth policy scenarios in Brisbane," *Land Use Policy*, vol. 64, pp. 38–55, May 2017.
- [14] D. Lu, K. Song, S. Zang, M. Jia, J. Du, and C. Ren, "The effect of urban expansion on urban surface temperature in Shenyang, China: An analysis with landsat imagery," *Environ. Model. Assessment*, vol. 20, pp. 197–210, Jun. 2015.
- [15] J. A. Voogt and T. R. Oke, "Thermal remote sensing of urban climates," *Remote Sens. Environ.*, vol. 86, pp. 370–384, Aug. 2003.
- [16] M. C. Anderson *et al.*, "The evaporative stress index as an indicator of agricultural drought in Brazil: An assessment based on crop yield impacts," *Remote Sens. Environ.*, vol. 174, pp. 82–99, Mar. 2016.
- [17] S. Park, J. J. Feddema, and S. L. Egbert, "Impacts of hydrologic soil properties on drought detection with MODIS thermal data," *Remote Sens. Environ.*, vol. 89, pp. 53–62, Jan. 2004.
- [18] S. Swain, B. D. Wardlow, S. Narumalani, T. Tadesse, and K. Callahan, "Assessment of vegetation response to drought in Nebraska using Terra-MODIS land surface temperature and normalized difference vegetation index," *GISci. Remote Sens.*, vol. 48, pp. 432–455, Jul. 2011.
- [19] H. T. Tran, J. B. Campbell, T. D. Tran, and H. T. Tran, "Monitoring drought vulnerability using multispectral indices observed from sequential remote sensing (Case study: Tuy Phong, Binh Thuan, Vietnam)," *GISci. Remote Sens.*, vol. 54, no. 2, pp. 167–184, Mar. 2017.
- [20] M. C. Anderson, C. Hain, B. Wardlow, A. Pimstein, J. R. Mecikalski, and W. P. Kustas, "Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States," *J. Climate*, vol. 24, no. 8, pp. 2025–2044, Apr. 2011.
- [21] S. Park, J. Im, E. Jang, and J. Rhee, "Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions," *Agricultural Forest Meteorol.*, vol. 216, pp. 157–169, Jan. 2016.
- [22] M. C. Anderson, J. M. Norman, J. R. Mecikalski, J. A. Otkin, and W. P. Kustas, "A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation," *J. Geophys. Res.-Atmos.*, vol. 112, May 2007, Art. no. D10.
- [23] E. Carpintero *et al.*, "Continuous evapotranspiration monitoring and water stress at watershed scale in a Mediterranean Oak Savanna," in *Proc. Remote Sens. Agriculture, Ecosyst., Hydrol. XVIII. Int. Soc. Opt. Photon.*, Edinburgh, U.K., Oct. 2016, vol. 9998, Art. no. 99980N.
- [24] S. Rahimi, M. A. Gholami Sefidkouhi, M. Raeni-Sarjaz, and M. Valipour, "Estimation of actual evapotranspiration by using MODIS images (a case study: Tajan catchment)," *Arch. Agronomy Soil Sci.*, vol. 61, no. 5, pp. 695–709, May 2015.
- [25] R. Tang and Z. L. Li, "Evaluation of two end-member-based models for regional land surface evapotranspiration estimation from MODIS data," *Agricultural Forest Meteorol.*, vol. 202, pp. 69–82, Mar. 2015.
- [26] T. Wang, J. Shi, H. Letu, Y. Ma, X. Li, and Y. Zheng, "Detection and removal of clouds and associated shadows in satellite imagery based on simulated radiance fields," *J. Geophys. Res. Atmos.*, vol. 124, no. 13, pp. 7207–7225, Jun. 2019.
- [27] M. Jin, "Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle: 2. Cloudy-pixel treatment," *J. Geophys. Res. Atmos.*, vol. 105, no. D3, pp. 4061–4076, Feb. 2000.
- [28] L. Lu, V. Venus, A. Skidmore, T. Wang, and G. Luo, "Estimating land-surface temperature under clouds using MSG/SEVIRI observations," *Int. J. Appl. Earth Observ. Geoinformat.*, vol. 13, no. 2, pp. 265–276, Apr. 2011.
- [29] W. Yu, J. Tan, M. Ma, X. Li, X. She, and Z. Song, "An effective similar-pixel reconstruction of the high-frequency cloud-covered areas of Southwest China," *Remote Sens.*, vol. 11, no. 3, Jan. 2019, Art. no. 336.
- [30] X. Zhang, J. Pang, and L. Li, "Estimation of land surface temperature under cloudy skies using combined diurnal solar radiation and surface temperature evolution," *Remote Sens.*, vol. 7, no. 1, pp. 905–921, Jan. 2015.
- [31] C. Zeng, D. Long, H. Shen, P. Wu, Y. Cui, and Y. Hong, "A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud," *ISPRS J. Photogramm.*, vol. 141, pp. 30–45, Jul. 2018.
- [32] G. Yang, W. Sun, H. Shen, X. Meng, and J. Li, "An integrated method for reconstructing daily MODIS land surface temperature data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 3, pp. 1026–1040, Mar. 2019.
- [33] J. Martins *et al.*, "An all-weather land surface temperature product based on MSG/SEVIRI observations," *Remote Sens.*, vol. 11, no. 24, Jan. 2019, Art. no. 3044.
- [34] P. Fu, Y. Xie, Q. Weng, S. Myint, K. Meacham-Hensold, and C. Bernacchi, "A physical model-based method for retrieving urban land surface temperatures under cloudy conditions," *Remote Sens. Environ.*, vol. 230, Sep. 2019, Art. no. 111191.
- [35] H. Gao, R. Fu, R. E. Dickinson, and R. I. N. Juárez, "A practical method for retrieving land surface temperature from AMSR-E over the amazon forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 193–199, Dec. 2007.
- [36] H. R. Shwetha and D. N. Kumar, "Prediction of high spatio-temporal resolution land surface temperature under cloudy conditions using microwave vegetation index and ANN," *ISPRS J. Photogramm.*, vol. 117, pp. 40–55, Jul. 2016.
- [37] S. B. Duan, Z. L. Li, and P. Leng, "A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data," *Remote Sens. Environ.*, vol. 195, pp. 107–117, Jun. 2017.
- [38] W. Tang, D. Xue, Z. Long, X. Zhang, and J. Zhou, "Near-real-time estimation of 1-km all-weather land surface temperature by integrating satellite passive microwave and thermal infrared observations," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, Apr. 2021, Art. no. 7001305.
- [39] X. Zhang, J. Zhou, S. Liang, L. Chai, D. Wang, and J. Liu, "Estimation of 1-km all-weather remotely sensed land surface temperature based on reconstructed spatial-seamless satellite passive microwave brightness temperature and thermal infrared data," *ISPRS J. Photogramm.*, vol. 167, pp. 321–344, Sep. 2020.
- [40] D. Long *et al.*, "Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach," *Remote Sens. Environ.*, vol. 246, Sep. 2020, Art. no. 111863.
- [41] C. M. Frey and C. Künzer, "Land surface temperature dynamics in the Upper Mekong Basin derived from MODIS time series," *Int. J. Remote Sens.*, vol. 35, no. 8, pp. 2780–2798, Apr. 2014.
- [42] M. Neteler, "Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data," *Remote Sens.*, vol. 2, no. 1, pp. 333–351, Jan. 2010.
- [43] M. Sarafanov, E. Kazakov, N. O. Nikitin, and A. V. Kalyuzhnaya, "A machine learning approach for remote sensing data gap-filling with open-source implementation: An example regarding land surface temperature, surface Albedo and NDVI," *Remote Sens.*, vol. 12, no. 23, Nov. 2020, Art. no. 3865.

- [44] D. J. Weiss, P. M. Atkinson, S. Bhatt, B. Mappin, S. I. Hay, and P. W. Gething, "An effective approach for gap-filling continental scale remotely sensed time-series," *ISPRS J. Photogramm.*, vol. 98, pp. 106–118, Dec. 2014.
- [45] Y. Xu and Y. Shen, "Reconstruction of the land surface temperature time series using harmonic analysis," *Comput. Geosci.*, vol. 61, pp. 126–132, Dec. 2013.
- [46] W. L. Crosson, M. Z. Al-Hamdan, S. N. Hemmings, and G. M. Wade, "A daily merged MODIS aqua-terra land surface temperature data set for the conterminous United States," *Remote Sens. Environ.*, vol. 119, pp. 315–324, Apr. 2012.
- [47] T. Wang *et al.*, "Recovering land surface temperature under cloudy skies considering the solar-cloud-satellite geometry: Application to MODIS and landsat-8 data," *J. Geophys. Res. Atmos.*, vol. 124, no. 6, pp. 3401–3416, Mar. 2019.
- [48] L. Ke, X. Ding, and C. Song, "Reconstruction of time-series MODIS LST in Central Qinghai-Tibet Plateau using geostatistical approach," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1602–1606, Jun. 2013.
- [49] X. M. Fan, H. G. Liu, G. H. Liu, and S. B. Li, "Reconstruction of MODIS land-surface temperature in a flat terrain and fragmented landscape," *Int. J. Remote Sens.*, vol. 35, no. 23, pp. 7857–7877, Dec. 2014.
- [50] P. Wu, Z. Yin, H. Yang, Y. Wu, and X. Ma, "Reconstructing geostationary satellite land surface temperature imagery based on a multiscale feature connected convolutional neural network," *Remote Sens.*, vol. 11, no. 3, Jan. 2019, Art. no. 300.
- [51] W. Zhao and S. B. Duan, "Reconstruction of daytime land surface temperatures under cloud-covered conditions using integrated MODIS/Terra land products and MSG geostationary satellite data," *Remote Sens. Environ.*, vol. 247, Sep. 2020, Art. no. 11931.
- [52] P. Bartkowiak, M. Castelli, and C. Notarnicola, "Downscaling land surface temperature from MODIS dataset with random forest approach over alpine vegetated areas," *Remote Sens.*, vol. 11, no. 11, Jun. 2019, Art. no. 1319.
- [53] G. Bertoldi, S. D. Chiesa, C. Notarnicola, L. Pasolli, G. Niedrist, and U. Tappeiner, "Estimation of soil moisture patterns in mountain grasslands by means of SAR RADARSAT2 images and hydrological modeling," *J. Hydrol.*, vol. 516, pp. 245–257, Aug. 2014.
- [54] G. Pastorello *et al.*, "The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data," *Sci. Data.*, vol. 7, no. 1, pp. 1–27, Jul. 2020.
- [55] A. K. Heidinger, I. Laszlo, C. C. Molling, and D. Tarpley, "Using SURFRAD to verify the NOAA single-channel land surface temperature algorithm," *J. Atmos. Ocean. Technol.*, vol. 30, no. 12, pp. 2868–2884, Dec. 2013.
- [56] Z. Xing *et al.*, "Estimation of daily mean land surface temperature at global scale using pairs of daytime and nighttime MODIS instantaneous observations," *ISPRS J. Photogramm.*, vol. 178, pp. 51–67, Aug. 2021.
- [57] G. Hulley, N. Malakar, and R. Freepartner, "Moderate resolution imaging spectroradiometer (MODIS) land surface temperature and emissivity product (MxD21) algorithm theoretical basis document collection-6," Pasadena, California, USA: JPL Publication, Dec. 2016, pp. 12–17.
- [58] K. Wang *et al.*, "Estimation of surface long wave radiation and broadband emissivity using moderate resolution imaging spectroradiometer (MODIS) land surface temperature/emissivity products," *J. Geophys. Res. Atmospheres*, vol. 110, Jun. 2005, Art. no. D11.
- [59] G. Hulley and J. Nickeson, "M\*D21 (MYD21/MOD21) and M\*D11 (MYD11/MYD21) land surface temperature and emissivity (LST&E) products status report," Aug. 2020.
- [60] K. Gallo, R. Hale, D. Tarpley, and Y. Yu, "Evaluation of the relationship between air and land surface temperature under clear-and cloudy-sky conditions," *J. Appl. Meteorol. Climatol.*, vol. 50, no. 3, pp. 767–775, Mar. 2011.
- [61] D. J. Mildrexler, M. Zhao, and S. W. Running, "A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests," *J. Geophys. Res. BioGeosci.*, vol. 116, Sep. 2011, Art. no. G3.
- [62] J. Silva, C. Ribeiro, and R. Guedes, "Roughness length classification of corine land cover classes," in *Proc. Eur. Wind Energy Conf.*, Milan, Italy, vol. 710, May 2007, Art. no. 110.
- [63] J. Cho, S. Miyazaki, P. J. F. Yeh, W. Kim, S. Kanae, and T. Oki, "Testing the hypothesis on the relationship between aerodynamic roughness length and albedo using vegetation structure parameters," *Int. J. Biometeorol.*, vol. 56, no. 2, pp. 411–418, May 2012.
- [64] L. Laiti *et al.*, "A solar atlas for the trentino region in the Alps: Quality control of surface radiation data," *Energy Procedia*, vol. 59, pp. 336–343, Jan. 2014.
- [65] A. Crespi, M. Matiu, G. Bertoldi, M. Petitta, and M. Zebisch, "A high-resolution gridded dataset of daily temperature and precipitation records (1980–2018) for trentino–South Tyrol (North-Eastern Italian Alps)," *Earth Syst. Sci. Data Discuss.*, vol. 13, no. 6, pp. 2801–2818, Jun. 2021.
- [66] M. New, M. Hulme, and P. Jones, "Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate," *J. Clim.*, vol. 13, no. 13, pp. 2217–2238, Jul. 2000.
- [67] C. Schaaf and Z. Wang, "MCD43A3 MODIS/Terra+Aqua BRDF/Albedo daily L3 Global – 500m V006," NASA EOSDIS Land Processes DAAC, NASA, WA, DC, USA, 2015, doi: [10.5067/MODIS/MCD43A3](https://doi.org/10.5067/MODIS/MCD43A3).
- [68] B. J. Choudhury and J. L. Monteith, "A four-layer model for the heat budget of homogeneous land surfaces," *Quart. J. Roy. Meteorological Soc.*, vol. 114, no. 480, pp. 373–398, Jan. 1988.
- [69] K. J. Schaudt and R. E. Dickinson, "An approach to deriving roughness length and zero-plane displacement height from satellite data, prototyped with BOREAS data," *Agricultural Forest Meteorol.*, vol. 104, no. 2, pp. 143–155, Aug. 2000.
- [70] R. Myneni, Y. Knyazikhin, and T. Park, "MCD15A3H MODIS/Terra+Aqua leaf area Index/FPAR 4-day L4 Global 500m SIN grid V006," NASA EOSDIS Land Processes DAAC, NASA, WA, DC, USA, 2015, doi: [10.5067/MODIS/MCD15A3H](https://doi.org/10.5067/MODIS/MCD15A3H).
- [71] Z. Wan, S. Hook, and G. Hulley, "MOD11A1 MODIS/Terra land surface temperature/emissivity daily L3 global 1km SIN grid V006," NASA EOSDIS Land Processes DAAC, NASA, WA, DC, USA, 2015, doi: [10.5067/MODIS/MOD11A1](https://doi.org/10.5067/MODIS/MOD11A1).
- [72] K. Didan, "MOD13Q1 MODIS/Terra vegetation indices 16-Day L3 global 250m SIN grid V006," NASA EOSDIS Land Processes DAAC, NASA, WA, DC, USA, 2015, doi: [10.5067/MODIS/MOD13Q1](https://doi.org/10.5067/MODIS/MOD13Q1).
- [73] LISS, "LISS-2013 – Land Use Information in South Tyrol: Update Harmonization with European Standards and Integration of Research Results," Autonome Provinz Bozen/Provincia autonoma di Bolzano, 28.0.1 Landeskartographie und Koordination der Geodaten, 2013.
- [74] C. Notarnicola *et al.*, "Snow cover maps from MODIS images at 250 m resolution, Part I: Algorithm description," *Remote Sens.*, vol. 5, no. 1, pp. 110–126, Jan. 2013.
- [75] R. Lecerf, L. Hubert-Moy, T. Corpetti, F. Baret, B. A. Latif, and H. Nicolas, "Estimating biophysical variables at 250 m with reconstructed EOS/MODIS time series to monitor fragmented landscapes," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Boston, MA, USA, vol. 2, pp. II–954, Jul. 2008.
- [76] M. Machwitz, U. Gessner, C. Conrad, U. Falk, J. Richters, and S. Dech, "Modelling the gross primary productivity of west Africa with the regional biomass model RBM+, using optimized 250 m MODIS FPAR and fractional vegetation cover information," *Int. J. Appl. Earth Obs.*, vol. 43, pp. 177–194, Dec. 2015.
- [77] M. Matiu, A. Jacob, and C. Notarnicola, "Daily MODIS snow cover maps for the European Alps from 2002 onwards at 250 m horizontal resolution along with a nearly cloud-free version," *Data*, vol. 5, no. 1, Mar. 2020.
- [78] F. Zambrano, M. Lillo-Saavedra, K. Verbist, and O. Lagos, "Sixteen years of agricultural drought assessment of the Biobío region in Chile using a 250 m resolution vegetation condition index (VCI)," *Remote Sens.*, vol. 8, no. 6, Jun. 2016, Art. no. 530.
- [79] A. Ignatov and G. Gutman, "Monthly mean diurnal cycles in surface temperatures over land for global climate studies," *J. Climate*, vol. 12, no. 7, pp. 1900–1910, Jul. 1999.
- [80] N. A. Engerer and F. P. Mills, "Validating nine clear sky radiation models in Australia," *Sol. Energy*, vol. 120, pp. 9–24, Oct. 2015.
- [81] J. Remund, L. Wald, M. Lefèvre, T. Ranchin, and J. Page, "Worldwide Linke turbidity information," in *Proc. Sol. World Congr. Int. Sol. Energy Soc.*, Göteborg, Sweden, vol. 400, Jun. 16, 2003, Art. no. 13.
- [82] C. Rigollier, O. Bauer, and L. Wald, "On the clear sky model of the ESRA—European solar radiation Atlas—With respect to the heliosat method," *Sol. Energy*, vol. 68, no. 1, pp. 33–48, Jan. 2000.
- [83] A. Benali, A. C. Carvalho, J. P. Nunes, N. Carvalhais, and A. Santos, "Estimating air surface temperature in Portugal using MODIS LST data," *Remote Sens. Environ.*, vol. 124, pp. 108–121, Sep. 2012.
- [84] D. Y. Kim and K. S. Han, "Remotely sensed retrieval of midday air temperature considering atmospheric and surface moisture conditions," *Int. J. Remote Sens.*, vol. 34, no. 1, pp. 247–263, Jan. 2013.
- [85] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

- [86] M. Kuhn and K. Johnson, *Applied Predictive Modeling*, vol. 26, 1st ed. New York, NY, USA: Springer, 2013, pp. 13–203.
- [87] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [88] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, “Do we need hundreds of classifiers to solve real world classification problems?,” *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 3133–3181, Jan. 2014.
- [89] S. Talukdar, P. Singha, S. Mahato, S. Pal, Y. A. Liou, and A. Rahman, “Land-use land-cover classification by machine learning classifiers for satellite observations—A review,” *Remote Sens.*, vol. 12, no. 7, Jan. 2020, Art. no. 1135.
- [90] H. Meyer and E. Pebesma, “Predicting into unknown space? Estimating the area of applicability of spatial prediction models,” *Methods Ecol. Evol.*, vol. 12, pp. 1620–1633, Jun. 2021.
- [91] J. F. Hair *et al.*, *Multivariate Data Analysis: A Global Perspective*, 7th ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2009.
- [92] H. Meyer, C. Reudenbach, S. Wöllauer, and T. Nauss, “Importance of spatial predictor variable selection in machine learning applications—Moving from data reproduction to spatial prediction,” *Ecol. Model.*, vol. 411, Nov. 2019, Art. no. 108815.
- [93] H. Meyer, C. Reudenbach, T. Hengl, M. Katurji, and T. Nauss, “Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation,” *Environ. Model. Softw.*, vol. 101, pp. 1–9, Mar. 2018.
- [94] D. R. Roberts *et al.*, “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure,” *Ecography*, vol. 40, no. 8, pp. 913–929, Aug. 2017.
- [95] P. Schratz, J. Muenchow, E. Iturrutxa, J. Richter, and A. Brenning, “Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data,” *Ecol. Model.*, vol. 406, pp. 109–120, Aug. 2019.
- [96] M. Kuhn, “Building predictive models in R using the caret package,” *J. Stat. Softw.*, vol. 28, no. 5, pp. 1–26, Nov. 2008.
- [97] M. Kuhn, “The caret package,” *J. Stat. Softw.*, vol. 28, Nov. 2009, Art. no. 5.
- [98] H. Meyer, C. Reudenbach, M. Ludwig, T. Nauss, and M. H. Meyer, “Package ‘CAST’,” *R package version 0.5.1*, Apr. 2021.
- [99] R. C. Team, “R: A language and environment for statistical computing,” *R Found. Statist. Comput.*, Vienna, Austria, 2013.
- [100] N. Bystriakova, M. Peregrym, R. H. Erkens, O. Bezsmertna, and H. Schneider, “Sampling bias in geographic and environmental space and its effect on the predictive power of species distribution models,” *Syst. Biodiversity*, vol. 10, no. 3, pp. 305–315, Aug. 2012.
- [101] R. Kadmon, O. Farber, and A. Danin, “Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models,” *Ecological Appl.*, vol. 14, no. 2, pp. 401–413, Apr. 2004.
- [102] M. Jin and R. E. Dickinson, “Land surface skin temperature climatology: Benefitting from the strengths of satellite observations,” *Environ. Res. Lett.*, vol. 5, no. 4, Nov. 2010, Art. no. 044004.
- [103] P. A. Schultz and M. S. Halpert, “Global analysis of the relationships among a vegetation index, precipitation and land surface temperature,” *Remote Sens.*, vol. 16, no. 15, pp. 2755–2777, Oct. 1995.
- [104] L. Busetto, M. Meroni, and R. Colombo, “Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series,” *Remote Sens. Environ.*, vol. 112, no. 1, pp. 118–131, Jan. 2008.
- [105] F. Gao, J. Masek, M. Schwaller, and F. Hall, “On the blending of the landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2207–2218, Aug. 2006.
- [106] P. D’odorico, Y. He, S. Collins, S. F. De Wekker, V. Engel, and J. D. Fuentes, “Vegetation–microclimate feedbacks in woodland–grassland ecotones,” *Glob. Ecol. Biogeography*, vol. 22, no. 4, pp. 364–379, Apr. 2013.
- [107] L. Congedo, L. Sallustio, M. Munafò, M. Ottaviano, D. Tonti, and M. Marchetti, “Copernicus high-resolution layers for land cover classification in Italy,” *J. Maps*, vol. 12, no. 5, pp. 1195–1205, Feb. 2016.
- [108] N. P. Majoji, C. M. Mannaerts, A. Ramoelo, R. Mathieu, and W. Verhoef, “Uncertainty and sensitivity analysis of a remote-sensing-based penman–Monteith model to meteorological and land surface input variables,” *Remote Sens.*, vol. 13, no. 5, Feb. 2021, Art. no. 882.
- [109] L. Sun *et al.*, “Reconstructing daily clear-sky land surface temperature for cloudy regions from MODIS data,” *Comput. Geosci.*, vol. 105, pp. 10–20, Aug. 2017.
- [110] K. Makowski, E. B. Jaeger, M. Chiacchio, M. Wild, T. Ewen, and A. Ohmura, “On the relationship between diurnal temperature range and surface solar radiation in Europe,” *J. Geophys. Res. Atmos.*, vol. 114, Apr. 2009, Art. no. D10.
- [111] X. Li, Y. Zhou, G. R. Asrar, and Z. Zhu, “Creating a seamless 1 km resolution daily land surface temperature dataset for urban and surrounding areas in the conterminous United States,” *Remote Sens. Environ.*, vol. 206, pp. 84–97, Mar. 2018.
- [112] W. Yu, M. Ma, X. Wang, and J. Tan, “Estimating the land-surface temperature of pixels covered by clouds in MODIS products,” *J. Appl. Remote Sens.*, vol. 8, no. 1, Nov. 2014, Art. no. 083525.
- [113] B. Jiang, S. Liang, J. Wang, and Z. Xiao, “Modeling MODIS LAI time series using three statistical methods,” *Remote Sens. Environ.*, vol. 114, no. 7, pp. 1432–1444, Jul. 2010.
- [114] E. G. Moody, M. D. King, S. Platnick, C. B. Schaaf, and F. Gao, “Spatially complete global spectral surface Albedos: Value-added datasets derived from Terra MODIS land products,” *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 144–158, Jan. 2005.
- [115] L. Pasolli *et al.*, “Retrieval of leaf area index in mountain grasslands in the Alps from MODIS satellite imagery,” *Remote Sens. Environ.*, vol. 165, pp. 159–174, Aug. 2015.
- [116] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, and N. Carvalhais, “Deep learning and process understanding for data-driven Earth system science,” *Nature*, vol. 566, no. 7743, pp. 195–204, Feb. 2019.
- [117] W. P. Tsai, M. Pan, K. Lawson, J. Liu, D. Feng, and C. Shen, “From parameter calibration to parameter learning: Revolutionizing large-scale geoscientific modeling with big data,” in *AGU Fall Meeting Abstr.*, vol. 20, pp. H166–0016, Dec. 2020.
- [118] S. See and J. Adie, “Challenges and opportunities for a hybrid modelling approach to earth system science,” in *Proc. CCF THPC*, Jul. 2021, pp. 1–10.
- [119] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, “Integrating physics-based modeling with machine learning: A survey,” Jul. 2021, *arXiv:2003.04919*.
- [120] H. R. Ghafarian Malamiri, I. Roustaei, H. Olafsson, H. Zare, and H. Zhang, “Gap-filling of MODIS time series land surface temperature (LST) products using singular spectrum analysis (SSA),” *Atmosphere*, vol. 9, no. 9, Aug. 2018, Art. no. 334.
- [121] M. Metz, D. Rocchini, and M. Neteler, “Surface temperatures at the continental scale: Tracking changes with remote sensing at unprecedented detail,” *Remote Sens.*, vol. 6, no. 5, pp. 3822–3840, Apr. 2014.

**Paulina Bartkowiak** (Student Member, IEEE) received the B.Sc. degree in geodesy and cartography and the M.Sc. in geodesy and cartography from the Warsaw University of Technology, Warsaw, Poland, in 2014 and 2016, respectively. She is currently working toward the Doctoral degree in environmental science with the Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milan, Italy, and with the Institute for Earth Observation, Eurac Research, Bolzano, Italy.

Her thesis is focused on development of a new satellite-based land surface temperature product for further evapotranspiration modeling in the Alps. Her research interests include optical and thermal remote sensing with main focus on image processing and machine learning.

Ms. Bartkowiak was the recipient of the REACH Prize for Ph.D. research funded by the European Space Agency and SpaceBrains Foundation, in 2020.

**Mariapina Castelli** received the master’s degree in environmental engineering from the University of Perugia in 2009, and the Ph.D degree in environmental engineering from the University of Trento, Trento, Italy, in 2015.

She is currently a Senior Researcher with the Eurac Institute for Earth Observation. Her main research interests include developing methods to support water resources management by using satellite data in water and energy balance models, focusing on defining proper processing lines to deal with the topographic complexity and land-cover heterogeneity of mountainous regions. Her current research interests include estimating evapotranspiration based by energy balance models, studying the impact of climate change on the water cycle, and developing innovative solutions to support drought risk management for mountain agriculture based on satellite data.

**Alice Crespi** received the M.Sc and Ph.D. degrees in physics from the University of Milan, Milan, Italy, in 2015 and 2018, respectively.

She is currently a Postdoc Researcher with the Institute for Earth Observation of Eurac Research. Her research is focused on the analysis of climate and meteorological data for the assessment of the spatio-temporal variability and trends of the climate signal, with a particular attention to climate extremes, and on the provision of effective climate data and tools supporting the risk assessment in mountain regions, which includes the implementation of interpolation techniques modeling the influence of orography on the spatial distribution of meteorological variables and the definition of proper postprocessing statistical approaches to derive climate model simulations at the local scale.

**Georg Niedrist** received the master's degree in ecology from the University of Innsbruck, Innsbruck, Austria, in 2006, and the Ph.D. degree in ecology from the Colorado State University, Fort Collins, CO, USA, in 2016.

In 2006, he joined Eurac Research. Beside "standard" research activities, he is technical coordinator of the LTER site Matsch/Mazia and Codeveloper of the terraXcube, a special research facility to study climate change effects on high alpine ecosystems. His research interests include mountain agriculture, biodiversity, and climate change.

**Damiano Zanotelli** received the joint Ph.D. degree in horticultural sciences from the Universities of Bologna and Bolzano, Bolzano, Italy, in 2012.

He is currently an Assistant Professor (tenure track RTD/B) with the Faculty of Science and Technology, Free University of Bozen-Bolzano, Bolzano, Italy. His main research interests include the assessment of carbon and water fluxes in the soil-plant-atmosphere continuum of woody agroecosystems in standard and stressed conditions.

Dr. Zanotelli was the recipient of Best Jung Researcher for the category "Fruticulture" by the Italian Society of Horticultural Sciences in September 2016.

**Roberto Colombo** received the master's degree in earth science from the University of Milan, Milan, Italy, in 1991.

He wrote his thesis on the use of remotely sensed data for geomorphological mapping of the inland delta of the Niger River in Mali. Roberto Colombo is professor of Remote Sensing at the Department of Environmental Sciences, University of Milano-Bicocca. His research interest is to develop remote sensing tools for quantitative estimation of land surface properties. He works with a wide range of Earth Observation data at different scales and geophysical methods, assimilating multi-source, multi-spectral and multi-temporal remote sensing data, from field spectroscopy to satellite level for modelling terrestrial and environmental dynamics, with focus on vegetation fluorescence and snow properties. Current activities include space mission concepts and definition, airborne campaigns and field cal/val strategies, engineering and design of proximal remote sensing instruments, algorithm development, environmental modelling and new applications.

**Claudia Notarnicola** (Member, IEEE) received the M.S. (*summa cum laude*) and the Ph.D. degrees in physics from the University of Bari, Bari, Italy, in 1995 and 2002, respectively.

She is currently the Vice-Head of the Institute for Earth Observation, Eurac Research, Bolzano, Italy, where she is also the Leader of a group dealing with remote sensing applications in the synthetic aperture radar (SAR) and optical domain for soil, snow, and vegetation monitoring, as well as integration of remotely sensed observations with models and ground measurements. She conducts research within the frameworks of several national and international projects. Among others, she was also involved in the Cassini - Huygens Mission for the application of an inversion procedure to the estimation of Titan surface parameters. Her research interests include biophysical parameter (soil moisture, vegetation, and snow) retrieval using optical and SAR images, optical and SAR data processing, data fusion, and electromagnetic models.

Dr. Notarnicola has been a Conference Chair for the SPIE International Conference on Active and Passive Microwave Remote Sensing for Environmental Monitoring since 2006. She was recipient of the South Tyrol Women in Science Award in 2020.

---

**Two source energy balance model of evapotranspiration with enhanced thermal MODIS observations over Alpine region (Paper 3)**

---

The last section of the thesis describes the practical use of the developed 250-m land surface product for two source energy balance modelling in the Matsch-Mazia Valley. In this regard, the performance of TSEB forced by satellite-based LST images, including both original and enhanced MODIS data is investigated, and then compared to on-site measurements along with benchmarked simulations of turbulent fluxes derived from eddy covariance systems at different types of grasslands for the entire year of 2017. This research study aims to understand the practical utility of remotely sensed thermal data for monitoring evapotranspiration estimates at larger scales for operational applications.

---

This work was submitted to *European Journal of Remote Sensing*. According to *Scopus* the journal received Q1 category in the field of Environmental Science in 2020.

---

# Two-source energy balance modeling of evapotranspiration with enhanced thermal MODIS observations over Alpine region

**Abstract:** Many satellite-derived evapotranspiration (ET) estimates rely on coarse resolution (CR) land surface temperature (LST) obtained from 1-km MODIS TIR bands. This affects prediction performance of ET with stronger impact in complex regions, such as the European Alps. Since most two-source energy balance (TSEB) models assume no thermal variability within pixel, a major challenge in TSEB modelling is related to cell grid heterogeneity. Given this limitation, we investigated the potential of kernel-driven downscaling to obtain 250-m Terra MODIS LST product for further estimating energy fluxes from TSEB model at Matsch-Mazia site in the Italian Alps. In this context, we exploited relationships between CR MODIS LST and fine-resolution (FR) explanatory variables with random forest algorithm. Due to reduced capabilities of univariate models in complex ecosystems, multi-source predictors were considered, including downwelling solar radiation, normalized difference vegetation index (NDVI), digital elevation model (DEM). The performance of the disaggregation was evaluated against ground-derived LST from two eddy covariance (EC) towers at Matsch-Mazia site. Validation showed that the average root mean square error (RMSE) and mean bias (MB) were about 10% and 33% reduced when compared to accuracy results obtained for 1-km MODIS datasets in the year 2017. In a second step, the produced LST grids were incorporated into TSEB model for estimating 250-m energy fluxes over the study site. First, the performance of TSEB driven by ground-derived temperatures was evaluated against measured fluxes by EC systems in 2017. The benchmark simulations for latent (LE) and sensible heat (H) yielded an averaged root mean square error (RMSE) of  $57 \text{ Wm}^{-2}$  and mean absolute bias (MB) of  $26 \text{ Wm}^{-2}$ . Secondly, the model estimates forced with satellite-based LSTs as main input, including original 1-km MODIS LST product and downscaled grids were validated against site measurements. Turbulent fluxes driven by 250-m surface temperatures resulted in RMSE equal to  $86 \text{ Wm}^{-2}$  and MB of  $55 \text{ Wm}^{-2}$ , which translated to 8% and 15% decrease in the respective errors when compared to TSEB estimates with coarse resolution MODIS LST data. Despite some limitations, mainly related to small-scale changes in landcover and topography that control LSTs and consequently affect TSEB-based ET estimates, the enhanced land surface temperature has potential for providing energy fluxes at finer spatial resolution in heterogenous ecosystems.

**Keywords:** evapotranspiration, land surface temperature, TSEB, kernel-driven downscaling, random forest

## 1. Introduction

Mountain ecosystems in the European Alps have been increasingly affected by drought events in recent years (Conedera et al., 2006; Fink et al., 2004; Gobiet et al., 2014; Lewińska et al., 2018). Due to changing climatic conditions, the Alps are vulnerable to deviations in water cycling, which is mainly observed in the context of natural vegetation and agricultural activities. One of the key processes of water cycle determined by interactions between atmosphere and Earth's surface is evapotranspiration (ET). ET is a crucial component for monitoring water loss from land surface, being the main driver of physical and biological processes for vegetation. An accurate estimation of terrestrial evapotranspiration is a key to help governance institutions and agricultural communities for improving sustainable agricultural production, drought forecasts, and also land degradation monitoring (Gao et al., 2014; Kundu et al., 2018; Park et al., 2016; Wang et al., 2021).

Precise determination of ET is possible with the use specific instruments, like lysimeters and eddy-covariance (EC) technique (Pastorello et al., 2020; Widmoser & Wohlfahrt, 2018). Moreover, due to high performance of EC systems, this method provides reference dataset for validating energy balance fluxes from other resources (Bastiaanssen et al., 1998b). However, the point-based nature of such data type and expensive maintenance of EC towers limit their practical application to field-oriented studies with relatively homogenous landscapes.

Development of evapotranspiration models driven by remote sensing data has made a significant contribution to assess land surface processes at larger scales. In particular, ET estimation using surface energy balance (SEB) modeling driven by satellite land surface temperature (LST) has been successfully applied in different fields (Bhattarai et al., 2017; French et al., 2015; Merlin et al., 2014; Senay et al., 2016). By contrast to other remotely sensed methods (Glenn et al., 2010; Nagler et al., 2009), only SEB models can minimize the need for meteorological input, since they do not require spatially distributed precipitation and soil moisture information, which is advantageous in highly heterogenous terrain of the Alps with sparsely distributed measurement stations (Castelli et al., 2018). The SEB models could be aggregated into one-source energy balance (OSEB) models and two-source energy balance (TSEB) models. While the methods of the first category, like SEBAL (Bastiaanssen

et al., 1998a), S-SEBI (Roerink et al., 2000), SEBS (Su et al., 2002), and METRIC (Allen et al., 2007), simplify the land surface to homogenous canopy layer by assuming a single observation of LST and aerodynamic resistance for the soil and vegetation; two-source models estimate fluxes by partitioning energy balance into separate soil and canopy sub-components using TSEB method (Choudhury et al., 1987; Norman et al., 1995) or ALEXI/DisALEXI approach (Anderson et al., 2011). In general, TSEB models exhibit larger potential for estimating turbulent fluxes over complex and non-uniform vegetation in mountain regions, as shown in other studies (Castelli et al., 2018; Elfarkh et al., 2020). In addition, they do not require more input parameters than those needed for OSEB models, including air temperature, wind speed and global shortwave radiation, being acquired by meteorological stations and reanalysis observations at regular basis (Hersbach et al., 2020; Tang et al., 2013). Furthermore, as reported by Timmermans et al. (2007) and later on by Yang et al. (2015) TSEB outputs were in the closest agreement with ground data from EC systems.

Land surface temperature observations from sun-synchronous NASA's MODIS (Moderate resolution Imaging Spectroradiometer) instrument have been applied in many studies to estimate evapotranspiration from TSEB models (Cammalleri et al., 2014; Guzinski et al., 2013; Zhang et al., 2021). This may be explained by its high temporal resolution (sub-daily) and long-term image collection (since 2000), which makes MODIS an appropriate sensor for routine tasks when compared to thermal infrared (TIR) imagery with fine spatial resolution (e.g., Landsat-8 TIRS with 100-m pixel size) with low temporal resolution of more than 15 days causing difficulties in monitoring environmental phenomena on more frequent basis (Wan & Dozier, 1996). The application of sub-daily MODIS data indeed improves the quality of continuous monitoring of ecosystems. However, its coarse spatial resolution may not be reliable over complex landscapes (Castelli et al., 2018; Guzinski & Nieto, 2019; Sharma et al., 2016). For example, Castelli et al. (2018) investigated TSEB performance with MODIS LST input and observed RMSEs of 100-172  $\text{Wm}^{-2}$  between model outputs (ALEXI/DisALEXI) driven by 1-km MODIS data and turbulent fluxes obtained from in-situ measurements. The high variability within 1-km MODIS LST pixels can be one main factor influencing TSEB performance. Until the planned launch of ESA's high spatial-temporal resolution TIR sensor, the next-future perspectives for remotely sensed TSEB estimates over Europe are constrained to coarse resolution LST data (Koetz et al., 2019).

Many different techniques have been proposed to improve spatial resolution of LST data for better understanding and modelling energy fluxes over heterogeneous ecosystems (Agam et al., 2007; Cammalleri et al., 2014; Gao et al., 2012; Liu et al., 2020; Mahour et al. 2017). Majority of these methods rely on establishing relationship between coarse resolution thermal data and fine-resolution (FR) explanatory variables degraded to coarse pixel size of TIR sensor, and then, applying the resulting geostatistical models to FR predictors for retrieval LST maps at enhanced spatial resolution. Although image fusion methods successfully improve temporal availability of high-resolution LST images, like Landsat TIRS and ISS ECOSTRESS (Sun et al., 2017; Yang et al., 2017) by exploiting (sub)daily observations from coarse resolution TIR scanners, the low temporal resolution of FR satellites and overcast conditions hamper their practical use for ET monitoring that requires regular and relatively frequent LST observations. In this context, enhancement method in spatial domain can be considered as one of the most commonly used approaches for thermal downscaling (Gao et al., 2016; Guzinski et al., 2020; Li et al., 2019; Liu et al., 2020; Maeda, 2014; Xia et al., 2019). The disaggregated datasets can be further incorporated into surface energy balance models for estimating evapotranspiration at the predictor-like spatial resolution.

To date, 1 km MODIS TIR imagery served as primary input for energy heat flux modelling in the Alps (Castelli, 2018). In this context, the main objective of this work was to investigate the effect of downscaled 250-m MODIS LST data on the performance of the Priestley-Taylor two-source energy balance (TSEB-PT) model over alpine grasslands. For this purpose, two different remotely sensed datasets as main TSEB inputs were explored: original daytime Terra MODIS LST and 250-m MODIS-based LST product. Next, satellite-based TSEB simulations were validated against ground data from two EC towers at Matsch-Mazia site, and next the potential impact of spatial heterogeneity within LST pixels was examined. This study provides valuable insights into understanding the performance of TSEB modeling with different LST inputs and gives future directions for input data requirements over highly complex regions such as the Alps. To the best authors' knowledge, the feasibility analysis of kernel-driven downscaling to force TSEB-PT model has never been performed in alpine ecosystems before.

## 2. Materials

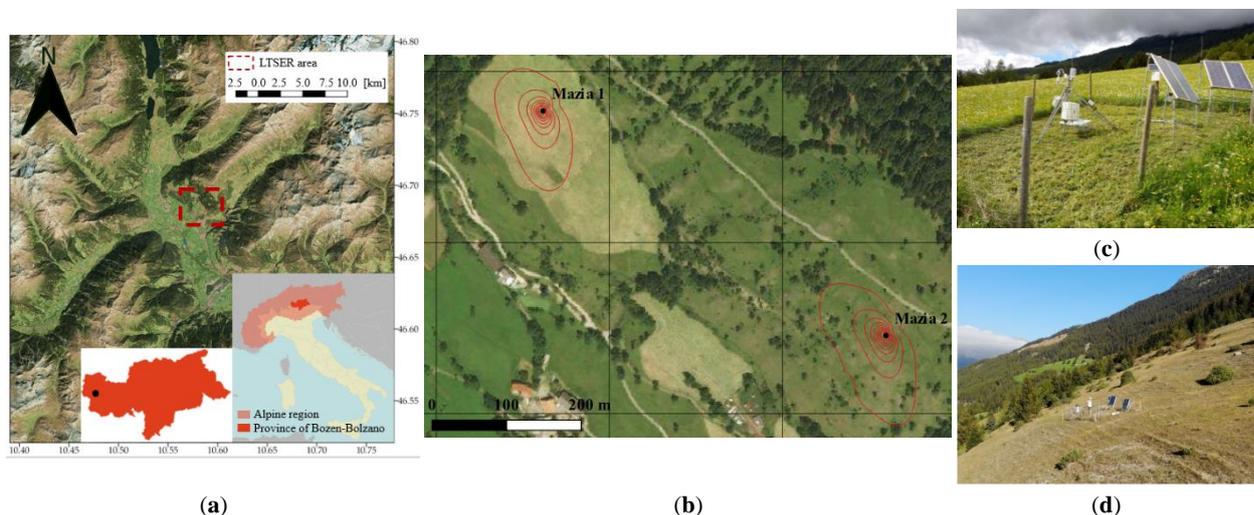
### 2.1. Study site

In this study, we focus on a grassland area of Matsch-Mazia Valley in the European Alps in Italy (Fig. 1a). Grassland ecosystems are one of the main landcovers in the world occupying nearly one-third of the global land surface (Lemaire et al., 2011). In the mountain regions they are recognized for their agricultural use, but also for their positive impacts on ecological aspects like high biodiversity of fauna and flora, carbon storage, and climate mitigation (Bengtsson et al., 2019). Thus, it is of high priority to monitor these ecosystems by providing good quality biophysical parameters that capture trends in spatio-temporal domain. Matsch-Mazia Valley is well-established study area of over 90 km<sup>2</sup> located at about 1600 m a.s.l., where diversified types of mountainous grasslands occur, including yield-oriented meadows and natural pasturelands. The region is a part of the Long-Term Socio-Ecological Research (LTSER) experiment (<http://lter.eurac.edu>) due to its unique Alpine landscape and untypical dry microclimate with a yearly average precipitation of nearly 525 mm in the period between 1920 and 2010 (Castelli et al., 2018). Due to harsh climatic conditions, this region is vulnerable to changes in water cycling, which can be mainly observed in the context of agricultural activities. Therefore, understanding evapotranspiration component as driving indicator of biophysical processes for vegetation is very essential for monitoring water availability in plants, but also for forecasting agricultural drought.

### 2.2. Data

#### 2.2.1. Ground measurements

Biometeorological measurements and eddy covariance (EC) data across the LTSER area were used in this study. Although the two EC towers are located in the close vicinity of each other (ca. 500 m), they comprise two grassland types: irrigated meadow (Mazia 1) and rainfed pasture (Mazia 2). Mazia 1 is homogeneously covered with seasonally mowed grass (usually two times per year), while the area of Mazia 2 belongs to natural pasture exposed to livestock and wildlife grazing as shown in Fig. 1b-d. Both sites are situated on the southwest oriented slopes.



**Fig. 1.** Overview of the study site: (a) General geographic location of the experimental area, (b) Locations of the EC towers (black points) in Matsch-Mazia Valley with corresponding isolines (in red) of computed 2-D flux footprints (Kljun et al., 2015) ranging from 10% to 90% of total contributing area. In the background black squares indicate spatial range of 250-m MODIS pixels. (c) and (d) EC towers situated in managed meadow (Mazia 1), and natural pasture (Mazia 2), respectively. Fig. 1a-b were created in QGIS using internal resources and open-source base layers from <http://kosmosnimki.ru> (Fig. 1a) and <https://www.bing.com/maps> (Fig. 1b). The pictures of the EC sites (Fig. 1c-d) were derived from the LTSER webpage (<https://browser.lter.eurac.edu>).

Two EC systems at meadow and pasture were equipped with identical instrumentation and methods to process raw in-situ data using the REdyProc R Package (Reichstein et al., 2014). They consisted of a three-dimensional sonic anemometer (CSAT3, Campbell Scientific) for measuring the three components of wind speed, and infrared

gas analyzer (LI-7500, LI-COR) to derive water vapor at a height of 2 m above ground surface. A four-component net radiometer (NR01, Hukseflux; CNR-1, Kipp and Zonen) served to register both downwelling and upwelling longwave ( $LW_{in}$ ,  $LW_{out}$ ) and shortwave radiation ( $SW_{in}$ ,  $SW_{out}$ ). In addition, each site had a soil thermocouple probe for soil temperature estimation (TCAV, Campbell Scientific), soil heat flux (G) plates located at a depth of 7.5 cm below the ground surface (Hukseflux), and standard instruments for registering air temperature (TA) and relative humidity (RH) (HMP45C, Campbell Scientific), precipitation (Young), and two-axis wind velocity and its direction (Windsonic, Gill). EC data of net radiation (Rn), latent heat flux (LE), sensible heat flux (H), and G were utilized for benchmark TSEB-PT simulations. Considering the inclined terrain of the study sites horizontally-measured observations of Rn were corrected for topography as described extensively in Wohlfahrt et al. (2016).

Even though TSEB is assumed to theoretically close the energy balance (see Section 3.1), in mountain regions with complex topography it is of higher risk to violate this assumption when compared to plain and homogenous terrain. The EC-related lack of closure and concepts for its correction are challenging aspects of the SEB modelling affecting its performance and uncertainty. The problem of energy imbalance is extensively discussed in the scientific community (Foken et al., 2011; Stoy et al., 2013; Wohlfahrt & Widmoser, 2013). To this end, the measured components were checked for the energy imbalance. First, we excluded all observations when absolute closure (i.e.,  $Rn - G - LE - H$ ) was greater than  $300 \text{ W m}^{-2}$  along with relative closure (i.e.,  $[Rn - G - LE - H]/Rn$ ) exceeding 0.4. In addition, the measurements were checked for energy balance closure ratio (i.e.,  $[H + LE]/[Rn - G]$ ) with minimum acceptable threshold of 0.7. Apart from those criteria, tower-derived turbulent fluxes needed to be forced with in-situ EC measurements. In this context, we explored residual- (*Residual*) and Bowen ratio-based (*Br*) approaches as shown in other studies (Burchard-Levine et al., 2021; Castelli et al., 2018; Guzinski et al., 2014; Twine et al., 2000). The applied corrections to turbulent fluxes are expressed as follows:

$$LE_{Residual} = Rn - G - H \quad (1)$$

$$H_{Br} = H + \Delta H \quad (2)$$

$$LE_{Br} = LE + \Delta LE \quad (3)$$

$$\Delta H = \frac{Br(LE + H) - H}{1 + Br} \quad (4)$$

$$\Delta LE = Rn - G - LE - H - \Delta H \quad (5)$$

where *Br* is the Bowen ratio between observed fluxes of H and LE (i.e.,  $Br = H/LE$ ). While the first method (referred as *Residual*) assigns the residuals to the in-situ latent heat flux records assuming retrieval errors for LE bigger than for H, the Bowen ratio closure (hereafter called *Br*) partitions the correction between the sensible and latent heat fluxes assuming that the ratio between observed H and LE is correctly registered by eddy covariance systems.

EC and micrometeorological observations, including H, TA, atmospheric pressure, friction velocity, and wind direction were utilized to compute flux footprint area for each tower. In this study, we applied a two-dimensional flux footprint model proposed by Kljun et al. (2015). The choice of the model was motivated by its simplicity and open-source function codes that are available online through <http://footprint.kljun.net>. The stationarity over the EC sites and related horizontal homogeneity of turbulence was set at monthly basis. The calculated footprint range was used for both tower-based benchmark simulations along with TSEB modelling forced by satellite-based LST inputs. Specifically, we used the footprint-weighted method to determine partial contributions of estimated turbulent fluxes from remotely sensed TSEB modelling at the moment of satellite overpass.

On-site measurements of Rn, LE, H and G served to investigate the performance of TSEB, while latent and sensible heat fluxes were used to validate the model driven by satellite-based land surface temperature. For TSEB benchmarking the observed soil heat flux was directly incorporated as input in order to access the model capabilities and minimize uncertainty of H and LE estimates related to G (Burchard-Levine et al., 2020). Considering lack of gridded data of G, during model runs forced by satellite-based LSTs the soil heat flux was estimated using the approach of Santanello & Friedl (2003). In addition, we exploited LW (both incoming and outgoing) to retrieve ground-derived LST for further validation of the surface temperatures from remote sensing observations. LST was derived from the Stefan-Boltzmann law:

$$LST = \left( \frac{LW_{out} - (1 - \varepsilon)LW_{in}}{\varepsilon\sigma} \right)^{0.25} \quad (6)$$

where  $LW_{in}$  is the downwelling longwave radiation;  $LW_{out}$  represents upwelling longwave radiation;  $\varepsilon$  denotes the surface broadband emissivity, and  $\sigma$  is the Stefan-Boltzmann constant ( $5.67 \times 10^{-8} \text{ W m}^{-2}\text{K}^{-4}$ ). The  $\varepsilon$  was derived from MODIS narrowband emissivities (i.e., M\*D21A1D bands 29, 31, and 32 (Hulley et al., 2016)) by exploiting empirical relationships given in Wang et al. (2005).

Meteorological and flux data were collected at 30-min time interval for the entire year of 2017 between April and October considering phenological cycle of vegetation at higher altitudes to capture periods with green canopy cover. Climate forcing and LST were derived at that temporal resolution, while biophysical properties of vegetation, like leaf area index (LAI), green ( $f_g$ ) and fractional vegetation cover ( $f_c$ ) due to lack of ground observations were estimated from remote sensing data as described in Section 2.2.2. In addition, instrumental set-up in Mazia 1 enabled to retrieve direct measurements of canopy height ( $h_c$ ), whereas at pasture site  $h_c$  was obtained using satellite observations according to the strategy of Anderson et al. (2007) using the following equation:

$$h_c = h_{min} + f_c(h_{max} - h_{min}) \quad (7)$$

where  $f_c$  indicates the fraction of canopy cover obtained from satellite data, while  $h_{min}$  and  $h_{max}$  represent observed minimum and maximum grass height during phenological vegetation cycle.

### 2.2.2. Satellite data

In this study, two Moderate Resolution Imaging Spectroradiometer (MODIS) datasets were collected for the entire year of 2017, including daytime LST (MOD11A1 Collection 6) and land surface emissivity (LSE) product (M\*D21A1D Collection 6). Land surface temperature maps with good quality assurance (QA) bits were used to build thermal downscaling regressions, while the LSE served to evaluate the performance of the disaggregated LSTs as explained in Section 3.2 in more detail. MOD11A1 has been widely used in many studies with a sufficient accuracy of 1 K over most landscapes (Wan & Dozier, 1996). Nonetheless, in this study, we applied physical-based M\*D21A1D data rather than simple classification-retrieved surface emissivity offered by M\*D11A1 (Nickeson, 2020).

Note that M\*D21A1D product was also applied for benchmarking TSEB simulations with ground-derived LST due to lack of in-situ emissivity measurements. Furthermore, original MODIS LST images served to compare downscaling results, and then they were used to model 1-km TSEB fluxes as an intercomparison with 250-m estimates.

Elevation was obtained from NASA's Global Digital Elevation Model (GDEM) acquired by ASTER sensor at 30 m spatial resolution. In addition, three ancillary DEM-derived products, including aspect, slope and heat load index were utilized to explain thermal variability within 250-m LST pixel (McCune, 2007).

To assess the impact of landcover on LST, the Normalized Difference Vegetation Index (NDVI) grids were obtained from 10-day PROBA-V NDVI composites at 300 m spatial resolution. Due to temporal differences between MOD11A1 and PROBA maps we assumed decadal consistency of vegetation content with respect to daily acquisitions of LST images. This means that NDVI granules with time-coincident observations to MODIS overpass were used for further downscaling. Along with canopy information, 100-m classification map from Copernicus CORINE land use land cover (LULC) served to derive areas covered by vegetation. Non-vegetated areas, like bare rocks, glaciers and urban areas were excluded from the analysis in order to minimize thermal variations of the land surface and at the same time provide stability of random forest model that is prone to high spatial heterogeneities and rare observations (e.g., extremely high and low temperatures).

Additionally, the sharpening models were forced with 250-m incoming solar radiation granules that were obtained by means of geostatistical downscaling applied to daily DSSF product acquired by MSG/SEVIRI instrument (<https://landsaf.ipma.pt>). On average, root mean square error (RMSE) was equal to  $2.64 \text{ MJ m}^{-2}\text{day}^{-1}$  when compared to in-situ measurements over the Alpine region (Bartkowiak et al., 2022).

In order to keep spatial consistency with original and downscaled output all gridded data were resampled to 1 km and 250-m pixel size using nearest neighbor resampling method accordingly for further establishing LST

regressions. The original MODIS data were obtained in MODIS Sinusoidal coordinate system (SR-ORG: 6974) that served as the main projection during the entire thermal downscaling process.

Owing to lack of ground-based observations of vegetation properties, TSEB model was forced by remote sensing-based inputs. Specifically, fractional vegetation cover ( $f_C$ ), green fraction of vegetation ( $f_g$ ) and leaf area index (LAI) were derived from Sentinel-2 MSI imagery (both 2A and 2B) at 20 m spatial resolution. These data were produced by Eurac research within the EO institute project ET\_int (<https://www.eurac.edu/en/institutes-centers/institute-for-earth-observation/projects/etint>) using S2ToolBox developed by Weiss and Baret (2016). For the simulations based on ground data, the variables were derived within the tower flux footprint area (see Section 2.2.1), while for satellite-based simulations,  $f_g$ ,  $f_C$ , and LAI were extracted in correspondence with LST pixels over the monthly footprints. Grids were inspected for the presence of clouds for extracting good quality pixels. Due to less frequent Sentinel-2 observations with an average 5-day repeat cycle, 20-m grid cells were linearly interpolated considering maximum temporal gap between observations of 14 days to derive spatial information at MODIS-like temporal resolution. Finally, the weighted average of Sentinel-2 based variables was calculated according to the area share within the 2-D footprint. TSEB simulations were performed only when input data covered at least 70% of the footprint area. Meanwhile, for flux predictions driven by satellite-based LSTs, daily  $f_C$ ,  $f_g$  and LAI estimates were averaged within LST pixel.

To give more general overview of the gridded datasets used in this study, Table 1 lists detailed information on the remotely sensed observations utilized for both the Priestley-Taylor two-source energy balance model and kernel-driven downscaling.

**Table 1.** An overview of remote sensing data used in this study.

Source dataset	Parameter	Pixel size	Short summary & derivative output
MOD11A1 <sup>1</sup>	Land surface temperature [K]	1 km	Terra MODIS LST (Collection 6) maps at clear-sky conditions based on quality assurance (QA) MODIS flags defined in Wan et al. (2015)
MOD21A1D <sup>1</sup> MYD21A1D <sup>1</sup>	Surface Emissivity bands: 29,31,32 [-]	1 km	Daily maps of narrowband surface emissivity derived from M <sup>*</sup> D21A1D (Collection 6) granules to obtain broadband surface emissivity according to Wang et al. (2005)
MSG/SEVIRI DSSF	Daily incoming shortwave radiation maps [MJ m <sup>-2</sup> day <sup>-1</sup> ]	250 m	Downscaled downwelling surface shortwave flux from original DSSF product ( <a href="https://landsaf.ipma.pt">https://landsaf.ipma.pt</a> ) derived through RK interpolation
PROBA-V	Normalized Difference Vegetation Index [-]	300 m	10-day NDVI composites accessed via Copernicus Global Land Service ( <a href="https://land.copernicus.eu">https://land.copernicus.eu</a> ). The grids were resampled to 250-m and 1km spatial resolution using nearest neighbor method
ASTER GDEM	Digital elevation model [m]	30 m	30-m elevation product from ASTER and its three derivative products: aspect, slope, heat load index were resampled to 1 km and 250 m spatial resolution

<sup>1</sup>data utilized for both thermal downscaling procedure and two-source energy balance modelling

### 3. Methods

#### 3.1. TSEB model

In this study, the Priestley-Taylor two-source energy balance model (TSEB-PT), proposed by Norman et al. (1995) and next enhanced by Kustas & Norman (1999), was applied at stations in the Matsch-Mazia Valley. In the TSEB-PT model The method treats the soil and canopy as two independent components and estimates latent and sensible heat fluxes for both of them using two-layer approach as follows:

$$Rn_S = LE_S + H_S + G \quad (8)$$

$$Rn_C = LE_C + H_C \quad (9)$$

where  $Rn$  is the net radiation ( $Wm^{-2}$ ),  $LE$  is the latent heat flux ( $Wm^{-2}$ ),  $H$  represents sensible heat flux ( $Wm^{-2}$ ),  $G$  is soil heat flux ( $Wm^{-2}$ ), and subscripts  $S$  and  $C$  denote soil and canopy components. The model assumes no impact

of other energy components, like photosynthesis and canopy heat storage. The net radiation sub-components are computed based on Campbell and Norman (1998), while  $H_C$  ( $H_S$ ) are derived from the gradient between canopy (soil) and air temperature (TA) at a reference height (Guzinski et al., 2020). The main remote sensing-based variables required by TSEB are land surface temperature (LST) that represents total contribution of soil and canopy, and fractional vegetation cover ( $f_C$ ) for partitioning the energy between canopy and soil. Since surface temperatures of soil ( $LST_S$ ) and vegetation ( $LST_C$ ) are unknown, the model splits LST into soil and canopy temperatures according to fractional vegetation content as a function of leaf area index (LAI), as given in Equation (10):

$$LST = (f_C LST_C^4 + (1 - f_C) LST_S^4)^{0.25} \quad (10)$$

TSEB-PT performs iterative procedure to compute  $LST_S$  and  $LST_C$  together with their corresponding soil and canopy sensible heat fluxes ( $H_S$ ,  $H_C$ ). The whole process to find  $LST_S$ ,  $H_S$ ,  $LST_C$ , and  $H_C$  starts with initial estimate of canopy transpiration ( $LE_C^{int}$ ) derived from Priestley-Taylor equation:

$$LE_C^{int} = \alpha_{PT} f_g R n_c \frac{\Delta}{\Delta + \gamma} \quad (11)$$

where  $\alpha_{PT}$  (-) denotes the Priestley-Taylor coefficient (Priestley & Taylor, 1972),  $f_g$  is the green fraction of vegetation (-),  $\Delta$  is the slope of the vapor pressure versus air temperature ( $\text{kPa K}^{-1}$ ), and  $\gamma$  is the psychrometric constant ( $\text{kPa K}^{-1}$ ). Next, with the first estimate of  $LE_C$ , sensible heat flux from vegetation is computed as a residual term of energy balance (Equation 9), and then  $LST_C$  is derived from the following formula:

$$LST_C = \frac{H_C r_C}{C_p \rho} + TA \quad (12)$$

where  $C_p$  is the air heat capacity ( $\text{J kg}^{-1} \text{K}^{-1}$ ),  $\rho$  indicates the air density ( $\text{kg m}^{-3}$ ),  $r_C$  is the aerodynamic resistance for heat transfer ( $\text{s m}^{-1}$ ), and TA is the air temperature at the measurement height (K).  $LST_S$  is obtained from Equation (10), and then soil sensible heat flux is derived as follows:

$$H_S = C_p \rho \frac{LST_S - TA}{r_s} \quad (13)$$

where  $r_s$  is the resistance for heat transfer in the boundary layer above soil ( $\text{s m}^{-1}$ ). Consequently, the latent heat flux from the soil is obtained as residual flux based on Equation (8), while providing the energy balance closure. Resistance term was based on the formulation proposed by Kustas and Norman (1999). The above described process (Equations 8-13) stops when the soil latent heat reaches non-negative value. Otherwise, the  $LE_C^{int}$  is iteratively modified by decreasing  $\alpha_{PT}$  (with an initial value of 1.26) until a physically realistic solution is found (Norman et al., 1995; Kustas & Norman, 1999).

The TSEB-PT model was applied in this work for estimating energy fluxes with ground-derived thermal data along with satellite-based LST inputs, including original Terra MODIS LST at 1 km resolution, downscaled MODIS LST at 250-m resolution, and high resolution observations from Landsat TIRS sensor.

In this study, we used open-source implementation of TSEB (pyTSEB) written in Python programming language. The package contains sub-components to calculate needed inputs for the model, like net radiation, Monin-Obukhov length, and the resistances for heat and momentum transport. More details on the source codes of pyTSEB is available online (<https://github.com/hectornieto/pyTSEB>).

### 3.2. Thermal downscaling model

In this study, thermal disaggregation is based on the kernel-driven modeling approach successfully used by many researchers to explain LST variability at finer spatial resolution (Dong et al., 2020). Considering the spatial heterogeneities and meteorological complexity of the study area, the method is a trade-off between availability of remote sensing predictors at fine spatial resolution (FR) and relatively frequent revisit time for ET monitoring at more regular basis. In this study, we propose a refined approach of Bartkowiak et al., (2019) by applying ancillary explanatory variables in the sharpening scheme (see Table 1). The whole downscaling process can be summarized with one general formula:

$$LST_{FR} = LST_{fR} + (LST_{CR} - LST_{fR}^{CR})^{FR} \quad (14)$$

where  $LST_{FR}$  denotes final downscaled LST data with residual correction applied,  $LST_{fR}$  is the indirect LST output after applying coarse resolution (CR) thermal model to the FR explanatory variables, and the term of  $(LST_{CR} - LST_{fR}^{CR})^{FR}$  indicates residual correction between coarse resolution surface temperatures and their corresponding upscaled fine resolution predictions resampled to spatial resolution of the kernels applied. As shown in previous studies (Bartkowiak et al., 2022; Bertoldi et al., 2010; He et al., 2019), driving forces of LST distribution in mountain regions may be related to many factors, including topography, vegetation content and climate forcing. To this end, selection of representative 250-m kernels was based on “the process-guided design” as an attempt to better explain LST variability in complex ecosystems (Mao et al., 2021). In this study, FR-LST were derived by exploiting relationship between 1-km clear-sky MOD11A1 and three main biophysical explanatory variables: NDVI, daily solar radiation, and DEM together with its derivative products such as aspect, slope and heat load index (Table 1).

The function relationship between 1-km MODIS LST data and predictors was obtained with random forest (RF) algorithm (Bartkowiak et al., 2019; Hutengs & Vohland, 2016; Njuki et al., 2020; Xu et al., 2020). That choice is reasoned by its non-parametric nature along with ability to capture non-linear relationships between Earth’s skin temperature and explanatory variables, that is especially important in complex landscapes (Zhao et al., 2019). Random forest belongs to statistical ensemble regression algorithm, which makes it less sensitive to multicollinearity and overfitting. RF generates multiple decision trees through bagging in order to create training subsets together with random feature selection for further splitting. In addition, the algorithm provides model performance error obtained during the training process. It is done through random sampling with replacement where about two-third of all observations (referred as *in-bag* samples) for each training set is used for creating individual decision tree to estimate an output (Breiman, 1996). For regression tasks final prediction is based on the average of the randomly generated ensemble of trees. In parallel, the remaining 33% of the records unseen by individual tree (called *out-of-bag* or *OBB* samples) serves to test the regression performance using model’s internal cross-validation (CV) approach (Santos et al., 2019; Zhao et al., 2019).

Although the RF approach reduces the risk of overfitting, final estimations from the ensemble decision trees are averaged, which introduces thermal biases for rare observations (e.g., significantly colder or warmer pixels when compared to majority values) towards dominant skin temperatures in the image. In particular, this is problematic over complex areas with spatial heterogeneities in conjunction with a global downscaling approach where all 1-km LST cells are used to build a model (Jeganathan et al., 2011). Hence, our study area was filtered with respect to landcover types considering MODIS pixels with minimum 75% green vegetation cover. In this study, we focus on vegetated areas, and thus, other classes were excluded from downscaling process. To minimize the effect of extreme LST values on random forest performance, pixels that contained any fraction of water bodies, glaciers, and bare rocks were masked. In addition, the uncertainties related to the scale effect, which results in differences in spatial resolution between original and downscaled images, were considered by applying a weighting factor for each CR-LST observation during the model learning process. In this regard, mean coefficient of variation (*cv*) was calculated within each 1-km pixel from medium-resolution predictors following the strategy of Kustas et al. (2003) and Gao et al. (2012). Next, *cvs* were inverted to derive final weights for the RF regressions. This means that heterogenous pixel with larger value of *cv* results in smaller weight, while for more homogenous cells bigger significance was assigned.

For purposes of this study, we selected “caret” package with in-built “ranger” functions developed by Wright and Ziegler (2015) available in R (Team R. Core, 2013) rather than commonly used “randomForest” implementation that is less efficient in handling big data (Liaw and Wiener, 2002). Moreover, the “ranger” contains weighting argument that allows assigning weights to individual observations according to user’s preferences.

### 3.3. Model evaluation

Quality assessment of TSEB-PT model described in Section 3.1 was based on reference data from EC systems. In this regard, benchmarked fluxes together with outputs driven by thermal MODIS observations (at both 1 km and 250 m spatial resolutions) were compared with ground measurements collected in 2017. In this study, the discrepancies between modelled and observed fluxes were evaluated by root mean square error (RMSE), coefficient of determination ( $R^2$ ), and mean bias (MB) using following formulas:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}; R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - y_{mean})^2}; MB = \frac{\sum_{i=1}^n (x_i - y_i)}{n} \quad (15)$$

where  $x_i$  denotes the measured value  $i$ ,  $y_i$  is the predicted value  $i$  of  $x_i$ , and  $n$  indicates number of observations.

As mentioned in Section 3.2, the random forest (RF) algorithm provides cross-validation statistics using OOB samples to evaluate model performance. Nevertheless, in order to strengthen validation process during training phase we applied user-controlled 5-fold CV strategy. To this end, training data for each MODIS acquisition were randomly partitioned into training and test sets with respect to LST distribution over image keeping 9/10 (90% of all data) and 1/10 ratio (10% of all data), respectively. This approach gives more objective overview on model performance, since in that way we evaluate model's ability to predict LST values within their full range. The robustness of daily RF regressions was evaluated by applying test samples to predict pixels unseen by model within whole training process. Average values of RMSE and  $R^2$  along with their corresponding standard deviations (SD) were calculated. Selection of sharpening models was based on coefficients of determination derived during the 5-fold CV and test phase. In case of low  $R^2$  values ( $R^2 \leq 0.4$ ) for both cross-validated simulations and predicted test data, RF models were excluded from further downscaling process.

Original MODIS LST and the downscaled data were validated against ground measurements from the EC systems. Similar to TSEB evaluation, the differences were determined by means of RMSE,  $R^2$ , and MB.

## 4. Results and Discussion

### 4.1. Thermal downscaling estimates

The performance of the daily RF downscaling regressions was evaluated in the 5-fold cross-validation approach, and then compared with test data unseen by models for the entire year of 2017. In general, RMSEs ( $R^2$  values) had similar accuracies for both test and training datasets yielding average scores of 2.12 K (0.60) and 2.07 K (0.61), respectively. Moreover, during training phase, we obtained considerably small differences between training folds with standard deviation of RMSE ( $SD_{RMSE}$ ) equal to 0.11 K and  $SD_{R^2}$  of 0.03. This confirms stability of the linking models to explain LST variability from cross-validated samples.

To evaluate the effectiveness of the implemented downscaling, we performed a quantitative analysis by comparing disaggregated LSTs with their time-coincident site observations. Due to observed jumps between neighboring 30-min records of longwave radiation, ground-derived LST data were linearly interpolated considering daytime Terra MODIS overpass time. Validation outcomes are shown in Table 2.

**Table 2.** Accuracy statistics obtained during validation with ground data. Note that values in brackets indicate results with good quality assurance flags (QA), while records without parentheses represent accuracy errors with the highest quality QA extracted (Wan et al., 2015).

Site	Pixel size	RMSE	$R^2$	MB
Mazia 1	1 km	2.87 (3.45)	0.65 (0.61)	-0.85 (0.67)
	250 m	2.78 (3.56)	0.66 (0.61)	-0.48 (1.06)
Mazia 2	1 km	5.04 (6.32)	0.79 (0.64)	4.22 (5.20)
	250 m	4.19 (5.35)	0.82 (0.69)	3.34 (4.17)

RMSE = root mean square error,  $R^2$  = coefficient of determination, MB = mean bias.

As presented in Table 2, better results were obtained with downscaled product rather than with 1-km MODIS LST data. On average, when including the highest quality QA flags, sharpening process improved RMSE (MB) by 10% (32%) translating into the mean absolute percentage difference of 11.5% when compared to ground observations. Larger errors appeared over pasture site (Mazia 2) yielding an overall RMSE of 5.04 K and 4.19 K

within CR and FR pixels, respectively. In contrast to low bias in Mazia 1 (MB = -0.48 K), for Mazia 2 we observed strong underestimation of 250-m LST exceeding 3.3 K, whilst having high correlation ( $R^2 = 0.82$ ) obtained (Table 2).

Such discrepancies as shown in Table 2, might have appeared due to several reasons. The first is related to terrain complexity of the study area that translates to spatial heterogeneities within both 1-km and 250-m LST pixels. High errors, in particular at the pasture site, may be explained by fragmented landscape with patched land-covers, including grazed grass, and sparsely distributed trees and alpine bushes (Fig. 1). In addition, as reported by Mildrexler (2011), grasslands are strongly influenced by complex interactions between surface energy balances and the respective landcover at field scales. The results are in line with findings of Bartkowiak et al. (2019) where local intercomparison of downscaled data with Landsat LSTs gave less satisfactory accuracy scores over grasslands and bushes yielding an average RMSE of about 3 K and 4 K, respectively.

On the other hand, the RF downscaling depends on thermal variability of 1-km MODIS LST data, which limits random forest to make predictions within that range (Meyer & Pebesma, 2021; Zhang et al., 2019). In addition, since residual correction applied to FR-LST predictions is based on their divergence from the original CR imagery, residuals do not reflect spatial heterogeneities within 250-m pixels. This could imply differences between in-situ and sharpened LST estimates. Furthermore, according to QA flags of the MOD11A1 product, the source maps are additionally affected by LST retrieval errors having a negative impact on downscaling model (Wan et al., 2015). Apart from that, spatial- and time-varying characteristics of land surface over the study area may not have been captured by biophysical explanatory variables such as 10-day NDVI and daily solar radiation products that do not reflect temporal variability of the land corresponding to LST models established at the time of MODIS overpass (Table 1). Those elements had negative impacts on land surface downscaling process over the study area.

#### 4.2. Tower-based TSEB-PT simulations

The TSEB output driven by ground data consists in half-hourly estimates of energy fluxes for soil and canopy components between sunrise and sunset time at irrigated meadow (Mazia 1) and natural pasture (Mazia 2). As mentioned previously in Section 3.3, the performance of TSEB-PT with on-site LST input was evaluated by means of average accuracy measures. Table 3 presents the mean errors obtained from performed simulations for all Mazia sites.

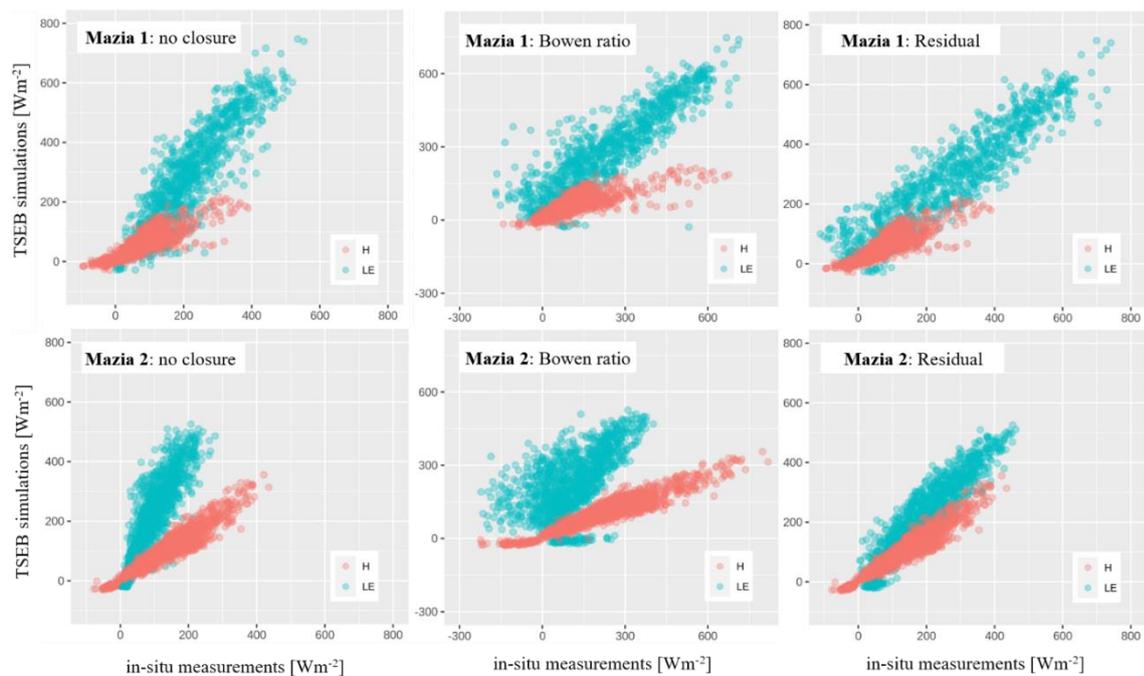
**Table 3.** Accuracy statistics for the benchmarked TSEB-PT simulations at irrigated meadow (Mazia 1) and rainfed pasture (Mazia 2). While values in parentheses indicate accuracy scores with all three quality checks included (i.e., energy balance closure ratio, and relative and absolute closure), remaining statistics were obtained after applying thresholds for relative and absolute closures (see Section 2.2.1). The model was run for the entire year of 2017 within months with green vegetation cover (April-October).

Site	Correction		Rn	LE	H	G	
Mazia 1	No closure	RMSE	60 (52)	110 (100)	49 (52)	52 (55)	
		R <sup>2</sup>	0.97 (0.97)	0.80 (0.81)	0.70 (0.71)	0.00 (0.00)	
		MB	33 (27)	-71 (-62)	23 (25)	29 (35)	
	Br	RMSE	-	99 (104)	100 (104)	-	
		R <sup>2</sup>	-	0.79 (0.80)	0.62 (0.63)	-	
		MB	-	-50 (-68)	55 (59)	-	
	Residual	RMSE	-	74 (72)	-	-	
		R <sup>2</sup>	-	0.86 (0.87)	-	-	
		MB	-	-18 (-34)	-	-	
	Mazia 2	No closure	RMSE	70 (69)	149 (132)	48 (60)	79 (77)
			R <sup>2</sup>	0.99 (0.99)	0.76 (0.74)	0.89 (0.87)	0.24 (0.19)
			MB	61 (61)	-124 (-114)	34 (50)	55 (59)
Br		RMSE	-	648 (193)	646 (196)	-	
		R <sup>2</sup>	-	0.00 (0.36)	0.08 (0.87)	-	
		MB	-	-99 (-168)	105 (170)	-	
Residual		RMSE	-	56 (56)	-	-	
		R <sup>2</sup>	-	0.87 (0.84)	-	-	
		MB	-	-28 (-47)	-	-	

Br = Bowen ratio, Residual = residual closure, RMSE = root mean square error, R<sup>2</sup> = coefficient of determination, MB = mean bias. RMSE and MB are given in Wm<sup>-2</sup>.

As shown in Table 3, the accuracy statistics for turbulent fluxes depended on the method to close the energy balance and the applied quality checks as described in Section 2.2.1. By analogy with Castelli et al. (2018), in the meadow (Mazia 1) the discrepancies between measured and modelled fluxes for H and LE were similar with two (i.e., relative and absolute closure) and three criteria (including additionally energy balance closure ratio) considered. On the other hand, for the pasture (Mazia 2) we observed a distorting effect of site observations when the aforementioned threshold 0.7 for the energy balance closure ratio was neglected. For that site *Br* residuals for latent heat (sensible heat) caused over 500% (400%) larger RMSE when compared to average values of the measured fluxes, as depicted by a scatterplot in Fig. 2. By applying the additional threshold, the absolute deviations from the averages reduced to 366% and 65% for LE and H, respectively. Nevertheless, regardless of the quality checks applied, for both sites the closest agreement between modelled and observed fluxes was obtained with residual closure (*Residual*) for LE and *no closure* correction for H. In Mazia 1 and Mazia 2 RMSE errors for latent heat (sensible heat) were of 72-74  $\text{Wm}^{-2}$  (49-52  $\text{Wm}^{-2}$ ) and 56  $\text{Wm}^{-2}$  (48-60  $\text{Wm}^{-2}$ ), respectively. For the remaining fluxes root mean square error ranged from 60  $\text{Wm}^{-2}$  to 70  $\text{Wm}^{-2}$  for  $R_n$ , and from 52  $\text{Wm}^{-2}$  to 79  $\text{Wm}^{-2}$  for G considering both study sites.

To sum up, LE performance improved when residual method was applied (Fig. 2), while Bowen ratio correction did not enhance the accuracy of turbulent fluxes, especially at the pasture site. On the contrary, the attribution of the *Br* to H and LE caused substantial increase in errors with over four times larger RMSE for latent heat flux when compared to statistics with *no closure*. The obtained accuracy results agree with other studies where authors estimated turbulent fluxes over Alpine grasslands (Castelli et al., 2018; Wohlfahrt et al., 2010).



**Fig. 2.** Scatterplots of the EC-based measurements of H and LE (X-axis) versus their corresponding outputs from TSEB-PT driven by in-situ LSTs for the entire year of 2017 (April-October). The presented results represent observations when thresholds for relative and absolute closures were considered.

### 4.3. Satellite-based TSEB-PT simulations

#### 4.3.1. TSEB with Terra MODIS LST

After the TSEB-PT benchmarking driven by ground measurements, we performed TSEB simulations at the plot scale using satellite LST data as a substitute for in-situ records. In this regard, TSEB fluxes forced with MODIS LSTs were derived for both original 1-km MODIS grids and 250-m downscaled maps. Table 4 shows a summary of accuracy statistics for TSEB obtained with coarse resolution and enhanced LST images. The stricter threshold of energy balance closure ratio did not improve the accuracy scores with satellite-based LST data. For this reason, we used the two previously mentioned quality checks (see Section 4.2). Note that in case of TSEB simulations driven by satellite-based LST, final estimates of energy fluxes were derived by weighting model

outputs according to spatial contribution of LST pixel within monthly tower footprints as explained before in Section 2.1.1.

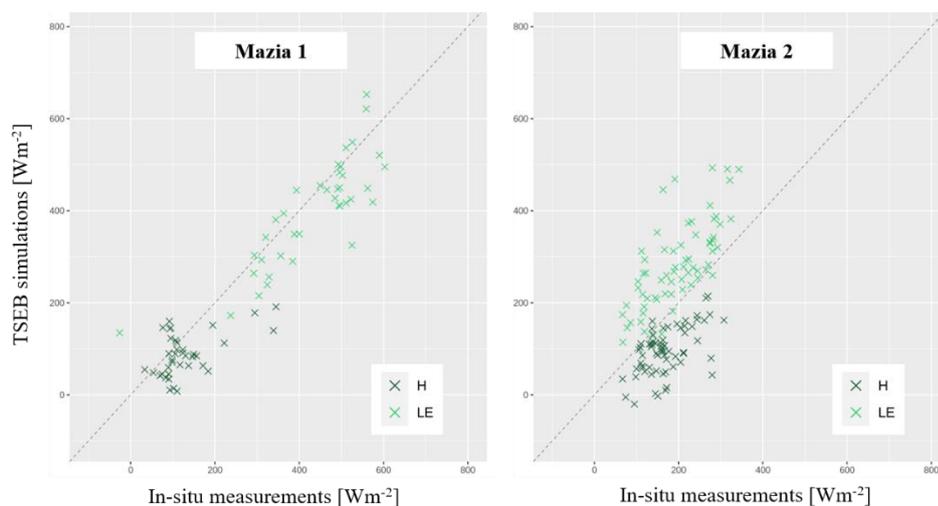
**Table 4.** An overview of the TSEB performance with satellite-based LST input considering 90% of total footprint area. All MOD11A1 QA bits decoded as “good quality”, as described in Wan et al. (2015), were considered in the flux estimation.

Site	LST		<i>No closure</i>				<i>Br</i>		<i>Residual</i>
			Rn	LE	H	G	LE	H	LE
Mazia 1	1 km	RMSE	91	112	73	50	101	145	102
		R <sup>2</sup>	0.94	0.58	0.28	0.05	0.66	0.28	0.66
		MB	85	-91	41	-29	20	94	73
	250 m	RMSE	80	145	72	44	91	144	76
		R <sup>2</sup>	0.93	0.59	0.29	0.10	0.70	0.29	0.70
		MB	72	-125	42	-3	-17	92	34
Mazia 2	1 km	RMSE	31	201	82	66	221	187	119
		R <sup>2</sup>	0.96	0.57	0.28	0.19	0.36	0.32	0.52
		MB	23	-187	63	59	-201	166	-98
	250 m	RMSE	39	188	88	56	213	200	108
		R <sup>2</sup>	0.96	0.57	0.31	0.15	0.37	0.33	0.51
		MB	34	-174	73	48	-194	180	-86

RMSE = root mean square error, R<sup>2</sup> = coefficient of determination, MB = mean bias. RMSE and MB are given in Wm<sup>-2</sup>.

As shown in Table 4, TSEB generally performed better with 250-m LST pixels than with original MODIS data. Similar to benchmarked simulations (Table 3), we obtained more satisfactory results for latent heat with residual correction applied. Consequently, by applying enhanced temperatures at the meadow site (Mazia 1), for LE we observed over 25% and 53% decrease in RMSE and MB, respectively. In case of Mazia 2, only marginally higher performance was obtained: RMSE for latent heat improved by 9%, while mean bias dropped about 12% when compared to simulations with 1-km MODIS LST data. For both sites sensible heat flux remained at the same level regardless of thermal data applied. While in Mazia 1 we observed nearly 1.4% improvement in RMSE using 250-m LST data, in the pasture the error for sensible heat flux increased by 7.3%.

As can be observed in Fig. 3, for latent heat flux we obtained a good agreement between TSEB outputs and in-situ measurements with R<sup>2</sup> of 0.70 and 0.51 over meadow (Mazia 1) and natural pasture (Mazia 2), respectively. Nevertheless, in Mazia 2 we noted relatively large overestimation of LE (MB = -86 Wm<sup>-2</sup>) and underestimation of H (MB = 73 Wm<sup>-2</sup>) when compared to the MB results for Mazia 1 (Table 4, Fig. 3) and benchmarked simulations (Table 3). In general, more consistent outcomes with MODIS LST data were obtained at the irrigated meadow site with barely 2 Wm<sup>-2</sup> difference in RMSE between TSEB forced by ground-derived LSTs (Table 3) and flux estimates with downscaled grids (Table 4).



**Fig. 3.** Scatterplots between turbulent fluxes from EC towers corrected for residual closure (*Residual*) and TSEB simulations with 250-m LST product. The comparison was based on observations corresponding to daytime Terra MODIS overpass time.

Considering the differences in the accuracy statistics between original and downscaled MODIS LST data, we evaluated the TSEB outputs forced by surface temperatures with the highest quality assurance (QA) flags, i.e., with emissivity and LST errors not exceeding 0.01 and 1 K, respectively (Table 5). The reason behind the selection of these observations was motivated by the dependency of the downscaling model on 1-km MODIS LST data (see Section 3.2). As presented in Table 5, the downscaled input improved the estimates of H (with *no closure*) in Mazia 1 translating to 15% and 30% decrease in RMSE and MB when compared to corresponding original MODIS LST grids. In addition, 250-m latent heat flux (with *Residual*) at that site yielded 10% smaller errors when compared to 1-km LST input (Table 5). Nonetheless, the accuracy scores for LE were less performant than the obtained statistics with all good quality flags included (Table 4). Meanwhile, in Mazia 2, regardless of the LST input used, we did not observe the accuracy improvement for the sensible heat flux (Table 5). Similar to outcomes in Table 4, the differences in estimation errors between model runs with 1-km and 250-m LST pixels at the pasture site remained at the same level.

**Table 5.** Summary of the TSEB evaluation measures with 1-km and 250-m MODIS LST inputs when only the highest quality QA flags were included (Wan et al., 2015).

Site	LST	<i>No closure</i>				<i>Br</i>		<i>Residual</i>	
		Rn	LE	H	G	LE	H	LE	
Mazia 1	1 km	RMSE	93	117	66	52	100	125	111
		R <sup>2</sup>	0.96	0.62	0.29	0.05	0.73	0.30	0.69
		MB	83	-88	37	-20	19	85	66
	250 m	RMSE	88	139	56	47	96	113	100
		R <sup>2</sup>	0.95	0.49	0.28	0.37	0.77	0.27	0.70
		MB	77	-113	26	5	1	71	46
Mazia 2	1 km	RMSE	30	210	79	70	216	175	120
		R <sup>2</sup>	0.97	0.65	0.24	0.27	0.39	0.29	0.57
		MB	25	-195	58	65	-197	157	-98
	250 m	RMSE	39	189	83	59	202	187	107
		R <sup>2</sup>	0.96	0.61	0.32	0.09	0.40	0.35	0.52
		MB	35	-174	68	49	-183	169	-82

RMSE = root mean square error, R<sup>2</sup> = coefficient of determination, MB = mean bias. RMSE and MB are given in Wm<sup>-2</sup>.

The overall accuracy of TSEB simulations for LE and H with MODIS LST differed between two sites. As depicted in Table 3 and Table 5, for Mazia 1, the TSEB-PT performance to estimate H with downscaled grids obtained similar RMSE and MB values (RMSE = 56 Wm<sup>-2</sup> and MB = 26 Wm<sup>-2</sup>) to the accuracy scores for the benchmarked fluxes (RMSE = 49 Wm<sup>-2</sup> and MB = 23 Wm<sup>-2</sup>). Meanwhile, larger deviations for sensible heat were observed in the pasture yielding an RMSE of 83 Wm<sup>-2</sup>, and as a result, translating to 80% bigger error than RMSE obtained from tower-based modeling as presented in Section 4.2. Moreover, regardless of MODIS LST inputs, the model led to strong underestimation of H (MB = 68 Wm<sup>-2</sup>) at that site. Similarly, LE estimates for Mazia 2 were less robust than for the meadow when compared to benchmarked fluxes in Table 3. While RMSE for latent heat increased on average by 33% in Mazia 1, for the pasture we obtained over 90% error growth with mean bias of -82 Wm<sup>-2</sup>. Despite downscaled data reduced deviations from ground-derived LST in Mazia 2 with lower MB than for MOD11A1 (Table 2), 250-m LST pixels still differed from in-situ data. As a consequence, this led to less performant accuracy statistics for turbulent fluxes (Table 5) with the greatest impact on H that is strongly influenced by surface temperature (Burchard-Levine et al., 2020; Timmermans et al., 2007).

Considering temporal differences in the model predictions between site simulations in Section 4.2. and TSEB-driven by satellite-based LST input, we compared absolute values of RMSE errors for 250-m energy fluxes (Table 4-5) against averaged in-situ measurements at the time of Terra MODIS overpass. Table 6 lists the obtained results.

As can be observed in Table 6, the most performant results were obtained with *Residual* method for both sites. Nonetheless, similar to the results in Table 4 and 5, the relative comparison for Mazia 1 confirmed better accuracy scores for latent heat than at Mazia 2 site with an average  $\Delta$ RMSE equal to 20% and 54%, respectively (Table 6). Meanwhile, although the absolute errors for H were higher at the pasture site (Table 4-5), the intercomparison between two stations in Table 6 revealed that  $\Delta$ RMSEs for both sites were at the same level yielding on average 52% share in the in-situ averages. Those discrepancies for sensible heat might be related to still existing land-cover heterogeneity within 250-m MODIS pixel (Fig. 1b). Consequently, the downscaled LST input was less

performant for modeling TSEB fluxes when compared to other studies (Guzinski et al., 2019; Liu et al., 2020). For example, Guzinski et al. (2019) reported 20% improvement for H estimates with sharpened LST, however, their downscaling model was driven by high-resolution explanatory variables from Landsat and Sentinel-2 observations, and then was applied to relatively plain areas in Denmark.

**Table 6.** A relative comparison of the accuracy results for 250-m TSEB estimates as a percentage share of their RMSEs (referred as  $\Delta$ RMSE) with respect to the averages of measured fluxes at the time of daytime Terra overpass. Values in brackets represent  $\Delta$ RMSE with regard to the statistics in Table 4, while the remaining records correspond to Table 5.

Site	Correction	Rn	LE	H	G
$\Delta$ RMSE					
Mazia 1	<i>No closure</i>	14% (13%)	52% (53%)	47% (56%)	78% (83%)
	<i>Br</i>	-	25% (24%)	68% (80%)	-
	<i>Residual</i>	-	23% (18%)	-	-
Mazia 2	<i>No closure</i>	8% (8%)	169% (173%)	51% (52%)	45% (43%)
	<i>Br</i>	-	196% (237%)	71% (72%)	-
	<i>Residual</i>	-	52% (55%)	-	-

#### 4. Conclusions

A large number of satellite-based evapotranspiration (ET) estimates rely on coarse resolution thermal remote sensing data derived from the MODIS TIR channels, which limits their accuracy over complex environments. Due to TSEB assumption of thermal homogeneity within pixel, a main challenge in ET modelling is related to mixed grid cells. In particular, this aspect is critical over heterogenous regions, like mountainous areas. To address this limitation, we exploited prediction capabilities of the Priestley-Taylor two-source energy balance model forced with original MODIS LST images and their corresponding downscaled product at 250 m spatial resolution. Our experimental area was located in alpine grasslands of Matsch-Mazia Valley in South Tyrol, which belongs to one of the driest areas in the European Alps. Two eddy covariance sites situated in irrigated meadow (Mazia 1) and rainfed pasture (Mazia 2) were used for ET modeling. First, they served to evaluate TSEB-PT performance with ground-derived land surface temperatures within their tower footprints. Next, the LST input was replaced by two MODIS datasets: 1-km MOD11A1 observations and enhanced LST maps. Results demonstrate that the sharpened LSTs were capable of improving retrievals of TSEB fluxes in the alpine grasslands when compared to original MODIS dataset.

The presented thermal disaggregation procedure provided added value for estimating energy fluxes at 250 m spatial resolution. The in-situ validation confirmed closer agreement with downscaled-like TSEB simulations rather than with 1-km MODIS-based model outputs. On average, RMSE and MB for 250-m turbulent fluxes were equal to  $86 \text{ Wm}^{-2}$  and  $55 \text{ Wm}^{-2}$ , respectively. Given the statistics obtained with coarse-resolution temperatures (Table 4-5), thermal downscaling translated to the accuracy improvement by 8% in RMSE and 15% for mean bias, averaged in the study area. The TSEB simulations are in line with other studies where authors estimated LE and H using two-source models with LSTs from MODIS and Landsat over mountain areas (Castelli et al., 2018; Elfarkh et al., 2020). Similar to Castelli et al. (2018), the discrepancies for turbulent fluxes were larger than Rn and G, being strongly affected by the correction method to close energy balance. Nevertheless, at the same time, errors for LE and H forced with 1-km MODIS LST were higher than ours with an average RMSE of  $134 \text{ Wm}^{-2}$ . In addition, our outcomes agree with the results of Guzinski et al. (2020) who studied the performance of TSEB-PT with sharpened Sentinel-3 SLSTR in lowland grasslands and obtained an average RMSE of  $85 \text{ Wm}^{-2}$  for turbulent fluxes.

Notably, it should be mentioned that TSEB model requires leaf area index and surface temperature information derived with high accuracy. In case of the first dataset, this condition was satisfied by applying 20-m Sentinel-2-based LAI with only good quality flags included. Even though for downscaling procedure only LST input with satisfactory QA bits was considered, the retrieval error in LST was not less than 2 K (Table 4-5). As a consequence, this may have translated to larger errors for turbulent fluxes, where 1 K error results in 17%

uncertainty in modelled H as reported by Burchard-Levine et al. (2020). The negative impact of inaccurate LSTs on flux estimates was also documented extensively in other studies (Cui et al., 2021; Gan and Gao, 2015; Kustas et al., 2012). Meanwhile, the retrieval errors for 1-km MODIS LST are propagated to the kernel-driven downscaling models that, in principle, are based on function relationship between 1-km LST pixels and fine-resolution explanatory variables. This limits the robustness of 250-m LST resulting in lower performance of TSEB simulations than obtained accuracies by other scientists. However, the documented accuracy scores were related to regional-oriented studies in China (Liu et al., 2020) or plain terrain with relatively homogenous landscapes (Guzinski et al., 2020). Despite the higher resolution of the downscaled product, high variability in terms of terrain features (i.e., elevation, aspect, and slope) and landcover was still present within 250-m pixel size (Fig. 1). The observed heterogeneities may have caused larger uncertainties in both thermal downscaling product and TSEB estimates.

TSEB-PT driven by downscaled LST input exhibits potential for estimating turbulent fluxes over vegetated areas in mountain regions, but it is still challenging to improve retrieval accuracy of the modelled estimates, mainly due to complex meteorological conditions, such as frequent cloudy-sky conditions, and small-scale heterogeneity in terms of topography and landcover. In fact, those factors have a negative impact on the accuracy retrieval of satellite-based TIR data and their coarse resolution translates to mixed spectral response from fragmented landscape. Despite increasing number of high-resolution TIR sensors such as Landsat-8 TIRS, ISS ECOSTRESS, and recently launched Landsat-9 TIRS-2, their acquisition frequency and related cloud cover are not satisfactory for routine ET monitoring. In this context, further improvements for thermal downscaling of MODIS and Sentinel-3SLSTR need to be investigated. As mentioned previously, spatial mismatch between tower flux footprints and 250-m pixels influenced the accuracy of modelled fluxes. Therefore, future research in this field should first investigate the effect of high-resolution LST data on TSEB-based fluxes in order to define acceptable spatial resolution for downscaling output. Although the documented feasibility studies confirmed accuracy improvement with thermal bands from Landsat over the Atlas mountains in Morocco (Elfarkh et al., 2020), an independent evaluation for the Alpine region is important due to differences in land characteristics between those two regions. Moreover, some modifications in the thermal disaggregation may enhance TSEB performance over mountainous areas. First, predictors should be derived at fine spatial resolution (< 250 m) with possibly high frequency, which will allow us to better explain LST variability at the time of satellite overpass. Next, the dependency of downscaled product on coarse resolution LST data will be reduced by applying more sophisticated methods for residual correction and building localized regression models considering similarities between spatial features over the study area.

## Acknowledgments

Mazia stations has been funded by the Autonomous Province of Bozen-Bolzano in the framework of the LTSER site.

## References

- 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(4), pp. 545-558.
- Allen, R. G., Tasumi, M., & Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Model. *Journal of irrigation and drainage engineering*, 133(4), pp. 380-394.
- Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., and Gao, F. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrology and Earth System Sciences*, 15(1), pp. 223-239.
- Anderson, M.C., Norman, J.M., Mecikalski, J.R., Otkin, J.A. and Kustas, W.P., 2007. A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *Journal of Geophysical Research: Atmospheres*, 112(D10).

- Bartkowiak, P., Castelli, M., Crespi, A., Niedrist, G., Zanotelli, D., Colombo, R., Notarnicola, C., 2022. MODIS Land Surface Temperature Reconstruction under Cloudy-sky Conditions at 250 m Spatial Resolution: Case Study of Vinschgau/Venosta Valley in the European Alps (resubmitted to *IEEE JSTARS* after minor revision).
- Bartkowiak, P., Castelli, M. and Notarnicola, C., 2019. Downscaling land surface temperature from MODIS dataset with random forest approach over alpine vegetated areas. *Remote Sensing*, 11(11), p. 1319.
- Bastiaanssen, W. G., Menenti, M., Feddes, R. A., & Holtslag, A. A. M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *Journal of hydrology*, 212, pp. 198-212.
- Bastiaanssen, W. G., Pelgrum, H., Wang, J., Ma, Y., Moreno, J. F., Roerink, G. J., & Van der Wal, T., 1998. A remote sensing surface energy balance algorithm for land (SEBAL): Part 2: Validation. *Journal of hydrology*, 212, 213-229.
- Bengtsson, J., Bullock, J.M., Egoh, B., Everson, C., Everson, T., O'Connor, T., O'Farrell, P.J., Smith, H.G. and Lindborg, R., 2019. Grasslands—more important for ecosystem services than you might think. *Ecosphere*, 10(2), p. e02582.
- Bertoldi, G., Notarnicola, C., Leitinger, G., Endrizzi, S., Zebisch, M., Della Chiesa, S. and Tappeiner, U., 2010. Topographical and ecohydrological controls on land surface temperature in an alpine catchment. *Ecohydrology: Ecosystems, Land and Water Process Interactions, Ecohydrogeomorphology*, 3(2), pp. 189-204.
- Bhattacharai, N., Wagle, P., Gowda, P.H. and Kakani, V. G., 2017. Utility of remote sensing-based surface energy balance models to track water stress in rain-fed switchgrass under dry and wet conditions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 133, pp. 128-141.
- Bindhu, V. M., Narasimhan, B. and Sudheer, K. P., 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, pp. 118-129.
- Bisquert, M., Sánchez, J. M. and Caselles, V., 2016. Evaluation of disaggregation methods for downscaling MODIS land surface temperature to Landsat spatial resolution in Barrax test site. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(4), pp. 1430-1438.
- Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32.
- Burchard-Levine, V., Nieto, H., Riaño, D., Migliavacca, M., El-Madany, T.S., Guzinski, R., Carrara, A. and Martín, M.P., 2021. The effect of pixel heterogeneity for remote sensing based retrievals of evapotranspiration in a semi-arid tree-grass ecosystem. *Remote Sensing of Environment*, 260, p. 112440.
- Burchard-Levine, V., Nieto, H., Riaño, D., Migliavacca, M., El-Madany, T.S., Perez-Priego, O., Carrara, A. and Martín, M.P., 2020. Seasonal adaptation of the thermal-based two-source energy balance model for estimating evapotranspiration in a semiarid tree-grass ecosystem. *Remote Sensing*, 12(6), p. 904.
- Cammalleri, C., Anderson, M.C., Gao, F., Hain, C.R. and Kustas, W.P., 2014. Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. *Agricultural and forest meteorology*, 186, pp.1-11.
- Campbell, G. S. and Norman, J. M., 2000. *An introduction to environmental biophysics*. Springer Science & Business Media.
- Castelli, M., Anderson, M. C., Yang, Y., Wohlfahrt, G., Bertoldi, G., Niedrist, G., Hammerle, A., Zhao, P., Zebisch, M. and Notarnicola, C., 2018. Two-source energy balance modeling of evapotranspiration in Alpine grasslands. *Remote Sensing of Environment*, 209, pp. 327-342.
- Choudhury, B. J., Idso, S. B., & Reginato, R. J. (1987). Analysis of an empirical model for soil heat flux under a growing wheat crop for estimating evaporation by an infrared-temperature based energy balance equation. *Agricultural and Forest Meteorology*, 39(4), 283-297.
- Conedera, M., Cesti, G., Pezzatti, G. B., Zumbrennen, T. and Spinedi, F., 2006. Lightning-induced fires in the Alpine region: An increasing problem. *Forest Ecology and Management*, 234(1), p. S68.
- Cui, Y., Song, L. and Fan, W., 2021. Generation of spatio-temporally continuous evapotranspiration and its components by coupling a two-source energy balance model and a deep neural network over the Heihe River Basin. *Journal of Hydrology*, 597, p. 126176.

- Dong, P., Gao, L., Zhan, W., Liu, Z., Li, J., Lai, J., Li, H., Huang, F., Tamang, S.K. and Zhao, L., 2020. Global comparison of diverse scaling factors and regression models for downscaling Landsat-8 thermal data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169, pp. 44-56.
- Elfarkh, J., Ezzahar, J., Er-Raki, S., Simonneaux, V., Ait Hssaine, B., Rachidi, S., Brut, A., Rivalland, V., Khabba, S., Chehbouni, A. and Jarlan, L., 2020. Multi-scale evaluation of the TSEB model over a complex agricultural landscape in Morocco. *Remote Sensing*, 12(7), p. 1181.
- Fink, A. H., Brücher, T., Krüger, A., Leckebusch, G. C., Pinto, J. G. and Ulbrich, U., 2004. The 2003 European summer heatwaves and drought—synoptic diagnosis and impacts. *Weather*, 59(8), pp. 209-216.
- Foken, T.; Aubinet, M.; Finnigan, J.J.; Leclerc, M.Y.; Mauder, M.; Paw, U.K.T. Results of a panel discussion about the energy balance closure correction for trace gases. *Bull. Am. Meteorol. Soc.* 2011, 92, ES13–ES18.
- French, A. N., Hunsaker, D. J. and Thorp, K. R., 2015. Remote sensing of evapotranspiration over cotton using the TSEB and METRIC energy balance models. *Remote Sensing of Environment*, 158, pp. 281-294.
- Gan, G. and Gao, Y., 2015. Estimating time series of land surface energy fluxes using optimized two source energy balance schemes: Model formulation, calibration, and validation. *Agricultural and forest meteorology*, 208, pp. 62-75
- Gao, F., Kustas, W. P. and Anderson, M. C., 2012. A data mining approach for sharpening thermal satellite imagery over land. *Remote Sensing*, 4(11), pp.3287-3319.
- Gao, Z., Wang, Q., Cao, X. and Gao, W., 2014. The responses of vegetation water content (EWT) and assessment of drought monitoring along a coastal region using remote sensing. *GIScience & remote sensing*, 51(1), pp. 1-16.
- Gao, L., Zhan, W., Huang, F., Quan, J., Lu, X., Wang, F., Ju, W. and Zhou, J., 2016. Localization or globalization? Determination of the optimal regression window for disaggregation of land surface temperature. *IEEE Transactions on Geoscience and Remote Sensing*, 55(1), pp. 477-490.
- Glenn, E.P., Nagler, P. L. and Huete, A. R., 2010. Vegetation index methods for estimating evapotranspiration by remote sensing. *Surveys in Geophysics*, 31(6), pp. 531-555.
- Gobiet, A., Kotlarski, S., Beniston, M., Heinrich, G., Rajczak, J. and Stoffel, M., 2014. 21st century climate change in the European Alps—A review. *Science of the Total Environment*, 493, pp. 1138-1151.
- Guzinski, R., Anderson, M. C., Kustas, W. P., Nieto, H. and Sandholt, I., 2013. Using a thermal-based two source energy balance model with time-differencing to estimate surface energy fluxes with day–night MODIS observations. *Hydrology and Earth System Sciences*, 17(7), pp. 2809-2825.
- Guzinski, R. and Nieto, H., 2019. Evaluating the feasibility of using Sentinel-2 and Sentinel-3 satellites for high-resolution evapotranspiration estimations. *Remote sensing of Environment*, 221, 157-172.
- Guzinski, R., Nieto, H., Jensen, R. and Mendiguren, G., 2014. Remotely sensed land-surface energy fluxes at sub-field scale in heterogeneous agricultural landscape and coniferous plantation. *Biogeosciences*, 11(18), pp.5021-5046.
- Guzinski, R., Nieto, H., Sandholt, I. and Karamitilios, G., 2020. Modelling high-resolution actual evapotranspiration through Sentinel-2 and Sentinel-3 data fusion. *Remote Sensing*, 12(9), p. 1433.
- He, J., Zhao, W., Li, A., Wen, F. and Yu, D., 2019. The impact of the terrain effect on land surface temperature variation based on Landsat-8 observations in mountainous areas. *International Journal of Remote Sensing*, 40(5-6), pp. 1808-1827.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., and Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999-2049.
- Hulley, G., Malakar, N. and Freepartner, R., 2016. Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature and emissivity product (MxD21) algorithm theoretical basis document collection-6. *JPL Publication*, pp. 12-17.
- Hutengs, C. and Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*, 178, pp.127-141.
- Jeganathan, C., Hamm, N. A., Mukherjee, S., Atkinson, P. M., Raju, P. L. N. and Dadhwal, V. K., 2011. Evaluating a thermal image sharpening model over a mixed agricultural landscape in India. *International Journal of Applied Earth Observation and Geoinformation*, 13(2), pp. 178-191.

- Kljun, N., Calanca, P., Rotach, M. W. and Schmid, H. P., 2015. A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP). *Geoscientific Model Development*, 8(11), pp. 3695-3713.
- Koetz, B., Berger, M., Blommaert, J., Del Bello, U., Drusch, M., Duca, R. and Hook, S. (2019). Copernicus High Spatio-Temporal Resolution Land Surface Temperature Mission: Mission Requirements Document. *ESA: Noordwijk, The Netherlands*.
- Kundu, S., Mondal, A., Khare, D., Hain, C. and Lakshmi, V., 2018. Projecting climate and land use change impacts on actual evapotranspiration for the Narmada river basin in central India in the future. *Remote Sensing*, 10(4), p.578.
- Kustas, W. P., Alfieri, J. G., Anderson, M. C., Colaizzi, P. D., Prueger, J.H., Evett, S. R., Neale, C. M., French, A. N., Hipps, L. E., Chávez, J. L. and Copeland, K. S., 2012. Evaluating the two-source energy balance model using local thermal and surface flux observations in a strongly advective irrigated agricultural area. *Advances in Water Resources*, 50, pp. 120-133.
- Kustas, W. P. and Norman, J. M., 1999. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. *Agricultural and Forest Meteorology*, 94(1), pp.13-29.
- Kustas, W. P., Norman, J. M., Anderson, M. C., and French, A. N., “Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship”, *Remote Sens. Environ.*, 85(4), pp. 429-440, Jun. 2003.
- Lemaire, G., Hodgson, J. and Chabbi, A. eds., 2011. *Grassland productivity and ecosystem services*. Cabi.
- Lewińska, K. E., Ivits, E., Schardt, M. and Zebisch, M., 2018. Drought impact on phenology and green biomass production of alpine mountain forest—case study of South Tyrol 2001–2012 inspected with MODIS time series. *Forests*, 9(2), p. 91.
- Li, W., Ni, L., Li, Z.L., Duan, S.B. and Wu, H., 2019. Evaluation of machine learning algorithms in spatial downscaling of MODIS land surface temperature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7), pp.2299-2307.
- Liaw, A. and Wiener, M., 2002. Classification and regression by randomForest. *R news*, 2(3), pp.18-22.
- Liu, K., Su, H., Li, X., and Chen, S., 2020. Development of a 250-m Downscaled Land Surface Temperature Data Set and Its Application to Improving Remotely Sensed Evapotranspiration Over Large Landscapes in Northern China, *IEEE Trans. Geosci. Remote Sens.*
- Maeda, E.E., 2014. Downscaling MODIS LST in the East African mountains using elevation gradient and land-cover information. *International Journal of Remote Sensing*, 35(9), pp.3094-3108.
- Mahour, M., Tolpekin, V., Stein, A., and Sharifi, A. (2017). A comparison of two downscaling procedures to increase the spatial resolution of mapping actual evapotranspiration. *ISPRS journal of photogrammetry and remote sensing*, 126, 56-67.
- Majazi, N. P., Mannaerts, C.M., Ramoelo, A., Mathieu, R. and Verhoef, W., 2021. Uncertainty and Sensitivity Analysis of a Remote-Sensing-Based Penman–Monteith Model to Meteorological and Land Surface Input Variables. *Remote Sensing*, 13(5), p.882.
- Mao, Q., Peng, J. and Wang, Y., 2021. Resolution Enhancement of Remotely Sensed Land Surface Temperature: Current Status and Perspectives. *Remote Sensing*, 13(7), p. 1306.
- McCune, B., 2007. Improved estimates of incident radiation and heat load using non-parametric regression against topographic variables. *Journal of Vegetation Science*, 18(5), pp. 751-754.
- Merlin, O., Chirouze, J., Olioso, A., Jarlan, L., Chehbouni, G. and Boulet, G., 2014. An image-based four-source surface energy balance model to estimate crop evapotranspiration from solar reflectance/thermal emission data (SEB-4S). *Agricultural and Forest Meteorology*, 184, pp. 188-203.
- Meyer, H. and Pebesma, E., 2021. Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, 12(9), pp.1620-1633.
- Mildrexler, D.J., Zhao, M. and Running, S.W., 2011. A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. *Journal of Geophysical Research: Biogeosciences*, 116(G3).
- Nagler, P. L., Morino, K., Murray, R. S., Osterberg, J. and Glenn, E. P., 2009. An empirical algorithm for estimating agricultural and riparian evapotranspiration using MODIS enhanced vegetation index and ground measurements of ET. I. Description of method. *Remote Sensing*, 1(4), pp. 1273-1297.

- Njuki, S. M., Mannaerts, C. M. and Su, Z., 2020. An Improved Approach for Downscaling Coarse-Resolution Thermal Data by Minimizing the Spatial Averaging Biases in Random Forest. *Remote Sensing*, 12(21), p. 3507.
- Park, S., Im, J., Jang, E. and Rhee, J., 2016. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agricultural and forest meteorology*, 216, pp.157-169.
- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M. and Isaac, P., 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific data*, 7(1), pp.1-27.
- Priestley, C. H. B. and TAYLOR, R. J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review*, 100(2), pp. 81-92.
- Reichstein, M., Moffat, A.M., Wutzler, T. and Sickel, K., 2014. REdDyProc: Data processing and plotting utilities of (half-) hourly eddy-covariance measurements. *R package version 0.6-0/r9*, 755.
- Roerink, G. J., Su, Z., & Menenti, M. (2000). S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 25(2), 147-157.
- Santanello Jr, J.A. and Friedl, M.A., 2003. Diurnal covariation in soil heat flux and net radiation. *Journal of Applied Meteorology*, 42(6), pp.851-862.
- Santos, F., Graw, V. and Bonilla, S., 2019. A geographically weighted random forest approach for evaluate forest change drivers in the Northern Ecuadorian Amazon. *PLoS one*, 14(12), p. e0226224.
- Senay, G. B., Friedrichs, M., Singh, R. K. and Velpuri, N. M., 2016. Evaluating Landsat 8 evapotranspiration for water use mapping in the Colorado River Basin. *Remote Sensing of Environment*, 185, pp. 171-185.
- Sharma, V., Kilic, A., & Irmak, S. (2016). Impact of scale/resolution on evapotranspiration from Landsat and MODIS images. *Water Resources Research*, 52(3), 1800-1819.
- Stoy, P.C., Mauder, M., Foken, T., Marcolla, B., Boegh, E., Ibrom, A., Arain, M.A., Arneth, A., Aurela, M., Bernhofer, C., et al., 2013. A data-driven analysis of energy balance closure across fluxnet research sites: the role of landscape scale heterogeneity. *Agric. For. Meteorol.* 171, 137–152.
- Su, Z. (2002). The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrology and earth system sciences*, 6(1), 85-100.
- Sun, L., Anderson, M. C., Gao, F., Hain, C., Alfieri, J. G., Sharifi, A., McCarty, G. W., Yang, Y., Yang, Y., Kustas, W. P. and McKee, L., 2017. Investigating water use over the C hoptank River Watershed using a multisatellite data fusion approach. *Water Resources Research*, 53(7), pp. 5298-5319.
- Tang, R., Li, Z.L., Jia, Y., Li, C., Chen, K.S., Sun, X. and Lou, J., 2013. Evaluating one-and two-source energy balance models in estimating surface evapotranspiration from Landsat-derived surface temperature and field measurements. *International Journal of Remote Sensing*, 34(9-10), pp.3299-3313.
- Team, R. C., 2013. R: A language and environment for statistical computing.
- Timmermans, W. J., Kustas, W. P., Anderson, M. C., & French, A. N. (2007). An intercomparison of the surface energy balance algorithm for land (SEBAL) and the two-source energy balance (TSEB) modeling schemes. *Remote Sensing of Environment*, 108(4), 369-384.
- Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P., Meyers, T. P., Prueger, J. H., Starks, P. J. and Wesely, M. L., 2000. Correcting eddy-covariance flux underestimates over a grassland. *Agricultural and forest meteorology*, 103(3), pp. 279-300.
- Wan, Z. and Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Transactions on geoscience and remote sensing*, 34(4), pp.892-905.
- Wan, Z., Hook, S. and Hulley, G., 2015. MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006, NASA EOSDIS Land Processes DAAC, NASA: Washington, DC, USA.
- Wang, L., Iddio, E. and Ewers, B., 2021. Introductory overview: Evapotranspiration (ET) models for controlled environment agriculture (CEA). *Computers and Electronics in Agriculture*, 190, p.106447.

- Wang, Q., Shi, W., Atkinson, P.M. and Zhao, Y., 2015. Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*, 166, pp.191-204.
- Wang, K., Wan, Z., Wang, P., Sparrow, M., Liu, J., Zhou, X. and Haginoya, S., 2005. Estimation of surface long wave radiation and broadband emissivity using Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature/emissivity products. *Journal of Geophysical Research: Atmospheres*, 110(D11).
- Weiss, M. and Baret, F., 2016. S2ToolBox Level 2 Products: LAI, FAPAR, FCOVER, Version 1.1. *ESA Contract nr 4000110612/14/I-BG*, p.52.
- Widmoser, P. and Wohlfahrt, G., 2018. Attributing the energy imbalance by concurrent lysimeter and eddy covariance evapotranspiration measurements. *Agricultural and forest meteorology*, 263, pp.287-291.
- Wohlfahrt, G., Albin, H., Georg, N., Katharina, S., Enrico, T., Peng, Z., 2016. On the energy balance closure and net radiation in complex terrain, *Agric. For. Meteorol.*, 226, 37–49.
- Wohlfahrt, G., Irschick, C., Thalinger, B., Hörtnagl, L., Obojes, N. and Hammerle, A., 2010. Insights from independent evapotranspiration estimates for closing the energy balance: a grassland case study. *Vadose Zone Journal*, 9(4), pp.1025-1033.
- Wohlfahrt, G. and Widmoser, P., 2013. Can an energy balance model provide additional constraints on how to close the energy imbalance? *Agricultural and forest meteorology*, 169, pp. 85-91.
- Wright, M.N. and Ziegler, A., 2015. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *arXiv preprint arXiv:1508.04409*.
- Xu, J., Zhang, F., Jiang, H., Hu, H., Zhong, K., Jing, W., Yang, J. and Jia, B., 2020. Downscaling ASTER land surface temperature over urban areas with machine learning-based area-to-point regression Kriging. *Remote Sensing*, 12(7), p.1082.
- Xu, S., Zhao, Q., Yin, K., He, G., Zhang, Z., Wang, G., Wen, M. and Zhang, N., 2021. Spatial Downscaling of Land Surface Temperature Based on a Multi-Factor Geographically Weighted Machine Learning Model. *Remote Sensing*, 13(6), p. 1186.
- Yang, Y., Anderson, M. C., Gao, F., Hain, C.R., Semmens, K. A., Kustas, W.P., Noormets, A., Wynne, R.H., Thomas, V.A. and Sun, G., 2017. Daily Landsat-scale evapotranspiration estimation over a forested landscape in North Carolina, USA, using multi-satellite data fusion. *Hydrology and Earth System Sciences*, 21(2), pp. 1017-1037.
- Yang, Y., Long, D., Guan, H., Liang, W., Simmons, C. and Batelaan, O., 2015. Comparison of three dual-source remote sensing evapotranspiration models during the MUSOEXE-12 campaign: Revisit of model physics. *Water resources research*, 51(5), pp. 3145-3165.
- Zhang, C., Long, D., Zhang, Y., Anderson, M.C., Kustas, W. P. and Yang, Y., 2021. A decadal (2008–2017) daily evapotranspiration data set of 1 km spatial resolution and spatial completeness across the North China Plain using TSEB and data fusion. *Remote Sensing of Environment*, 262, p.112519.
- Zhang, H., Nettleton, D. and Zhu, Z., 2019. Regression-enhanced random forests. *arXiv preprint arXiv:1904.10416*.
- Zhao, W., Duan, S.B., Li, A. and Yin, G., 2019. A practical method for reducing terrain effect on land surface temperature using random forest regression. *Remote sensing of environment*, 221, pp.635-649.
- Zhao, W., Wu, H., Yin, G. and Duan, S.B., 2019. Normalization of the temporal effect on the MODIS land surface temperature product using random forest regression. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, pp.109-118.



This thesis explores the potential of the developed land surface temperature (LST) product for estimating fine-resolution (FR) energy fluxes using satellite-based two-source energy balance approach. The study area is located in the agricultural landscape of South Tyrol, an Alpine region in northern Italy, that stands out due its complex topography and landcover heterogeneity.

The research study starts with the implementation of spatial enhancement method for 1-km Terra MODIS LST data. In Section 3 data-driven downscaling with random forest (RF) was tested for predicting subpixel land surface temperature over vegetated ecosystems in the Province of Bozen-Bolzano. Considering the first research question in Section 2, kernel-driven method was found as a sound solution for regions with complex landscapes due to availability of fine-resolution explanatory variables and high revisit time of MODIS LST scenes that reduces chances for image acquisitions with persistent cloud cover. In this regard, the author exploited empirical relationship between coarse-resolution LST data and FR predictors, including digital elevation model (DEM) and normalized difference vegetation index (NDVI), which partially answers the third question in Section 2. For model creation three different pixel selection approaches were considered, including: all vegetated pixels, grid cells with 90% threshold for canopy contribution, and pixels with minimum 75% homogeneity for one biome class. The last two models achieved slightly better results when compared to the first one, yielding an average Root Mean Square Error (Mean Absolute Error) of 2.3 K (1.8 K) and 2.2 K (1.7 K) with all LST pixels included. The disaggregated maps were more performant than original MODIS data translating to 20% reduction of the RMSE error with regard to statistics obtained for 1-km LST data (the second research question).

Along with RF sharpening of LST, Section 4 presents a novel method to reconstruct missing MODIS LST imagery at 250 m resolution under long-term overcast conditions. The reconstruction was performed with two steps: LST modelling from site-based variables and applying the fitted models to cloudy-sky grids. Answering the fourth research question regarding modelling approach, the method integrates data-driven analysis with physical-based assumptions of European Solar Radiation Atlas to derive cloudy-sky observations for establishing LST models. Considering complex land-atmosphere interactions in mountain regions, correlations between local LST and air temperature in conjunction with other biophysical variables, like incoming solar radiation, albedo, and leaf area index (LAI) were investigated for three separate model groups: forest, grassland, and permanent crops. Cloudy-sky LSTs over forest were predicted with air temperature, while grasslands and permanent crops, being more influenced by near-surface physical interactions, were additionally explained by solar radiation and biophysical variables such as LAI and albedo (the third research question). The regressions were applied to 250-m gridded predictors considering similarities between training data and new targets in predictor variable space, which responds to the prediction part of the fourth research question. The gap-filling experiment was promising in terms of accuracy results yielding an average RMSE of 2.12 K and  $R^2$  of 0.84 when compared to ground-derived LSTs.

Section 5 concludes the thesis by applying the enhanced MODIS LST data to the Priestley-Taylor two-source energy balance (TSEB-PT) model for estimating 250-m energy fluxes in alpine grasslands. In this context, the kernel-driven downscaling method (Section 3) was refined by incorporating other potential predictors, like downwelling solar radiation, heat load index, aspect and slope, as an attempt to better explain LST variability at MODIS's subpixel scale. The performance of the thermal sharpening

was validated with local LSTs derived from longwave radiation measurements at two eddy covariance (EC) systems at Matsch-Mazia site. On average, 250-m LST grids improved accuracy statistics resulting in 10% and 30% decrease in RMSE and mean bias (MB) when compared to evaluation measures for original MODIS data. However, the field-based analysis gave stricter overview on the downscaling capabilities and related heterogeneity within 250-m grid cells. The obtained errors were larger than accuracy scores with Landsat images (degraded to 250 m spatial resolution), as presented in Section 3. Next, in a second step of this work, the author aimed to answer to the fifth research question listed in Section 2. Given satisfactory results from model benchmarking with an overall RMSE of  $58 \text{ Wm}^{-2}$  for latent (LE) and sensible (H) heat, the performance of TSEB-PT was evaluated by incorporating satellite-based thermal forcing using both original MODIS data and enhanced LST product. In general, LE and H driven by 250-m surface temperatures yielded an RMSE of  $86 \text{ Wm}^{-2}$  and a MB of  $55 \text{ Wm}^{-2}$ , which caused 8% and 15% decrease in the respective errors when compared to coarse resolution temperatures.

Considering the last research question (Section 2), the validation results demonstrate that the downscaled maps are able to provide better results for estimating TSEB-derived fluxes when compared to 1-km source images. Nonetheless, the statistics scores obtained when validating LST predictions from downscaling model do not translate into more performant TSEB-PT simulations. In fact, the two-source model still estimates sensible heat flux with considerable errors, which relates them to land surface temperature input. These discrepancies might have been caused by spatial mismatch between thermal conditions of the land surface within 250-m LST grid cell and flux tower footprints that represent about 30% of total pixel area. Although downscaled product was modelled with predictors depicting vegetation content (NDVI) and relief complexity (i.e., elevation, slope, aspect), 250-m spatial resolution and low frequency of image acquisitions may not have captured terrain heterogeneity and its temporal variability. In addition, since kernel-driven sharpening is based on 1-km LSTs, it is strongly influenced by uncertainties of original MODIS data which are usually between 1 K and 2 K over the study area.

The thesis aims to provide a novel approach for generating enhanced land surface temperature for monitoring and sustainable management of the water resources in mountain areas. The envisaged product will help to better understand complex land-atmosphere processes and tackle climate change (CC) at regional scales by exploiting thermal conditions of the Earth's skin at 250 m resolution. Having spatially continuous set of sub-kilometer remote sensing data, it would be possible to monitor multi-temporal trends with respect to land-cover changes, the CC, and water cycling. Moreover, the findings will be key to help governance institutions and agricultural community for improving sustainable agricultural production, drought predictions and also land degradation monitoring. In parallel, time-series of the 250-m maps will temporally complement the planned Copernicus LSTM mission as an alternative to the 1-km thermal observations from Sentinel-3 SLSTR and Aqua/Terra MODIS instruments.

Producing high quality daily TIR-based LST data at fine-resolution in heterogenous ecosystems, like mountain regions remains highly challenging, and it is still under active debate among the scientific community (Wu et al., 2021). Even though open-source high spatiotemporal resolution TIR missions are at the advance level of preparation, they are not planned to be launched soon. To this end, time-series of the enhanced LST maps are highly needed to temporally complement the future thermal missions as an alternative to the coarse-resolution observations from MODIS or SLSTR instrument.

Despite improved results with the developed 250-m LST product, the estimates of TSEB-PT can be further enhanced by applying more sophisticated approaches and algorithms for retrieval subpixel skin temperatures from coarse resolution TIR imagery. LST modelling over complex landscapes could be further performed at multi-scale level considering explanatory variables with smaller pixel size such as those derived from Landsat or Sentinel-2 observations (Guzinski & Nieto, 2019). Moreover, thermal downscaling might be additionally forced by other meaningful predictors acquired at more frequent basis, such as sub-hourly air temperature grids or 5-day Sentinel-2 observations in order to capture time-induced thermal variability of the surface (Singh et al., 2019). In addition, some documented limitations of random forest should be reduced like the averaging effect of prediction trees towards majority values of an explaining variable (Zhang et al., 2019). Moreover, additional work shall an implementation of more advanced residual correction for fine resolution LST pixels (Njuki et al., 2020).

The proposed reconstruction method under long-term overcast conditions gives important insights about reliable retrieval of cloudy-sky LSTs and proves the incapability of geostatistical interpolation methods to predict those values from neighboring clear-sky pixels. The method depicts cooling impact of thick clouds that limit amount of downwelling solar radiation reaching the Earth. A topography-induced analysis showed that regardless of vegetation type, land surface temperatures under cloudy skies are smaller than for clear-sky observations (see Appendix A1). These findings give future directions for the scientific community into development of new remote sensing-based reconstruction methods that consider that aspect. To this aim, physical-based approaches exploiting coarse resolution data ( $\geq 1$  km) from geostationary satellites, like MSG SEVIRI, along with passive microwave instruments and the ECMWF ERA5 reanalysis datasets can be combined with field-scale observations from meteorological stations for estimating cloud-covered LST maps at sub-kilometer resolution (Long et al., 2020; Martins et al., 2019; Shwetha et al. 2016).

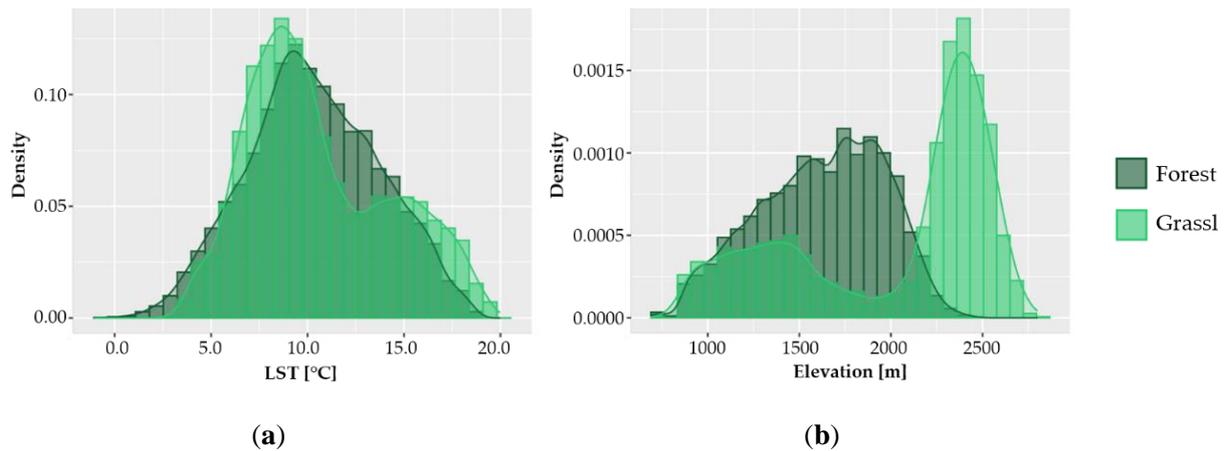
With the recent development of hybrid approaches in the field of Earth system sciences, combining physical- and data-driven modelling opens up new possibilities towards learning and better understanding natural processes with deep learning approaches at more user-controlled manner (Reichstein et al., 2019; Willard et al., 2020). In this context, future research should be focused on parametrizations of physical processes and their augmenting using deep learning models that exploit patterns and associations between sequential data over time. This would be useful not only for site-based LST reconstructions for both long-term and short-term cloudiness at fine-spatial resolution, but also for coupling the TSEB model with deep neural networks as shown in Cui et al. (2021).

As mentioned previously, evapotranspiration forced by thermal remote sensing observations provide valuable information on water cycle for hydrological and agricultural applications. However, estimating high-resolution ET at larger scales is still challenging, which is mainly related to scarcity of open-source FR TIR instruments and cloud-free VNIR observations. Nevertheless, the recent advent of high-

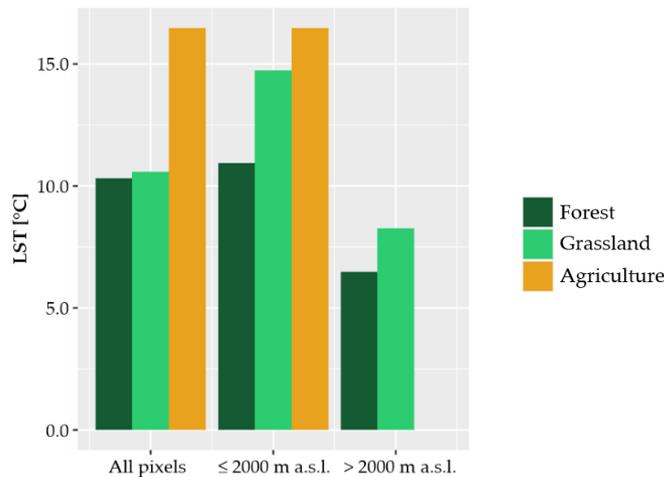
resolution constellations from Sentinel-2 (2A and 2B with planned launch of S2C-D ) and Landsat 8-9 combined with Sentinel-3 SLSTR opens new possibilities for estimating radiometric temperatures with pixel size smaller than 250 m. The need to consolidate the efforts towards better understanding and modeling hydrological cycle at both 1 km and 100 m spatial resolution, including ET component, has been underlined within the framework of the ongoing ESA's 4DMED-Hydrology project initiated in November 2021. In this context, multi-scale ET mapping from TSEB approach was suggested as solid solution over diversified landscape of the Mediterranean region.

**A1. Thermal variability of the reconstruction results over alpine vegetation**

The reconstructed LST images under long-term cloudy sky conditions were additionally analyzed in terms of the topographical impact on LST variability over different vegetation biomes as given in Section 4. Considering terrain variability of grasslands and forests in the study area, Fig. A1.1 shows distribution of surface temperatures and altitude for all reconstructed images.



**Fig. A1.1.** Density plots of (a) reconstructed MODIS LST pixels and (b) their corresponding elevation located in forest and grassland habitats.



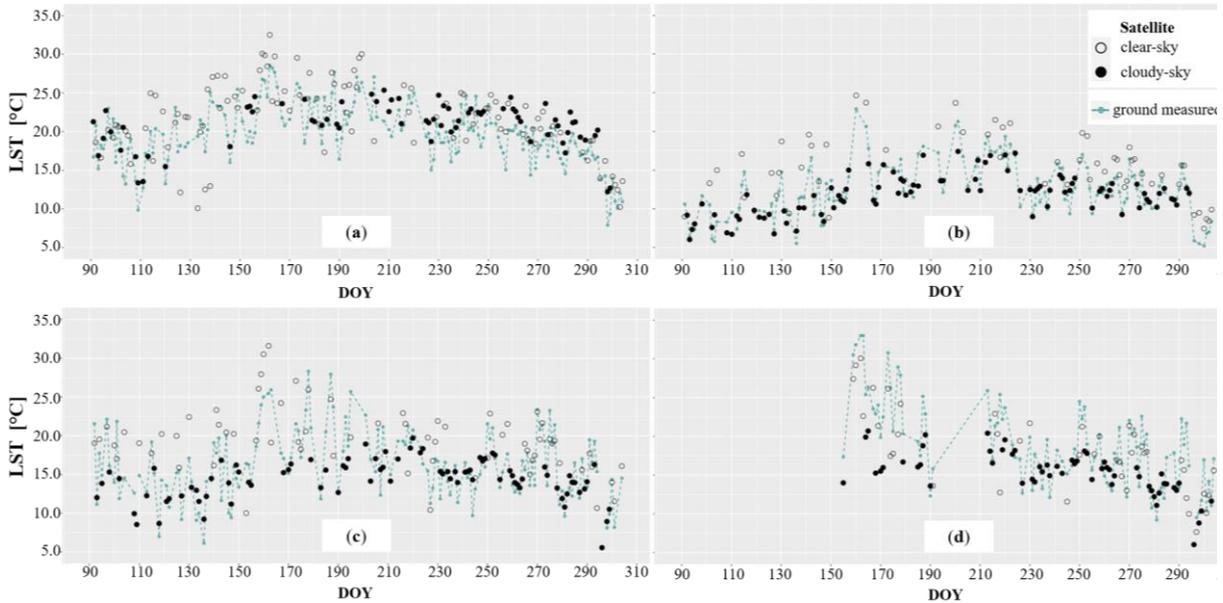
**Fig. A1.2.** Mean values of LST pixels under cloudy skies for the vegetation groups considering different elevation splits.

The obvious altitudinal control on LST was visible for grassland and forest with greater terrain amplitudes than in the areas covered by permanent crops (Fig. A1.2). While both LST and elevation grid cells had Gaussian distribution in the forest, for grassland divergence from a normality was noticed for both plots. Furthermore, if we split grassland pixels according to elevation, we observed increased  $LST_{mean}$  for all altitudinal groups with maximum growth of 39% (Fig. A1.2) when compared to statistics obtained for the entire study area (Tab. A1.1). In comparison, forest had the lowest mean temperatures that varied with topography, however, less pixels at high altitudes (> 2000 m a.s.l.) resulted in the global  $LST_{mean}$  comparable to thermal conditions in the grassland (Tab. A1.1).

Group	Elevation [m a.s.l.] (Standard deviation) [m]	LST [°C] (Standard deviation) [°C]
Permanent crops	887 (77)	16.46 (2.99)
Forest	1620 (341)	10.35 (3.41)
Grassland	2012 (546)	10.59 (3.65)

**Tab. A1.1.** Global analysis of mean and standard deviation obtained from elevation and LST pixels under cloudy skies for the analyzed vegetation groups, averaged in the study area considering all reconstructed MODIS grids.

In parallel with global analysis, a local comparison between original clear-sky MODIS LST (non-reconstructed), LST derived from field measurements, and reconstructed cloudy-sky LST in 2014 was carried out (Fig. A1.3). Due to limited data availability in 2014, in Fig. A1.3 we present timeseries of LST based on four meteorological stations located at permanent crops, forest and grasslands.



**Fig. A1.3.** Intercomparison of reconstructed LST values with both in-situ data and clear-sky MODIS LST pixels for the entire year of 2014 (April-October) collected at the four meteorological stations: (a) Caldarò (240 m a.s.l.), (b) Lavarone (1349 m a.s.l.), (c) Mazia 1 (1450 m a.s.l.), (d) Mazia 2 (1550 m a.s.l.). While green lines with points represent average ground-derived LST, the reconstructions are shown as black squares, and unfilled squares indicate original MODIS LST. Ground-derived LSTs in Mazia 2 were not available in spring resulting in a lack of timeseries observations, as shown in Fig. A1.3d.

As shown in Fig. A1.3, the timeseries from the ground, reconstructed and original clear-sky 1-km MODIS observations follow the seasonal LST patterns with the highest values in the summer months. In general, clear-sky MODIS LST values were higher than their corresponding ground-derived surface temperatures with the mean difference of 1.4°C for all stations. Similarly, lower LSTs were observed for the gap-filled pixels when compared to the original cloud-free satellite observations. As shown in Fig. A1.3, temperature distributions differed at local scale. Temporal variations between the reconstructions and in-situ data throughout the year can be observed for Caldarò site (Fig. A1.3a) leading to substantial biases between observations. Apart from this station, LSTs from ground measurements were generally larger or had practically no bias when compared to the gap-filling results under cloudy-sky-conditions (Fig. A1.3b-d).

- Agam, N., Kustas, W. P., Anderson, M. C., Li, F. and Neale, C. M., **2007**. A vegetation index based technique for spatial sharpening of thermal imagery. *Remote Sens. Environ.*, 107(4), pp. 545-558.
- Allen, R., Irmak, A., Trezza, R., Hendrickx, J. M., Bastiaanssen, W. and Kjaersgaard, J., **2011**. Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrol. Processes*, 25(26), pp. 4011-4027.
- Allen, R., Tasumi, M. and Trezza, R., **2007**. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-Model. *J. Irrig. Drain. Eng.*, 133(4), pp. 380-394.
- Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P. and Mecikalski, J. R., **1997**. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sens. Environ.*, 60(2), pp. 195-216.
- Anderson, M. C., Yang, Y., Xue, J., Knipper, K. R., Yang, Y., Gao, F., Hain, C. R., Kustas, W. P., Cawse-Nicholson, K., Hulley, G., and Fisher, J. B., **2021**. Interoperability of ECOSTRESS and Landsat for mapping evapotranspiration time series at sub-field scales. *Remote Sens. Environ.*, 252, p. 112189.
- Baldrige, A. M., Hook, S. J., Grove, C. I., Rivera, G., **2009**. The ASTER spectral library version 2.0. *Remote Sens. Environ.*, 114, pp. 711-715.
- Bastiaanssen, W. G., Menenti, M., Feddes, R. A. and Holtslag, A. A. M., **1998**. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *J. Hydrol.*, 212, pp. 198-212.
- Belgiu, M., and Stein, A., **2019**. Spatiotemporal image fusion in remote sensing. *Remote Sens.*, 11(7), p. 818.
- Bellvert, J., Jofre-Čekalović, C., Pelechá, A., Mata, M. and Nieto, H., **2020**. Feasibility of using the two-source energy balance model (TSEB) with Sentinel-2 and Sentinel-3 images to analyze the spatio-temporal variability of vine water status in a vineyard. *Remote Sens.*, 12(14), p. 2299.
- Berk, A., Anderson, G. P., Acharya, P. K., Bernstein, L. S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S. M., Chetwynd, J. H., Hoke, M. L., Lockwood, R. B., Gardner, J. A., Cooley, T. W., Borel, C. C., and Lewis, P. E. **2005**. MODTRAN 5: A Reformulated Atmospheric Band Model with Auxiliary Species and Practical Multiple Scattering Options: Update. In S.S. Sylvania, and P.E. Lewis (Eds.), in Proc SPIE, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI (pp. 662-667). Bellingham, WA: Proceedings of SPIE.
- Budzier, H. and Gerlach, G., **2011**. Thermal infrared sensors: theory, optimization and practice. John Wiley & Sons.
- Burchard-Levine, V., Nieto, H., Riaño, D., Migliavacca, M., El-Madany, T.S., Guzinski, R., Carrara, A. and Martín, M. P., **2021**. The effect of pixel heterogeneity for remote sensing based retrievals of evapotranspiration in a semi-arid tree-grass ecosystem. *Remote Sens. Environ.*, 260, p. 112440.
- Burchard-Levine, V., Nieto, H., Riaño, D., Migliavacca, M., El-Madany, T.S., Perez-Priego, O., Carrara, A. and Martín, M. P., **2019**. Adapting the thermal-based two-source energy balance model to estimate energy fluxes in a complex tree-grass ecosystem. *Hydrol. Earth Syst. Sci. Discuss.*, pp. 1-37.
- Castelli M., **2021**. Evapotranspiration Changes over the European Alps: Consistency of Trends and Their Drivers between the MOD16 and SSEBop Algorithms. *Remote Sens.*, 13(21), p. 4316.
- Castelli, M., Anderson, M. C., Yang, Y., Wohlfahrt, G., Bertoldi, G., Niedrist, G., Notarnicola, C., **2018**. Two-source energy balance modeling of evapotranspiration in Alpine grasslands. *Remote Sens. Environ.*, 209, pp. 327-342.
- Cui, Y., Song, L. and Fan, W., **2021**. Generation of spatio-temporally continuous evapotranspiration and its components by coupling a two-source energy balance model and a deep neural network over the Heihe River Basin. *J. Hydrol.*, 597, p. 126176.

- Czajkowski, K. P., Goward, S. N., Mulhern, T., Goetz, S. J., Walz, A., Shirey, D., Stadler, S., Prince, S. D. and Dubayah, R.O., **2004**. Estimating environmental variables using thermal remote sensing. In *Thermal remote sensing in land surface processes* (pp. 10-32).
- Della Chiesa, S., Bertoldi, G., Niedrist, G., Obojes, N., Endrizzi, S., Albertson, J. D., Wohlfahrt, G., Hörtnagl, L. and Tappeiner, U., **2014**. Modelling changes in grassland hydrological cycling along an elevational gradient in the Alps. *Ecohydrology*, 7(6), pp. 1453-1473.
- Dong, P., Gao, L., Zhan, W., Liu, Z., Li, J., Lai, J., Li, H., Huang, F., Tamang, S.K. and Zhao, L., **2020**. Global comparison of diverse scaling factors and regression models for downscaling Landsat-8 thermal data. *ISPRS J. Photogramm. Remote Sens.*, 169, pp. 44-56.
- Duan, S. B., Li, Z. L. and Leng, P., **2017**. A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data. *Remote Sens. Environ.*, 195, pp. 107-117.
- Elfarkh, J., Ezzahar, J., Er-Raki, S., Simonneaux, V., Ait Hssaine, B., Rachidi, S., Brut, A., Rivalland, V., Khabba, S., Chehbouni, A. and Jarlan, L., **2020**. Multi-scale evaluation of the TSEB model over a complex agricultural landscape in Morocco. *Remote Sens.*, 12(7), p. 1181.
- Fisher, J. B., Lee, B., Purdy, A. J., Halverson, G. H., Dohlen, M. B., Cawse-Nicholson, K., Wang, A., Anderson, R. G., Aragon, B., Arain, M. A. and Baldocchi, D. D., **2020**. ECOSTRESS: NASA's next generation mission to measure evapotranspiration from the International Space Station. *Water Resour. Res.*, 56(4), p. e2019WR026058.
- Foken, T., Aubinet, M. and Leuning, R., **2012**. The eddy covariance method. In *Eddy covariance* (pp. 1-19). Springer, Dordrecht.
- Gallo, K., Hale, R., Tarpley, D. and Yu, Y., **2011**. Evaluation of the relationship between air and land surface temperature under clear-and cloudy-sky conditions. *J. Appl. Meteorol. Climatol.*, 50(3), pp. 767-775.
- Gao, F., Kustas, W. P. and Anderson, M. C., **2012**. A data mining approach for sharpening thermal satellite imagery over land. *Remote Sens.*, 4(11), pp. 3287-3319.
- Gillespie, A. R., Rokugawa, S., Hook, S. J., Matsunaga, T. and Kahle, A. B., **1999**. Temperature/emissivity separation algorithm theoretical basis document, version 2.4. ATBD contract NAS5-31372, NASA.
- Good, E. J., **2016**. An in situ-based analysis of the relationship between land surface “skin” and screen-level air temperatures. *J. Geophys. Res. Atmos.*, 121(15), pp. 8801-8819.
- Guzinski, R., Anderson, M. C., Kustas, W. P., Nieto, H. and Sandholt, I., **2013**. Using a thermal-based two source energy balance model with time-differencing to estimate surface energy fluxes with day–night MODIS observations. *Hydrol. Earth Syst. Sci.*, 17(7), pp. 2809-2825.
- Guzinski, R. and Nieto, H., **2019**. Evaluating the feasibility of using Sentinel-2 and Sentinel-3 satellites for high-resolution evapotranspiration estimations. *Remote Sens. Environ.*, 221, pp. 157-172.
- Hulley, G. C., and Ghent, D., eds., **2019**. Taking the temperature of the Earth: steps towards integrated understanding of variability and change. *Elsevier*.
- Hulley, G. C., Hook, S. J., Abbott, E., Malakar, N., Islam, T., Abrams, M., **2015**. The ASTER Global Emissivity Dataset (ASTER GED): mapping Earth’s emissivity at 100 meter spatial scale. *Geophys. Res. Lett.*, 42, pp. 7966–7976.
- Hutengs, C.; Vohland, M., **2016**. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sens. Environ.*, 178, pp. 127–141.
- Jia, L., Xi, G., Liu, S., Huang, C., Yan, Y. and Liu, G., **2009**. Regional estimation of daily to annual regional evapotranspiration with MODIS data in the Yellow River Delta wetland. *Hydrol. Earth Syst. Sci.*, 13(10), pp. 1775-1787.
- Jiménez-Muñoz, J. C. and Sobrino, J. A., **2003**. A generalized single channel method for retrieving land surface temperature from remote sensing data. *J. Geophys. Res.*, 108.

- Jiménez-Muñoz, J. C.; Sobrino, J. A., **2008**. Split-window coefficients for land surface temperature retrieval from low-resolution thermal infrared sensors. *IEEE Geosci. Remote Sens. Lett.*, v. 5, pp. 806-809.
- Jimenez-Munoz, J. C., Sobrino, J. A., **2010**. A single-channel algorithm for land surface temperature retrieval from ASTER data. *IEEE Geosci. Remote Sens. Lett.*, 7, pp. 176–179.
- Kalma, J. D., McVicar, T. R. and McCabe, M. F., **2008**. Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surv. Geophys.*, 29(4), pp. 421-469.
- Kljun, N., Calanca, P., Rotach, M.W. and Schmid, H.P., **2015**. A simple two-dimensional parameterization for Flux Footprint Prediction (FFP). *Geosci. Model Dev.*, 8(11), pp. 3695-3713.
- Kustas, W. P. and Anderson, M., **2009**. Advances in thermal infrared remote sensing for land surface modeling. *Agric. For. Meteorol.*, 149(12), pp. 2071-2081.
- Kustas, W. P., Anderson, M. C., Cammalleri, C. and Alfieri, J. G., **2013**. Utility of a thermal-based two-source energy balance model for estimating surface fluxes over complex landscapes. *Procedia Environ. Sci.*, 19, pp. 224-230.
- Kustas, W. P. and Norman, J. M., **1999**. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. *Agric. For. Meteorol.*, 94(1), pp. 13-29.
- Kustas, W. P. and Norman, J. M., **2000**. A two-source energy balance approach using directional radiometric temperature observations for sparse canopy covered surfaces. *Agron. J.*, 92(5), pp. 847-854.
- Kustas, W. P., Norman, J. M., Anderson, M. C. and French, A. N., **2003**. Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship. *Remote Sens. Environ.*, 85(4), pp. 429-440.
- Künzer, C. and Dech, S. eds., **2013**. Thermal infrared remote sensing: sensors, methods, applications (Vol. 17). Springer Science & Business Media.
- Li, W., Ni, L., Li, Z. L., Duan, S. B. and Wu, H., **2019**. Evaluation of machine learning algorithms in spatial downscaling of MODIS land surface temperature. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 12(7), pp. 2299-2307.
- Li, Z. L., Tang, B H., Wu, h., Ren, H., Yan, G., Wan, Z., Trigo, I., Sobrino, J., **2013**. Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.*, v. 131, n. 15, pp. 14-37.
- Liang, S. and Wang, J. eds., **2019**. Advanced remote sensing: terrestrial information extraction and applications. Academic Press.
- Liu, K., Su, H., Li, X. and Chen, S., **2020**. Development of a 250-m Downscaled Land Surface Temperature Data Set and Its Application to Improving Remotely Sensed Evapotranspiration Over Large Landscapes in Northern China. *IEEE Trans. Geosci. Remote Sens.*
- Long, D., Yan, L., Bai, L., Zhang, C., Li, X., Lei, H., Yang, H., Tian, F., Zeng, C., Meng, X. and Shi, C., **2020**. Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach. *Remote Sens. Environ.*, 246, p. 111863.
- Maeda, E. E., **2014**. Downscaling MODIS LST in the East African mountains using elevation gradient and land-cover information. *Int. J. Remote Sens.*, 35, pp. 3094-3108.
- Mao, Q., Peng, J. and Wang, Y., **2021**. Resolution Enhancement of Remotely Sensed Land Surface Temperature: Current Status and Perspectives. *Remote Sens.*, 13(7), p. 1306.
- Martins, J., Trigo, I. F., Ghilain, N., Jimenez, C., Göttsche, F. M., Ermida, S. L., Olesen, F. S., Gellens-Meulenberghs, F. and Arboleda, A., **2019**. An all-weather land surface temperature product based on MSG/SEVIRI observations. *Remote Sens.*, 11(24), p. 3044.

- Mildrexler, D. J., Zhao, M., Cohen, W. B., Running, S. W., Song, X., Jones, M. O., **2018**. Thermal anomalies detect critical global land surface changes. *J. Appl. Meteorol. Climatol.*, 57, pp. 391–397.
- Neteler, M., **2010**. Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data. *Remote Sens.*, 2, pp. 333–351.
- Nieto, H., Kustas, W. P., Torres-Rúa, A., Alfieri, J. G., Gao, F., Anderson, M. C., White, W. A., Song, L., del Mar Alsina, M., Prueger, J. H. and McKee, M., **2019**. Evaluation of TSEB turbulent fluxes using different methods for the retrieval of soil and canopy component temperatures from UAV thermal and multispectral imagery. *Irrigation science*, 37(3), pp. 389-406.
- Njuki, S. M., Mannaerts, C. M. and Su, Z., **2020**. An Improved Approach for Downscaling Coarse-Resolution Thermal Data by Minimizing the Spatial Averaging Biases in Random Forest. *Remote Sens.*, 12(21), p. 3507.
- Norman, J. M., and Becker, F., **1995**. Terminology in thermal infrared remote sensing of natural surfaces. *Agric. For. Meteorol.*, 77(3-4), pp. 153-166.
- Norman, J. M., Kustas, W. P., and Humes, K. S., **1995**. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agric. For. Meteorol.*, 77(3-4), pp. 263-293.
- Priestley, C. H. B. and TAYLOR, R. J., **1972**. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review*, 100(2), pp. 81-92.
- Rahimi, S., Gholami Sefidkouhi, M. A., Raeini-Sarjaz, M. and Valipour, M., **2015**. Estimation of actual evapotranspiration by using MODIS images (a case study: Tajan catchment). *Arch. Agron. Soil Sci.*, 61(5), pp. 695-709.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J. and Carvalhais, N., **2019**. Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), pp. 195-204.
- Rigollier, C., Bauer, O. and Wald, L., **2000**. On the clear sky model of the ESRA—European Solar Radiation Atlas—with respect to the Heliosat method. *Solar energy*, 68(1), pp. 33-48.
- Ruhoff, A. L., Paz, A. R., Collischonn, W., Aragao, L. E., Rocha, H. R. and Malhi, Y. S., **2012**. A MODIS-based energy balance to estimate evapotranspiration for clear-sky days in Brazilian tropical savannas. *Remote Sens.*, 4(3), pp. 703-725.
- Schotland, R. M., **1955**. The measurement of wind velocity by sonic means. *J. Meteorol.*, 12(4), pp. 386-390.
- Semmens, K. A., Anderson, M. C., Kustas, W. P., Gao, F., Alfieri, J. G., McKee, L., Prueger, J. H., Hain, C. R., Cammalleri, C., Yang, Y. and Xia, T., **2016**. Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach. *Remote Sens. Environ.*, 185, pp. 155-170.
- Sheffield, J., Wood, E. F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., and Verbist, K., **2018**. Satellite remote sensing for water resources management: Potential for supporting sustainable development in data-poor regions. *Water Resour. Res.*, 54(12), pp. 9724-9758.
- Singh, S., Bhardwaj, A., Singh, A., Sam, L., Shekhar, M., Martín-Torres, F. J. and Zorzano, M.P., **2019**. Quantifying the congruence between air and land surface temperatures for various climatic and elevation zones of Western Himalaya. *Remote Sens.*, 11(24), p. 2889.
- Sobrino, J. A., Jimenez-Munoz, J. C., Soria, G., Romaguera, M., Guanter, L., Moreno, J., Plaza, A., Martinez, P., **2008**. Land surface emissivity retrieval from different VNIR and TIR sensors. *IEEE Trans. Geosci. Remote Sens.*, 46, pp. 316–327.
- Su, Z., **2002**. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.*, 6(1), pp. 85-100.
- Swain, S., Wardlow, B. D., Narumalani, S., Tadesse, T. and Callahan, K., **2011**. Assessment of vegetation response to drought in Nebraska using Terra-MODIS land surface temperature and normalized difference vegetation index. *GISci. Remote Sens.*, 48(3), pp. 432-455.

- Wan, Z.; Hook, S.; Hulley, G., **2015**. MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006; NASA EOSDIS LP DAAC; NASA: Washington, DC, USA.
- Wan, Z., Wang, P. and Li, X., **2004**. Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA. *Int. J. Remote Sens.*, 25(1), pp. 61-72.
- Weng, Q., Fu, P. and Gao, F., **2014**. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sens. Environ.*, 145, pp. 55-67.
- West, H., Quinn, N. and Horswell, M., **2019**. Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities. *Remote Sens. Environ.*, 232, p. 111291.
- Willard, J., Jia, X., Xu, S., Steinbach, M. and Kumar, V., **2020**. Integrating physics-based modeling with machine learning: A survey. *arXiv preprint arXiv:2003.04919*, 1(1), pp.1-34.
- Wu, H. and Li, W., **2019**. Downscaling land surface temperatures using a random forest regression model with multipity predictor variables. *IEEE Access*, 7, pp. 21904-21916.
- Wu, P., Yin, Z., Zeng, C., Duan, S. B., Göttsche, F. M., Li, X., Ma, X., Yang, H. and Shen, H., **2021**. Spatially Continuous and High-Resolution Land Surface Temperature Product Generation: A Review of Reconstruction and Spatiotemporal Fusion Techniques. *IEEE Geosci. Remote Sens. Mag.*
- Xia, H., Chen, Y., Li, Y. and Quan, J., **2019**. Combining kernel-driven and fusion-based methods to generate daily high-spatial-resolution land surface temperatures. *Remote Sens. Environ.*, 224, pp. 259-274.
- Xu, S., Cheng, J. and Zhang, Q., **2019**. Reconstructing all-weather land surface temperature using the bayesian maximum entropy method over the Tibetan plateau and Heihe river basin. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 12(9), pp. 3307-3316.
- Yang, Y., Cao, C., Pan, X., Li, X. and Zhu, X., **2017**. Downscaling land surface temperature in an arid area by using multiple remote sensing indices with random forest regression. *Remote Sens.*, 9(8), p. 789.
- Yu, W., Tan, J., Ma, M., Li, X., She, X. and Song, Z., **2019**. An effective similar-pixel reconstruction of the high-frequency cloud-covered areas of Southwest China. *Remote Sens.*, 11(3), p. 336.
- Zhang, H., Nettleton, D. and Zhu, Z., **2019**. Regression-enhanced random forests. *arXiv preprint arXiv:1904.10416*.
- Zhang, G., Xiao, X., Dong, J., Kou, W., Jin, C., Qin, Y., Zhou, Y., Wang, J., Menarguez, M.A. and Biradar, C., **2015**. Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data. *ISPRS J. Photogramm. Remote Sens.*, 106, pp. 157-171.
- Zhang, X., Zhou, J., Göttsche, F. M., Zhan, W., Liu, S. and Cao, R., **2019**. A method based on temporal component decomposition for estimating 1-km all-weather land surface temperature by merging satellite thermal infrared and passive microwave observations. *IEEE Trans. Geosci. Remote Sens.*, 57(7), pp. 4670-4691.
- Zhao, W. and Duan, S. B., **2020**. Reconstruction of daytime land surface temperatures under cloud-covered conditions using integrated MODIS/Terra land products and MSG geostationary satellite data. *Remote Sens. Environ.*, 247, p. 111931.
- Zhou, Y., Xiao, X., Zhang, G., Wagle, P., Bajgain, R., Dong, J., Jin, C., Basara, J. B., Anderson, M. C., Hain, C. and Otkin, J. A., **2017**. Quantifying agricultural drought in tallgrass prairie region in the US Southern Great Plains through analysis of a water-related vegetation index from MODIS images. *Agric. For. Meteorol.*, 246, pp. 111-122.
- Zhu, X., Chen, J., Gao, F., Chen, X. and Masek, J. G., **2010**. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.*, 114(11), pp. 2610-2623.