Article

Not peer-reviewed version

Downscaling Land Surface Temperature at Farm-Scale Using Multi-Model Approaches over Diverse Agro Climatic Zones

Debasish Roy , Bappa Das , Pooja Rathore , Priyabrata Santra , Shovik Deb , Bimal K. Bhattacharya , Ajit Govind , Raguveer Singh Jatav , Deepak Sethi , Tridiv Ghosh , Joydeep Mukherjee , Vinay Kumar Sehgal , Sheshakumar Goroshi , Debashis Chakraborty *

Posted Date: 5 January 2024

doi: 10.20944/preprints202401.0413.v1

Keywords: LST downscaling; TsHARP; TPS; Random forest; spatial downscaling, agro-climate

Article

Downscaling Land Surface Temperature at Farm-Scale Using Multi-Model Approaches over Diverse Agro Climatic Zones

Debasish Roy [1], Bappa Das [1,#], Pooja Rathore [1], Priyabrata Santra [2], Shovik Deb [3], Bimal Kumar Bhattacharya [4], Ajit Govind [5], Raghuveer Jatav [1], Deepak Sethi [1], Tridiv Ghosh [1], Joydeep Mukherjee [1], Vinay Kumar Seghal [1], Sheshakumar Goroshi [6] and Debashis Chakraborty [1,*,##]

[1] Division of Agricultural Physics, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India
[2] Natural Resource Management, ICAR-Central Arid Zone Research Institute, Jodhpur, India
[3] Department of Soil Science and Agricultural Chemistry, Uttar Banga Krishi Viswavidyalaya, Cooch Behar, 736165, West Bengal, India
[4] Space Application Centre-ISRO, Ahmedabad, Gujarat 380015, India
[5] ICARDA, 2 Port Said, Victoria Sq, Ismail El-Shaaer Building, Maadi, Cairo, Egypt
[6] India Meteorological Department, New Delhi, 110003, India
[#] Present Address: ICAR-Central Coastal Agricultural Research Institute, Old Goa-403402, Goa, India
[##] Present Address: CIMMYT, India Office, NASC Complex, DPS Marg, New Delhi 110012
* Correspondence: author: D.CHAKRABORTY@cgiar.org; Tel:- +91-8826743644

Abstract: Land surface temperature (LST) is a critical parameter for land surface and atmospheric interactions. However, the applicability of current land surface temperature estimates for field-level hydrological, agricultural, and ecological operations is still challenging due to their coarse spatiotemporal resolution. In the current article, we have downscaled 100 m LST to 10 m by using high-resolution Sentinel-1,2 optical-microwave data and three different models- 1) Thermal Sharpening (TsHARP); 2) Thin Plate Spline (TPS); and 3) Random Forest (RF). The extensive analysis was performed at agricultural farms in the semi-arid (IARI) region of India during the winter and summer seasons of 2020-21 and 2021-22, arid (CAZRI) and per-humid (UBKV) region (2021-2022). The calibration accuracy of the RF model was shown to better agreement with the coefficient of determination (R^2) of 0.982, root means square error (RMSE) 0.181 and normalized root means square error (RMSE) 0.061 with their lower values of standard errors over three diverse agro-climatic zones. The findings demonstrate that the validation accuracy of models' varied according to the agroclimatic zones, although RF and TPS consistently outperformed as compared to the TsHARP models.

Keywords: LST downscaling; TsHARP; TPS; Random forest; spatial downscaling; agro-climate

1. Introduction

Land surface temperature (LST) is the skin temperature of the earth's surface which is required to study processes such as urban moisture and drought (Wan, Wang and Li, 2004), monitor spatiotemporal dynamics of urban heat islands (Nichol, 2005; Huang and Wang, 2019), and growth status of vegetation (Julien and Sobrino, 2009). Thermal imaging and retrieval of LST variables from space-borne sensors get significant interest in numerous practical applications in remote sensing and the scientific community. LST directs many biophysical processes near the land-atmosphere interface and is a key component in environmental modelling model surface energy fluxes, soil moisture estimation (Sandholt, Rasmussen and Andersen, 2002; Merlin et al., 2010), and drought and vegetation dynamics (Wan, Wang and Li, 2004; Casper and Vohland, 2008; Julien and Sobrino, 2009; Karnieli et al., 2010) at local to global scales. Many studies reported on the importance and aspects of LST on evapotranspiration (Seney et al., 2019), forest fire monitoring (Maffei, Alfieri and Menenti, 2018), geological and geothermal studies (Coolbaugh et al., 2007; Eskandari, De Rosa and Amini, 2015; Mia, Fujimitsu and Nishijima, 2018; Sekertekin and Arslan, 2019). Traditionally, LST is

estimated from meteorological stations based on radiance measurements for point locations (Becker and Li, 1995; Hale et al., 2011) but more recently, thermal remote sensing-based approaches primarily in the Thermal Infrared (TIR) region of the Electromagnetic (EM) spectrum have been used for LST retrieval (Zhao and Li, 2013; Petropoulos et al., 2020; Kamaraj et al., 2021). TIR sensors map the radiation emitted by the earth in a range of 8–15 μm and measure the radiation emitted by the ground to determine the surface temperature. Satellite-based TIR bands contribute both spatiotemporal coverage at local or global scales and are more cost-efficient and accessible as compared to the traditional methods (Gao et al., 2013).

LST parameters are highly dependent on atmospheric parameters such as relative humidity and near-surface air temperature (Mallick et al., 2009) and sensors parameters like spectral range, viewing geometry, and surface emissivity (Dash et al., 2001; Ghent et al., 2019; Liu et al., 2019). Many researchers consider both, atmospheric conditions and surface emissivity as the most important parameters to retrieve LST from remote sensing (Becker and Li, 1990; Gillespie et al., 1998; Zhihao Qin, Karnieli and Berliner, 2001; Jiménez-Muñoz and Sobrino, 2003; Yu, Guo and Wu, 2014). Land Surface Emissivity (LSE) is another dynamic parameter that greatly affects the accuracy of the LST estimation (Li, Wu, et al., 2013; Neinavaz, Skidmore and Darvishzadeh, 2020; Sekertekin and Bonafoni, 2020). The algorithms developed for LST estimation are generally named based on the number of bands used. For example, single-channel or mono-window algorithms use only one thermal band whereas split window requires two TIR bands (Qin et al., 2004; Sobrino and Romaguera, 2004; Meng et al., 2023). Landsat 5 (TM), Landsat 7 ETM+ and Landsat 8 OLI/TIRS data have been used for LST retrieval (Sekertekin, 2019) using Radiative Transfer Equation (RTE) (Becker and Li, 1990; Gillespie et al., 1998; Jiménez-Muñoz and Sobrino, 2003), Single Channel Algorithm (SCA) (Jiménez-Muñoz and Sobrino, 2009) and Mono Window Algorithm (MWA) (Zhihao Qin, Karnieli and Berliner, 2001).

The LST obtained from the current RS-based TIR bands has limited applications due to their coarser resolution (Khanal, Fulton and Shearer, 2017; Gao et al., 2021). Landsat-7, 8 provide LST at medium spatial resolution (~100 m) but a low temporal resolution (16-day interval), whereas, MODIS offers LST from both the Terra and Aqua sensors daily but with a low spatial resolution (~1000 m). NASA's ASTER provides thermal data at a spatial resolution of ~90 m at limited access. Therefore, there is a significant increase in the interest within the space community to enhance the spatial resolution of land surface temperature (LST) maps (Zhan et al., 2013). Spatial downscaling is a strategy that is frequently used to improve the resolution of LST by integrating coarse resolution with high-resolution predictor variables (Atkinson, 2013). Statistics-based models are built on statistical relationships between LST and other surface characteristic data based on ancillary information such as reflectance data in the visible, near-infrared (NIR), and shortwave infrared (SWIR) bands, digital elevation model (DEM) data, and spectral indices available at relatively high spatial resolutions (Xu et al., 2020; Pu and Bonafoni, 2023). Thermal image sharpening (TsHARP) and Thin Plate Spline (TPS) are examples of classical statistical models. TsHARP is a linear regression-based model that uses vegetation indices such as NDVI or fractional vegetation cover as a scale predictor variable for LST sharpening. TPS on the other hand is a powerful and precise deterministic non-geostatistical surface fitting process that adjusts a set of functions to interpolate source point data (coarse resolution LST) while minimizing the smoothness term (Rawat, Sehgal and Ray, 2019). While these models are simplistic in approach however their performance may be limited under complex heterogeneous environments or additional predictor variables. Lately, machine learning (ML) techniques applying advanced nonlinear methods are being tested in LST downscaling (Ebrahimy and Azadbakht, 2019; Xu et al., 2020, 2021). The merits of ML approaches over statistical models include very few assumptions, easy implementation, handling complex and nonlinear relationship downscaled parameter and independent variables and greater reliability.

In this pretext, the current aims at improving the resolution of LST by spatial downscaling of coarser LST using three different models including the statistical and ML-based- TsHARP, TPS and Random Forest (RF) for practical applications. Overall, the main objectives of the study were -1) downscaling Landsat 8 TIR-based LST at 100 m resolution to plot level resolution at 10 m; 2)

 doi:10.20944/preprints202401.0413.v1

comparing the efficiency of the models over different agro-climatic zones, and 3) understanding the LST-NDVI contrasts. The results presented in the study can facilitate the applicability of LSTs in a variety of research domains such as agriculture, hydrology, meteorology, and forestry.

2. Materials and methods

2.1. Study Area

The proposed study was conducted over three experimental sites - 1) ICAR-Indian Agricultural Research Institute (IARI), New Delhi; 2) Central Arid Zone Research Institute (CAZRI), Rajasthan, and 3) Uttar Banga Krishi Viswavidyalaya, West Bengal located at semi-arid, arid, and per-humid agro-climatic regions respectively **(Figure 1).**

2.1.1. Semi-Arid, New Delhi (IARI)

The agricultural farm of the IARI, is situated at New Delhi between 28° 37′ 22.8″ N to 28° 38′ 58.92″ N and 77° 8′ 42.36″ to 77° 10′ 27.84″ E. The farm's total cropped area is 242 ha. The soil of the IARI farm is coarse loamy to fine loamy in texture. The farm has a semi-arid climate with an annual average rainfall of 765 mm, out of which 82.5% (632 mm) is received from June to September through the southwest monsoon. The rest of the rainfall comes during the post-monsoon period through western disturbances characterized by cloudy and foggy conditions over a few days. Wheat crop is dominant during the winter period followed by mustard, gram, and lentil.

2.1.2. Arid, Jodhpur (CAZRI)

The CAZRI agricultural farm is located at Jodhpur city in western Rajasthan extending from 26° 15′ 48.38″N to 26°14′ 26.08″N and 73° 0′ 38.90″ E to 72° 59′ 9.99″ E. The farm is spread across 241 ha which is dominated by agroforestry followed by crops and barren land. The climate is classified as arid with an annual mean temperature is 26.5 °C, mean annual rainfall is 500 mm, and potential evapotranspiration (PET) is 1400-2000 mm.

2.1.3. Per-humid, Coochbehar (UBKV)

The study was also conducted at UBKV, Coochbehar district of West Bengal, India**.** It is extended from 26o 23' to 26o 2 N and 89o 23' E. The bio-climate of the area is humid to per-humid (Bhattacharyya et al., 2008). The area receives 2400 - 3700 mm of rainfall per year, with 70% of that falling between May and September. The average annual maximum and minimum temperature range from 37° to 5° Celsius (Coochbehar district Website, 2020).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

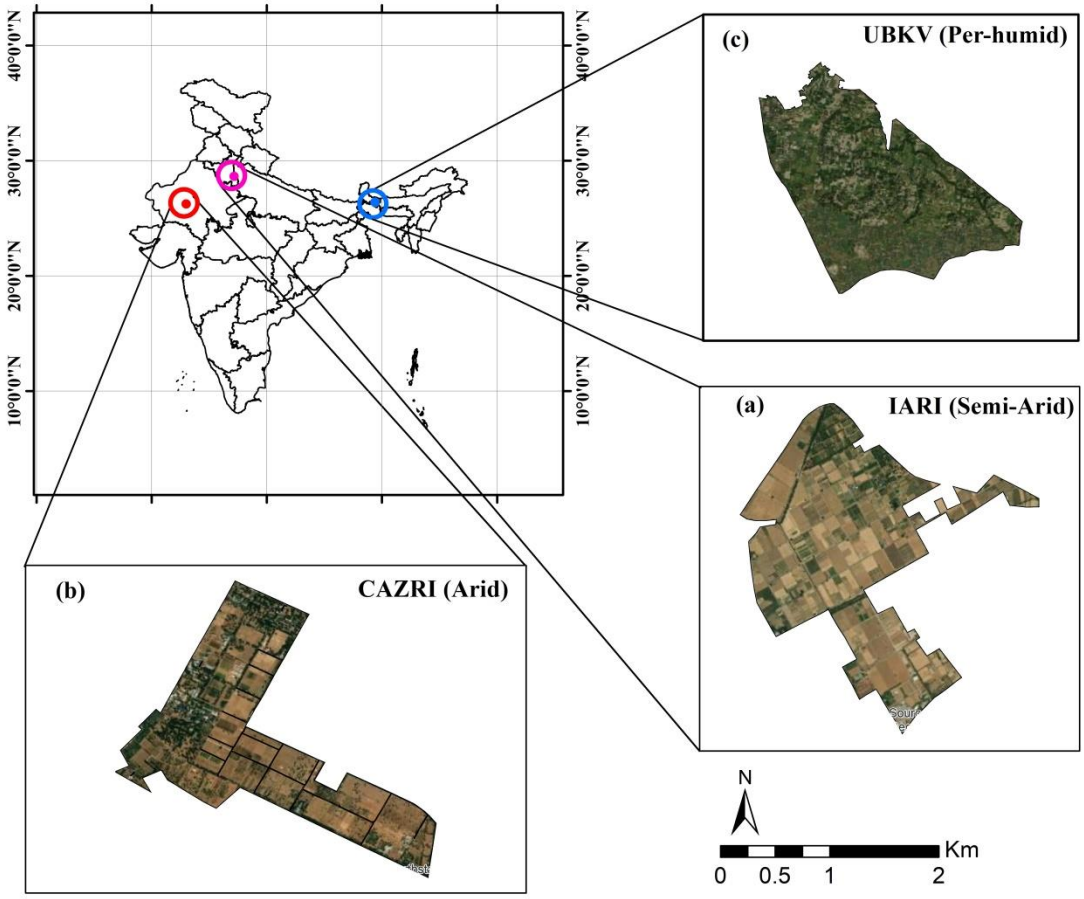


Figure 1. Spatial extent of study area- a) IARI, New Delhi; b) CAZRI, Jodhpur and c) UBKV, coochbehar.

2.2. Satellite data

2.2.1. Optical-thermal-SAR data

The NASA Landsat-8 (L8) products were acquired from the United States of Geological Survey (USGS) Earth Explorer (https://ers.cr.usgs.gov) and ESA Level-1 Ground Range Detected (GRD) C-band synthetic aperture radar (SAR) Sentinel-1 data products in the Interferometric Wide (IW) swath was acquired from the ESA's Sentinel scientific data hub (https://scihub.copernicus.eu/dhus/#/home). Sentinel-2 Level-2 products which provide bottom-of-atmosphere (BOA) reflectance images at 10 m, 20 m and ~10 days temporal resolution was acquired from the same platform. Landsat-8 has 11 spectral bands in the optical and thermal regions with 30 to 100 m spatial and 16 days temporal resolution. For the study, we used red, NIR, from Operational Land Imager (OLI), and thermal bands (Band 10; Thermal Infrared Sensor, TIRS, 10.6-11.19 µm). For IARI, New Delhi, the data were acquired from July 2020 to April 2022 spanning over both kharif and rabi seasons whereas for CAZRI, Jodhpur and UBKV, Coochbehar only the rabi season was covered. A set of total 75 level-1C S2 and Level-1 L8 cloud-free images were selected for this study. The detail on satellite sensors, data products, and dates of acquisition is listed in Table 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

Table 1. Details of satellite data used in the study.

Test sites	Satellite	Sensor specifications	Data products (Resolution)	Date of acquisition
IARI, New Delhi	Sentinel-1	IW beam (VV, VH)	Backscatter coefficient (10 m)	3 July, 13 Sep, 25 Sep (Kharif, 2020); 31 Oct, 24 Nov, 18 Dec, 27 Jan, 20 Feb, 24 Mar (Rabi, 2020-2021); 04 June, 02 July (Kharif, 2021); 19 Nov, 13 Dec, 29 Dec, 15 Feb, 19 Mar, 04 April (Rabi, 2021-2022)
	Sentinel-2	Red band (0.665 µm); NIR band (0.842 µm)	NDVI (10 m)	3 July, 11 Sep, 01 Oct (Kharif-2020); 31 Oct, 20 Nov, 20 Dec; 29 Jan, 18 Feb, and 25 Mar (Rabi, 2020-2021); 08 June, 08 July (2021); 15 Nov, 15 Dec, 25 Dec, 13 Feb, 15 Mar, 04 April (2021-2022)
	Landsat-8	TIRS1 (10.895µm)	LST (100 m)	03 July,12 Sep, 28 Sep, (Kharif, 2020); 30 Oct, 24 Nov, 17 Dec, 27 Jan, 19 Feb, and 23 Mar (Rabi, 2020-2021); 04 June, 06 July (Kharif, 2021); 18 Nov, 13 Dec, 29 Dec, 15 Feb, 19 Mar, 04 April (Rabi, 2021-2022)
CAZRI, Jodhpur	Sentinel-1	IW beam (VV, VH)	Backscatter coefficient (10 m)	05 Mar, 22 Mar, 18 Dec (2021), 20 Feb (2022)
	Sentinel-2	Red band (0.665 µm); NIR band (0.842 µm)	NDVI (10 m)	03 Mar, 23 Mar, 18 Dec (2021), 20 Feb (2022)
	Landsat-8	TIRS1 (10.895 µm)	LST (100 m)	03 Mar, 23 Mar, 18 Dec (2021), 20 Feb (2022)
UBKV, Cooch Behar	Sentinel-1	IW beam (VV, VH)	Backscatter coefficient (10 m)	08 Mar, 24 Mar, 05 Dec, 21 Dec (2021)
	Sentinel-2	Red band (0.665 µm); NIR band (0.842 µm)	NDVI (10 m)	05 Mar, 25 Mar, 5 Dec, 20 Dec (2021)
	Landsat-8	TIRS1 (10.895 µm)	LST (100 m)	08 Mar, 24 Mar, 5 Dec, 21 Dec (2021)

2.3. LST retrieval

LST was retrieved from Landsat-8 band-10 using Mono Window Algorithm (MWA) as proposed by (Z. Qin, Karnieli and Berliner, 2001). MWA is named as such due to its use of single thermal band data for LST retrieval. In contrast to the several atmospheric parameters required in their second correction method, the mono-window algorithm only requires two atmospheric parameters i.e. transmittance and mean atmospheric temperature (Hurtado, Vidal and Caselles, 1996). The LST was retrieved using the "LST" package (Das *et al.*, 2021) in R software version 4.0.2 (R Core Team, 2020).

2.4. LST downscaling

2.4.1. Thermal Sharpening (TsHARP)

In the present study, TsHARP model was used to downscale Landsat-8 thermal data at 100 m spatial resolution to 10 m using red (10 m) and NIR (10 m) reflectance bands obtained from sentinel-2A (Figure 2). TsHARP implies that the exclusive relationship exists between LST and VI using RS data at multiple spatial resolutions (Rawat, Sehgal and Ray, 2019; Pu, 2021). In the first step, TsHARP establish an empirical relationship between coarse resolution NDVI and LST-

$$LST_{coarse} = f_{reg}(VI) + \delta_{reg} \quad (1)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

$$= a * NDVI_{coarse} + b + \delta_{reg}$$

where f_{reg} is *the* regression function, VI is NDVI, coarse suffix represents poor resolution image (here, coarse = 30 m SR), a and b are regression parameters, δ_{reg} is the residual field at poor resolution image pixel (30 m). The regression equation was applied over high-resolution NDVI images (10 m) and high-resolution LST (10 m) was calculated -

$$LST_{fine\_reg} = a * NDVI_{fine} + b \quad (2)$$

Finally, high-resolution LST from TsHARP was obtained by adding residual field (δreg) behind of regression estimate -

$$LST_{TsHARP} = LST_{fine\_reg} + \delta_{reg}$$

$$= LST_{fine\_reg} + LST_{coarse} - LST_{high=reg}$$

$$= LST_{coarse} + a[NDVI_{low} - NDVI_{high}] \quad (3)$$

2.4.2. Thin Plate Spline (TPS)

TPS is a deterministic non-geostatistical surface fitting process or in simple terms, it is an interpolation method for known point data sets, used for downscaling of geo-data in raster format. The fundamental principle of the TPS technique is based on the spatial dependence of geo-data (Bindu *et al.*, 2013). A general form of the TPS model is given by (Rawat, Sehgal and Ray, 2019)as -

$$f_{TPS} = a_0 + a_1 x + a_2 y + \frac{1}{2}\sum_{i}^{n} b_i r_i^2 \, log r_i^2 \quad (4)$$

$$\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} b_i \, x_i = \sum_{i=1}^{n} b_i \, y_i = 0 \quad (5)$$

where, $\boldsymbol{r_i^2} = (\boldsymbol{x} - \boldsymbol{x_i})^2 + (\boldsymbol{y} - \boldsymbol{y_i})^2$ and parameters are obtained by satisfying the following conditions-

$$f_{low\_TPS} = LST_{coarse} \quad (6)$$

$$LST_{TPS\_high} = f_{high\_TPS} \quad (7)$$

2.4.3. Random Forest (RF)

RF is an ensemble approach that assimilates decision trees and bagging models (Breiman, 2001). Based on the interpretation of the study region, the 35 foremost drivers of LST were considered for the RF model (Supplementary Table. S1). The models were adjusted in command to select the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

grouping of the hyperparameters that would give the highest precision. The hyperparameters of the ML models were optimized through 10-fold cross-validation with 5 repetitions using "caret" (Kuhn, 2008), "ranger" (Wright and Ziegler, 2015), "randomForest" (https://cran.rproject.org web/packages\\randomForest/index.html) packages in R version 4.0.2. Feature selection was performed to eliminate unsuitable variables. It ensures that model training is faster and less complicated, with a higher likelihood (Yan et al., 2019). The variable importance for the RF model was tested by the "varImp" function using the "caret" package in RStudio. For downscaling, the input predictors at fine resolution were first aggregated to match the coarse resolution LST data (equivalent to the ~100 m LST of Landsat-8 OLI/TIRS). The statistical relationship between the LST and predictor variables were developed using random forest algorithm as -

$$LST_{100m} = f(S_i, \delta_{VV}, \delta_{VH}) \quad (8)$$

where, S_i is an i[th] variable derived from optical data (Sentinel-2A) and δ_{VV}, δ_{VH} variable derived from SAR data (Sentinel-1), f is a non-linear function fitted by the random forest.

$$\Delta LST_{100m} = LST_{100m} - PLST_{100m} \quad (9)$$

$$PLST_{10m} = PLST_{10m} + \Delta LST_{100m} \quad (10)$$

The details in the description of the linear regression model were successively applied with a finer scale to calculate LST at a high resolution (Hutengs and Vohland, 2016).

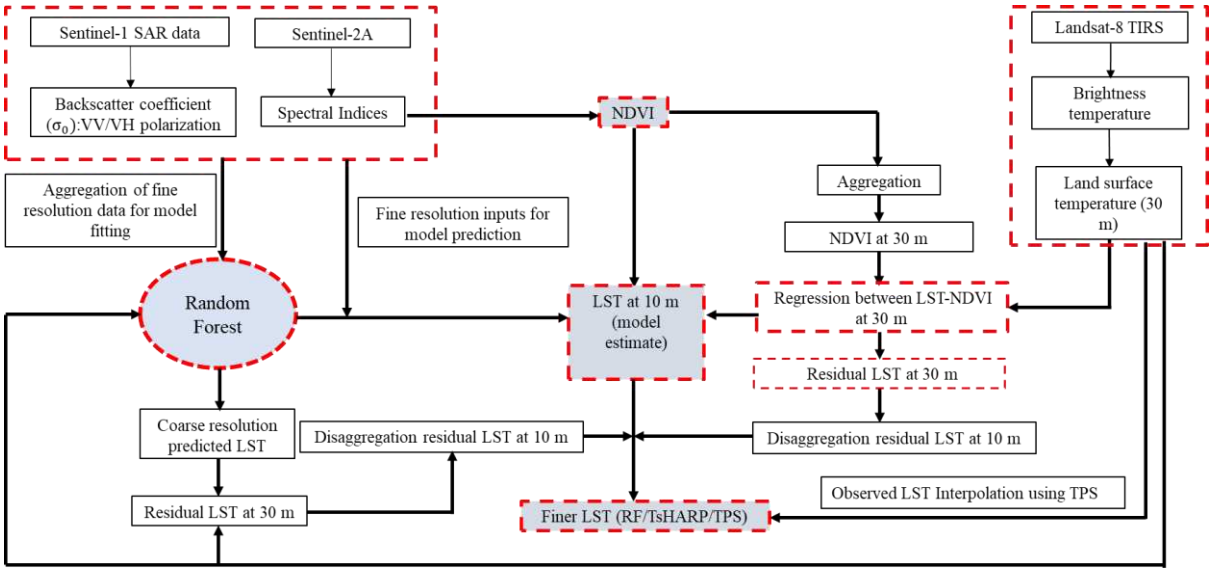


Figure 2. Downscaling methodology of Landsat-8 LST data at 10 m resolution using TsHARP, TPS and RF.

2.4.4. Accuracy evaluation.

Three standard matrices were used to evaluate the model performance coefficient of determination (R^2) (Eq. 11), root means square error (RMSE) (Eq. 12) and Mean Bias Error (MBE) (Eq. 13) given as -

$$\mathbf{R^2} = \left(\frac{\frac{1}{n}\sum_{i=1}^{n}(LST_P - \overline{LST_P})(LST_O - \overline{LST_O})}{\sigma_P \sigma_O} \right) \quad (11)$$

$$\mathbf{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(LST_O - LST_P)^2} \quad (12)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

$$\mathbf{MBE}= \frac{\sum_i^n (LST_O - LST_P)}{n} \quad (13)$$

LST_P is model output; $\overline{LST_P}$ and σ_P are the mean and standard deviation of model outputs respectively, LST_O is observations; $\overline{LST_O}$ and σ_O are the mean and standard deviation of observations, respectively. R^2 values close to 1 and RMSE close to 0 indicate better model performance.

2.5. Measuring parameter sensitivity

Conducting a sensitivity analysis is a crucial step not only in the pursuit of optimization but also in enhancing the predictive capabilities of the machine learning model. Accumulated local effects (ALE) plots were used to demonstrate each predictor's impact on LST results, including any nonlinear response. Because ALEs more accurately reflect interactions between predictors, they were used for the evaluation of main effects rather than conventional partial dependency (PD) plots. The split of training data has an impact on ALE plots, much like it does on VIPs. Results from RF model run are therefore reported as 95th percentile ranges. It is challenging and time-consuming to understand second-order ALE plots, necessitating concurrent evaluation of the main-effect ALE plots. Because PD plots incorporate interaction effects and allow for easier interpretation, they were utilized to evaluate second-order interactions. Although every pairwise interaction was displayed, only those with a significant impact were considered for study.

3. Results

3.1. LST calibration with RF, TsHARP, and TPS

The calibration performance of each model (RF, TsHARP and TPS) evaluated using the statistical parameters- R2, RMSE and nRMSE are summarized in Supplementary Table S2. The LST was calibrated for each observation date in terms of R2, RMSE, and nRMSE, and it showed significant fluctuation for each observation date. The results showed that RF and TPS were the best calibrated for each agroclimatic zone- semi-arid, arid, and per humid followed by TsHARP. TPS achieved its highest performance in the arid region (CAZRI), where the R2, RMSE, and nRMSE values ranged from 0.97 to 0.997, 0.17 to 0.359, and 0.054 to 0.114, respectively. TPS showed highest mean R2 (0.993) for arid zone (CAZRI) while the lowest mean RMSE (0.105) and nRMSE (0.035) was obtained for per-humid region (UBKV). The statistics shown by TPS were consistently high for other sites as well. The RF performance was close, with a high R2 for all sites (0.980 - 0.983). RF generated relatively higher RMSE (0.305 - 0.439) and nRMSE (0.097- 0.13) for the arid site (CAZRI) as compared to other sites. However, when all observation dates were merged, the RF model outperformed the TsHARP and TPS models in terms of SD and SE values. Overall, the RF model demonstrated good and consistent calibration proficiency in the LST downscaling across all three agro-climatic regions.

3.2. Variable importance and sensitivity analysis

The sensitivity analysis performed in the study investigated the impact of selected variables on LST downscaling (Supplementary Figure S1). The variable importance and ALE plot sensitivity of predictor variables included in the process at each experiment site are demonstrated in Fig 3. The trends in variable importance were consistent between the semi-arid zone (IARI) and the arid zone (CAZRI). B11 (SWIR-1) was found to be the most important variable in LST downscaling at both these sites followed by B5, B1 and MNDWI (Figure 3a&b). These variables have a strong effect on LST prediction. The prediction tends to increase with increase in B11 until band value reaches 0.3 and 0.42 at IARI and CAZRI respectively. The relationship between prediction and B1 was discovered to be highly varied, with both being positively and negatively associated (0.05- 0.06 and 0.08- 0.1 at IARI; 0.06 - 0.8 and 0.16 onwards at CAZRI). MNDWI has a negative effect overall i.e. higher the MNDWI, the lower the prediction (Supplementary Figure S1 b).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

However, the degree of different parameter's sensitivity and importance varied significantly at per- humid area (UBKV) highlighting the difference in crop types, soil moisture availability and land use patterns. The coastal aerosol band (B1) was found to be most critical driver variable in LST downscaling at UBKV (Figure 3c) (Supplementary Figure S1c). B3, B9, B6, and B2 are among the other top variables that significantly affected the overall LST estimations at UBKV. B3 and B9 (peaked at 0.2), with their gradual and sharp changes, respectively, have mixed (positive and negative) effects. Between 0.175 and 0.19 band levels, B6 exhibited a significant negative impact on LST prediction.

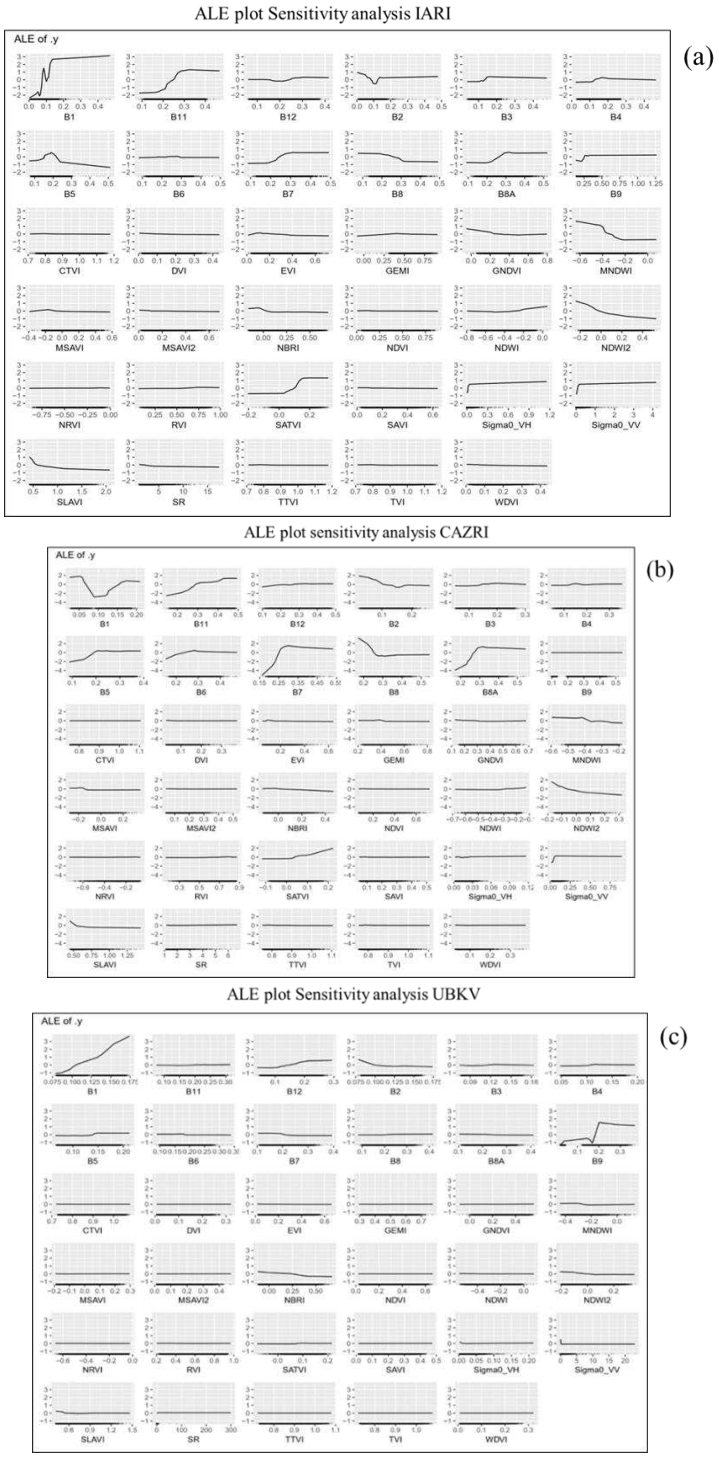


Figure 3. ALE plots for sensitivity analysis of random forest downscaled LST over a) IARI site at pre-monsoon to the post-monsoon period covering both inter and intra-seasonal time scale b) CAZRI and c) UBKV farm scale.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

3.3. Downscaled LST validation with MODIS data

Due to the unavailability of ground-observed LST data (use of a thermal camera in the field mounted on a pole could not be used for satellite LST validation; LST of the pixel of the camera and the corresponding pixel of the satellite showed large variations and disagreement), LST downscaled to 10 m was aggregated to 1000 m to match with MODIS LST (MOD11A1 and MYD11A1) of the same resolution, and was compared for all the sites. The statistical parameters -coefficient of determination, R2; Root Mean Square Error, RMSE', and normalized RMSE for evaluating the accuracy of the models used for downscaling of LST are presented in Table 2 and Supplementary Figure S2. Overall, the RF provided the best predictability in terms of RMSE and nRMSE, compared to the TPS and TsHARP, while R2 values were nearly similar among the three approaches. It is interesting to note that, the machine learning approach (RF) and the TPS and TsHARP methods in the downscaled LST on different dates have quite varied RMSE and nRMSE values.

The downscaled and aggregated LST followed the trends with the MOD11A1 and MYD11A1 LST, although the values were underestimated in relation to both MODIS Aqua and MODIS Terra satellites. The regression line had R2= 0.82 (Aqua) and R2= 0.94 (Terra), but it showed a large deviation from the 1:1 line. At the semi-arid site (IARI) the machine learning based RF performed the best in terms of R2, RMSE, and nRMSE. The RF model had R2 values ranging between 0.44 and 0.93, RMSE between 0.80 and 11.47 K, and nRMSE between 0.27 and 3.84 across all dates of observation (Table 3). R2, RMSE and nRMSE for TsHARP model ranged between 0.39 to 0.93, 0.90 to 11.41 K, and 0.30 to 3.82 respectively, while for TPS the R2, RMSE and nRMSE varied between 0.38 to 0.92, 0.90 to 11.41 K, and 0.30 to 3.82 respectively. It could be noted that the RMSE values on September 12, 2020 at the IARI (semi-arid) location were significantly higher (11.41-11.47 K) for all three downscaling approaches.

At CAZRI, all three models showed a greater agreement (mean R2 >0.84) between the downscaled and reference LST (Table 3). RF method validation had R2 values between 0.75 and 0.91, RMSE between 1.15 and 12.80 K, and nRMSE between 0.36 and 4.30. On the other hand, the TPS method had R2 ranging from 0.75 to 0.92, RMSE 1.10-12.76 K, and the nRMSE 0.35-4.29, whereas the TsHARP method showed R2, RMSE, and nRMSE ranging between 0.75 and 0.89, 1.15 K and 12.94 K, and 0.49 and 4.56, respectively. The RMSE for all three methods were found to be high (12.76 - 12.94) on 20th February 2022 at the CAZRI site. For the UBKV site, the validation R2 in the case of the RF method ranged from 0.57-0.70, RMSE as 0.31- 3.13 K, and nRMSE as 0.11 -1.05. Similarly, in the case of TsHARP, the R2, RMSE, and nRMSE were varying between 0.62 and 0.72, 0.48 K and 3.06 K, and 0.16 and 1.03. The validation R2 for the TPS method ranged between 0.59 and 0.72, RMSE between 0.47 K and 3.10 K, and nRMSE between 0.16 and 1.04.

Table 2. Validation of downscaled aggregated LST with observed LST at 1000 m scale using MODIS terra product.

	RF			TsHARP			TPS		
Date	**R2**	**RMSE**	**nRMSE**	**R2**	**RMSE**	**nRMSE**	**R2**	**RMSE**	**nRMSE**
					IARI				
12.09.2020	0.68	11.47	3.84	0.66	11.41	3.82	0.66	11.41	3.82
28.09.2020	0.59	6.66	2.18	0.39	6.61	2.17	0.38	6.69	2.16
30.10.2021	0.83	4.38	1.45	0.77	4.36	1.45	0.78	4.33	1.44
24.11.2020	0.63	3.23	1.09	0.43	3.29	1.11	0.43	3.30	1.11
19.02.2021	0.55	1.59	0.53	0.57	1.71	0.57	0.56	1.71	0.57
23.03.2021	0.56	5.35	1.75	0.52	5.60	1.83	0.51	5.60	1.83
18.11.2021	0.44	1.27	0.43	0.44	1.26	0.42	0.43	1.24	0.42
15.02.2022	0.53	0.80	0.27	0.57	0.90	0.30	0.55	0.90	0.30
11.03.2022	0.58	1.25	0.57	0.52	2.03	0.68	0.48	2.02	0.68
19.03.2022	0.70	2.04	0.66	0.47	2.16	0.70	0.61	2.08	0.68
04.04.2022	0.93	6.77	2.26	0.93	6.72	2.20	0.92	6.71	2.19
Mean	0.64	4.07	1.37	0.57	4.18	1.39	0.57	4.18	1.38

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

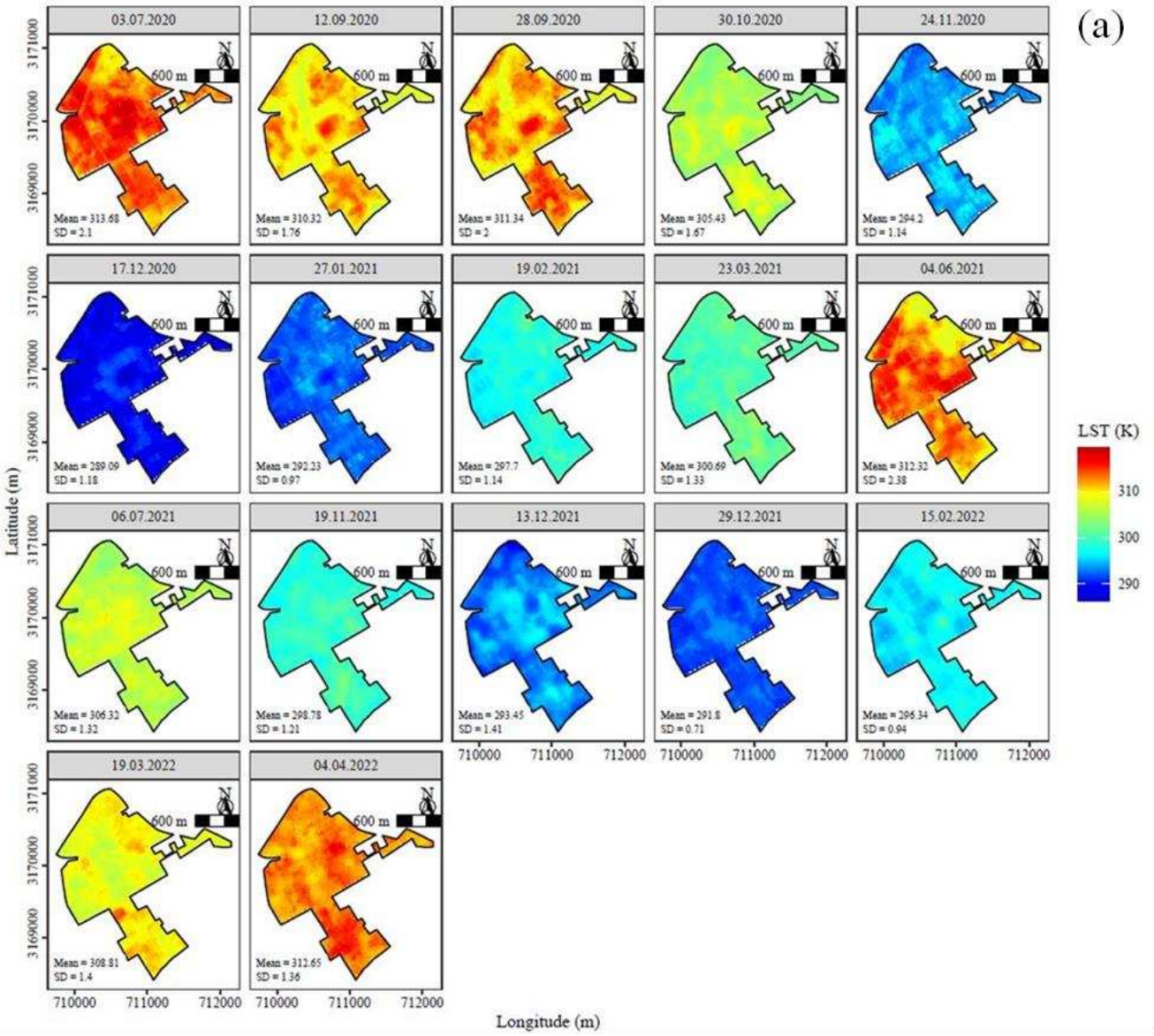
SD	4.52	1.24	1.26	3.54	1.31	1.31	3.56	1.31	1.30
SE	1.36	0.37	0.38	1.07	0.40	0.39	1.07	0.39	0.39
CAZRI									
05.03.2021	0.88	2.82	0.90	0.88	2.77	0.89	0.88	2.75	0.88
21.03.2021	0.75	1.15	0.36	0.75	1.15	0.36	0.75	1.10	0.35
18.12.2021	0.91	1.56	0.52	0.89	1.54	0.52	0.92	1.53	0.52
20.02.2022	0.85	12.80	4.30	0.85	12.80	4.30	0.84	12.76	4.29
Mean	0.85	4.58	1.52	0.84	4.56	1.52	0.85	4.54	1.51
SD	11.90	0.83	0.82	12.81	0.82	0.81	11.43	0.82	0.81
SE	5.95	0.41	0.41	6.41	0.41	0.41	5.71	0.41	0.40
UBKV									
08.03.2021	0.70	0.31	0.11	0.72	0.48	0.16	0.72	0.47	0.16
24.03.2021	0.57	3.13	1.05	0.62	3.06	1.03	0.59	3.10	1.04
Mean	0.64	1.72	0.58	0.67	1.77	0.59	0.66	1.78	0.60
SD	7.04	0.86	0.86	9.48	0.97	0.97	6.91	0.96	0.96
SE	4.98	0.61	0.61	6.70	0.69	0.69	4.89	0.68	0.68

3.4. LST maps at 30 and 10 m resolutions

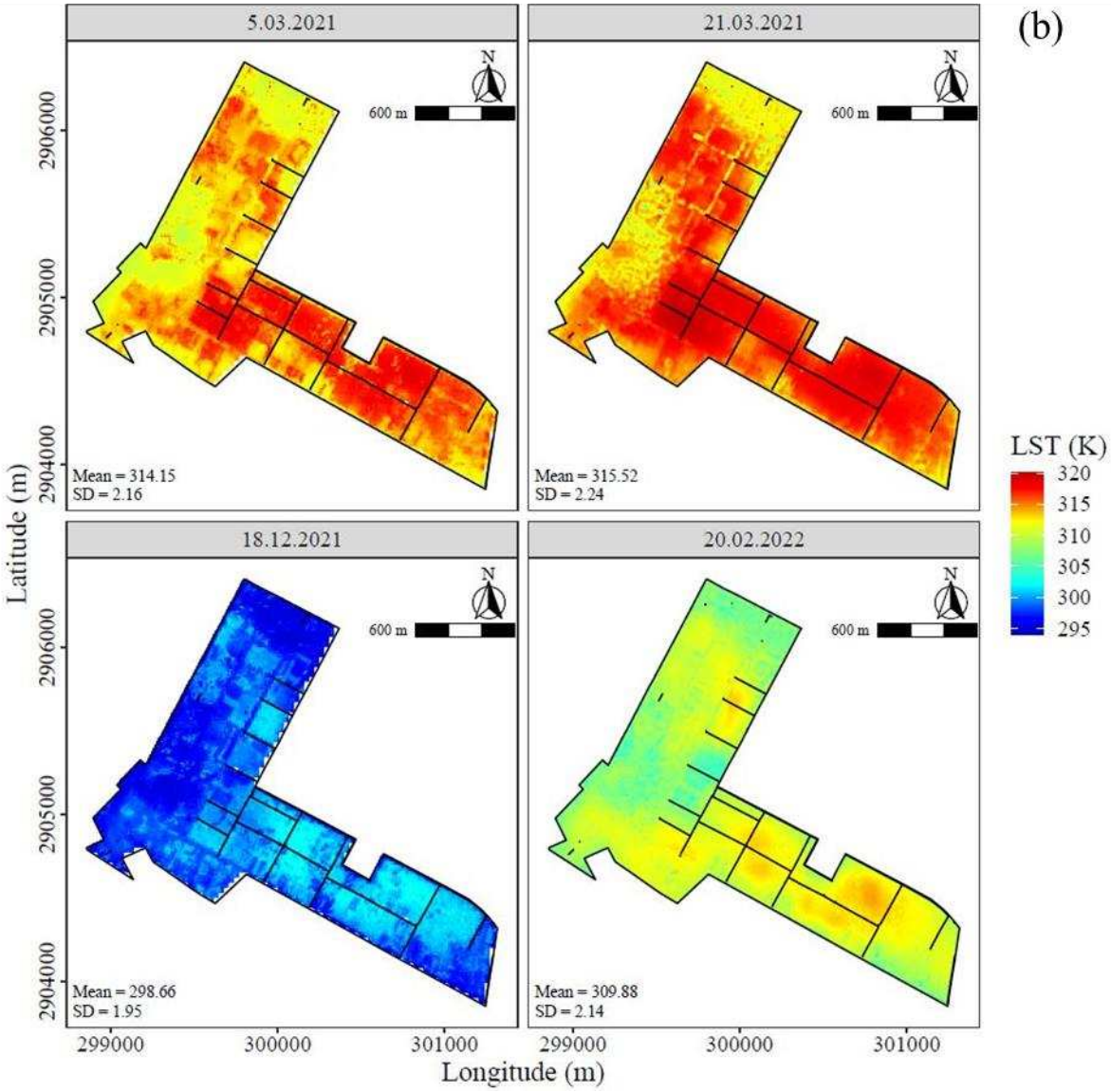
A simultaneous display of downscaled LST maps at 10 m and the Landsat-8 LST (the 'reference') at 30 m scales showed similar patterns and spatial distributions (Figure 4; Supplementary Figures S3 - S7). The LST values could effectively pick up the temperature trends over the season and were evident at both resolutions. At the IARI farm, the image captured LST values as low as 16°C (December 17, 2020) and as high as 41°C (July 3, 2020) (Figure 4a; Supplementary Figures S3 - S5). However, the reference LST had a fuzzy appearance and was homogeneous in distribution. Although the RF and TsHARP-generated 10 m maps appeared similar, the RF had more information, indicating the possibility of analysing plot-scale variations in temperature and identifying between crops and adjacent areas. The performance of TPS was compromised, and there was little improvement in image clarity (Supplementary Figure S4). For instance, the reference map and TPS maps both have hazy appearances in the north and west parts of the IARI farm on March 23, 2021, however RF exhibits higher distinction and clarity. It should be noted here that indistinctiveness was inevitable in RF and TsHARP methods which required residual correction procedures. In *kharif*, when the LST variations across space and time were perceptible, RF produced a better resolution in maps compared to the other two methods.

Similarly, over CAZRI, the mean LST values of images were similar over a specific date, irrespective of the methods, and were also similar to the reference map (Figure 4b; Supplementary Figures S6 - S8). The CAZRI farm is diverse in land use and represents a mixed agricultural farm comprising cropland agroforestry, pasture, and bare land. The pixels appearing in red showed high values while blue pixels referred to low LST. Here also, LST maps derived from RF and TsHARP methods were similar as compared to the TPS and could distinguish the heterogeneity of land use. The reference image had a smoothing effect and was relatively uniform in appearance. It occurred that the centre and south-eastern blocks of CAZRI have greater LST values than the north-western area, which is where crops were grown. For the per-humid region (UBKV), the downscaled LST maps had a homogenous appearance in all the approaches (Figure 4c; Supplementary Figures S9 – S11). The UBKV farm, is predominantly a lowland area with rice crops where the mean LST varied over a narrow range (24 - 29°C). However, the south-eastern block of the farm area had a considerably higher LST value (27°C; March 8, 2021) as it featured orchards, horticultural crops, rubber plantations, etc. intermingled with the rice fields. The north-western block was low-lying paddy areas showing comparatively lower values (~21°C) and was more distinguishable in RF maps.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

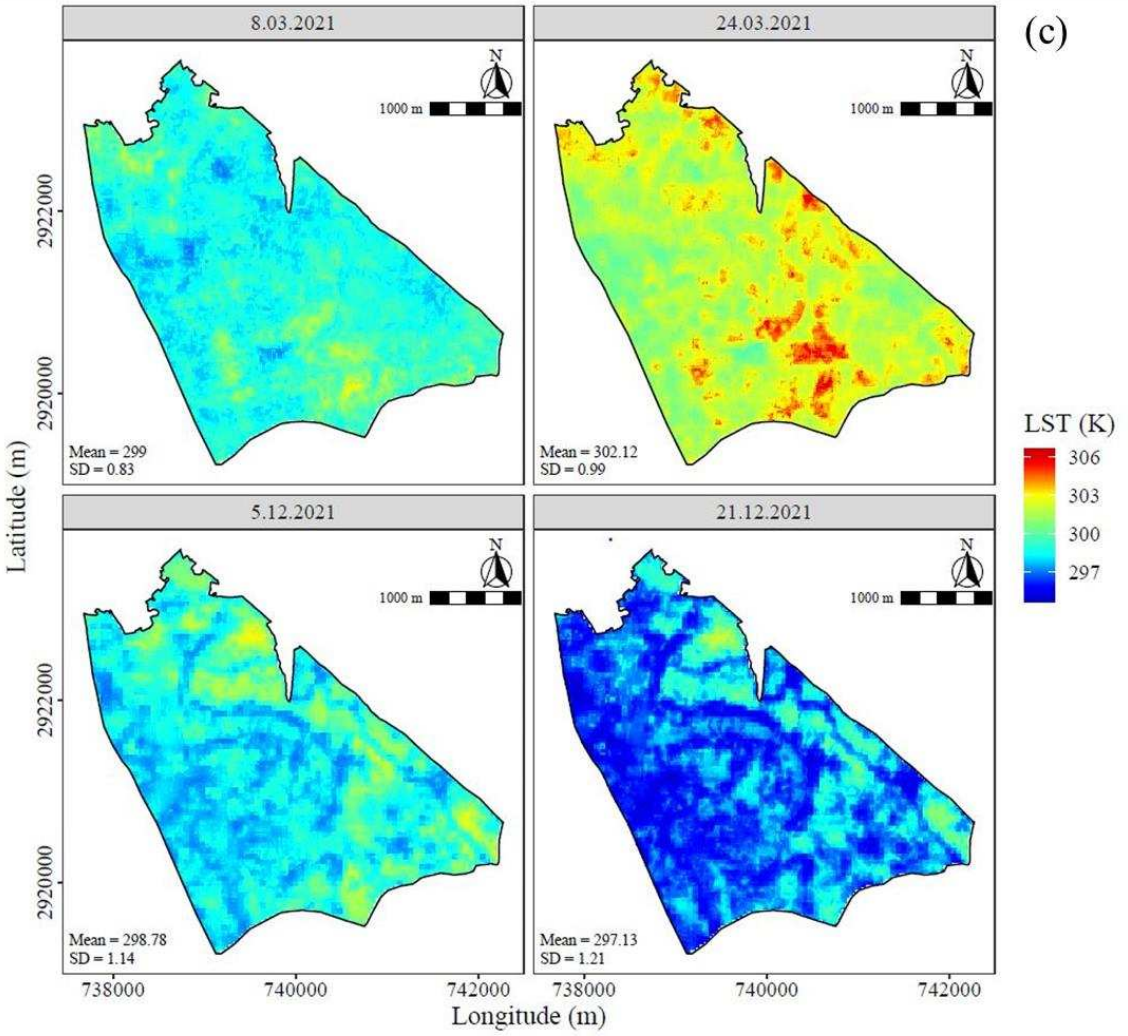


Figure 4. Spatial extent of random forest downscaled LST (10 m) over (a) IARI site at pre-monsoon to the post-monsoon period covering both inter and intra-seasonal time scale (b) CAZRI and (c) UBKV farm scale.

3.5. LST-NDVI profile distribution

The LST and NDVI profile distribution of IARI (averaged over the farm area) for multi date satellite passes demonstrated the inter- and intra-seasonal variations of these parameters for two consecutive years: rabi 2020-21 and 2021-22, and kharif 2020 and 2021 (Figure 5). The NDVI images of all three sites are also presented to understand the spatiotemporal changes in vegetation growth and vigour in the respective farms (Supplementary Figures S12-S14). The trends are typically reversed to the LST values. At the IARI site, blue colours represent cropland with high values of NDVI (>0.5), indicating high levels of green biomass. The red tone, on the contrary, represents the lower values of NDVI corresponding to lesser biomass vigour or in the extreme case, bare land with no or sparse vegetation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 January 2024 doi:10.20944/preprints202401.0413.v1

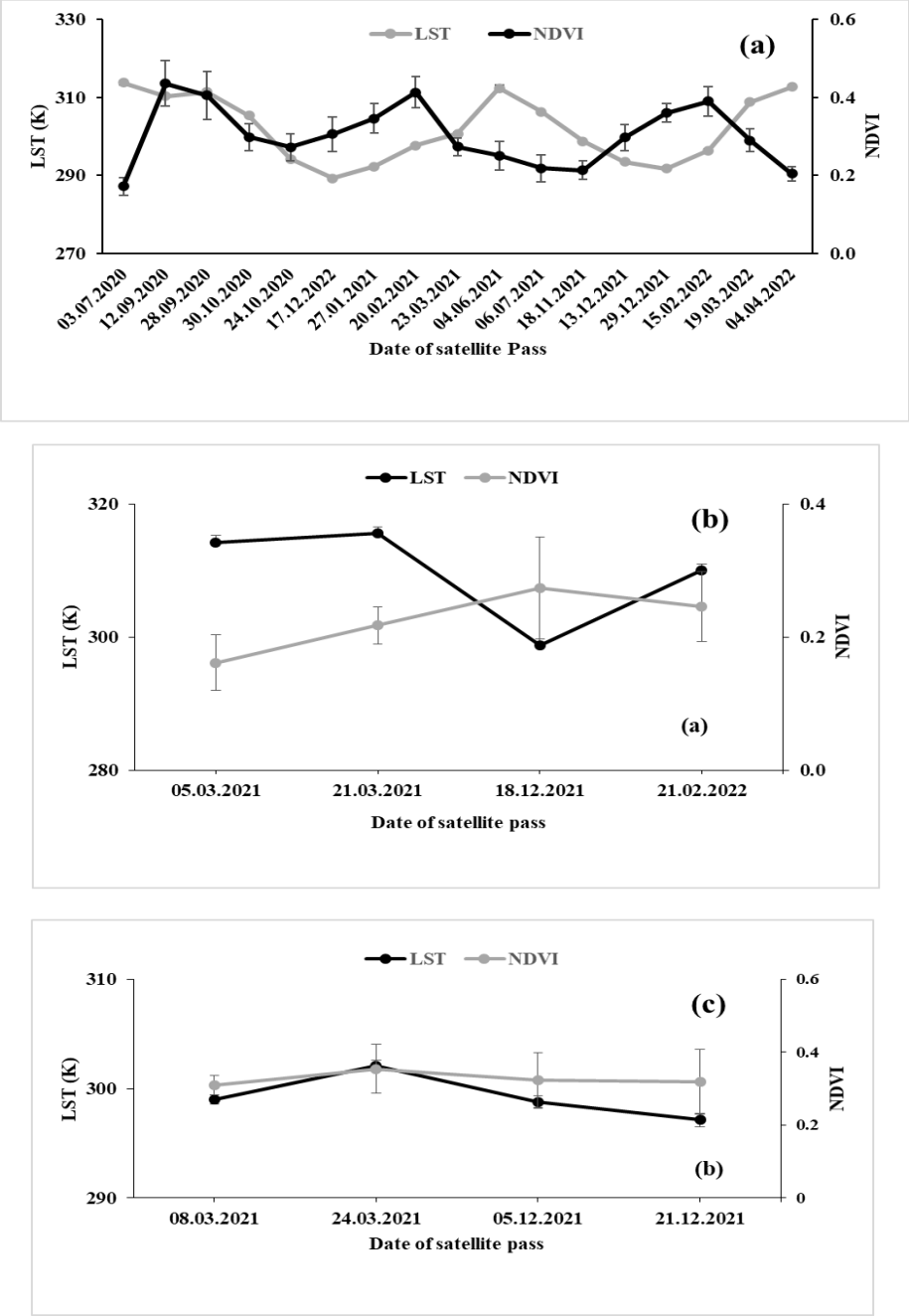


Figure 5. Mean LST-NDVI distribution over a) IARI site at pre-monsoon to the post-monsoon period covering both inter and intra-seasonal time scale b) CAZRI and c) UBKV.

For IARI farm, on July 3, 2020, during the pre-monsoon season, the image displayed the mean maximum LST (313.7 K), while on October 30, 2020, during the post-monsoon season, it displayed 305.3 K (Figure 5a). The LST was lowered by 3.4 and 2.3 K on September 12 and 28, 2020 respectively in monsoon compared to the pre-monsoon period. During winters, the LST decreased sharply and was recorded 24.4 and 21.5 K lower on December 17, 2020 and January 27, 2021 respectively. The mean LST of the farm did not vary significantly across dates, regardless of the downscaling techniques used. Since there were only two dates available for the monsoon 2021, the LST may not have been adequately reflected over the study area, but the mean LST (June 4, 2021) followed a pattern that was comparable to that of the monsoon 2020. In the case of NDVI, the mean value was 0.17 in the pre-monsoon period (July 3, 2020) which increased to 0.44 during end of monsoon season (September 12, 2020) and decreased again (0.30) during the post-monsoon season (October 30, 2020). An increasing NDVI followed by a decreasing trend was the typical characteristic of crop growth

 doi:10.20944/preprints202401.0413.v1

both in the *kharif* and *rabi* seasons owing to phenological changes. This was more explicit in the *rabi* (winter) season owing to a large number of available images.

The mean value of LST over CAZRI was higher on March 21, 2021 (315.6 K) but lower on December 18, 2021 (298.5 K) (Figure 5b). The mean NDVI value was lower on March 5, 2021 (0.16) and higher on December 18, 2021 (0.27) at CAZRI. The lower values of NDVI at the central and south-eastern blocks in the CAZRI farm are dominated by bare soil and pasture land with low biomass vigour (Supplementary Figure S13). However, the north-western part has higher NDVI due to crops and agroforestry. The CAZRI farm had a variety of crops, grasslands, agroforestry, and under different soil moisture (adequately irrigated, limited irrigated, and non-irrigated plots) and followed heterogeneous cropping patterns as compared to semi-arid and per-humid regions. At UBKV site also, LST was higher on March 24, 2021 (302.1 K), while a lower value was recorded on December 5 and 21, 2021 (298.7 and 297.1 K, respectively) (Figure 5c). At UBKV farm, higher NDVI (deep blue patches) at the north-western block indicates to *boro* (winter) rice cultivation (November-May). The red-marked areas at the south-eastern block with low NDVI can be linked to limited biomass or fallow conditions. The NDVI at the UBKV site remained consistent over the dates, likely due to the presence of rice crops in the field.

4. Discussion

4.1. Downscaling of land Surface Temperature: Comparison of methods

Due to the high variability of earth's surface parameters, a priori knowledge of parameters like atmosphere, land surface emissivity (LSE), meteorological conditions, sensor specifications, and satellite-based Land Surface Temperature (LST) retrieval remains a difficult process (Trigo *et al.*, 2008; Li, Tang, *et al.*, 2013; Zhou *et al.*, 2015; Cao *et al.*, 2019). Some studies used Artificial Neural Networks (ANN) and Support Vector Machines (SVM) to retrieve LST and surface soil moisture (Bindhu, Narasimhan and Sudheer, 2013; Santi *et al.*, 2018). However the use of RF algorithm makes empirical models easy to interpret. This study looked at the performance of RF, TsHARP, and TPS in LST downscaling (10 m) at three different agroclimatic zones: semi-arid (IARI), arid (CAZRI), and per-humid (UBKV) having different land cover patterns.

Across locations, the RF algorithm was found to be more sensitive including clear distinction of boundary lines between agriculture and bare soil or soil with minimal vegetation. This study revealed that the downscaling of LST to 10 m resolution by using an RF model encapsulated larger details of the land surface. This was due to the fact that the RF model used 35 multi-covariates as inputs, ranging from bare to complete vegetation data. Considering the findings of the present study and relevant evidence (Yang *et al.*, 2017; Xu *et al.*, 2021), It is possible to argue that RF makes substantial use of information on land surface and exhibits the relative value of several independent variables, such as the visible, near-infrared (NIR), and shortwave (SWIR) bands.

TPS demonstrated good accuracy performance in the calibration and validation for LST downscaling at all three sites in the current study. The findings are consistent with Chen and Li (2012) and Rawat, Sehgal and Ray (2019) which describe TPS as a prevalent technique for spatial interpolating point data sources. Despite, there was a loss of details of downscaled LST using the TPS model. The LST image was smoothed (levelled off) as a result of the land surface information not being taken into account when the LST was downscaled using TPS.

Compared to TPS, the linear classical model (TsHARP) could capture clear boundary conditions mostly in the presence of vegetation. However, when the ground was either fallow or crops reached physiological maturity, employing NDVI as the sole independent variable during LST downscaling by TsHARP provided better uniformity. (Xu et al., 2020) applied TsHARP for downscaling of ASTER 90 m thermal data to 10 m LST with the help of NDVI derived from sentinel-2 and obtained good estimates over the vegetated area (R2=0.853; RMSE=0.806). In our results, the accuracy is highly improved with R2 >0.96 and RMSE <0.18 during calibration at all sites. However, overestimation of downscaled LST at lower temperatures and an underestimation at higher temperatures generated errors in TsHARP as compared to RF. The inability to anticipate extremely high or extremely low

 doi:10.20944/preprints202401.0413.v1

temperatures could be due to a lack of sufficient training data samples (Hutengs and Vohland, 2016), although it was unlikely in this study. Hence, it's suggested that models employing global variables should include the entire LST (Land Surface Temperature) map. This approach might make conventional models less susceptible to fluctuations in local temperature data. Furthermore, among different land cover types, the predicted LST at all three multi-locational farms by using TsHARP showed lower accuracy and consistency with reference LST as compared to the RF approach. Similar results were reported by (Hutengs and Vohland, 2016; Xu et al., 2020; Guo, Hu and Schlink, 2022).

4.2. Influence of predictor variables on LST downscaling

SWIR-1 (B11) was identified as the most sensitive variable for LST downscaling in the relatively drier areas- semi-arid (IARI) and arid (CAZRI). The SWIR reflectance was higher in bare soil where the moisture content is low, while the NIR reflectance was higher where healthy vegetation prevails. Since soil moisture is crucial for maintaining LST, it serves as a reliable indication of land surface temperature conditions (Hulley, Hook and Baldridge, 2010; Colliander *et al.*, 2017). SWIR has provided a strong basis for monitoring SM in specially in semi-arid regions (Fensholt and Sandholt, 2003). The coastal aerosol band (B1) from Sentinel-2, on the other hand, had a higher ranking for the per-humid (UBKV) region due to its stronger penetration in standing water in rice fields (the farm has low-lying topography and rice farming under flooded conditions) (Brodu, 2017; Poursanidis *et al.*, 2019).

Similarly, the red edge band (B5), was also identified as one of the key factor in LST prediction for its ability to describe health status of crops that in turn control the moisture condition (Liu, Qian and Yue, 2021). The outcomes of the study reassure the use red-edge band in drought monitoring in bare soil and low vegetation-covered areas. The water-based indices such as MNDWI (Modified Normalized Difference Water Index) is also found to be useful to the LST downscaling for their sensitivity to seasonal contrast changes of evaporation, precipitation, moisture content, air temperature, etc (Subhanil and Govil, 2021). MNDWI is the modified NDWI that uses green and SWIR bands for monitoring water quality and features and can eliminate the influence of shadow in theory (Li *et al.*, 2021).

SATVI also played an important role in the LST predictions as it can efficiently map photosynthesizing biomass, plant litter, while compensating for varying soil brightness and surface conditions. The sigma nought of vertical-horizontal and vertical-vertical polarization (Sigma0_VH and Sigma0_VV) also had significant impacts on LST downscaling as is indirectly linked to soil water availability (Ulaby, Batlivala and Dobson, 1978; Das et al., 2023).

4.3. Sources of Uncertainties in LST Estimation

Three factors primarily contribute to errors in LST estimation (Xu *et al.*, 2021). The spectrum variations resulting from the differing acquisition dates of Landsat-8 multispectral, Sentinel-1 SAR, and Sentinel-2 multispectral images (lapse) could be the prime source of LST downscaling inaccuracy. Secondly, in the LST retrieval approach, the aggregation processing of the initial (satellite-retrieved) LST data and the explanatory factors produced inadvertent errors. Sentinel-1 and 2 covariates for LST (Landsat-8) downscaling had inherent temporal resolution differences of a few image acquisition dates which were two or three days earlier or later. The third element might have been caused by residual calibration, which was a key strategy for removing output uncertainty from the models. However, the model output uncertainty could not be eliminated by the coarse-resolution residual bilinear resampling technique. As a result, certain LST inaccuracies were produced during the residual calibration procedures.

5. Conclusions

The study evaluated the performance of three models- RF, TsHARP and TPS for downscaling LST (100 m) to plot level resolution (~ 10 m) using Sentinel-1, 2 derived covariates at three different agro-climatic sites. The study aimed at determining whether the ML-based multi-parameter

nonlinear regression model (RF) could outperform the traditional single-parameter-based model (TsHARP) or non-parameter-based model (TPS) in a predominately agricultural land cover. RF model was developed using thirty-five spectral indices as covariates including optical and SWIR bands, as well as backscatter coefficients. The model was calibrated by tuning and optimising the hyperparameters through 10-fold cross-validation for two consecutive rabi and kharif seasons, 2020-21 and 2020-22.

The variable importance analysis was carried out to improve modelling effectiveness and minimize noise. The findings revealed that B1 (Coastal aerosol) is the most relevant metric for the per-humid region whereas B11 (SWIR1) is a key parameter for the semi-arid (IARI) and arid (CAZRI) zones. In addition, B3 (green band in Sentinel 2), B5 and water based index- MNDWI and other variables strongly influenced LST downscaling. The calibration findings revealed that RF generated the highest mean R^2 and the lowest SE for RMSE and nRMSE at all three sites. The random forest (RF) approach provided the best predictability in terms of R^2, RMSE and nRMSE for semi-arid (IARI) and per humid (UBKV) but for arid region (CAZRI), the performance of TPS was marginally better.

An important limitation of the study that could contribute to errors in LST downscaling was varying acquisition dates of Landsat-8 multispectral, Sentinel-1 SAR, and Sentinel-2 multispectral images result in spectrum variations. Based on the evaluation of downscaling accuracy for all three models across all experiment sites, it is reasonable to conclude that RF is a successful and flexible approach that can be employed across a wide range of agro-climates and land cover. The study's findings confirm that LST downscaling at the plot scale can be accomplished using stand-alone and straightforward procedures.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Figure S1. Variable importance in downscaling LST by RF model (a) IARI (b) CAZRI and (c) UBKV. Figure S2. Validation of downscaled-and-aggregated LST from LANDSAT-8 with MODIS Aqua and Terra LST products at 1000 m resolution [1:1 line is shown as dotted line]. Figure S3. Spatial extent of observed LST (30 m) over IARI farm scale in pre- and post-monsoon period, 2020-2022.. Figure S4 Spatial extent of TPS downscaled LST (10 m) over IARI farm scale pre- and post-monsoon period, 2020-2022.. Figure S5 Spatial extent of TsHARP downscaled LST (10 m) over IARI farm scale in the scale pre and post-monsoon period, 2020-2022.. Figure S6 Spatial extent of observed LST (30 m) over CAZRI site.. Figures S7 Spatial extent of TPS downscaled LST (10 m) over CAZRI site. Figure S8 Spatial extent of TsHARP downscaled LST (10 m) over CAZRI site. Figure S9 Spatial extent of observed LST (30 m) over UBKV site. Figure S10 Spatial extent of TPS downscaled LST (10 m) over UBKV site Figure S11 Spatial extent of TsHARP downscaled LST (10 m) over UBKV site. Figure S12. NDVI maps at IARI (semi-arid) site, covered the pre-monsoon to post-monsoon period from 2020-2022. Figure S13 NDVI maps at 10 m over CAZRI (arid) site 2021-2022. Figure S14 NDVI maps at 10 m over UBKV (Per-humid) site 2021. Table S1. Details of satellite covariates from sentinel-1 and 2A used in the study. Table S2. Calibration between downscaled aggregated LST using RF, TsHARP and TPS with observed LST (Landsat-8/9).

Author Contributions: Conceptualization, Bimal K. Bhattacharya and Debashis Chakraborty; Methodology, Debasish Roy, Bappa Das and Pooja Rathore; Formal analysis, Bappa Das, Raguveer Singh Jatav, Deepak Sethi and Tridiv Ghosh; Investigation, Debasish Roy, Bappa Das, Priyabrata Santra, Shovik Deb, Raguveer Singh Jatav, Joydeep Mukherjee and Vinay Kumar Sehgal; Resources, Priyabrata Santra, Shovik Deb, Joydeep Mukherjee, Vinay Kumar Sehgal and Sheshakumar Goroshi; Writing – original draft, Debasish Roy; Writing – review & editing, Pooja Rathore and Ajit Govind; Supervision, Bimal K. Bhattacharya and Debashis Chakraborty; Project administration, Debashis Chakraborty. All authors have read and agreed to the published version of the manuscript.

Funding: research was funded by Indian Council of Agricultural Research, National Fellow..

Data Availability Statement: We encourage all authors of articles published in MDPI journals to share their research data. In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Where no new data were created, or where data is unavailable due to privacy or ethical restrictions, a statement is still required. Suggested Data Availability Statements are available in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics.

Acknowledgments: We thank the Indian Council of Agricultural Research (ICAR), Government of India for the financial support under the National Fellow Scheme. We acknowledge the staff-time and the technical inputs

from CGIAR Research Initiative F2RCWANA. The first author thanks the Post Graduate School, IARI for the fellowship support during the study. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the associated and/or supporting institutions/funders.

Conflicts of Interest: The authors declare no conflict of interest.

References

Atkinson, P.M. (2013) 'Downscaling in remote sensing', *International Journal of Applied Earth Observation and Geoinformation*, 22, pp. 106–114.
Becker, F. and Li, Z.-L. (1990) 'Towards a local split window method over land surfaces', *Remote Sensing*, 11(3), pp. 369–393.
Becker, F. and Li, Z. (1995) 'Surface temperature and emissivity at various scales: Definition, measurement and related problems', *Remote sensing reviews*, 12(3–4), pp. 225–253.
Bhattacharyya, T., Pal, D. K., Chandran, P., Ray, S. K., Mandal, C., & Telpande, B. (2008). Soil carbon storage capacity as a tool to prioritize areas for carbon sequestration. Current science, 482-494
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.
Bindu, S. *et al.* (2013) 'Development of transfer function model and analysis of the peaking capacitor and switch', *IEEE Transactions on Plasma Science*, 41(8), pp. 2415–2420.
Breiman, L. (2001) 'Random forests', *Machine learning*, 45(1), pp. 5–32.
Brodu, N. (2017) 'Super-resolving multiresolution images with band-independent geometry of multispectral pixels', *IEEE Transactions on Geoscience and Remote Sensing*, 55(8), pp. 4610–4617.
Cao, B. *et al.* (2019) 'A review of earth surface thermal radiation directionality observing and modeling: Historical development, current status and perspectives', *Remote Sensing of Environment*, 232, p. 111304.
Casper, M.C. and Vohland, M. (2008) 'Validation of a large scale hydrological model with data fields retrieved from reflective and thermal optical remote sensing data–A case study for the Upper Rhine Valley', *Physics and Chemistry of the Earth, Parts A/B/C*, 33(17–18), pp. 1061–1067.
Chen, C. and Li, Y. (2012) 'A robust method of thin plate spline and its application to DEM construction', *Computers & Geosciences*, 48, pp. 9–16.
Colliander, A. *et al.* (2017) 'Spatial downscaling of SMAP soil moisture using MODIS land surface temperature and NDVI during SMAPVEX15', *IEEE Geoscience and Remote Sensing Letters*, 14(11), pp. 2107–2111.
Coolbaugh, M.F. *et al.* (2007) 'Detection of geothermal anomalies using advanced spaceborne thermal emission and reflection radiometer (ASTER) thermal infrared images at Bradys Hot Springs, Nevada, USA', *Remote Sensing of Environment*, 106(3), pp. 350–359.
Das, B. *et al.* (2021) 'LST: Land Surface Temperature Retrieval for Landsat 8'.
Dash, P. *et al.* (2001) 'Retrieval of land surface temperature and emissivity from satellite data: physics, theoretical limitations and current methods', *Journal of the Indian Society of Remote Sensing*, 29, pp. 23–30.
Ebrahimy, H. and Azadbakht, M. (2019) 'Downscaling MODIS land surface temperature over a heterogeneous area: An investigation of machine learning techniques, feature selection, and impacts of mixed pixels', *Computers & geosciences*, 124, pp. 93–102.
Eskandari, A., De Rosa, R. and Amini, S. (2015) 'Remote sensing of Damavand volcano (Iran) using Landsat imagery: Implications for the volcano dynamics', *Journal of Volcanology and Geothermal Research*, 306, pp. 41–57.
Fensholt, R. and Sandholt, I. (2003) 'Derivation of a shortwave infrared water stress index from MODIS near- and shortwave infrared data in a semiarid environment', *Remote Sensing of Environment*, 87(1), pp. 111–121.
Gao, C. *et al.* (2013) 'A generalized split-window algorithm for land surface temperature estimation from MSG-2/SEVIRI data', *International journal of remote sensing*, 34(12), pp. 4182–4199.
Gao, Z. *et al.* (2021) 'A two-step integrated MLP-GTWR method to estimate 1 Km land surface temperature with complete spatial coverage in humid, cloudy regions', *Remote Sensing*, 13(5), p. 971.
Ghent, D. *et al.* (2019) 'A new approach to defining uncertainties for MODIS land surface temperature', *Remote Sensing*, 11(9), p. 1021.
Gillespie, A. *et al.* (1998) 'A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images', *IEEE transactions on geoscience and remote sensing*, 36(4), pp. 1113–1126.
Guo, F., Hu, D. and Schlink, U. (2022) 'A new nonlinear method for downscaling land surface temperature by integrating guided and Gaussian filtering', *Remote Sensing of Environment*, 271, p. 112915.
Hale, R.C. *et al.* (2011) 'Characterization of variability at in situ locations for calibration/validation of satellite-derived land surface temperature data', *Remote Sensing Letters*, 2(1), pp. 41–50.

Huang, X. and Wang, Y. (2019) 'Investigating the effects of 3D urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: A case study of Wuhan, Central China', *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, pp. 119–131.
Hulley, G.C., Hook, S.J. and Baldridge, A.M. (2010) 'Investigating the effects of soil moisture on thermal infrared land surface temperature and emissivity using satellite retrievals and laboratory measurements', *Remote Sensing of Environment*, 114(7), pp. 1480–1493.
Hurtado, E., Vidal, A. and Caselles, V. (1996) 'Comparison of two atmospheric correction methods for Landsat TM thermal band', *International Journal of Remote Sensing*, 17(2), pp. 237–247.
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.
Jiménez-Muñoz, J.C. and Sobrino, J.A. (2009) 'A single-channel algorithm for land-surface temperature retrieval from ASTER data', *IEEE Geoscience and Remote Sensing Letters*, 7(1), pp. 176–179.
Jiménez-Muñoz, J.C. and Sobrino, J.A. (2003) 'A generalized single-channel method for retrieving land surface temperature from remote sensing data', *Journal of geophysical research: atmospheres*, 108(D22).
Julien, Y. and Sobrino, J.A. (2009) 'The Yearly Land Cover Dynamics (YLCD) method: An analysis of global vegetation from NDVI and LST parameters', *Remote sensing of environment*, 113(2), pp. 329–334.
Kamaraj, N.P. *et al.* (2021) 'Detecting heat-inducing urban built-up surface material with multi remote sensing datasets using reflectance and emission spectroscopy', *Remote Sensing of Environment*, 264, p. 112591.
Karnieli, A. *et al.* (2010) 'Use of NDVI and land surface temperature for drought assessment: Merits and limitations', *Journal of climate*, 23(3), pp. 618–633.
Khanal, S., Fulton, J. and Shearer, S. (2017) 'An overview of current and potential applications of thermal remote sensing in precision agriculture', *Computers and Electronics in Agriculture*, 139, pp. 22–32.
Kuhn, M. (2008) 'Building Predictive Models in R Using the caret Package', *Journal of Statistical Software*, 28(5). Available at: https://doi.org/10.18637/jss.v028.i05.
Li, L. *et al.* (2021) 'A novel surface water index using local background information for long term and large-scale Landsat images', *ISPRS Journal of Photogrammetry and Remote Sensing*, 172, pp. 59–78.
Li, Z.-L., Wu, H., *et al.* (2013) 'Land surface emissivity retrieval from satellite data', *International Journal of Remote Sensing*, 34(9–10), pp. 3084–3127.
Li, Z.-L., Tang, B.-H., *et al.* (2013) 'Satellite-derived land surface temperature: Current status and perspectives', *Remote sensing of environment*, 131, pp. 14–37.
Liu, Y. *et al.* (2019) 'Enterprise LST algorithm development and its evaluation with NOAA 20 data', *Remote Sensing*, 11(17), p. 2003.
Liu, Y., Qian, J. and Yue, H. (2021) 'Comprehensive Evaluation of Sentinel-2 Red Edge and Shortwave-Infrared Bands to Estimate Soil Moisture', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, pp. 7448–7465. Available at: https://doi.org/10.1109/JSTARS.2021.3098513.
Maffei, C., Alfieri, S.M. and Menenti, M. (2018) 'Relating spatiotemporal patterns of forest fires burned area and duration to diurnal land surface temperature anomalies', *Remote Sensing*, 10(11), p. 1777.
Mallick, K. *et al.* (2009) 'Latent heat flux estimation in clear sky days over Indian agroecosystems using noontime satellite remote sensing data', *Agricultural and Forest Meteorology*, 149(10), pp. 1646–1665.
Meng, X. *et al.* (2023) 'Estimating hourly Land Surface Temperature from FY-4A AGRI using an Explicitly Emissivity Dependent Split-Window Algorithm', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* [Preprint].
Merlin, O. *et al.* (2010) 'An improved algorithm for disaggregating microwave-derived soil moisture based on red, near-infrared and thermal-infrared data', *Remote Sensing of Environment*, 114(10), pp. 2305–2316.
Mia, M.B., Fujimitsu, Y. and Nishijima, J. (2018) 'Monitoring of thermal activity at the Hatchobaru–Otake geothermal area in Japan using multi-source satellite images—With comparisons of methods, and solar and seasonal effects', *Remote Sensing*, 10(9), p. 1430.
Neinavaz, E., Skidmore, A.K. and Darvishzadeh, R. (2020) 'Effects of prediction accuracy of the proportion of vegetation cover on land surface emissivity and temperature using the NDVI threshold method', *International Journal of Applied Earth Observation and Geoinformation*, 85, p. 101984.
Nichol, J. (2005) 'Remote sensing of urban heat islands by day and night', *Photogrammetric engineering and remote sensing*, 71(5), pp. 613–621.
Petropoulos, G.P. *et al.* (2020) 'Evaluating the capabilities of optical/TIR imaging sensing systems for quantifying soil water content', *Geocarto International*, 35(5), pp. 494–511.
Poursanidis, D. *et al.* (2019) 'On the use of Sentinel-2 for coastal habitat mapping and satellite-derived bathymetry estimation using downscaled coastal aerosol band', *International Journal of Applied Earth Observation and Geoinformation*, 80, pp. 58–70.
Pu, R. (2021) 'Assessing scaling effect in downscaling land surface temperature in a heterogenous urban environment', *International Journal of Applied Earth Observation and Geoinformation*, 96, p. 102256.
Pu, R. and Bonafoni, S. (2023) 'Thermal infrared remote sensing data downscaling investigations: An overview on current status and perspectives', *Remote Sensing Applications: Society and Environment*, p. 100921.

Qin, Z. *et al.* (2004) 'Comparison of split window algorithms for land surface temperature retrieval from NOAA-AVHRR data', in *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 3740–3743.
Qin, Zhihao, Karnieli, A. and Berliner, P. (2001) 'A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region', *International journal of remote sensing*, 22(18), pp. 3719–3746.
Qin, Z., Karnieli, A. and Berliner, P. (2001) 'A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region', *International Journal of Remote Sensing*, 22(18), pp. 3719–3746. Available at: https://doi.org/10.1080/01431160010006971.
R Core Team (2020) 'R: A Language and Environment for Statistical Computing'. Vienna, Austria.
Rawat, K.S., Sehgal, V.K. and Ray, S.S. (2019) 'Downscaling of MODIS thermal imagery', *The Egyptian Journal of Remote Sensing and Space Science*, 22(1), pp. 49–58.
Sandholt, I., Rasmussen, K. and Andersen, J. (2002) 'A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status', *Remote Sensing of environment*, 79(2–3), pp. 213–224.
Santi, E. *et al.* (2018) 'On the synergy of SMAP, AMSR2 AND SENTINEL-1 for retrieving soil moisture', *International journal of applied earth observation and geoinformation*, 65, pp. 114–123.
Sekertekin, A. (2019) 'Validation of physical radiative transfer equation-based land surface temperature using Landsat 8 satellite imagery and SURFRAD in-situ measurements', *Journal of Atmospheric and Solar-Terrestrial Physics*, 196, p. 105161.
Sekertekin, A. and Arslan, N. (2019) 'Monitoring thermal anomaly and radiative heat flux using thermal infrared satellite imagery–A case study at Tuzla geothermal region', *Geothermics*, 78, pp. 243–254.
Sekertekin, A. and Bonafoni, S. (2020) 'Land surface temperature retrieval from Landsat 5, 7, and 8 over rural areas: Assessment of different retrieval algorithms and emissivity models and toolbox implementation', *Remote sensing*, 12(2), p. 294.
Sobrino, J.A. and Romaguera, M. (2004) 'Land surface temperature retrieval from MSG1-SEVIRI data', *Remote sensing of environment*, 92(2), pp. 247–254.
Subhanil, G. and Govil, H. (2021) 'Relationship between land surface temperature and normalized difference water index on various land surfaces: A seasonal analysis', *International journal of engineering and geosciences*, 6(3), pp. 165–173.
Trigo, I.F. *et al.* (2008) 'An assessment of remotely sensed land surface temperature', *Journal of Geophysical Research: Atmospheres*, 113(D17).
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', *International journal of remote sensing*, 25(1), pp. 61–72.
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* [Preprint].
Xu, J. *et al.* (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, N. *et al.* (2021) 'Changes in the urban surface thermal environment of a Chinese coastal city revealed by downscaling MODIS LST with random forest algorithm', *Journal of Meteorological Research*, 35(5), pp. 759–774.
Yan, H., Song, X., Tian, K., Gao, J., Li, Q., Xiong, Y., & Min, S. (2019). A modification of the bootstrapping soft shrinkage approach for spectral variable selection in the issue of over-fitting, model accuracy and variable selection credibility. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 210, 362-371.
Yang, J. *et al.* (2017) 'Assessing the impacts of urbanization-associated green space on urban land surface temperature: A case study of Dalian, China', *Urban Forestry & Urban Greening*, 22, pp. 1–10.
Yu, X., Guo, X. and Wu, Z. (2014) 'Land surface temperature retrieval from Landsat 8 TIRS—Comparison between radiative transfer equation-based method, split window algorithm and single channel method', *Remote sensing*, 6(10), pp. 9829–9852.
Zhan, W. *et al.* (2013) 'Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats', *Remote Sensing of Environment*, 131, pp. 119–139.
Zhao, W. and Li, Z.-L. (2013) 'Sensitivity study of soil moisture on the temporal evolution of surface temperature over bare surfaces', *International Journal of Remote Sensing*, 34(9–10), pp. 3314–3331.
Zhou, J. *et al.* (2015) 'Developing a temporally land cover-based look-up table (TL-LUT) method for estimating land surface temperature based on AMSR-E data over the Chinese landmass', *International Journal of Applied Earth Observation and Geoinformation*, 34, pp. 35–50.