Estimation of Minimum Land Surface Temperature Using MODIS LST Product for the Himalayan region

Smarika Kulshrestha[1] and RAAJ Ramsankaran[1, 2]

[1] Inter Disciplinary Program (IDP) in Climate Studies, Indian Institute of Technology Bombay, Mumbai-400076, India

[2] Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai-400076, India
E-mail: ramsankaran@civil.iitb.ac.in

KEY WORDS: Land surface Temperature, MODIS, Validation

ABSTRACT: Surface temperature distribution is a governing meteorological factor for studies investigating the hydrological and climatic behavior of a basin. In the Himalayan mountainous region, it is difficult to obtain the high resolution spatial records of the surface temperature because of the scarce and widely scattered meteorological stations. In this study, the potential of MODIS (Moderate Resolution Imaging Spectroradiometer) land surface temperature (LST) products such as MOD11A2 and MYD11A2 Collection 6 has been investigated for estimating accurate 8-Day minimum ground temperature (8DayMinT) values over Western Himalayan region. The forward, backward and stepwise variable selection methods have been employed for 8 major variables (including: LST products (four variables), altitude, latitude, longitude, and Julian day) and regression models have been formulated for 8DayMinT estimations. The results show that the regional topography explains most of the differences between the MODIS LST and the ground temperature records derived from the 7 climate stations in the Himalayan region. The best results for 8DayMinT estimation has been achieved when a combination of all the 4 LST products of TERRA and AQUA, along with altitude, latitude and longitude data is employed.

1. Introduction

The ground temperature (TG) that is usually measured at about 2m from the land surface as a point data with the help of automatic weather stations is an extremely important meteorological driver. Its application includes a wide range of areas like agriculture, hydrology, ecology, environment and climate change. In regions like Himalayan Mountains, there is a sparse network of ground-based weather stations mostly because of the inaccessibility of these areas. The interpolation methods to estimate temperature might not give the desired accuracy in such regions because of a high variability of terrain elevation (Vancutsem *et al*., 2010).

Remote sensing data offers a huge potential to overcome this limitation with the availability of various satellite based Land Surface Temperature (LST) products, such as from Advanced Very High Resolution Radiometer (AVHRR), Visible Infrared Imaging Radiometer Suite (VIIRS), Advanced Along-Track Scanning Radiometer (AATSR), Moderate Resolution Imaging Spectroradiometer (MODIS), and Spinning Enhanced Visible and Infrared Imager (SEVIRI), at a high spatial and temporal resolution often at free of cost (Liu *et al*., 2015; Noi *et al*, 2016; Zou *et al*., 2014; Zhu *et al*., 2013). Among these, the MODIS sensor on board AQUA and TERRA satellites can provide daily LST data with a high temporal (four times per day, TERRA LST day-time, TERRA LST night-time, AQUA LST day-time, AQUA LST night-time, with overpass local time at around 10:30 a.m., 10:30 p.m., 1:30 a.m. and 1:30 p.m., respectively) and very high spatial resolution (1 km). Researchers from different regions of the world have stated that there is a strong linear correlation between MODIS LST and T_G over many regions, e.g., in Africa (Vancutsem *et al*., 2010), in Portugal (Benali *et al*., 2012), over the U.S. and Canada (Zeng *et al*., 2015; Hachem *et al*., 2012, Gallo *et al*., 2011) and in China (Zou *et al*., 2014; Zhu *et al*., 2013).

Zakšek and Homscheidt (2009) have presented a review of the methods used for estimation of TG based on LST. They have been divided into 4 types: simple statistical approaches, advanced statistical approaches, the temperature-vegetation index (TVX) approach and energy-balance approaches. Simple statistical approaches are usually based on a simple linear regression between the LST and TG. The difference between LST and T_G is a strong function of the surface characteristics and the atmospheric conditions (Noi *et al*, 2016). Zhang *et al*. (2016) concluded that for daily TG estimation, TERRA LST and AQUA LST give the same results. Benali *et al*. (2012) showed that the combination of AQUA LST day-time and LST night-time gives a better accuracy of T_G-max and T_G-min estimation, respectively than TERRAs in Portugal. In contrast, Zhu *et al*. (2013) stated that TERRA LST day-time and TERRA night-time, were better than AQUA LST day-time and night-time for TG estimation. Noi *et al*. (2016) concluded that the AQUA LST night-time is the best predictor for both TG-max and TG-min estimation and best estimates of TG were obtained using a multiple regression approach using all the 4 LST products. These differences arise mainly on account of the geographical location of these studies affecting the relationship between LST products and also the time-period of the study. In some regions, the difference between LST and TG is high

(Lai *et al*., 2012; Gallo *et al*., 2011), whereas in some regions it is comparatively small (Noi *et al*., 2016). The detailed information of this difference, as well as the possible causes of this difference, are still limited and needs to be studied.

There are few studies over the Himalayan region that have used satellite-based LSTs to derive ground temperature (T_G) (Zhou *et al*., 2016; Rafiq *et al*., 2014), however, an integrated approach using both TERRA and AQUA LST of day-time and night-time for T_G estimation (using all four MODIS LST datasets) has not been employed in these studies. Therefore, in order to better understand the association between T_G and MODIS LST products, in this study, the relationship has been analyzed between T_G and both TERRA LST and AQUA LST products for the Himalayan region which has scarce availability of the meteorological data. This study has been limited to the estimation of 8-day minimum ground temperature (8DayMinT) based on the 8-day AQUA and TERRA MODIS LST products, mainly for the purpose of employing the 8DayMinT to correct the MODIS obtained snow-cover products (MOD10A2 and MYD10A2).

2. Study Area and Datasets

2.1 Study Area

The study area (Fig. 1) is a part of the Satluj Basin limited to the Indian Territory of the Western Himalayas. The total geographical area of Satluj Basin up to Bhakra dam is about 56,980 km^2, of which about 37,153 km^2 lies in Tibet. The remaining about 19,827 km^2 area lies in the Indian Territory. The Indian part of the Satluj basin is elongated in shape. Elevation of the catchment varies widely from about 500 m to 7,000 m above msl, although only a very small area exists above 6,000 m. The mean elevation of the basin is about 3,600m. The gradient is very steep near its source and gradually reduces downstream. Owing to the large differences in seasonal temperatures and a great range of elevation in the catchment, (BBMB, 1988), this basin is a representative of the mountainous Western Himalayan basins. Therefore, it has been chosen as the study site.

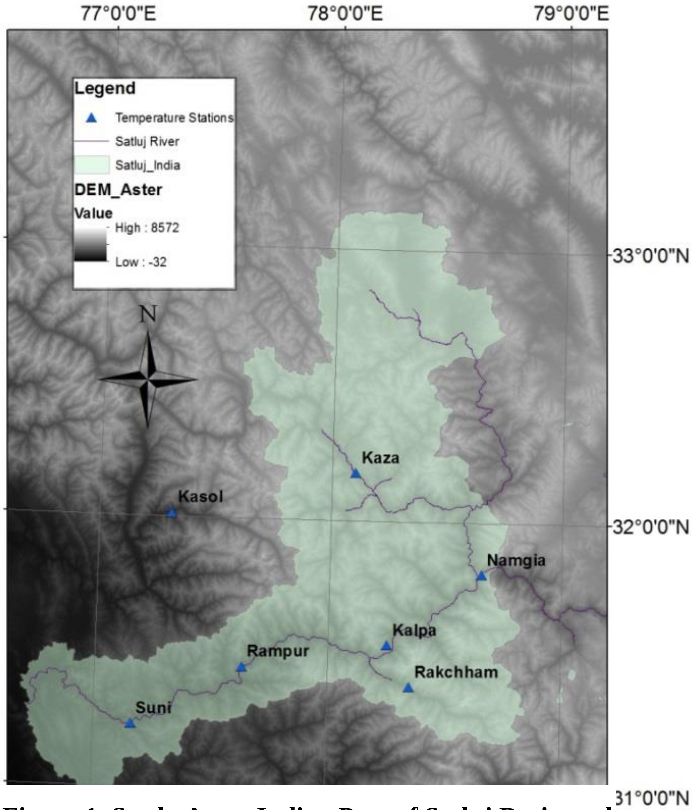
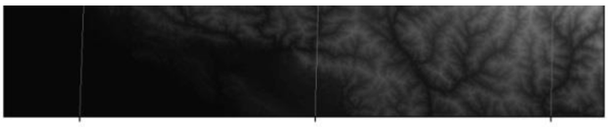


Figure 1. Study Area: Indian Part of Satluj Basin and the location of the meteorological stations.



2.2 Meteorological Data

Figure 1 also shows the locations of 7 meteorological stations located in this region. These stations are monitored by Bhakra Beas Management Board (BBMB) and Irrigation and Public Health (IPH) Himachal Pradesh. These stations provide data of daily minimum and maximum air temperature. The location and elevation of these stations are shown in the Table. 1. The analysis in this study has been done for the year 2006. The daily observed minimum T_G, has been used to identify the 8DayMinT as the minimum value of the observed minimum T_G in a static 8-day window, corresponding to the 8-Day MODIS LST window. Therefore there are 46 values of 8DayMinT for every meteorological station in a year.

Table 1. Location and Altitude for the available Meteorological Stations

S.No.	Station Name	Latitude (°)	Longitude (°)	Altitude (m)
1	Kaza	32.1763	78.1044	3541
2	Rakchham	31.3919	78.3543	3131
3	Namgia	31.8101	78.6563	2843

4	Kalpa	31.5439	78.2554	2731
5	Kasol	32.0117	77.3146	1580
6	Rampur	31.4517	77.6330	972
7	Suni	31.2303	77.1642	765

2.3 Remote Sensing Data

MODIS LST products (h24v06, Collection 6) such as MOD11A2 and MYD11A2 8-day land surface temperature and emissivity from the TERRA and AQUA satellites, respectively covering the Western Himalayas for the year 2006 have been used in this study. The 8-day data are generated from the daily 1-kilometer LST products (MOD11A1 and MYD11A1) and stored on a 1-kilometer Sinusoidal grid as an average value of daily LSTs during an 8-day period. There exist 4 LST data records per day, two from the TERRA satellites and two others from the AQUA satellites, which pass over the study site (local solar time) around 10:30 a.m., 10:30 p.m. and 1:45 a.m., 1:45 p.m., respectively. These times correspond well with the daily maximum and minimum T_G. In total, there are 184 images (in HDF format, for the year 2006) downloaded from the Level-1 and Atmosphere Archive and Distribution System (LAADS) Distributed Active Archive Center (DAAC) (https://ladsweb.modaps.eosdis.nasa.gov/).

2.4 DEM

Elevation data is obtained from ASTER Global DEM. This data is available from the U.S. Geological Survey (USGS) with a spatial resolution of 30 m. These altitude data were resized to 1-km resolution using the nearest neighbor resampling technique in order to be associated with MODIS LST data.

3. Methodology

A multivariate linear regression (MLR) analysis has been used for estimation of 8-Day minimum ground temperature from MODIS LST products (both TERRA and AQUA). MLR approach is chosen because it can be applied for sparse meteorological station networks like in the Himalayas. Although some interpolation methods may give a higher accuracy results, they are not possible for regions with poorly-distributed station networks or physical models that require an unreliable amount of input data (Collins, 1995). In addition to the 4 LST products (MODIS Aqua, Terra – Day and Night), the other predictors employed to describe the relationship between LST and T_G are altitude, latitude, longitude, and Julian day. Altitude (Alt) derived from the DEM is an important predictor for temperature since higher elevations are associated with lower temperatures. Latitude and Longitude are the location parameters that influence the spatial variation of the temperature estimates. Julian day (JD) serves as a proxy for the fraction of solar energy absorption during the day and emitted energy during the night, influencing the diurnal amplitude of T_G throughout the year.

The complete methodology is presented in Fig. 2. All the MODIS data were downloaded from USGS in HDF format (Hierarchical Data Format), with each file containing 12 data layers, the MODIS Reprojection Tool (MRT) has been used to extract the corresponding bands (LST_Day_1km and LST_Night_1km) from both MOD11A2 and MYD11A2. MODIS LST day and night was extracted from MOD11A2 and MYD11A2 data for the pixels in which the meteorological stations are located. The pixels that are lightly or moderately cloud contaminated will create errors in the clear-sky LST (Wan, 2006) and hence, all unrealistic LST data that had values greater than 100°C and below −50°C has been removed.

In order to predict the most significant predictors out of the 8 parameters (LST products (four variables), altitude, latitude, longitude, and Julian day), as well as their order of significance, the forward, backward and stepwise variable selection method proposed by Noi *et al.* (2016) has been used. Forward selection starts with no variable in the model (intercept only model). In the next steps, the most significant variables are added to the model one by one. The process stops when all of the variables not in the model have a p-value greater than 0.15. Backward elimination starts with the model including all variables. In the next step, the least significant variable will be removed. The procedure continues until all of the variables in the model have p-values less than or equal to 0.15. The stepwise method adds or removes a variable in each step, depending on its p-value. This process continues until all variables within the model have a p-value ≤ 0.15, and all of the variables that were not in the model have a p-value >0.15.

Afterwards, different models have been formulated for 8DayMinT estimation using the identified important predictors. The model parameters for each model have been estimated using 70% dataset for the year 2006 and these models have been tested using the remaining 30% data for the year 2006. These model evaluation has been done using the coefficient of determination (r^2), the root mean square error (RMSE) and the mean absolute error (MAE). Because RMSE is very sensitive to outliers, MAE was chosen as an additional measure of the model quality.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(T_{a,i} - T_{es,i})^2} \quad (1)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|T_{a,i} - T_{es,i}\right| \quad (2)$$

Where $T_{a,i}$ is the observed land surface temperature from weather stations and $T_{es,i}$ is the corresponding land surface temperature estimated using the linear regression analysis methods.

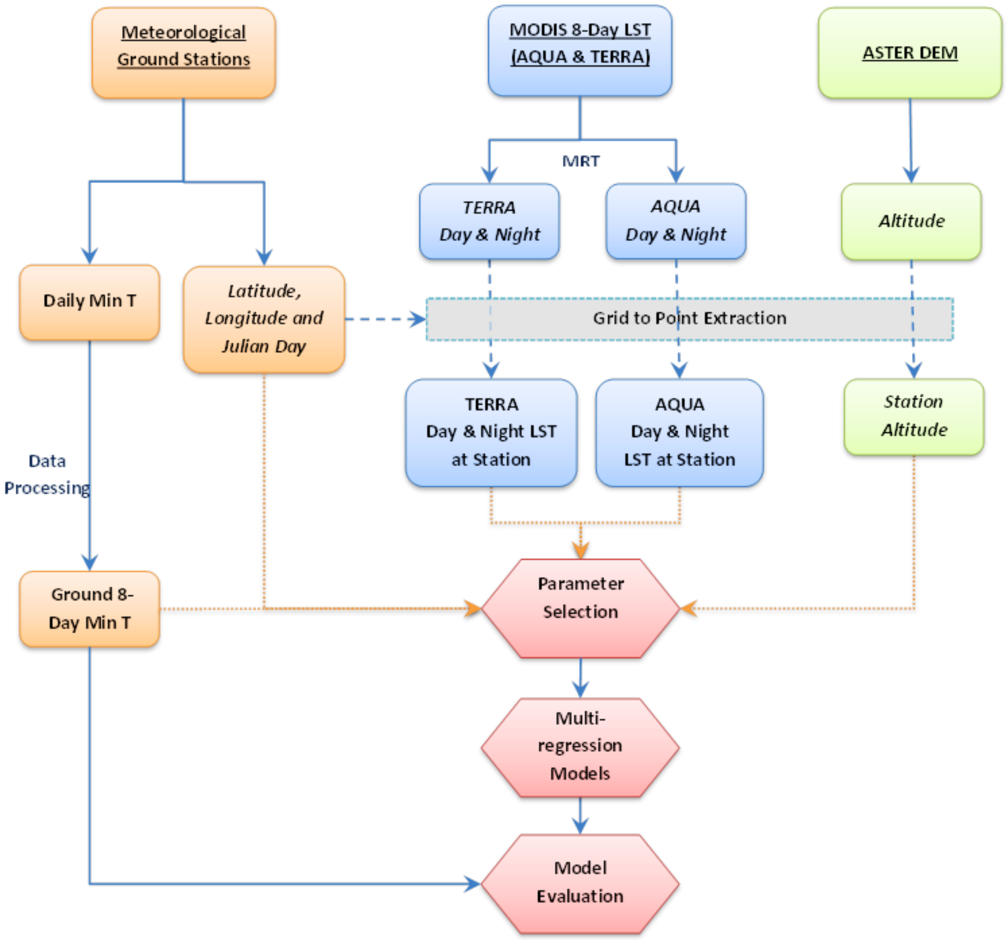


Figure 2. Methodology Flowchart

4. Results and Discussions

4.1 Variable Selection and Model Development

The results of forward, backward and stepwise selection indicates that the most significant predictors in their order of significance are: Terra_Night LST, Aqua_Night LST, Latitude, Aqua_Day LST, Longitude, Altitude, and Terra_Day LST. The Julian Day is found to be an insignificant predictor for MLR based estimation of 8-Day min Temperature because of its very low p-value.

Based on the above analysis, the following models have been formulated for the 8DayMinT estimation:

Model 1: 8DayMinT = a×Terra_Night + b (3)

Model 2: 8DayMinT = a×Aqua_Night + b (4)

Model 3: 8DayMinT = a×Terra_Night + b×Aqua_Night + c (5)

Model 4: 8DayMinT = a×Terra_Night + b×Aqua_Night + c×Lat + d (6)

Model 5: 8DayMinT = a×Terra_Night + b×Aqua_Night + c×Lat + d×Aqua_Day + e (7)

Model 6: 8DayMinT = a×Terra_Night + b×Aqua_Night + c×Lat + d×Aqua_Day + e×Lon + f (8)

Model 7: 8DayMinT = a×Terra_Night + b×Aqua_Night + c×Lat + d×Aqua_Day + e×Lon + f×Alt + g (9)

Model 8: 8DayMinT = a×Terra_Night + b×Aqua_Night + c×Lat + d×Aqua_Day + e×Lon + f×Alt + g×Terra_Day +h (10)

Where, a, b, c, d, e, f, g, h are the respective model coefficients.

4.2 8-Day T-min Estimation

The parameters for all the 8 models are determined when the Models 1-8 were applied to the calibration dataset (70% data for the year 2006). The results of this analysis is shown in Table 2.

Table 2. Parameters of the Models for 8DayMinT estimation obtained during model development

		Estimate	Std. Error	t-value	p-value
Model 1	(Intercept)	-0.5920	0.3322	-1.7820	0.0759
	Terra_Night	0.9292	0.0279	33.2570	0.0000
Model 2	(Intercept)	1.1047	0.3330	3.3170	0.0010
	Aqua_Night	0.8978	0.0301	29.8310	0.0000
Model 3	(Intercept)	-0.2473	0.3504	-0.7060	0.4809
	Terra_Night	0.6878	0.0908	7.5730	0.0000
	Aqua_Night	0.2503	0.0897	2.7890	0.0057
Model 4	(Intercept)	-81.2692	28.3368	-2.8680	0.0045
	Terra_Night	0.6850	0.0896	7.6470	0.0000
	Aqua_Night	0.2824	0.0892	3.1650	0.0017
	Latitude	2.5570	0.8942	2.8590	0.0046
Model 5	(Intercept)	-103.5669	29.2011	-3.5470	0.0005
	Terra_Night	0.7917	0.0970	8.1610	0.0000
	Aqua_Night	0.2793	0.0881	3.1680	0.0017
	Latitude	3.3182	0.9278	3.5760	0.0004
	Aqua_Day	-0.1156	0.0431	-2.6860	0.0077
Model 6	(Intercept)	-354.7507	55.2476	-6.4210	0.0000
	Terra_Night	1.0040	0.1007	9.9690	0.0000
	Aqua_Night	0.3492	0.0849	4.1140	0.0001
	Latitude	3.9378	0.8903	4.4230	0.0000
	Aqua_Day	-0.2719	0.0506	-5.3750	0.0000
	Longitude	2.9971	0.5699	5.2590	0.0000
Model 7	(Intercept)	-416.8000	74.2400	-5.6140	0.0000
	Terra_Night	0.9666	0.1050	9.2100	0.0000
	Aqua_Night	0.3475	0.0848	4.0970	0.0001
	Latitude	4.4330	0.9736	4.5530	0.0000
	Aqua_Day	-0.2560	0.0521	-4.9110	0.0000
	Longitude	3.6130	0.7527	4.8000	0.0000
	Altitude	-0.0007	0.0006	-1.2500	0.2120
Model 8	(Intercept)	-423.6000	74.3500	-5.6980	0.0000
	Terra_Night	0.9235	0.1101	8.3850	0.0000
	Aqua_Night	0.3242	0.0867	3.7410	0.0002
	Latitude	4.7970	1.0130	4.7330	0.0000
	Aqua_Day	-0.3214	0.0732	-4.3940	0.0000
	Longitude	3.5570	0.7530	4.7240	0.0000
	Altitude	-0.0010	0.0006	-1.6110	0.1084
	Terra_Day	0.1086	0.0853	1.2740	0.2038

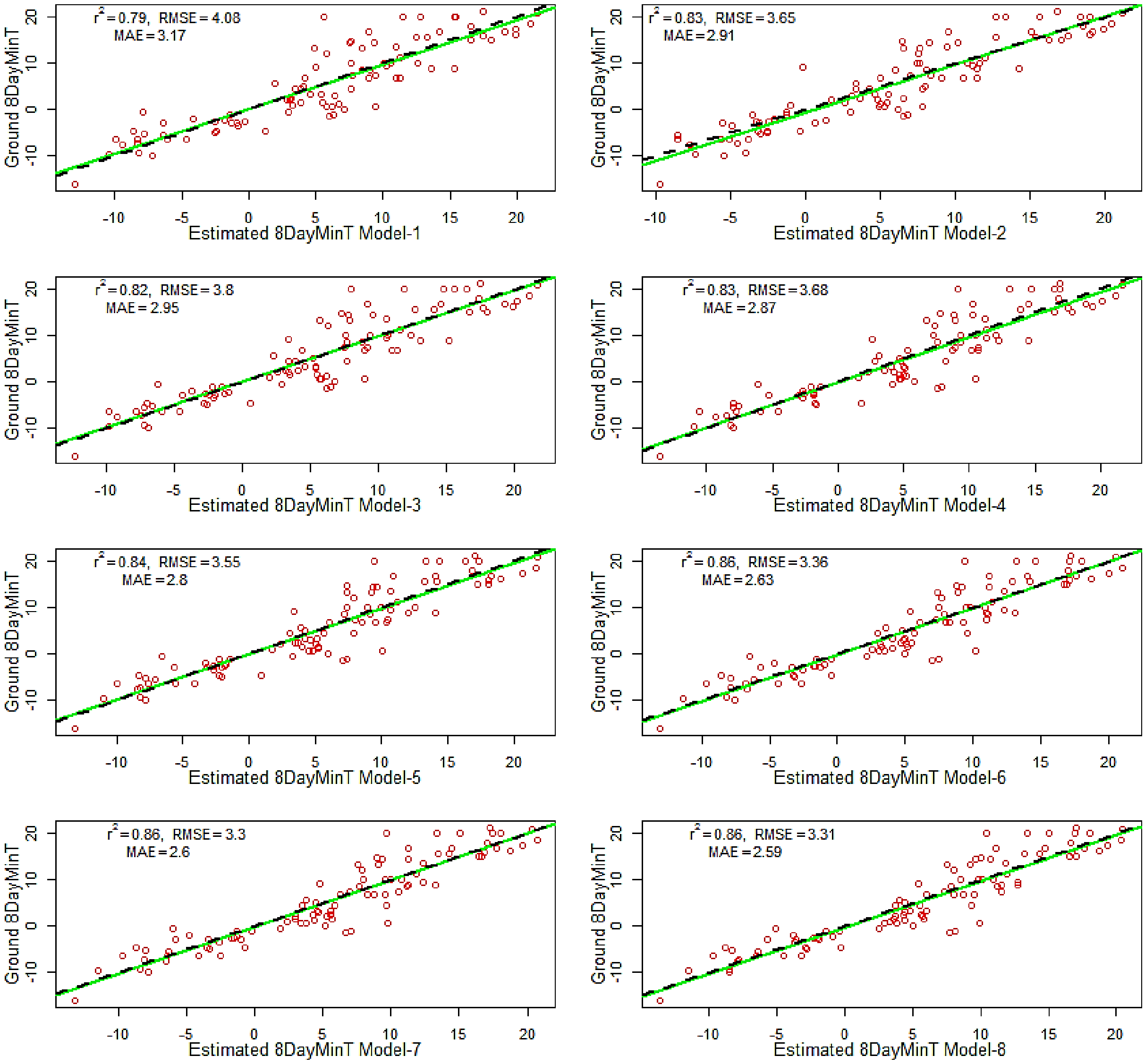


Figure 3.The correlation between estimated and observed 8DayMinT using the models 1 to 8 for test dataset. The values of r^2, RMSE and MAE is also shown. The dashed black shows the 1:1 line. The solid green line corresponds to the regression line.

Table 3. Summary of the model performance for estimation of 8DayMinT during testing

Statistics	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
r^2	0.79	0.83	0.82	0.83	0.84	.86	.86	.86
RMSE	4.08	3.65	3.8	3.68	3.55	3.36	3.3	3.31
MAE	3.17	2.91	2.95	2.87	2.8	2.63	2.6	2.59

Figure 3 and Table 3 gives the summary of results when all the models are applied to the validation/test dataset. , i.e., remaining 30% data for the year 2006. The relationship between ground observed and estimated 8-day T Min for the validation dataset has been analyzed using the coefficient of determination (r^2), root mean square error (RMSE) and mean absolute error (MAE) for each model.

From Figure 3 and Table 3, it is observed that all models give almost similar results of 8-DayTmin estimation. However, there is a slight improvement in the accuracy from (r^2 = 0.79, RMSE = 4.08, MAE = 3.17) when using only 1 variable, i.e. Terra Night (Model-1) to (r^2 = 0.86, RMSE = 3.31, MAE = 2.59) when using 7 variables in Model 8.

5. Summary and Conclusions

In this study, the relationship between 8DayMinT and four LST products (Terra_Day, Terra_Night, Aqua_Day, Aqua_Day) along with 4 auxiliary variables namely, latitude, longitude, altitude and Julian Day has been analyzed and discussed. Important variables out of these 8 have been identified based on backward, forward and stepwise selection methods. Based on the identified variables which have significant relationship with the 8DayMinT, multiple linear regression analysis has been carried out to estimate 8DayMinT using eight different models. The best estimates of 8DayMinT were achieved when all four LSTs have been combined with location parameters and altitude of the station (Model 8). Hence, it can be concluded that, in order to achieve the best results in terms of T_G estimation (in this case for 8DayMinT), other variables like longitude, latitude, and altitude, should be taken into consideration and put into the models. However, to check the consistency of the obtained results further analysis needs to be carried out for different locations and for a longer time period.

6. References

BBMB. (1988) Snow hydrology studies in India with particular reference to the Satluj and Beas catchments. In: Proc.Workshop on Snow Hydrology (Manali, India, 23–26 November, 1988).

Benali, A., Carvalho, A. C., Nunes, J. P., Carvalhais, N., and Santos, A. (2012). Estimating air surface temperature in Portugal using MODIS LST data. Remote Sensing of Environment, 124, 108–121. https://doi.org/10.1016/j.rse.2012.04.024

Collins, F.C., 1995, A Comparison of Spatial Interpolation Techniques in Temperature Estimation, Retrieved Sep 2, 2017, from http://www.ncgia.ucsb.edu/SANTA_FE_CD-ROM/sf_papers/collins_fred/collins.html

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. Journal of Applied Meteorology and Climatology, 50(3), 767–775. https://doi.org/10.1175/2010JAMC2460.1

Hachem, S., Duguay, C. R., and Allard, M. (2012). Comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain. Cryosphere, 6(1), 51–69. https://doi.org/10.5194/tc-6-51-2012

Lai, Y.-J., Li, C.-F., Lin, P.-H., Wey, T.-H., and Chang, C.-S. (2012). Comparison of MODIS land surface temperature and ground-based observed air temperature in complex topography. International Journal of Remote Sensing, 33(24), 7685–7702. https://doi.org/10.1080/01431161.2012.700422

Laskin, D. N., Montaghi, A., Nielsen, S. E., and McDermid, G. J. (2016). Estimating understory temperatures using MODIS LST in mixed cordilleran forests. Remote Sensing, 8(8), 1–21. https://doi.org/10.3390/rs8080658

Liu, Y., Yu, Y., Yu, P., Göttsche, F. M. and Trigo, I. F. (2015). Quality assessment of S-NPP VIIRS land surface temperature product. Remote Sensing, vol. 7, no. 9, pp. 12215–12241.

Noi, P., Kappas, M., and Degener, J. (2016). Estimating Daily Maximum and Minimum Land Air Surface Temperature Using MODIS Land Surface Temperature Data and Ground Truth Data in Northern Vietnam. Remote Sensing, 8(12), 1002. https://doi.org/10.3390/rs8121002

Rafiq, M., Rashid, I., and Romshoo, S. A. (2014) Proceedings of the 16th International Association for Mathematical Geosciences - Geostatistical and Geospatial Approaches for the Characterization of Natural Resources in the Environment: Challenges, Processes and Strategies, IAMG 2014, Pages 420-423

The Report of MOD11A1 and MYD11A1 Validation Accuracy. Available online: http://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD11 (accessed on 2 September 2017)

Vancutsem, C., Ceccato, P., Dinku, T., and Connor, S. J. (2010). Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. Remote Sensing of Environment, 114(2), 449–465. https://doi.org/10.1016/j.rse.2009.10.002

Wan, Z., 2006, MODIS Land Surface Temperature Products Users' Guide, Retrieved on 4 September 2017, from http://www.icess.ucsb.edu/modis/LstUsrGuide/MODIS_LST_products_Users_guide_C5.pdf,

Wan, Z., S. H. (2015). MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/modis/mod11a2.006

Wang, W., Liang, S., and Meyers, T. (2008). Validating MODIS land surface temperature products using long-term nighttime ground measurements. Remote Sensing of Environment, 112(3), 623–635. https://doi.org/10.1016/j.rse.2007.05.024

Zakšek, K., and Schroedter-Homscheidt, M. (2009). Parameterization of air temperature in high temporal and spatial resolution from a combination of the SEVIRI and MODIS instruments. ISPRS Journal of Photogrammetry and Remote Sensing, 64(4), 414–421. https://doi.org/10.1016/j.isprsjprs.2009.02.006

Zeng, L., Wardlow, B. D., Tadesse, T., Shan, J., Hayes, M. J., Li, D., and Xiang, D. (2015). Estimation of daily air temperature based on MODIS land surface temperature products over the Corn Belt in the US. Remote Sensing, 7(1), 951–970. https://doi.org/10.3390/rs70100951

Zhang, H., Zhang, F., Ye, M., Che, T., and Zhang, G. (2016). Estimating daily air temperatures over the Tibetan Plateau by dynamically integrating MODIS LST data. Journal of Geophysical Research: Atmospheres, 121(19), 11425–11441. https://doi.org/10.1002/2016JD025154

Zhou, W., Peng, B., Shi, J., Dhital, Y. P., Wang, T., Ji, D., Zhao, T.,Yao, P., Cui, Y., Shi, L.,Yao, R., and Liu, C. (2016). Estimating daytime surface air temperature using multi-source remote sensing and climate reanalysis data at glacierized basins: a case study at Langtang valley, Nepal. International Geoscience and Remote Sensing Symposium (IGARSS), (c), 4143–4146.

Zhu, W., Lű, A., and Jia, S. (2013). Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. Remote Sensing of Environment, 130, 62–73. https://doi.org/10.1016/j.rse.2012.10.034

Zou, D., Zhao, L., Wu, T., Wu, X., Pang, Q., and Wang, Z. (2014). Modeling ground surface temperature by means of remote sensing data in high-altitude areas: test in the central Tibetan Plateau with application of moderate-resolution imaging spectroradiometer Terra/Aqua land surface temperature and ground-based infrared. Journal of Applied Remote Sensing, 8(1), 83516. https://doi.org/10.1117/1.JRS.8.083516