Spatially adaptive estimation of multi-layer soil temperature at a daily time-step across China during 2010-2020

Xuetong Wang[1, 2], Liang He[3, *], Peng Li[1, 2], Jiageng Ma[4], Yu Shi[5], Qi Tian[1, 2], Gang Zhao[2, 6], Jianqiang He[7], Hao Feng[2], Hao Shi[8, 9, *], Qiang Yu[2, *]

[1] College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China

[2] State Key Laboratory of Soil and Water Conservation and Desertification Control, Northwest A&F University, Yangling 712100, China

[3] National Meteorological Center, Beijing, 100081, China

[4] Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, PR China

[5] Institute of Carbon Neutrality, Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China

[6] College of Soil and Water Conservation Science and Engineering, Northwest A&F University, Yangling, Shaanxi, 712100, China

[7] Key Laboratory for Agricultural Soil and Water Engineering in Arid Area of Ministry of Education, Northwest A&F University, Yangling 712100, China

[8] State Key Laboratory for Ecological Security of Regions and Cities, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China

[9] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China

Correspondence:

Liang He (heliang_hello@163.com)

Hao Shi (haoshi@rcees.ac.cn)

Qiang Yu (yuq@nwafu.edu.cn)

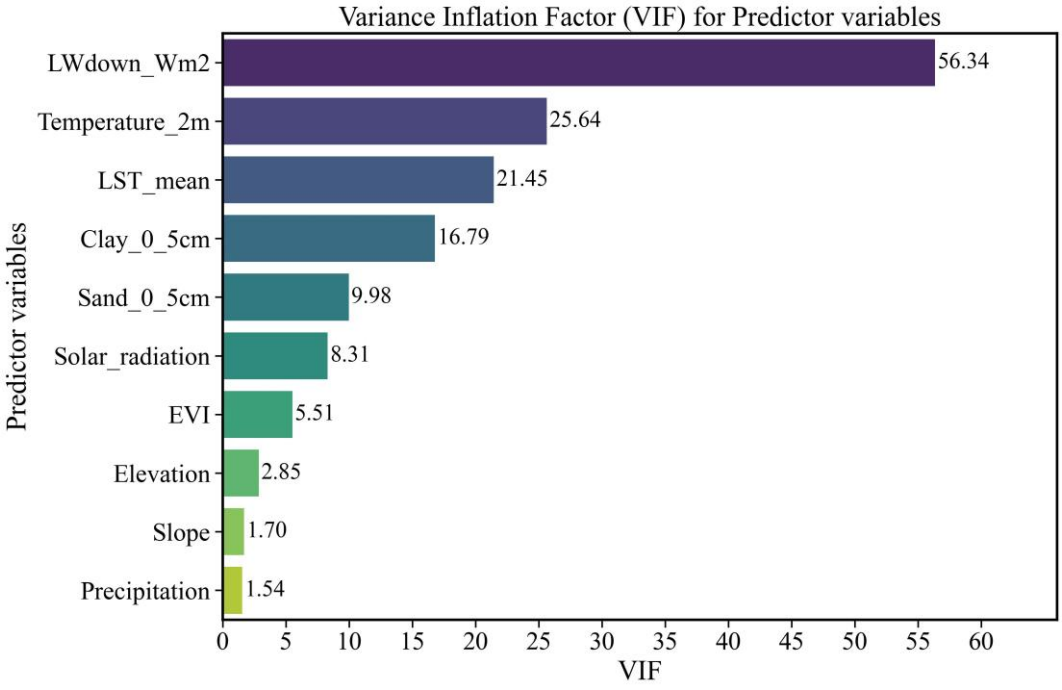

Figure S1. Variance Inflation Factor (VIF) of predictor variables (with LWD)

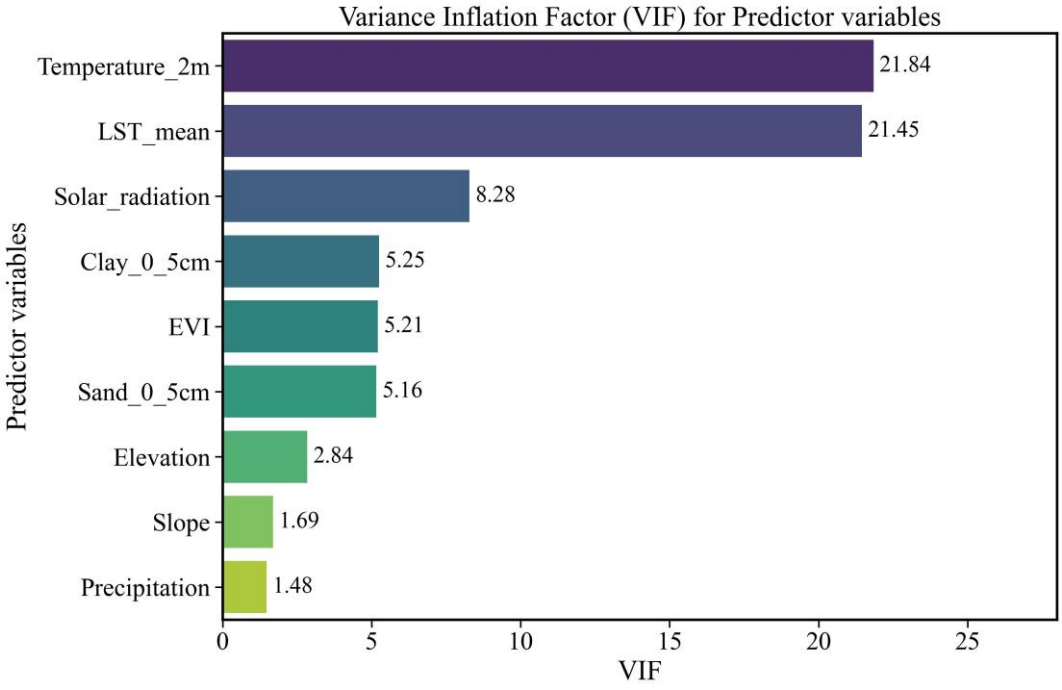

Figure S2. Variance Inflation Factor (VIF) of predictor variables

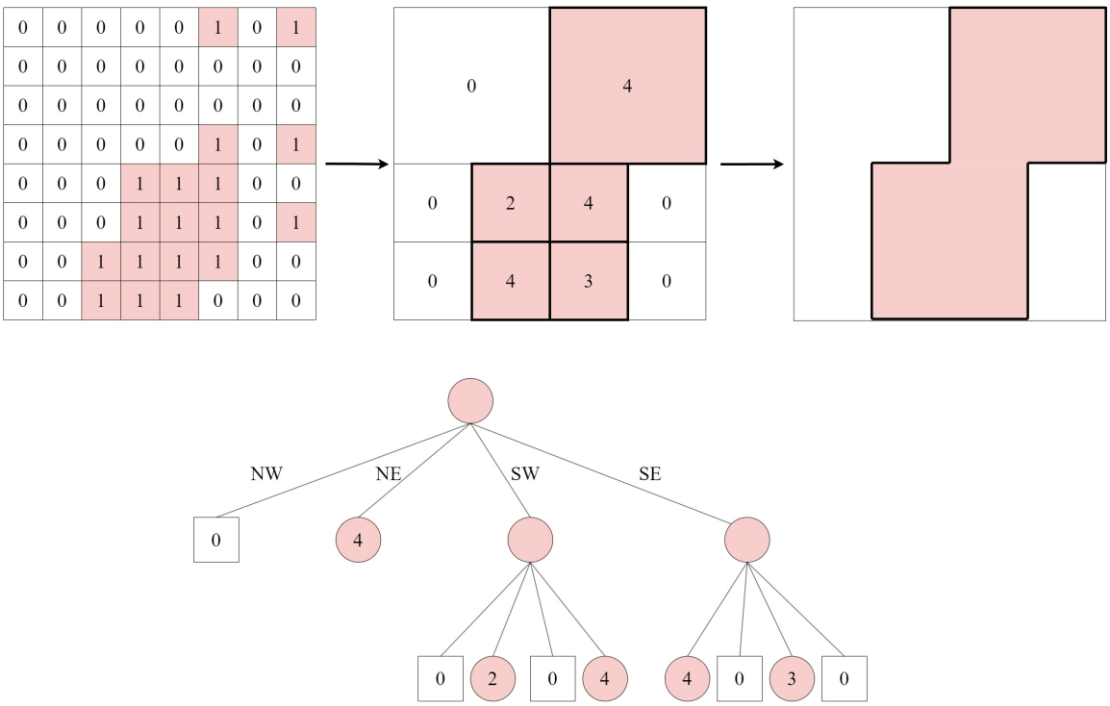

Figure S3. Quadtree spatial index principle. (Northwest (NW), Northeast (NE), Southwest (SW), Southeast (SE) refer to the four quadrants into which the quadtree divides the two-dimensional space.)

We conducted a systematic evaluation of the partitioning performance under different thresholds using three key metrics: the coefficient of variation (CV) of point count, the CV of point density, and the total number of grid cells. The CV of point count was used to evaluate the balance of sample distribution across spatial units under different thresholds. Point density was defined as the number of observation stations within a grid cell divided by its area. A lower CV of point density indicates that the partitioning effectively adjusted grid size according to local station density—i.e., producing smaller grids in dense regions and larger grids in sparse areas—thus reflecting a more adaptive spatial division. Conversely, a higher CV suggests that the partitioning failed to capture the spatial heterogeneity of station density. Therefore, the CV of point density serves as a key indicator of the spatial adaptivity of the quadtree partitioning. The total number of grids corresponds to the number of local models to be trained, and thus indirectly reflects the computational and time cost associated with model training.

As shown in Figure S3 (a–c), we systematically evaluated quadtree performance under a series of point-count thresholds (10, 30, 50, 70, 90): Figure S3a shows that the CV of point count drops rapidly with increasing threshold, indicating improved balance in sample allocation across grids. However, this trend levels off beyond threshold = 30, suggesting diminishing returns. Thus, threshold 30 marks an optimal trade-off. Figure S3b shows a notable inflection point in the CV of point density near threshold = 30. Although not the global minimum, this point represents an optimal trade-off where grid subdivision sufficiently reflects sample density variation without causing over- or under-segmentation—thereby capturing spatial adaptivity effectively. Figure S3c shows that the number of grid cells decreases rapidly as the threshold increases, leading to substantial computational savings. However, the rate of reduction slows considerably beyond threshold = 30, indicating limited additional benefit from further increases.

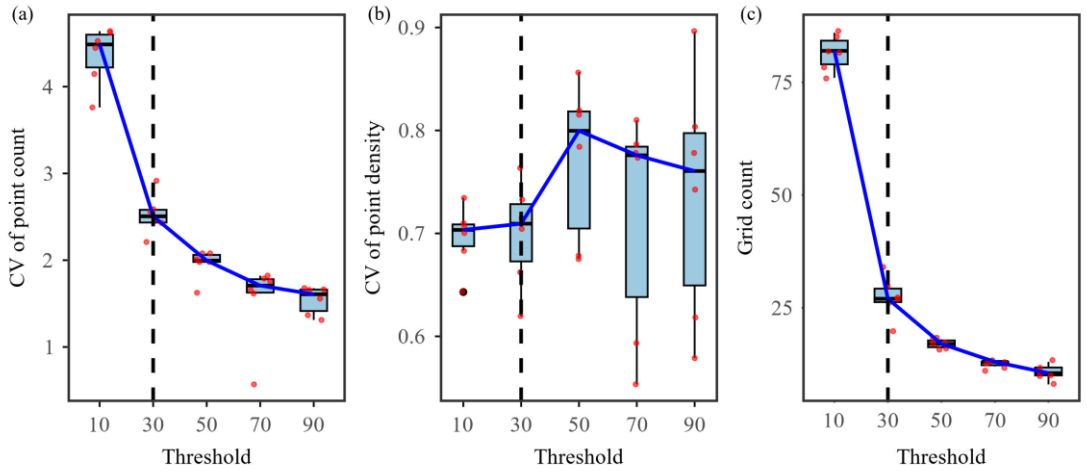

Figure S4. Performance evaluation of quadtree partitioning under different point-count thresholds. (a) Coefficient of variation (CV) of point count across spatial units. (b) CV of point density (point count per unit area). (c) Total number of generated grid cells. Dashed vertical line indicates the selected threshold of 30.

Table.S1 Candidate values of hyperparameters in XGBoost.

hyperparameters	candidate value		
	Start	End	Step
n_estimators	60	200	20
max_depth	5	15	1
learning rate	0.1	1	0.1

Table.S2 Metadata of flux tower soil temperature observations used for validation

Site name	Ecosystem type	Depth (cm)	Time series
Baotianman Forest Station	Forest	0,5,20	2010-2014
Changling Rice Paddy Station	Cropland	5,10,20	2018-2020
Daan Cropland Station	Cropland	0,5,10,15,20	2017-2020
Damao Grassland Station	Grassland	0,5,10,15,20,40	2017-2020
Danzhou Rubber Plantation Station	Forest	5,10,20	2010
Haibei Alpine Meadow Station	Grassland	5,10,15,20,40	2015-2020
Haibei Shrubland Station	Grassland	0,5,20,40	2016-2018
Huzhong Boreal Forest Station	Forest	5,10,20	2014-2018
Jinzhou Cropland Station	Cropland	5,10,15,20,40	2011-2014
Lijiang Alpine Meadow Station	Grassland	5,10,15,20,40	2013-2020
Maoershan Forest Station	Forest	5	2016-2018
Panjin Reed Wetland Station	Wetland	10,20,40	2018-2020
Qianyanzhou Plantation Forest Station	Forest	5,10,20	2011-2015
Ruoergai Alpine Wetland Station	Wetland	0,5,10,20	2013-2020
Sanjiangyuan Alpine Grassland Station	Grassland	0,5,15	2013-2015
Taoyuan Cropland Station	Cropland	5,10,15,20,40	2010-2014
Xishuangbanna Rubber Plantation Station	Forest	0,5,20	2010-2014
Yuanjiang Dry-Hot Valley Savanna Station	Grassland	5,10,20,40	2013-2015

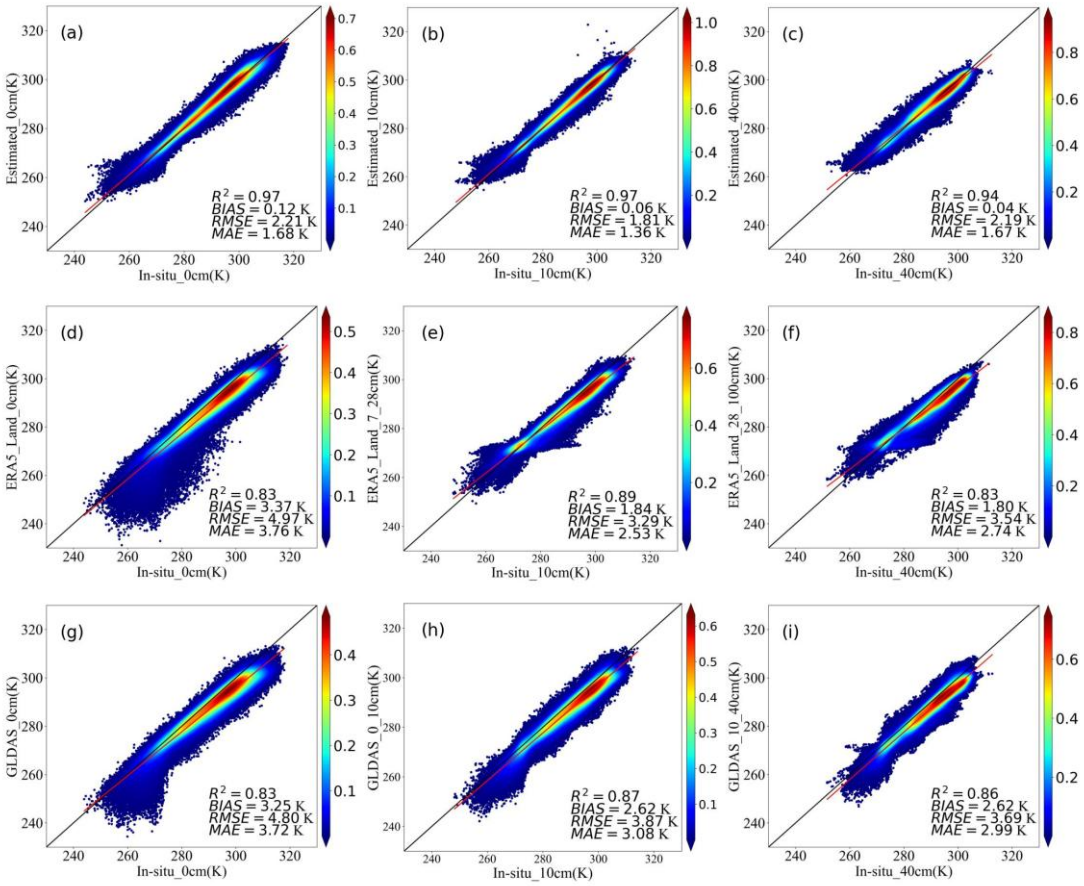

Figure S5. Scatter density plot comparing the accuracy of different products (e.g., 0, 10, and 40 cm)

55

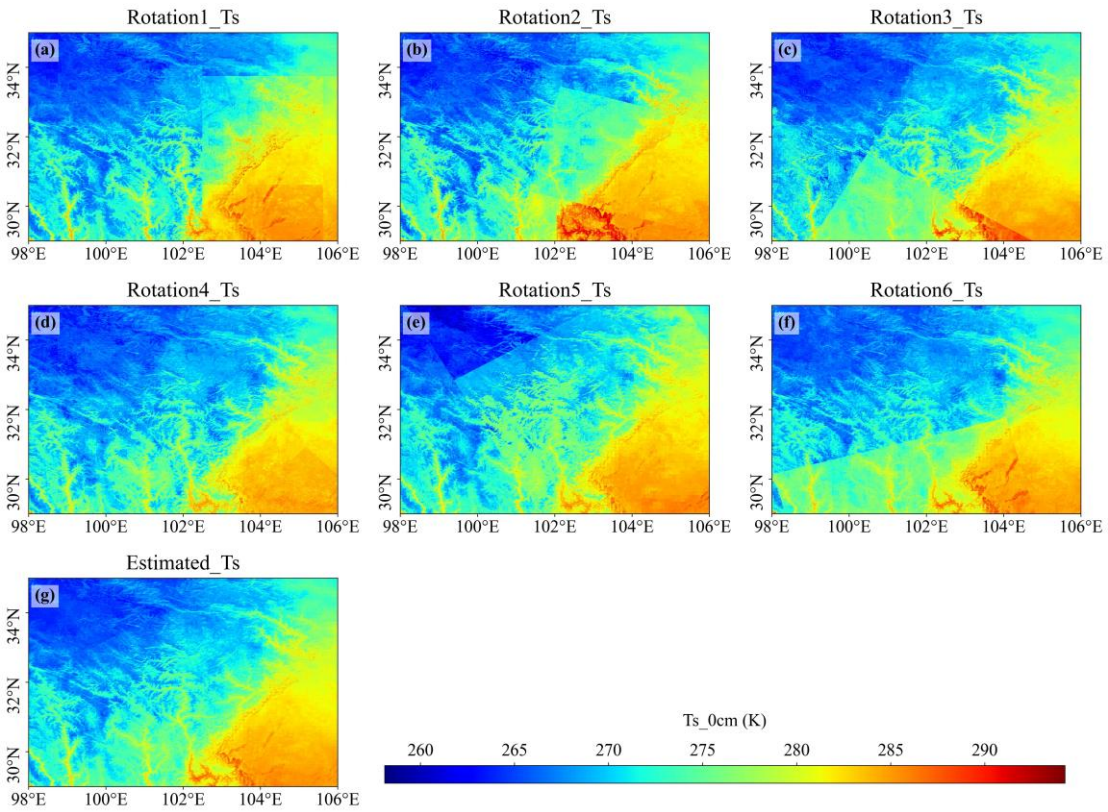

Figure S6. The difference in spatial performance after modeling different rotations separately (Rotation1-Rotation 6) and the average value of all different rotations (Estimated_T_s) within the same sub-region (29°N–35°N, 98°E–106°E)

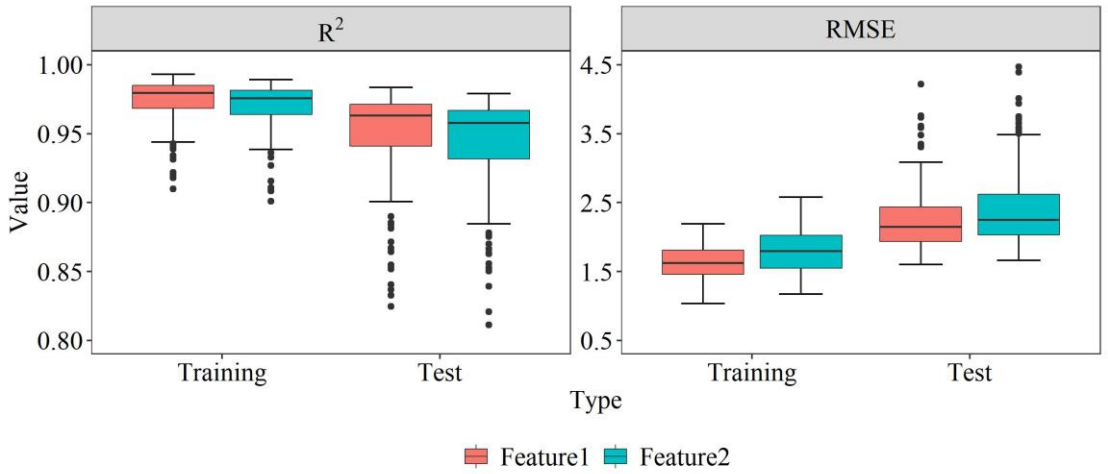

Figure S7. Comparison of Modeling Accuracy with Different Feature Variables (Feature1 represents using both air temperature and LST together with other feature variables, while Feature 2 represents using only air temperature together with other feature variables)

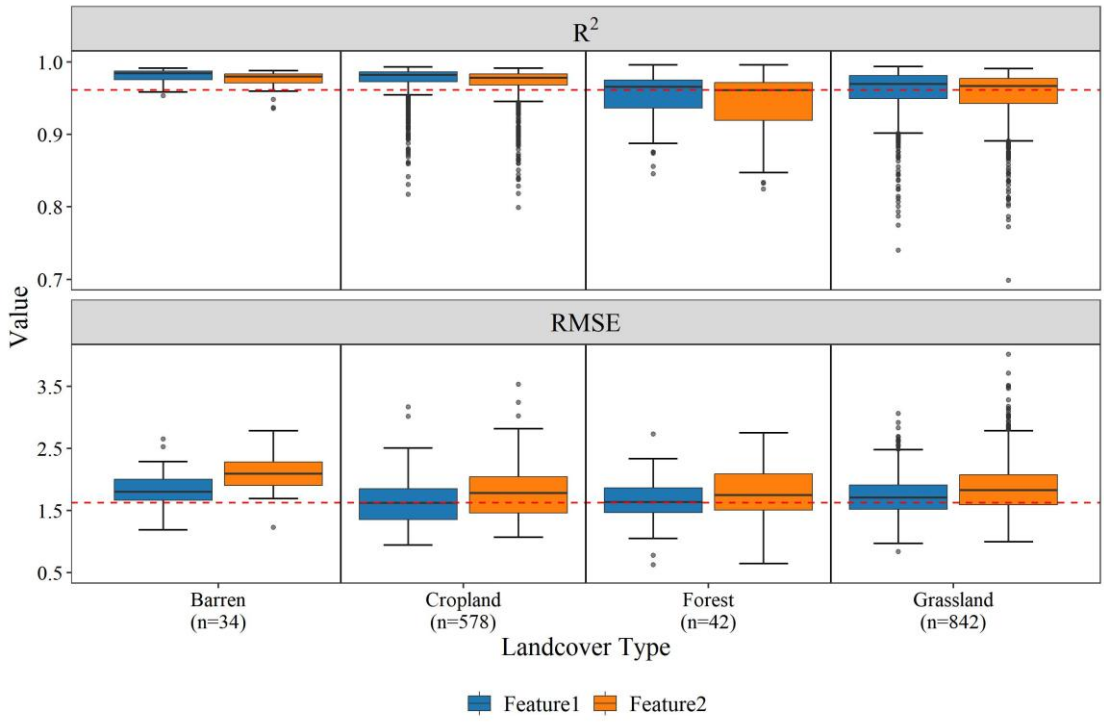

Figure S8. Differences in model accuracy across land cover types under different feature variable combinations. (Feature1 represents using both air temperature and LST together with other feature variables, while Feature 2 represents using only air temperature together with other feature variables)