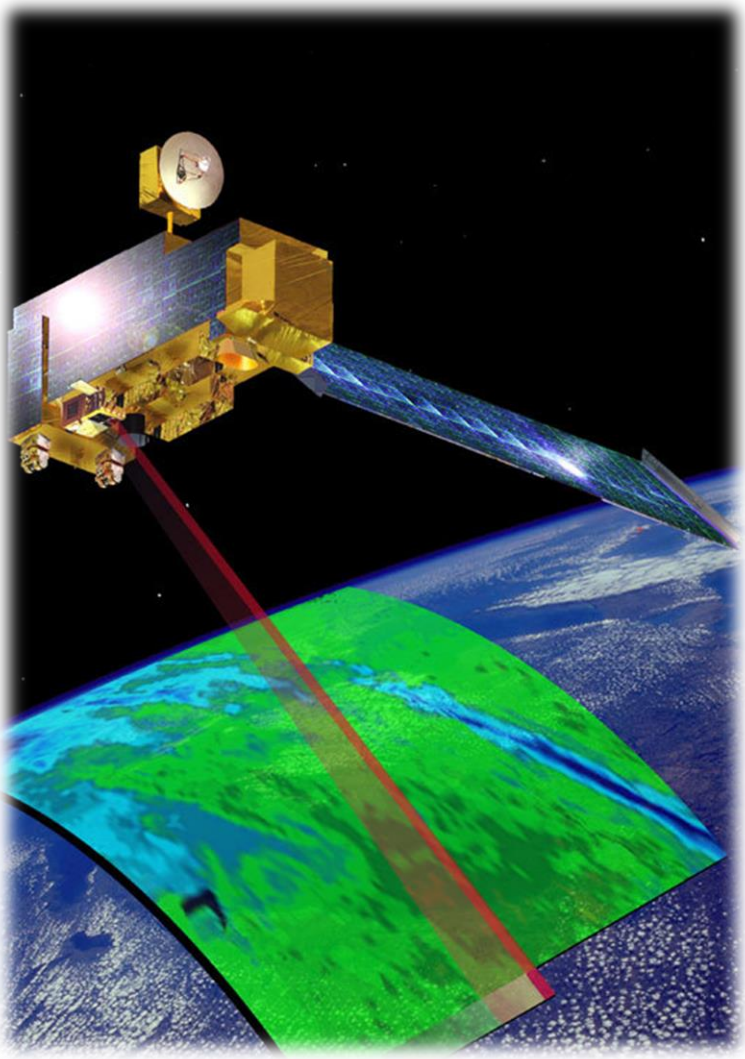


ISTANBUL TECHNICAL UNIVERSITY

Special Topics in Remote Sensing
Final Paper
Prof. Dr. Elif Sertel
Res. Assis. Samet Aksoy

Air Temperature Regression in Four States of the USA

Umut Bediroğlu
umut.bediroglu@hotmail.com

Table of Content

1. INTRODUCTION

Changing climate, agriculture, hydrology, and biodiversity, for example, all require precise air temperature data at varying temporal and geographical resolutions (Vogt et al., 1997; Monestiez et al., 2001; Yang et al., 2013; Ge et al., 2014). In furthermore, temperature of the air is utilized to investigate the influence of air temperature on plant photosynthesis or respiration. For illustration, the leaf temperature required to achieve the greatest net photosynthesis rate is in the 20–30°C range. As a result, environmental factors linked with this energy balance (light intensity, temperature, humidity, air flow rate, and so on) impact leaf temperature. As a result, this affects photosynthesis and growth. The most important element of the climate is the reason why humidity, precipitation, pressure and wind are formed from other elements. At this point, the importance of the work done, the air temperatures of the regions can be estimated with remote sensing and indirectly information about the rate of photosynthesis or respiration in plants can be obtained. By means of remote sensing of the air temperature apart from the local data, the air temperature data can be transformed into a continuous data instead of being a point data. Many articles have been written on the use of different algorithms and different data in order to predict the air temperature correctly, and the most accurate method has been tried to be found in different areas.

Shen et al. extensively used deep learning to map air temperature using satellite images and base station inputs. China was chosen as the venue for the study. The study used MODIS LST and MOD13A2 for NDVI data, MODIS Annual Terra and Aqua combined MODIS land use / cover product (MCD12Q1) of 2015 for land cover, and SRTM for elevation. Data was fed into the computer, and the results were achieved by using a deep belief network as a deep learning approach (2020). Xu et al. investigated the feasibility of forecasting month air temperature using machine learning algorithms using remote sensing data. The study's focal point is the Tibetan Plateau. Ten deep learning methods were applied with the data to correlate the air temperature and environmental parameters supplied by MODIS. The Cubist methodology was chosen as the best method since it has the lowest prediction error and also the lowest variation among errors achieved with higher overall values and errors obtained with poorer quality samples (2018). Adab et al. employed four forms of quick linear data structures to predict soil moisture in their study (ANN, SVM, RF, and EN). As datasets, Landsat-8 satellite data, soil samples, CHIRPS, and SMAP data were employed. The region of Khorasan-Razavi in Iran was chosen as the location for the investigation. According to the findings obtained from the RF, SVR, ANN, and EN models, the RF model beat the other three approaches for estimating soil

moisture (2020). Melinho et al. used MODIS images and air temperature measurements collected at meteorological stations in Morocco's Souses basin to train six machine learning algorithms on monthly and weekly air temperature predictions. It exhibits the capacity of machine learning algorithms and remote sensing data to forecast temperature using effective models. Aqua and Terra sensors are suitable for LST. In Morocco's Souses basin, the Cubist and RF ML algorithms were the most successful in forecasting air temperature (2022). Wang et al. (2010) assessed the efficacy of two regression studies and three spatial interpolation approaches in forecasting monthly air temperature in China. The findings of many GWR models were considered to be the most accurate, and hence adequate for predicting monthly air temperature on a broader scale than the MLR model. Kriging produced the best results of any interpolation method (2017).

Satellite images and terrestrial station data from four state regions in the United States (Washington, Oregon, Montana, and Idaho) were used in this study; regression was performed using four different algorithms: Random Forest (RFR), Support Vector Regression-Radial Basis Function (SVR-RBF), Support Vector Regression-Linear (SVR-Linear), and smileCart. It is aimed to determine the most accurate algorithm as a result of the accuracy values of the regressions. In addition to the monthly average data of 1102 local stations in line with the study; One-month average temperature data from MODIS and Landsat-8 satellites were used.

1.1 Data and Study Area

The research area comprises of four states in the United States (USA), namely Washington, Oregon, Montana, and Idaho (Figure 1.1.1).

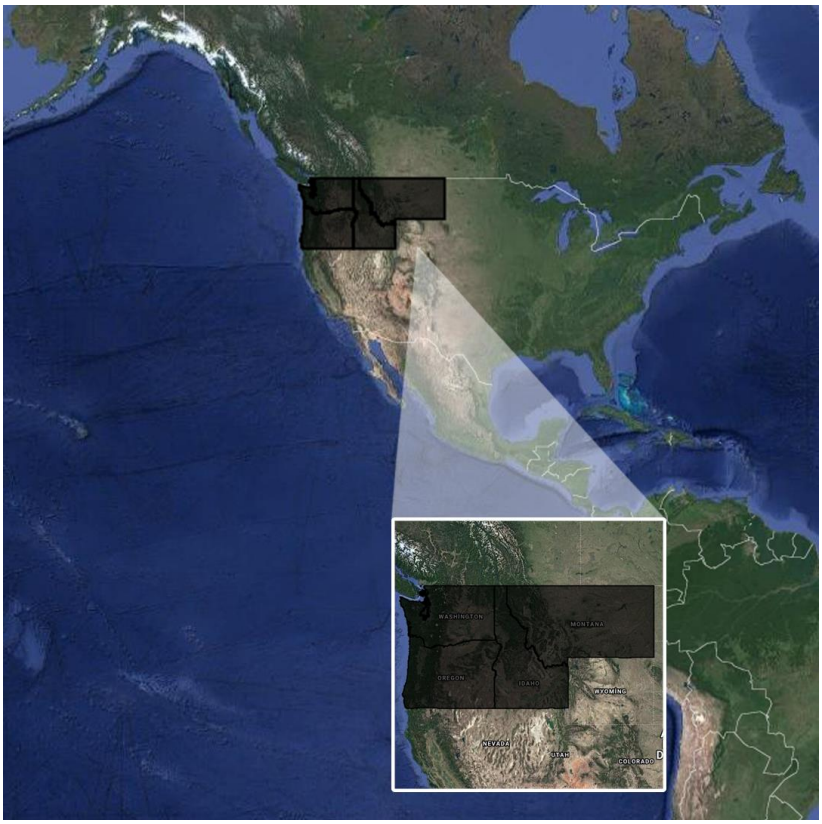


Figure 1.1.1 : Study area

In addition to terrestrial station data, data from two satellites were used in the study (Figure 1.1.2). The National Centres for Environmental Information open source data access site was used for terrestrial station data. March 2021 monthly average temperature (Kelvin) data of 1102 terrestrial stations homogeneously distributed over the site were used. The spatial distribution of the terrestrial data is important to create a more accurate estimation algorithm, so care has been taken to ensure that the data is in a homogeneous station distribution. The presence of clustered data in a certain area may increase the coefficient of the data belonging to that region and cause the estimation results to be incorrect.

Stations Air Temperature **MODIS Terra, LST Day (1 km)** **Landsat 8 OLI/TIRS, Band-10 (30 m)**

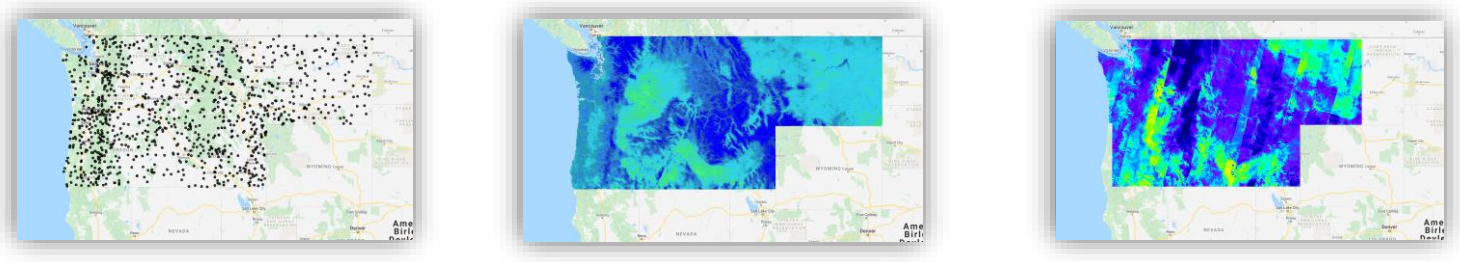


Figure 1.1.2 : Data to be used

Two satellites were utilized in the satellite images: MODIS Terra 1 Land Surface Temperature and Emissivity Daily Global - 1km data and Landsat 8 Level 2, Collection 2, Tier 2 – 30m data (Table 1.1.1).

SENSOR	Landsat 8 OLI/TIRS	MODIS Terra
Launch	2013	1999
Sensor Altitude (km)	705	705
Temporal Resolution (day)	16	1
Swath (nadir)	185 km	2330 km (cross track) by 10 km (along track at nadir)
Number of Spectral Bands	9	36

Table 1.1.1 : Satellite features

In these remote sensing images, the monthly average temperature (Kelvin) for March 2021 was obtained. Satellite images have two different resolution images; The lowest resolution was MODIS with 1 kilometre and Landsat-8 with 30 meters resolution. In this way, the region was viewed from satellites with both a wide perspective and more detailed information. Investigating whether these data contradicted or correlated with each other was another key point for the accuracy of the study's data. According to Kim et al, the term "land surface temperature" refers to a fundamental physical feature that is linked to numerous ecological, hydrological, and atmospheric phenomena. LST and near-surface air temperature have a close link (2013).

2. METHOD

Satellite imagery and terrestrial station data from four US state regions (Washington, Oregon, Montana, Idaho); It has been tried to estimate with Random Forest (RFR), Support Vector Regression-Radial Basis Function (SVR-RBF), Support Vector Regression-Linear (SVR-Linear) and smileCart algorithms. During the learning of the algorithms, 1102 terrestrial station data was divided into 881 trains and 221 validations. Thanks to this, the temperatures taught

and predicted to the algorithm were tested with values that the algorithm had never seen. Three different studies were conducted with these split terrestrial station data; the first one is teaching with Landast-8 and terrestrial station data, the second one is teaching with MODIS and terrestrial station data, and the last one is teaching with the use of three data together. By using these different data combinations, it is aimed to make inferences and connections between the algorithms and data, with teaching.

3. CONCLUSIONS

With the regressions performed, estimation maps were obtained, and the obtained maps were displayed as a result of the classification. The Coefficient of Determination (R^2) and Root Mean Square Error (RMSE) values resulting from the regressions enabled us to make sense of the algorithm's accuracy at this point, not only visually, but also numerically (Table 3.1). Here, R^2 is expected to converge to 1 and RMSE to 0. In the regressions carried out in this direction, smileCart is the algorithm with the most accurate prediction result in all data combinations, while the best regression has emerged when three data are together. When a satellite image and terrestrial station data are used, the regression with Landsat-8 has more erroneous predictions than the MODIS satellite image. When Landsat-8 entered into regression with MODIS satellite images, the most accurate prediction results and values were obtained.

Satellite Image	RFR				SVR - RBF				SVR - LINEAR				smileCart			
	Train		Validation		Train		Validation		Train		Validation		Train		Validation	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
Landsat 8 Level 2, Collection 2 - MODIS	0.82	2.22	0.82	2.29	0.65	3.08	0.63	3.20	0.52	3.61	0.53	3.63	0.92	1.51	0.91	1.53
Landsat 8 Level 2, Collection 2	0.55	3.48	0.60	3.51	0.17	4.62	0.19	4.79	0.11	4.81	0.14	4.96	0.70	2.77	0.77	2.56
MODIS	0.77	2.44	0.74	2.61	0.60	3.21	0.57	3.35	0.53	16.45	0.52	16.57	0.85	1.98	0.83	2.11

Table 3.1 : R2 and RMSE Values of Regression

REFERENCES

Vogt JV, Viau AA, Paquet F. 1997. Mapping regional air temperature fields using satellite-derived surface skin temperatures. *Int. J. Climatol.* 17: 1559– 1579.

Monestiez P, Courault D, AllardF Ruget D. 2001. Spatial interpolation of air temperature using environmental context: application to a crop model. *Environ. Ecol. Stat.* 8: 297– 309.

Yang J, Tan C, Zhang T. 2013. Spatial and temporal variations in air temperature and precipitation in the Chinese Himalayas during the 1971–2007. *Int. J. Climatol.* 33: 2622– 2632.

Shen, H., Jiang, Y., Li, T., Cheng, Q., Zeng, C., & Zhang, L. (2020). Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. *Remote Sensing of Environment*, *240*, 111692.

Xu, Y., Knudby, A., Shen, Y., & Liu, Y. (2018). Mapping monthly air temperature in the Tibetan Plateau from MODIS data based on machine learning methods. *IEEE journal of selected topics in applied earth observations and remote sensing*, *11*(2), 345-354.

Adab, H., Morbidelli, R., Saltalippi, C., Moradian, M., & Ghalhari, G. A. F. (2020). Machine learning to estimate surface soil moisture from remote sensing data. *Water*, *12*(11), 3223.

Meliho, M., Khattabi, A., Zejli, D., Orlando, C. A., & Dansou, C. E. (2022). Artificial intelligence and remote sensing for spatial prediction of daily air temperature: case study of Souss watershed of Morocco. *Geo-spatial Information Science*, 1-15.

Wang, M., He, G., Zhang, Z., Wang, G., Zhang, Z., Cao, X., ... & Liu, X. (2017). Comparison of spatial interpolation and regression analysis models for an estimation of monthly near surface air temperature in China. *Remote Sensing*, *9*(12), 1278.

Kim, D. Y., & Han, K. S. (2013). Remotely sensed retrieval of midday air temperature considering atmospheric and surface moisture conditions. *International Journal of Remote Sensing*, *34*(1), 247-263.

Source Code

https://code.earthengine.google.com/de6b52ce3b004063ffe5243f115d86c4

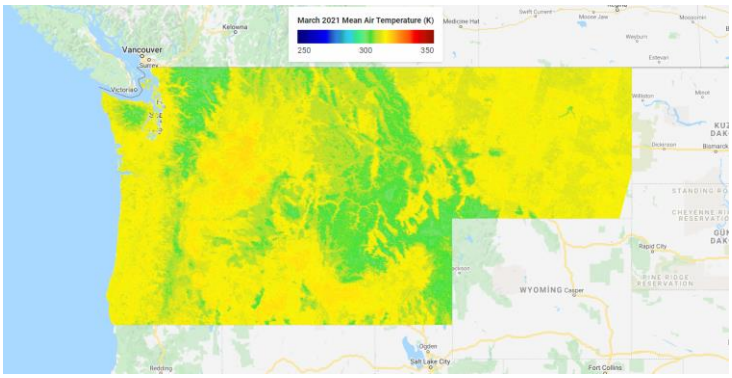
https://code.earthengine.google.com/d1284baf94c2c4da3652c37eaeca9266

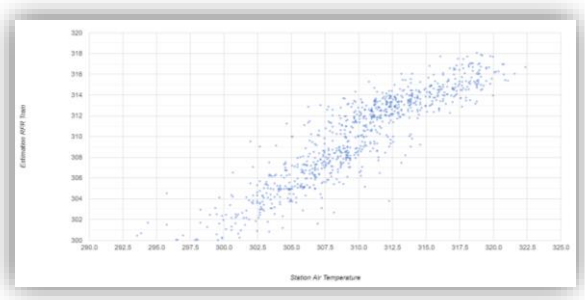
https://code.earthengine.google.com/eed01603d9e28101ce359f1e76f57e99

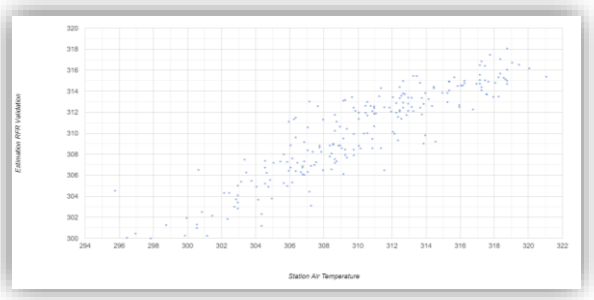
APPENDIX

Regression with Landsat-8 and MODIS

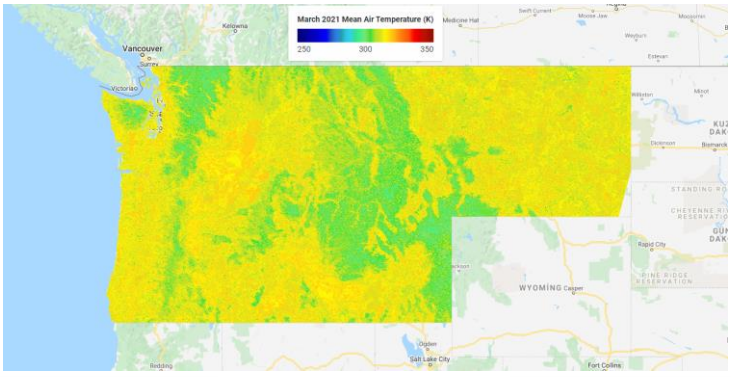
Classification RFR

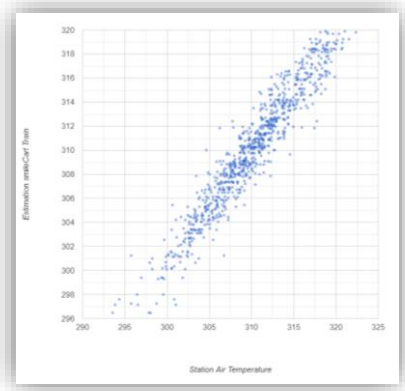


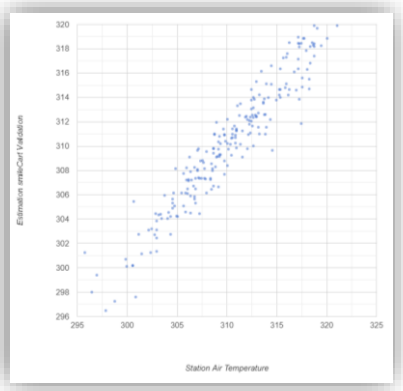




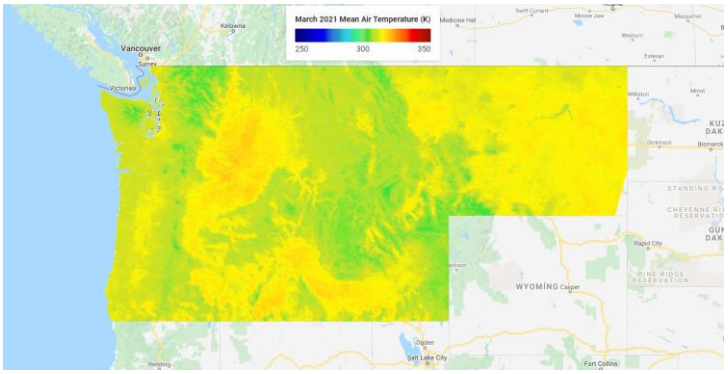
Classification smileCart

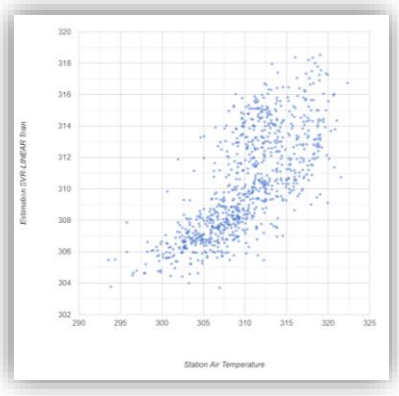


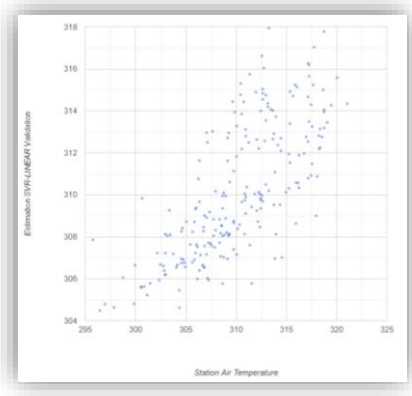




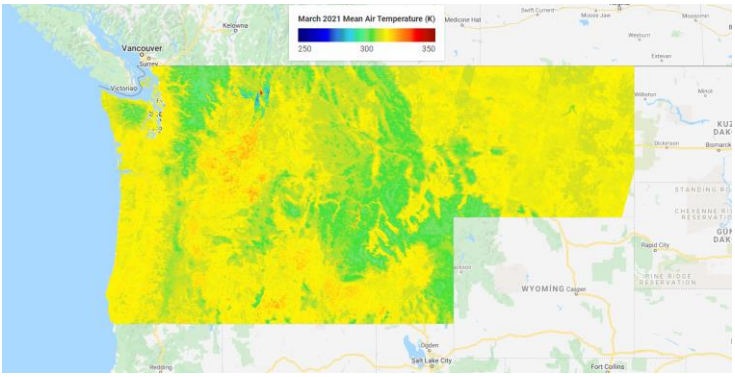
Classification SVR-LINEAR

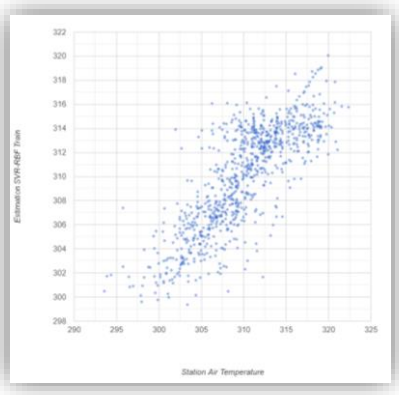


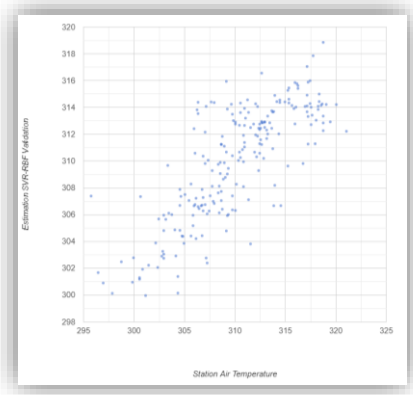




Classification SVR-RBF

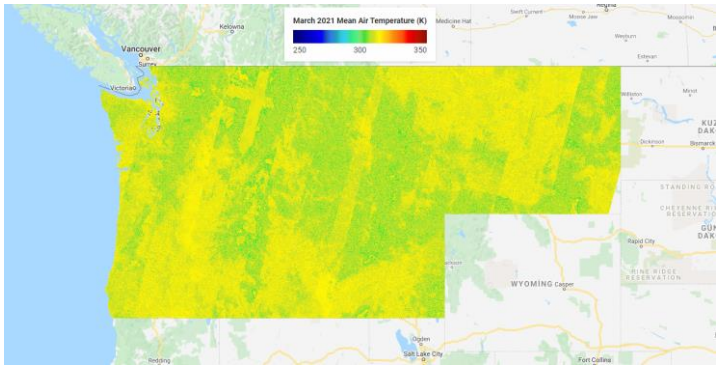


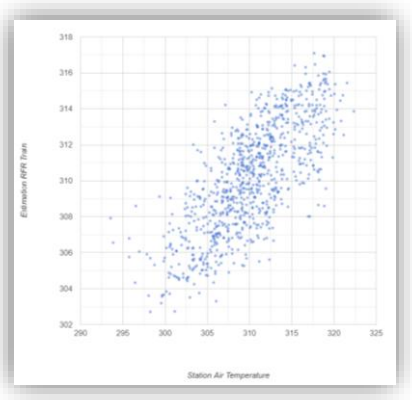


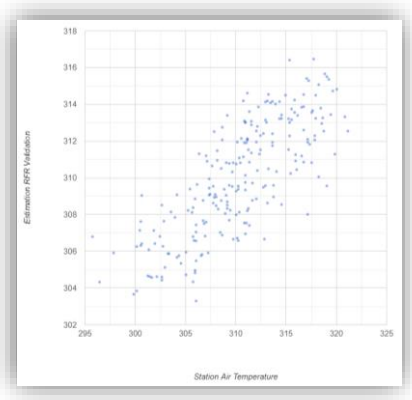


Regression with Landsat-8

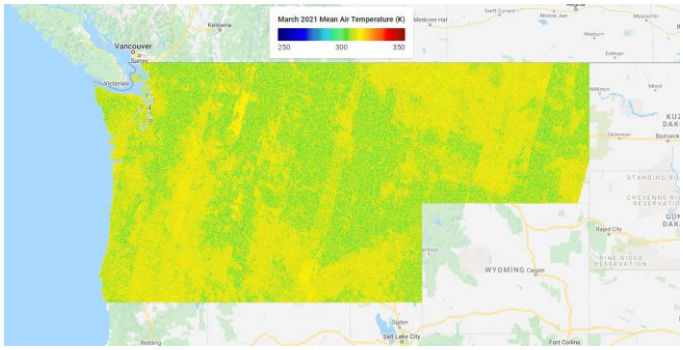
Classification RFR

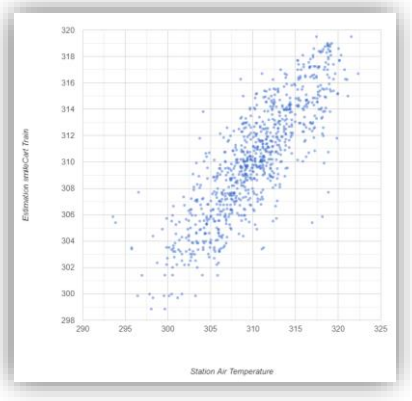


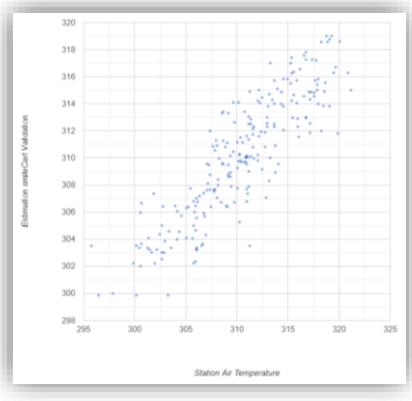




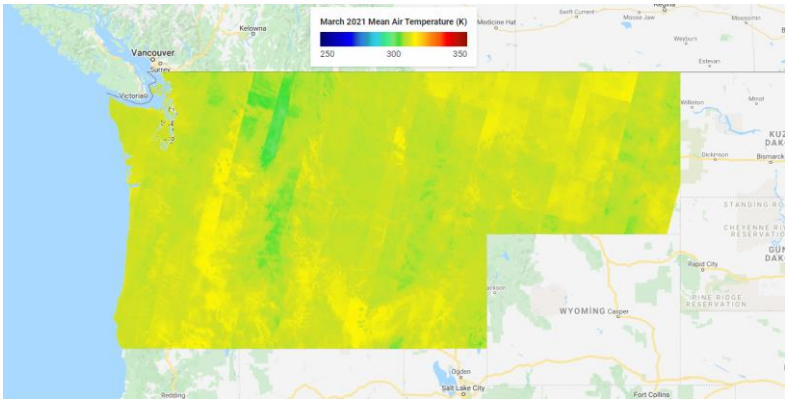
Classification smileCart

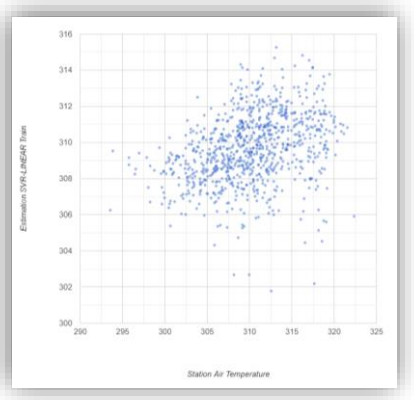


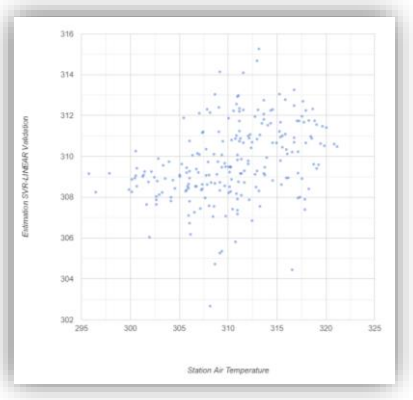




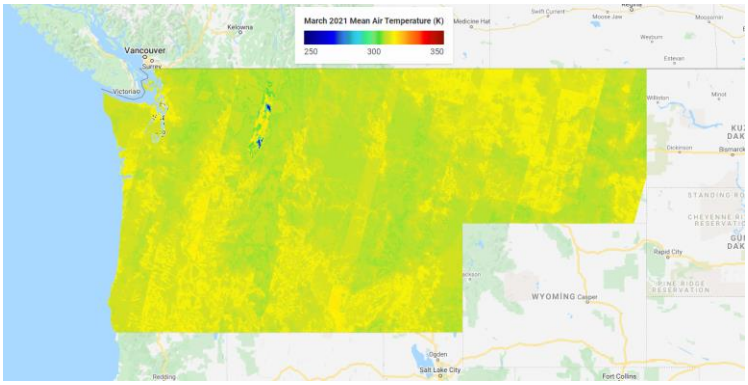
Classification SVR-LINEAR

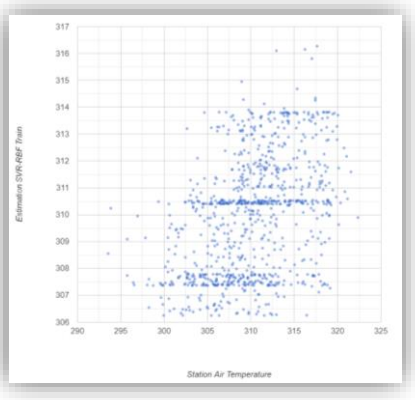


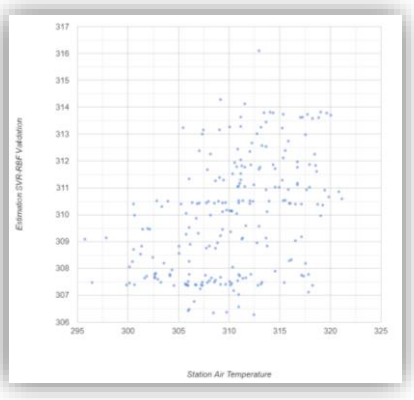




Classification SVR-RBF

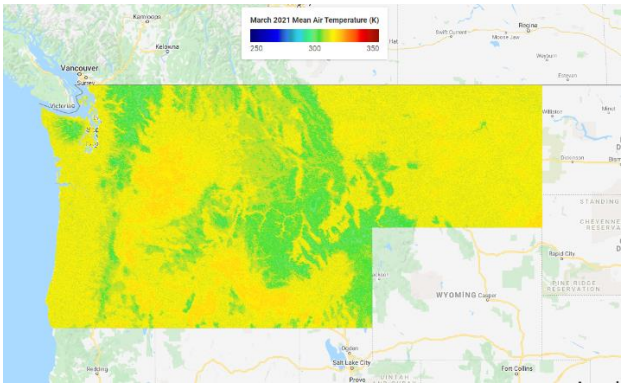


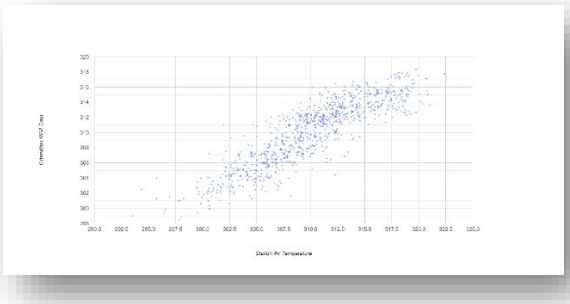


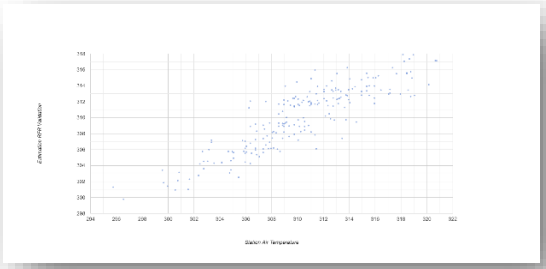


Regression with MODIS

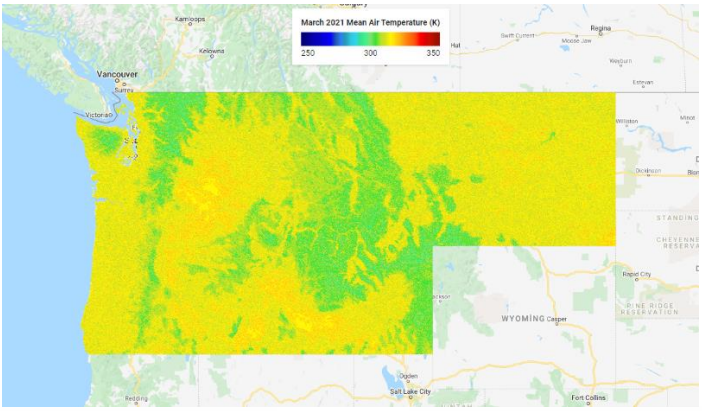
Classification RFR

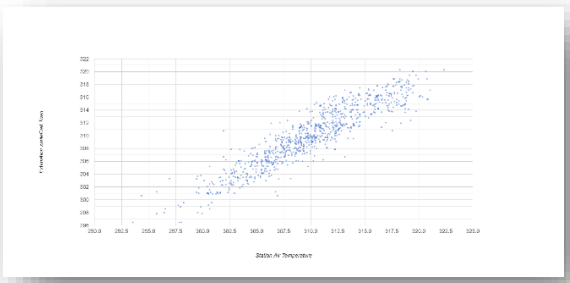


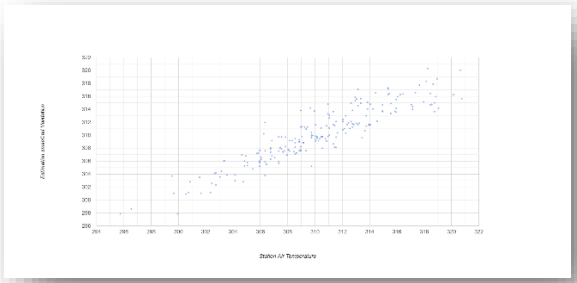




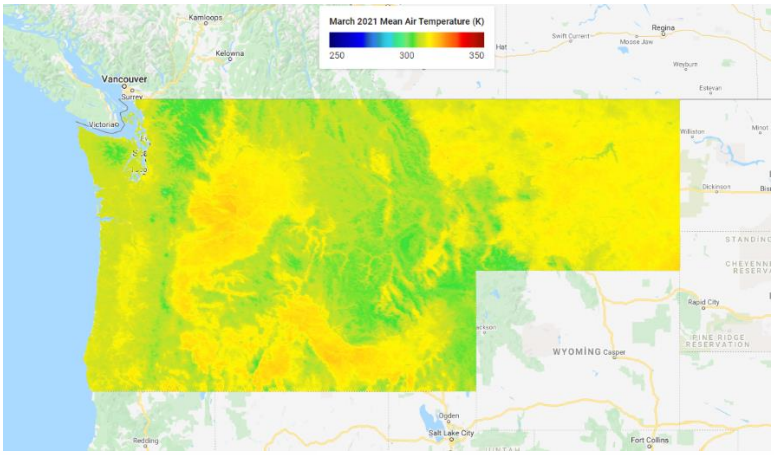
Classification smileCart

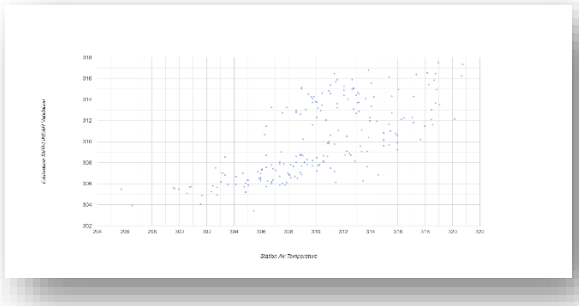


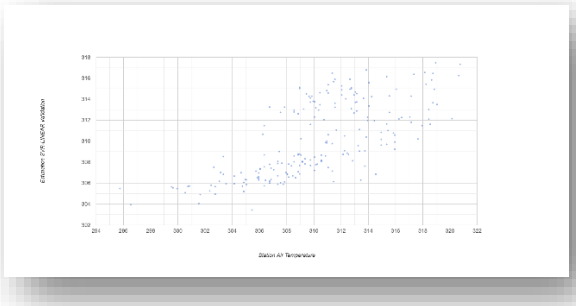




Classification SVR-LINEAR







Classification SVR-RBF