Turkish Journal of Engineering
https://dergipark.org.tr/en/pub/tuje
e-ISSN 2587-1366

Prediction of elevation points using three different heuristic regression techniques

Vahdettin Demir*[1], Ramazan Doğu[1]

[1]KTO Karatay University, Department of Civil Engineering, Türkiye, vahdettin.demir@karatay.edu.tr, rdogu27@gmail.com

Cite this study: Demir, V., & Doğu, R. (2024). Prediction of elevation points using three different heuristic regression techniques. Turkish Journal of Engineering, 8 (1), 56-64

Keywords
M5-Tree
MARS
LSSVR
Elevation Point
Google Earth Pro

Research Article
DOI:10.31127/tuje.1257847

Received:28.02.2023
Revised: 14.04.2023
Accepted:17.04.2023
Published:15.09.2023

Abstract

The aim of this study is to estimate the digital elevation model, which is the most important data of the projects and needed in the engineering project, using latitude and longitude information of the elevation points and three different heuristic regression techniques. As the study area, an area with mid-level elevations, located in the Marmara region, and covering a part of the intersection of Edirne, Kırklareli and Tekirdağ provinces was chosen. In the study, the estimations were investigated for three different sized areas, and these areas are square areas with the dimensions of 1x1 km, 10x10 km and 100x100 km, respectively. A total of 3500 elevation points were used in the study, and this number is constant in all areas, and 60% of these points were used in the testing phase and 40% in the training phase. The models used in the study are M5 model tree (M5-tree), multivariate adaptive regression curves (MARS) and Least Square Support Vector Regression (LSSVR). The results of the models were evaluated according to three different comparison criteria. These, coefficient of determination (R^2), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were used. When the modeling results are examined; M5-Tree regression method gave the best results (1), LSSVR method was better than MARS methods (2), The most successful input data was found in datasets using X and Y coordinates information, and the worst results were found in datasets using X coordinates (3). As the study area increased, the model performance did not improve (4). The least error was obtained in the modeling of 1x1 km area, and the highest R^2 was obtained from the modeling of 10x10 km area (5). It was concluded that the M5-tree method is a very successful method in elevation modeling.

1. Introduction

The data of the heights (elevation) of the land points are used in many areas. However, measuring all points in the field is difficult and costly. Therefore, various mathematical-statistical methods and more modern techniques such as machine learning are used to estimate elevation points [1].

Disaster risk assessment, agriculture, forestry, watershed management, urban and rural planning, transportation planning, etc. numerous fields make use of numerical models created from land locations. These models serve as the foundation for engineering study projects [2-4]. Information about the surface and the subsurface can be processed, analyzed, and visually presented using digital surface models [5].

Important studies in the literature in recent years; Demir and Keskin [1], estimated elevations in Samsun Mert River Basin using X and Y coordinate information and three different Artificial Neural Networks and IDW and Kriging interpolation techniques. Demir and Çubukçu [5], estimated elevation points in a similar study area (Samsun Mert River Basin) using M5 model tree (M5-tree) and multivariate adaptive regression curves (MARS) heuristic regression methods. The results were compared with the regression methods. In the literature on surface modeling, there are also studies on the use of regression or artificial neural network methods on mathematical-theoretical surfaces [6–8].

For other important studies in the literature, a search was made on the Scopus database with the keywords "machine AND learning, AND elevation AND point AND estimation" and 27 studies were found. The relationship map of the keywords in these studies was obtained in the VOSwiewer software (Figure 1).

In Figure 1, it is seen that methods such as deep learning and artificial neural networks and keywords such as remote sensing, 3D point cloud, classification is more prominent. Especially deep learning and artificial neural networks methods and similar machine learning methods are used successfully in solving many engineering problems [9–20]. In addition, when the years of these publications are examined, it is seen that they are between 2020 and 2022. It is seen that this situation is among the studies that have been researched in recent years and the keywords researched are included in current studies. This study differs from the literature in terms of the changing field of the study area.

In this study, elevation estimations were made using latitude and longitude information of points obtained by remote sensing and three different heuristic regression techniques (M5-tree, MARS and LSSVR). A region with moderate (flat-mountainous area with mid-level elevations) heights (Z) was chosen as the study area. Predictions were made separately for three different sized areas and compared.

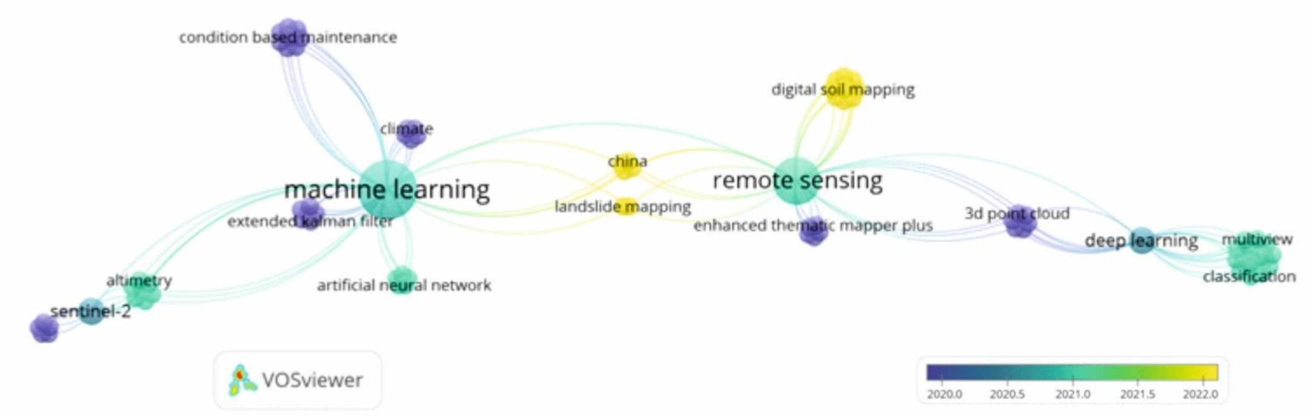

Figure 1. The relationship of the keywords of the similar studies.

2. Material and Method

The study area is located in the Marmara region of Türkiye, the region at the intersection of Kırklareli, Edirne and Tekirdağ provinces was selected, and data were obtained for three different sized areas. These areas are 1x1 km, 10x10 km and 100x100 km square areas. The data were obtained with the help of Google-Earth Pro. The following study can be examined for the methodology used in obtaining the data [21]. The study area is shown in Figure 2.

In Figure 2, areas of three different sizes are represented by square polygons. In choosing this area as the model area, the distinction in the classification of heights was considered. The height points within these square areas are shown in Figure 3. Hassan et al. [22], areas with a height difference of up to 0.06-5 meters are considered as flat-mountainous. In this study, the estimations were made in a flat-mountainous region.

In Figure 3, it is seen that the points are randomly distributed. The reason why this distribution is preferred is to ensure that the models give unbiased and non-memorizing predictions. Points that go out of the study area are not included in the modeling. Statistical information about the data is given in Table 1-3. The flowcart of the study is shown in Figure 4.

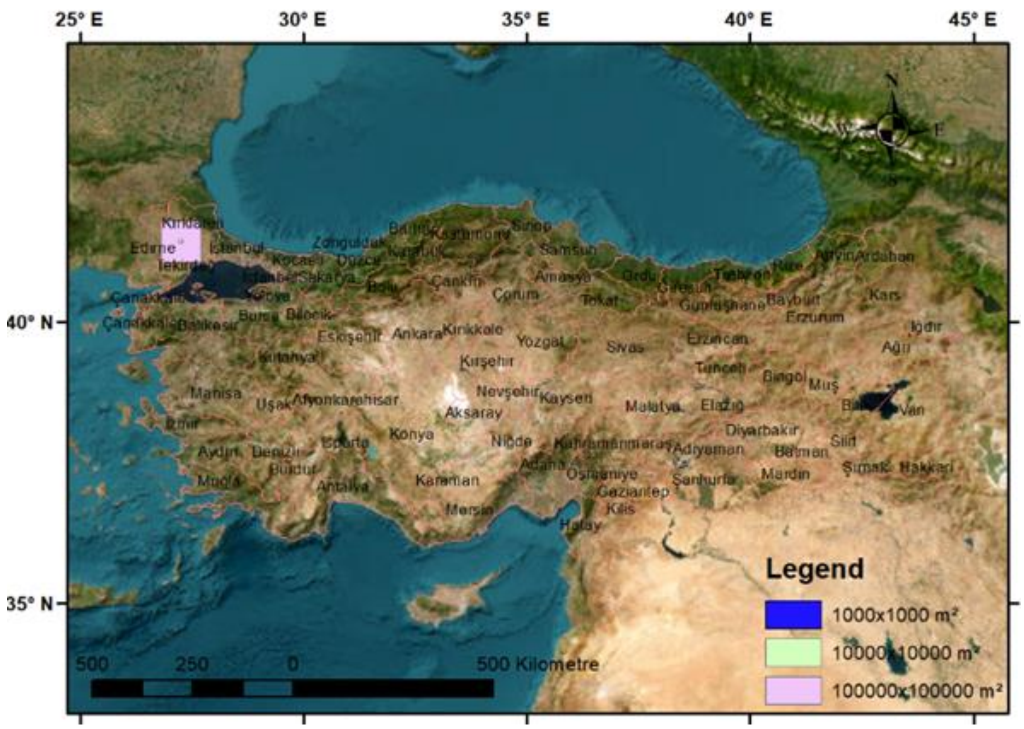

Figure 2. Study area.

Figure 3. Elevation points.

Table 1. Statistical indicators of model data for 1x1 km area.

Data Set	Training			Testing		
Variable	Latitude	Longitude	H	Latitude	Longitude	H
Number of Data	2100	2100	2100	1400	1400	1400
Maximum Value (m)	27.25	41.373	37.888	27.25	41.373	37.505
Minimum Value (m)	27.241	41.366	33	27.241	41.366	33
Average (m)	27.245	41.369	35.27	27.245	41.369	35.269
Standard Deviation	0.003	0.002	0.834	0.003	0.002	0.829
Skewness Coefficient	0.022	0.16	-0.083	-0.011	0.075	-0.092

Table 2. Statistical indicators of model data for 10x10 km area.

Data Set	Training			Testing		
Variable	Latitude	Longitude	H	Latitude	Longitude	H
Number of Data	2100	2100	2100	1400	1400	1400
Maximum Value (m)	27.296	41.416	109.000	27.296	41.413	97.504
Minimum Value (m)	27.198	41.332	31.092	27.196	41.332	31.759
Average (m)	27.245	41.365	52.114	27.246	41.366	52.698
Standard Deviation	0.028	0.021	14.188	0.028	0.021	14.123
Skewness Coefficient	-0.005	0.265	0.722	-0.064	0.197	0.587

Table 3. Statistical indicators of model data for 100x100 km area.

Data Set	Training			Testing		
Variable	Latitude	Longitude	H	Latitude	Longitude	H
Number of Data	2100	2100	2100	1400	1400	1400
Maximum Value (m)	27.739	41.797	644.837	27.729	41.777	462.810
Minimum Value (m)	26.743	40.995	16.022	26.748	41.001	17.000
Average (m)	27.262	41.316	131.044	27.249	41.309	128.260
Standard Deviation	0.273	0.210	69.874	0.277	0.207	66.079
Skewness Coefficient	-0.124	0.372	1.581	-0.066	0.410	1.304

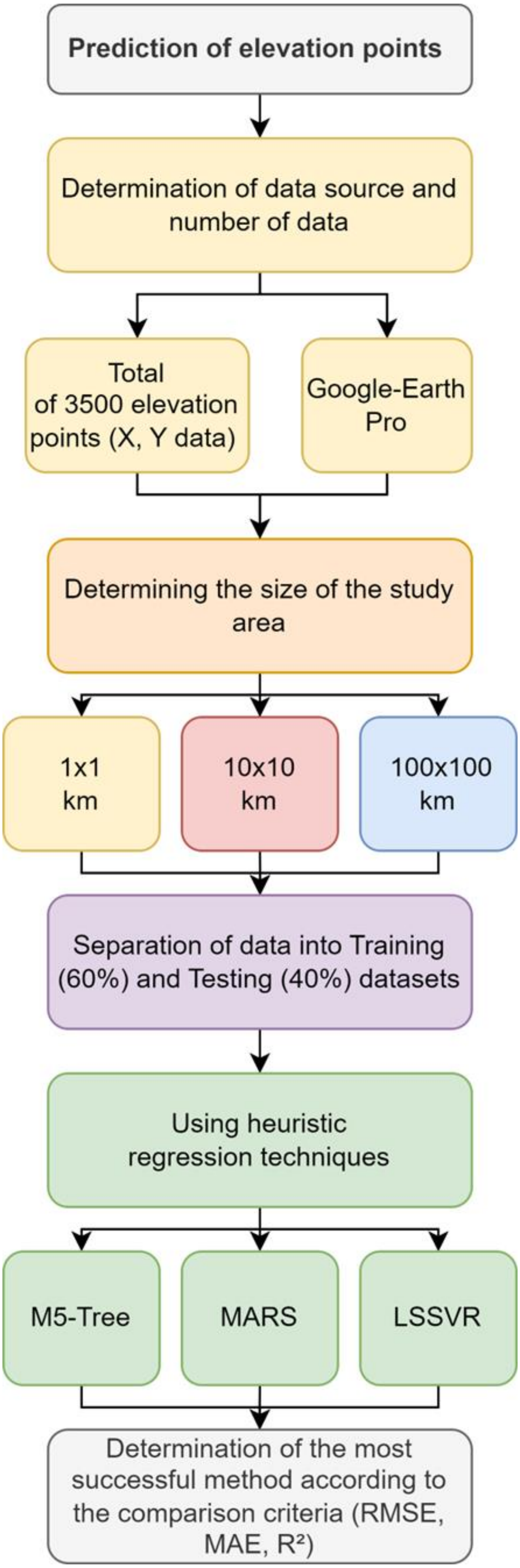

Figure 4. Flow chart of the study.

2.1. M5-Tree

The M5 model tree algorithm is a new regression method developed by Quinlan in 1992 [23]. The M5 model tree is better than other decision tree models used for categorical data. The model also gives successful predictions in numerical data [24].

The M5 model fits the model in two steps. The data are divided into sets in the first stage and created a decision tree. The splitting of the decision tree is based on calculating the predicted reduction in this error as a result of evaluating each attribute at the node and utilizing the standard deviations of the class values that reach a node as measurements of the error at the nodes [25]. The formulation of the standard deviation reduction (SDR) is as follows.

The formulation of the standard deviation reduction (SDR) is shown in Equation 1.

$$SDR = sd(T) - \sum \frac{|T_i|}{|T|} sd(T_i) \tag{1}$$

In Equation 1, T_i is the subset of examples that have the i^{th} possible outcome of the set, SD is the standard deviation, and T is a set of examples that reach the node.

2.2. Multivariate Adaptive Regression Splines (MARS)

The MARS is a type of regression analysis developed by Friedman [26]. This method is one of the non-parametric regression techniques, which is an extension of linear models.

It explains the complex nonlinear relationship between the model, estimation method and dependent variables. The MARS algorithm consists of two steps, forward and backward. It selects a set of suitable input variables with the forward step algorithm [27]. With the backward step algorithm, it eliminates unnecessary variables in the pre-selected set. This method also increases the accuracy of the predictions. The function is drawn from variable X to the new variable Y by two base functions or both variable values defined at the deviation point across the input range in Equation 2-3 [28].

$$Y = \max(0, X - c) \tag{2}$$

$$Y = \max(0, c - X) \tag{3}$$

Here c represents the threshold (lower limit) value. MARS model is used especially in financial affairs management system, time series data in engineering and in many fields [5,29–33].

2.3. Least Square Support Vector Regression (LSSVR)

LSSVR is an extended version of the support vector regression (SVR) model by Suykens and Vandewalle [34]. In this study, the optimal mapping function between inputs and outputs of LSSVR is used to estimate with statistically randomly distributed x and y values for z values. It performs this operation with a nonlinear relationship function with a multidimensional feature space. The regression function can be formulated in Equation 4.

$$y(x) = w^{1}\varphi(x) + b \tag{4}$$

Here y is the value obtained in x, w is the coefficient vector, φ is the mapping function, b is the bias term obtained by minimizing the upper bound of the generalization error [34].

3. Results

In all areas, 3500 elevation points were used for the study; 60% of these points were used for testing, and 40% were used for training. This amount remains constant across all places. The M5-tree, MARS, and LSSVR models were employed in the study. The models' outputs were assessed using three different comparison metrics. These included the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2) in Equation 5-7. In addition, three different input combinations were tried in the modeling: (i) X (1 input); (ii) Y (1 input); (iii) X, Y (2 inputs). Model performance is evaluated as more successful with the RMSE and MAE values approaching the minimum and the R^2 value approaching 1.

The observed and predicted height in the above equations is denoted by Z. N stands for the amount of data. The training and testing results of the three models are given in Table 4. The flow chart of the study is given below. Figure 5 and Figure 6 shows the scatter plots of the most successful training and test results for each method.

$$RMSE = \frac{1}{n}\sum_{i=1}^{n}\sqrt{\left(Z_{\text{predicted}} - Z_{measured}\right)^2} \tag{5}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Z_{\text{predicted}} - Z_{measured}\right| \tag{6}$$

$$R^2 = \frac{\sum_{i=1}^{n}\left(Z_{i\,\text{measured}} - \overline{Z_{i\,\text{measured}}}\right)^2 \cdot \left(Z_{i\,\text{predicted}} - \overline{Z_{\text{predicted}}}\right)^2}{\sum_{i=1}^{n}\left(Z_{i\,\text{measured}} - \overline{Z_{i\,\text{measured}}}\right)^2 \cdot \sum_{i=1}^{n}\left(Z_{i\,\text{predicted}} - \overline{Z_{\text{predicted}}}\right)^2} \tag{7}$$

Table 4. Results of the training and test phase.

Model	Region/Area	Input	Training			Testing		
			RMSE	MAE	R^2	RMSE	MAE	R^2
M5-Tree	1x1 km	X	0.702	0.548	0.292	0.854	0.689	0.043
		Y	0.572	0.422	0.530	0.745	0.576	0.237
		X and Y	0.187	0.116	**0.950**	0.312	0.199	**0.861**
	10x10 km	X	12.434	9.564	0.232	15.006	11.915	0.008
		Y	7.574	5.243	0.715	9.430	6.765	0.566
		X and Y	2.400	1.502	**0.971**	4.198	2.770	**0.913**
	100x100 km	X	58.617	41.965	0.296	69.241	51.917	0.041
		Y	43.657	29.724	0.610	54.212	39.644	0.368
		X and Y	16.415	10.672	**0.945**	25.032	18.244	**0.857**
MARS	1x1 km	X	0.787	0.646	0.111	0.793	0.654	0.086
		Y	0.722	0.583	0.251	0.723	0.594	0.239
		X and Y	0.695	0.561	**0.307**	0.686	0.557	**0.316**
	10x10 km	X	13.980	11.180	0.029	13.890	11.276	0.035
		Y	9.004	6.848	0.597	8.950	6.941	0.599
		X and Y	8.703	6.615	**0.624**	8.588	6.675	**0.631**
	100x100 km	X	65.767	49.845	0.114	63.067	48.677	0.092
		Y	52.698	39.518	0.431	51.263	39.277	0.400
		X and Y	47.154	33.864	**0.545**	46.824	34.069	**0.503**
LSSVR	1x1 km	X	0.807	0.664	0.065	0.802	0.664	0.065
		Y	0.747	0.604	0.198	0.739	0.606	0.206
		X and Y	0.560	0.440	**0.552**	0.568	0.452	**0.533**
	10x10 km	X	14.036	11.256	0.022	13.842	11.254	0.049
		Y	9.093	6.947	0.589	9.019	7.023	0.592
		X and Y	5.013	3.622	**0.875**	5.174	3.748	**0.866**
	100x100 km	X	66.069	50.381	0.106	63.238	48.993	0.087
		Y	53.081	39.725	0.423	51.678	39.675	0.392
		X and Y	27.196	20.750	**0.849**	27.954	21.387	**0.822**

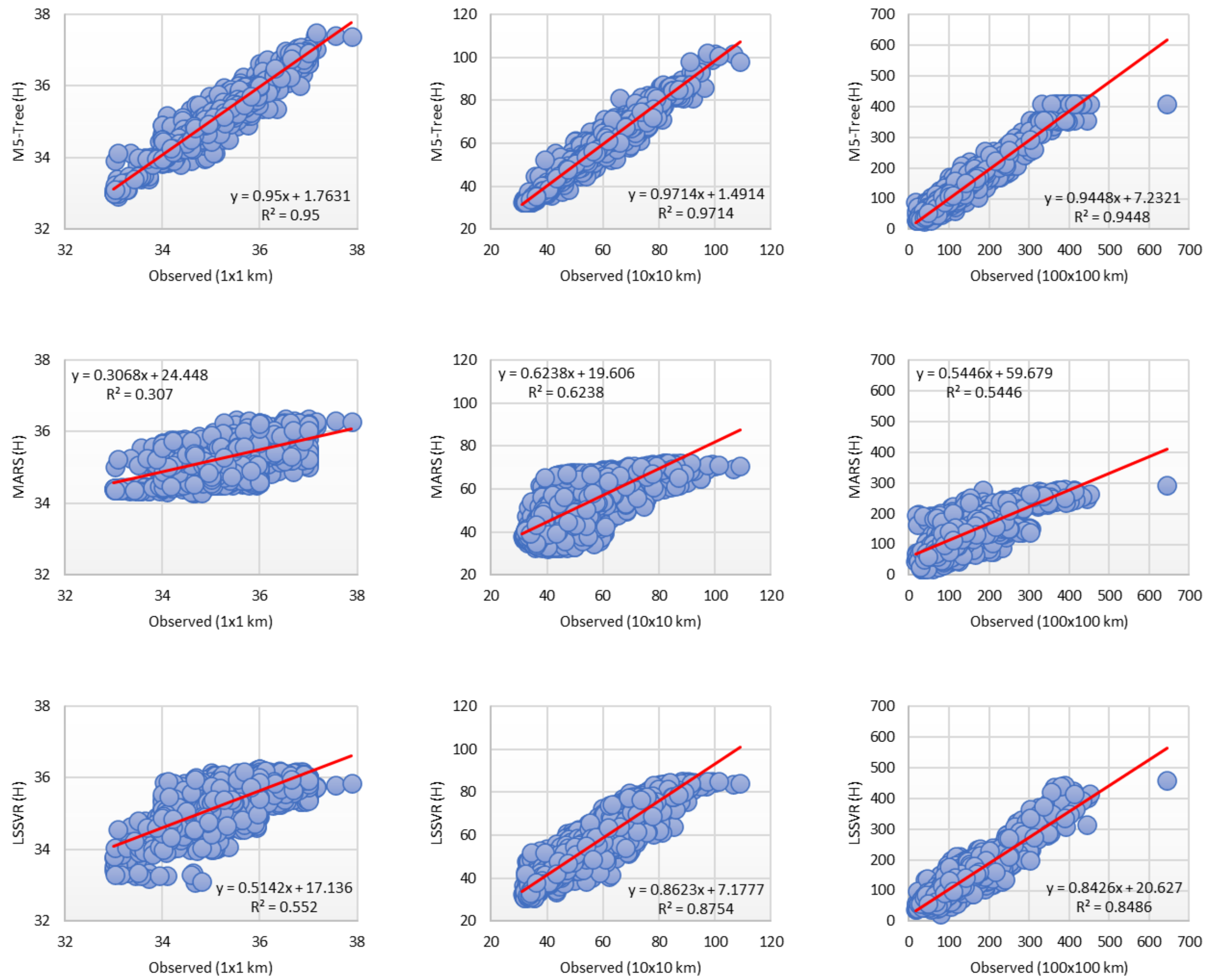

Figure 5. Scatter plots of best models for training phase.

In Table 4, the results of the most successful training phase were obtained in the data package using two inputs. The lowest error values and the highest R^2 value were obtained in the M5-tree method during the training phase, followed by LSSVR, and the more unsuccessful model was determined as MARS. Considering X and Y coordinate inputs, more successful results were observed in input sets using Y coordinate information. When the results are evaluated in terms of areas, the most successful results were observed in areas of 10x10 km. According to the coefficient of determination, this situation is similar in all models. According to the RMSE criterion, the models that give the least error are the models in which points in 1x1 km areas are used. The reason for this is that the points are closer to each other than other areas. As a result, the training has been more successful. Depending on the area growth, the prediction performance decreases as the points move away from each other.

When the results in the test phase were examined, the most successful results (considering the RMSE) were observed in the M5-tree method as in the training phase, in 1x1 km areas and in models using X and Y input data sets. Then, LSSVR and MARS methods made successful predictions. The highest coefficient of determination was observed in areas of 10x10 km, as in the training phase.

4. Discussion

As a result of the modeling, the most successful results were obtained in the 1x1 km area where the area is the least. In this model, both X and Y coordinates information are used. In addition, it has been observed that Y coordinates information gives more successful results in the study area than X coordinates information. In this study, modeling was done on fixed points (3500 units) but increasing sizes. It is seen that the error value increases depending on the increasing areas. For this reason, for more successful results in larger areas, either by increasing the ratios of the training-test datasets or with new measurements, the points should be added.

The effect of the number of data can be investigated by changing the training and testing rates. However, in any case, the increase in the number of training data sets means not that the model performance will always increase [35]. In addition, with point compaction, lower resolution raster data can be compressed with similar models and higher resolution (pixel size) models can be obtained.

Models were run on a computer with 12th Gen Intel(R) Core (TM), i7-12700H, 2.30 GHz, 64 GB RAM and 6 GB graphics card and the modeling times were compared, MARS and M5-tree yielded modeling results quite recently (average of 5 sec). But the LSSVR model took a lot of time (average 1 hour). The kernels of LSSVR contain two modification parameters (γ, α). In order to obtain the optimum of these parameters, cycles were established for these parameters from 1 to 100 and it was observed that the most successful results (100,1) were generally observed in coefficient pairs. Therefore, the method took more time. As a result of the study, it has been determined that MARS and M5-Tree are more advantageous because they do not have any model parameters. Although the results of the LSSVR model are close to the M5-tree, faster and more successful results can be obtained with a hybrid optimization technique.

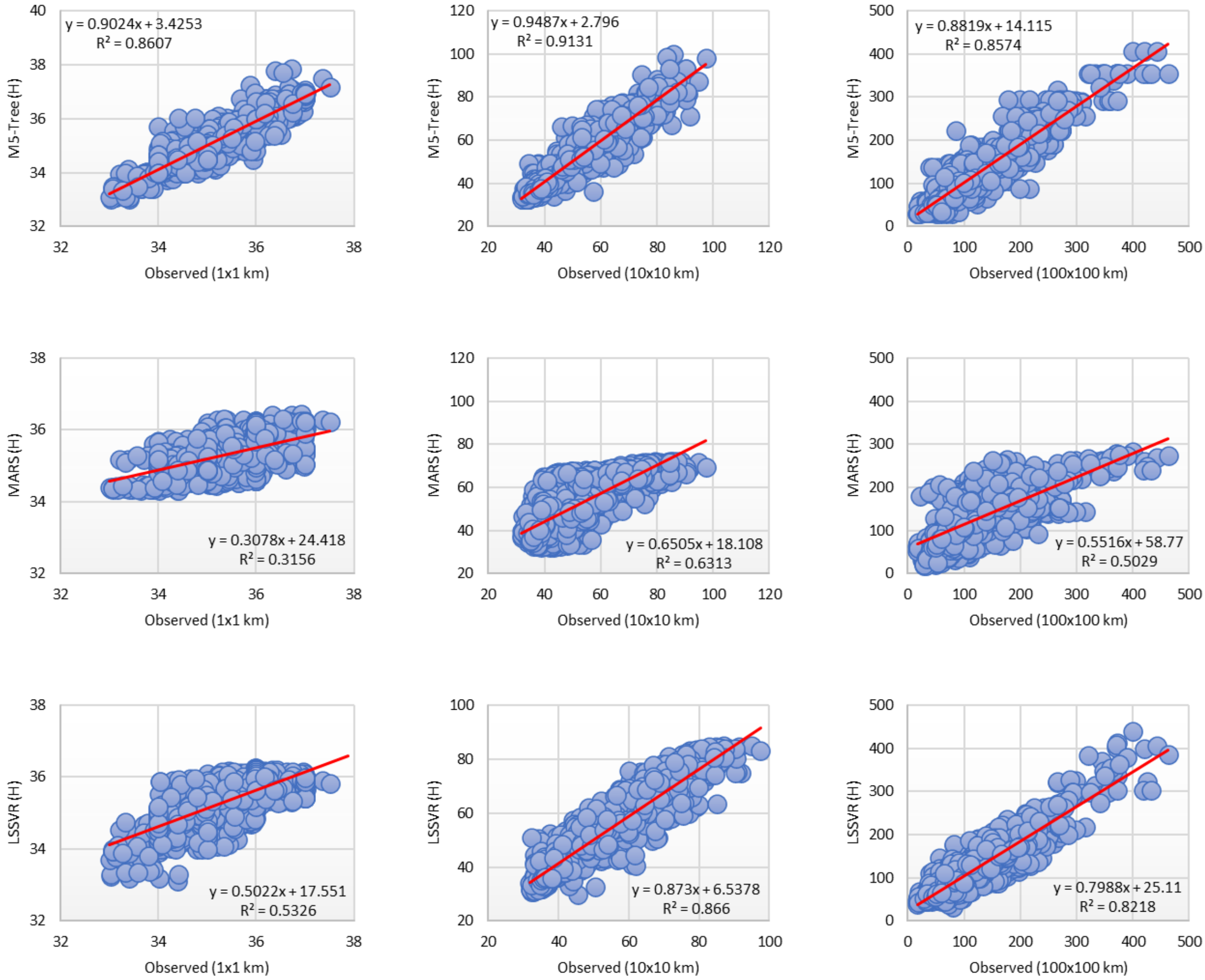

Figure 6. Scatter plots of best models for testing phase.

5. Conclusion

In this study, the estimation of the elevation points in a flat-mountainous area where a fixed number of data is randomly distributed over three different sized areas was performed. Three different models were used in the study: M5-tree, MARS and LSSVR. The training and testing rates in the models are 60% and 40%, respectively, and the performance of three different input types in the models was investigated. These are 1-only X coordinates, 2-only Y coordinates, and 3-both X and Y coordinates information are used. In the study, a total of 3500 points belonging to the fields were obtained from the Google earth pro database and the study areas are 1x1 km, 10x10 km and 100x100 km, respectively. When the results are examined;

- The most successful results in models were obtained as a result of using 2-input data sets.
- Models using Y location information are more successful than models using X location information. Therefore, it is important to minimize these coordinate errors.
- Increasing the area did not increase the model performance.
- Although the coefficient of determination is highest in areas of 10x10 km, the lowest errors were detected in areas of 1x1 km.
- As a result of the study, the most successful method is M5-Tree, followed by LSSVR and MARS methods.

The limits of this study are as follows, using the same number of data sets in areas of different sizes, making predictions and comparisons using three different heuristic regression techniques, using X and Y location information as input data in the models, providing the data source in Google earth pro software with remote sensing techniques.

In future studies, results for different fields will be investigated using different training and testing rates and different methods. In addition, the effect of the change in the topography of the study area on the performance will be investigated.

Acknowledgement

The authors thank KTO Karatay University and TÜBİTAK.

Funding

This study is supported by the application numbered **1919B012107905** within the scope of the **2209-A** University Students Research Projects Support Program 2021-2, carried out by the **TÜBİTAK** Scientist Support Programs Presidency (BİDEB).

Author contributions

Vahdettin Demir: Writing-Reviewing, Methodology, Application, Editing, **Ramazan Doğu:** Methodology, Application.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Demir, V., & Ulke Keskin, A. (2020). Height Modelling with Artificial Neural Networks (Samsun_Mert River Basin). Gazi Journal of Engineering Sciences, 6(1), 54-61. https://dx.doi.org/10.30855/gmbd.2020.01.05
2. Sahin, İ., Yakar, M. (2008). Farklı kaynaklardan elde edilen sayısal yükseklik modellerinin ortofoto doğruluğuna etkilerinin araştırılması. Harita Dergisi, 74(140), 45-59.
3. Yakar, M. (2009). Digital elevation model generation by robotic total station instrument. Experimental Techniques, 33, 52-59. https://doi.org/10.1111/j.1747-1567.2008.00375.x
4. Yakar, M., Yilmaz, H. M., & Yurt, K. (2010). The effect of grid resolution in defining terrain surface. Experimental Techniques, 34, 23-29. https://doi.org/10.1111/j.1747-1567.2009.00553.x
5. Demir, V., & Çubukçu, E. A. (2021). Sezgisel Regresyon Teknikleri ile Sayısal Yükseklik Modellenmesi. Avrupa Bilim ve Teknoloji Dergisi, (24), 484-488. https://doi.org/10.31590/ejosat.916012
6. Çakır, L. (2013) Sayısal Yükseklik Modellerinde Polinomlar ve Yapay Sinir Ağları Yöntemlerinin Karşılaştırılması. In Proceedings of the Türkiye Ulusal Fotogrametri ve Uzaktan Algılama Birliği VII. Teknik Sempozyumu (TUFUAB'2013), 23-25 Mayıs 2013, 1–4, Trabzon, Türkiye.
7. Çakır, L. (2015) Sayısal Yükseklik Modellerinde Klasik ve Esnek Hesaplama Yöntemlerinin Karşılaştırılması. In Proceedings of the TMMOB Harita ve Kadastro Mühendisleri Odası, 15. Türkiye Harita Bilimsel ve Teknik Kurultayı, 25--28 Mart 2015, 1-6, Ankara, Türkiye
8. Konakoglu, B., Cakır, L., & Gökalp, E. (2016). 2D coordinate transformation using artificial neural networks. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 183-186. https://doi.org/10.5194/isprs-archives-XLII-2-W1-183-2016
9. Biyik, M. Y., Atik, M. E., & Duran, Z. (2023). Deep learning-based vehicle detection from orthophoto and spatial accuracy analysis. International Journal of Engineering and Geosciences, 8(2), 138-145. https://doi.org/10.26833/ijeg.1080624
10. Çubukçu, E. A., Demir, V., & Sevimli, M. F. (2023). Modeling of annual maximum flows with geographic data components and artificial neural networks. International Journal of Engineering and Geosciences, 8(2), 200-211. https://doi.org/10.26833/ijeg.1125412
11. Demiryege, İ., & Ulukavak, M. (2022). Derin öğrenme tabanlı iyonosferik TEC tahmini. Geomatik, 7(2), 80-87. https://doi.org/10.29128/geomatik.870773
12. Demirgül, T., Demir, V., & Sevimli, M. F. (2023). Model-Ağacı (M5-tree) yaklaşımı ile HELIOSAT tabanlı güneş radyasyonu tahmini. Geomatik, 8(2), 124-135. https://doi.org/10.29128/geomatik.1137687
13. Kotan, B., & Erener, A. (2023). PM10, SO2 hava kirleticilerinin çoklu doğrusal regresyon ve yapay sinir ağları ile sezonsal tahmini. Geomatik, 8(2), 163-179. https://doi.org/10.29128/geomatik.1158565
14. Tasdemir, S., & Ozkan, I. A. (2019). ANN approach for estimation of cow weight depending on photogrammetric body dimensions. International Journal of Engineering and Geosciences, 4(1), 36-44. https://doi.org/10.26833/ijeg.427531
15. Uncuoglu, E., Citakoglu, H., Latifoglu, L., Bayram, S., Laman, M., Ilkentapar, M., & Oner, A. A. (2022). Comparison of neural network, Gaussian regression, support vector machine, long short-term memory, multi-gene genetic programming, and M5 Trees methods for solving civil engineering problems. Applied Soft Computing, 129, 109623. https://doi.org/10.1016/j.asoc.2022.109623
16. Bayram, S., & Çıtakoğlu, H. (2023). Modeling monthly reference evapotranspiration process in Turkey: application of machine learning methods. Environmental Monitoring and Assessment, 195(1), 67. https://doi.org/10.1007/s10661-022-10662-z
17. Zeybekoglu, U. (2018). Forecasting of Annual Mean Rainfall Using Artificial Neural Network and Wavelet Components: Case of Study Sinop Forecasting of Annual Mean Rainfall Using Artificial Neural Network and Wavelet Components: Case of Study Sinop. In Proceedings of the 1. International Technological Sciences and Design Symposium, 1700-1709, Giresun Türkiye.

18. Hezarani, A. B., Zeybekoğlu, U., & Keskin, A. Ü. (2021). Hydrological and meteorological drought forecasting for the Yesilirmak river basin, Turkey. Sürdürülebilir Mühendislik Uygulamaları ve Teknolojik Gelişmeler Dergisi, 4(2), 121-135. https://doi.org/10.51764/smutgd.993792
19. Öztürk, A., Allahverdi, N., & Saday, F. (2022). Application of artificial intelligence methods for bovine gender prediction. Turkish Journal of Engineering, 6(1), 54-62. https://doi.org/10.31127/tuje.807019
20. Gülgün, O. D., & Hamza, E. R. O. L. (2020). Classification performance comparisons of deep learning models in pneumonia diagnosis using chest x-ray images. Turkish Journal of Engineering, 4(3), 129-141. https://doi.org/10.31127/tuje.652358
21. Demir, V., & Doğu, R. (2022). Creating digital elevation model with Google Earth Pro. Advanced Engineering Days (AED), 4, 78-80.
22. Hassan, O., Elnazeer, E., & Zomrawi, N. (2015). Application of Artificial Neural Network for Height Modelling. International Journal of Recent and Innovation Trends in Computing and Communication, 3(3), 1374-1377.
23. Quinlan, J. R. (1992) Learning with Continuous Classes. Proceedings of Australian Joint Conference on Artificial Intelligence, Hobart 16-18 November 1992, 343-348.
24. Mitchell, T. M. (1997). Machine learning. McGraw-Hill Science, ISBN: 0070428077
25. Srivastava, R., Tiwari, A. N., & Giri, V. K. (2019). Solar radiation forecasting using MARS, CART, M5, and random forest model: A case study for India. Heliyon, 5(10), e02692. https://doi.org/10.1016/j.heliyon.2019.e02692
26. Friedman, J. H. (1991). Multivariate adaptive regression splines. The annals of statistics, 19(1), 1-67. https://doi.org/10.1214/aos/1176347963
27. De Andrés, J., Lorca, P., de Cos Juez, F. J., & Sánchez-Lasheras, F. (2011). Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). Expert Systems with Applications, 38(3), 1866-1875. https://doi.org/10.1016/j.eswa.2010.07.117
28. Sharda, V. N., Patel, R. M., Prasher, S. O., Ojasvi, P. R., & Prakash, C. (2006). Modeling runoff from middle Himalayan watersheds employing artificial intelligence techniques. Agricultural water management, 83(3), 233-242. https://doi.org/10.1016/j.agwat.2006.01.003
29. Yaseen, Z. M., Kisi, O., & Demir, V. (2016). Enhancing long-term streamflow forecasting and predicting using periodicity data component: application of artificial intelligence. Water resources management, 30, 4125-4151. https://doi.org/10.1007/s11269-016-1408-5
30. Adnan, R. M., Petroselli, A., Heddam, S., Santos, C. A. G., & Kisi, O. (2021). Comparison of different methodologies for rainfall–runoff modeling: machine learning vs conceptual approach. Natural Hazards, 105, 2987-3011. https://doi.org/10.1007/s11069-020-04438-2
31. Kisi, O., Parmar, K. S., Soni, K., & Demir, V. (2017). Modeling of air pollutants using least square support vector regression, multivariate adaptive regression spline, and M5 model tree models. Air Quality, Atmosphere & Health, 10, 873-883. https://doi.org/10.1007/s11869-017-0477-9
32. Bera, P., Prasher, S. O., Patel, R. M., Madani, A., Lacroix, R., Gaynor, J. D., ... & Kim, S. H. (2006). Application of MARS in simulating pesticide concentrations in soil. Transactions of the ASABE, 49(1), 297-307. https://doi.org/10.13031/2013.20228
33. Sephton, P. (2001). Forecasting recessions: can we do better on MARS?. Federal Reserve Bank of St. Louis, 83, 39-49.
34. Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural processing letters, 9, 293-300. https://doi.org/10.1023/A:1018628609742
35. Çubukçu, E. A., Demir, V., & Sevimli, M. F. (2022). Digital elevation modeling using artificial neural networks, deterministic and geostatistical interpolation methods. Turkish Journal of Engineering, 6(3), 199-205. https://doi.org/10.31127/tuje.889570