BEÜ Fen Bilimleri Dergisi
8 (4), 1327-1337, 2019

BEU Journal of Science
8 (4), 1327-1337, 2019

*Araştırma Makalesi / Research Article*

# Monthly Soil Temperature Modeling Using Gene Expression Programming

Erkan DİKMEN[1], Kazım KUMAŞ[2], Arzu ŞENCAN ŞAHİN[1*]

[1]*Isparta University of Applied Sciences, Faculty of Technology, Department of Mechanical Engineering, Isparta*
[2]*Burdur Mehmet Akif Ersoy University, Department of Electrical and Energy, Burdur*
*(ORCID: 0000-0002-6804-8612) (ORCID: 0000-0002-2348-4664) (ORCID: 0000-0001-8519-4788)*

**Abstract**
Soil temperature is a critical variable controlling below-ground processes for global and continental carbon budgets. However, there are an insufficient number of climatic stations monitoring soil temperature. In this study, GEP model was used for estimation of monthly soil temperature using air temperature, depth, relative humidity and solar radiation data for the Antalya, Isparta, and Burdur in Turkey. This model was tested using measured meteorological data. The values of $R^2$ between observed and predicted soil temperatures ranged from 0.95 to 0.97. Predictions with GEP model show good agreement with actual soil temperature measurements. New equations are presented for calculation of soil temperatures at different depths. The GEP-based formulations are very practical to predict soil temperature. Soil temperature prediction with GEP model is helpful in various processes, including agricultural decision, heating or cooling of buildings and ground-source heat pump applications.

**Keywords**: GEP, soil temperature, meteorological data, modeling.

# Gen İfade Programlama Kullanılarak Aylık Toprak Sıcaklığının Modellenmesi

**Öz**
Toprak sıcaklığı, küresel ve karasal karbon bütçeleri için yer altı süreçlerini kontrol eden kritik bir değişkendir. Ancak, toprak sıcaklığını izleyen az sayıda iklim istasyonu vardır. Bu çalışmada, Antalya, Isparta ve Burdur illeri için hava sıcaklığı, derinlik, bağıl nem ve güneş ışınımı verileri yardımıyla aylık toprak sıcaklığının tahmini için GEP modeli kullanılmıştır. Bu model ölçülen meteorolojik veriler kullanılarak test edilmiştir. Ölçülen ve tahmin edilen toprak sıcaklıkları arasındaki $R^2$ değerleri 0.95 ila 0.97 arasında değişmiştir. GEP modeli ile yapılan tahminler, gerçek toprak sıcaklığı ölçümleriyle iyi bir uyum göstermektedir. Farklı derinliklerde toprak sıcaklıklarının hesaplanması için yeni denklemler sunulmuştur. GEP modelinden elde edilen denklemler, toprak sıcaklığını tahmin etmek için çok pratiktir. GEP modeli ile toprak sıcaklığı tahmini, tarımsal uygulamalar, binaların ısıtılması veya soğutulması ve toprak kaynaklı ısı pompası uygulamaları gibi işlemlerde oldukça yardımcı olacaktır.

**Anahtar kelimeler:** GEP, toprak sıcaklığı, meteorolojik veri, modelleme.

## 1. Introduction

Antalya, Isparta and Burdur are located in the Mediterranean Region. Agricultural processes in these cities are especially important. Accurate soil temperature predictions can dramatically affect the decision making process of the agricultural crops. Although soil temperature is a significant for agricultural and ground-source heat pump applications, there are not routinely soil temperature values in meteorological stations. There are many studies about soil temperature prediction in literature. Gao et al. [1] presented the revised force-restore technic for soil temperature estimation. Citakoglu [2] carried out comparison

[*]Sorumlu yazar: arzusencan@sdu.edu.tr
*Geliş Tarihi: 14.02.2019, Kabul Tarihi: 16.10.2019*

of artificial neural network (ANN), neuro-fuzzy (ANFIS), and multiple linear regression methods for estimation of soil temperatures in Turkey. Talaee [3] estimated daily soil temperature using neuro-fuzzy method in Iran. Behmanesh and Mehdizadeh [4] have estimated the soil temperature by gene expression programming (GEP) and ANN. They used air temperatures, relative humidity, wind speed, sunshine hours and extraterrestrial radiation as input. Kermani [5] analyzed the performance of ANN and multiple linear regression models for prediction of soil temperature. Kim and Singh [6] used multilayer perceptron and ANFIS for predicting daily soil temperature in Illinois. Kisi et al. [7] predicted soil temperatures at various depths by different neural networks methods. Mihalakakou [8] used deterministic model and neural network model for estimating soil surface temperature profiles. Bilgili [9] developed artificial neural network models to estimate monthly soil temperature by using monthly meteorological variables in Adana. Kişi et al. [10] compared neural computing methods for predicting monthly soil temperatures in Mersin. Wu et al. [11] performed spatiotemporal estimating of monthly soil temperature using ANN.

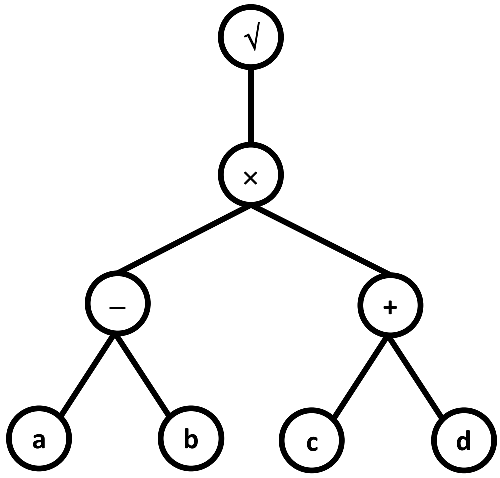
As seen above, soft computing techniques can be used for predicting of soil temperature. But, studies about estimation of soil temperature with GEP model are very limited. In this study, the GEP model was applied for predicting of soil temperature depending on three meteorological variables (air temperature, relative humidity and solar radiation) and depth for Antalya, Isparta and Burdur in Turkey. The performance of the GEP model was compared with the measured soil temperature values. Obtained mathematical equations from the GEP model can be easily used for predicting of soil temperature.

## 2. Materials and Methods

### 2.1. GEP Model Development

GEP is an evolutionary algorithm and was proposed by Ferreira [12]. The algorithm is based on the chromosomes and the expression trees.

The chromosome consists of a linear, symbolic string of fixed length composed of one or more genes. Each chromosome is comprised of genes that are translated into an expression tree to solve a given problem. An expression tree and mathematical expression is seen in Figure 1. Detailed information about GEP can be found in the References [12-15].




$$\sqrt{(a-b)\times(c+d)}$$

**Figure 1.** Expression tree diagram

The monthly weather data of the Antalya, Isparta and Burdur stations operated by the Turkish State Meteorological Service were used for the data set of GEP model. The location of the Antalya, Isparta and Burdur cities are shown in Figure 2. The data set is taken for the 17 year (2000–2016) monthly values of air temperature, relative humidity and solar radiation and soil temperature at different depths (5,10, 20, 50, and 100 cm).


**Figure 2.** The map of location of the stations in Turkey

In this work, GEP model was used for estimation of monthly soil temperature using air temperature, depth, relative humidity and solar radiation data for the Antalya, Isparta and Burdur. Various GEP parameters were employed for obtaining the excellent topology. The optimum GEP parameters for estimating of monthly soil temperature for the Antalya, Isparta, and Burdur are presented in Table 1. GeneXpro program for modeling was used.

**Table 1.** Parameters for prediction of the soil temperature of the stations

| **Parameters of GEP models** | **Stations** | | |
|---|---|---|---|
| | **Antalya** | **Isparta** | **Burdur** |
| Generations Number | 101537 | 86258 | 36470 |
| Chromosomes Number | 50 | 50 | 50 |
| Genes Number | 3 | 3 | 3 |
| Head size | 8 | 8 | 8 |
| Function set | + , – , * , / , power, √ , $10^x$, ln, sin, cos, tan, 1/x | + , – , * , / , power, √ , $10^x$, ln, log, sin, cos, tan, 1/x | + , – , * , / , power, √ , $10^x$, ln, sin, cos, tan, 1/x |
| $R^2$ | 0.9617 | 0.9763 | 0.9550 |

Root-mean-squared error (RMSE), mean absolute percentage error (MAPE) and R-square ($R^2$) were used for evaluating the accuracy of the GEP model. The RMSE, MAPE and $R^2$ can be expressed as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(m_i - p_i)^2}{n}} \tag{1}$$

$$MAPE = \frac{1}{n}\left[\frac{\sum_{i=1}^{n}|m_i - p_i|}{\sum_{i=1}^{n} m_i} x100\right] \tag{2}$$

$$R^2 = \frac{\left(n\sum m_i p_i - \sum m_i \sum p_i\right)^2}{\left(n\sum m_i^2 - \left(\sum m_i\right)^2\right)\left(n\sum p_i^2 - \left(\sum p_i\right)^2\right)} \tag{3}$$

where m is the measured soil temperature, p is the predicted soil temperature and n is total number of data.

## 3. Application and Results

Mathematical equations of monthly soil temperatures for the best results using GEP model are performed. These simple equations can be used for the estimation of the monthly soil temperatures in Antalya, Isparta and Burdur cities, Turkey. The corresponding equations for predicting monthly soil temperatures in Antalya, Isparta and Burdur cities from the best GEP model are presented as Equations (4–6), respectively:

$$T_s = \sin\left(\frac{R_h+M-T_a}{\frac{h}{R_a}+T_a}\right) + T_a + \sqrt{\frac{T_a-\frac{T_a}{M}}{\frac{R_h}{R_a}\times\frac{h}{T_a}}} \tag{4}$$

$$T_s = \frac{M}{\left(\frac{R_h\times M}{h\times R_a}\right)+\frac{R_a}{M}} + \left(\cos\left(\frac{M}{R_a}\right)\left(\frac{T_a}{h}-\frac{h}{R_h}\right)\right) + \frac{h}{\ln(M\times T_a)+R_a+T_a+(M\times T_a)} + T_a \tag{5}$$

$$T_s = \frac{\sqrt{h\times M}\times\frac{h}{R_a}}{M+R_h+(T_a\times R_a)} + \frac{1}{e^{\left(\frac{h}{R_a-R_h}-\cos R_h+T_a\right)}} + T_a \tag{6}$$

The regression curves of the monthly soil temperatures in Antalya, Isparta and Burdur are given in Figures 3-5. It can be seen from Figures 3-5 that the value of correlation coefficients is very high.

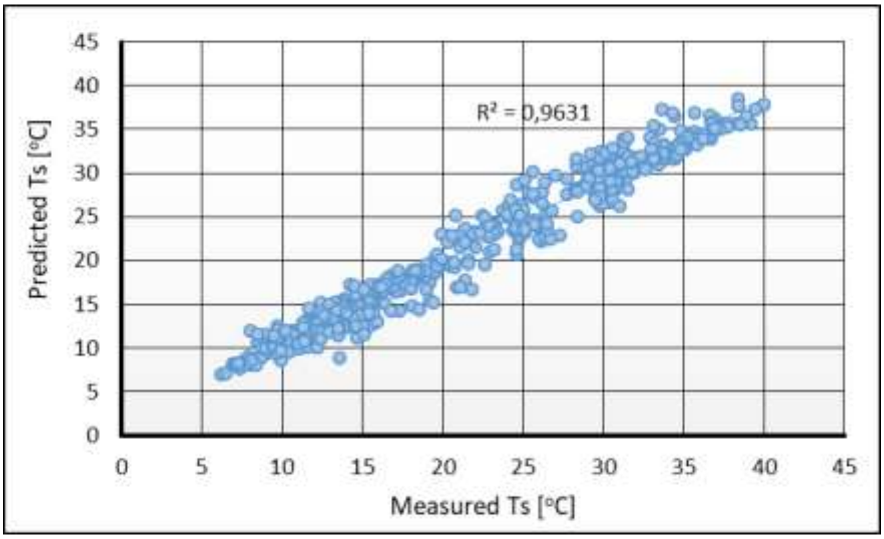



**Figure 3.** The correlation of the observed and predicted monthly soil temperature of the Antalya

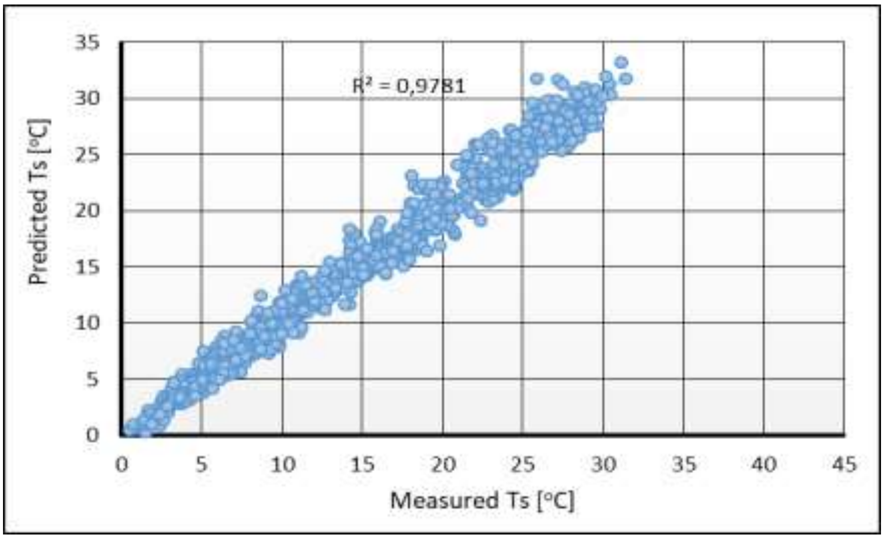



**Figure 4.** The correlation of the observed and predicted monthly soil temperature of the Isparta

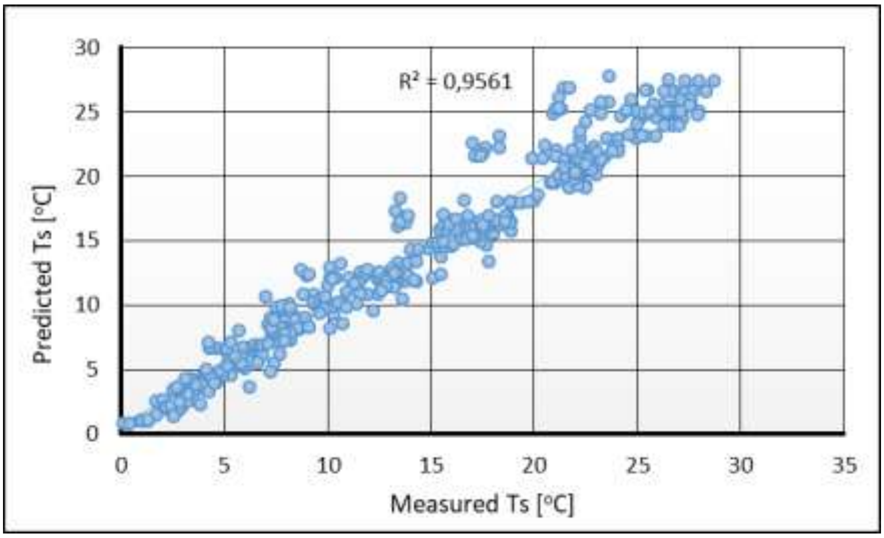



**Figure 5.** The correlation of the observed and predicted monthly soil temperature of the Burdur

The performance values of the GEP model, such as RMSE, MAPE and $R^2$ are given in Table 2. The performance values of the GEP model as seen in Table 2 are very satisfactory.

**Table 2.** Performance evaluation for predicting monthly soil temperature of the stations

| Statistical parameters | Stations | | |
|---|---|---|---|
| | Antalya | Isparta | Burdur |
| MAPE | 0.45248 | 0.17994 | 0.53234 |
| RMSE | 3.07284 | 1.58698 | 2.92649 |
| $R^2$ | 0.9617 | 0.9763 | 0.9550 |

The monthly soil temperatures for 2007 were estimated using Eqs. (4-6). Figures 6-8 show comparisons the measured and predicted monthly soil temperature values using GEP for different stations. As seen in Figs. 6-8, the predicted soil temperature values from the GEP model agree with the measured soil temperature values.

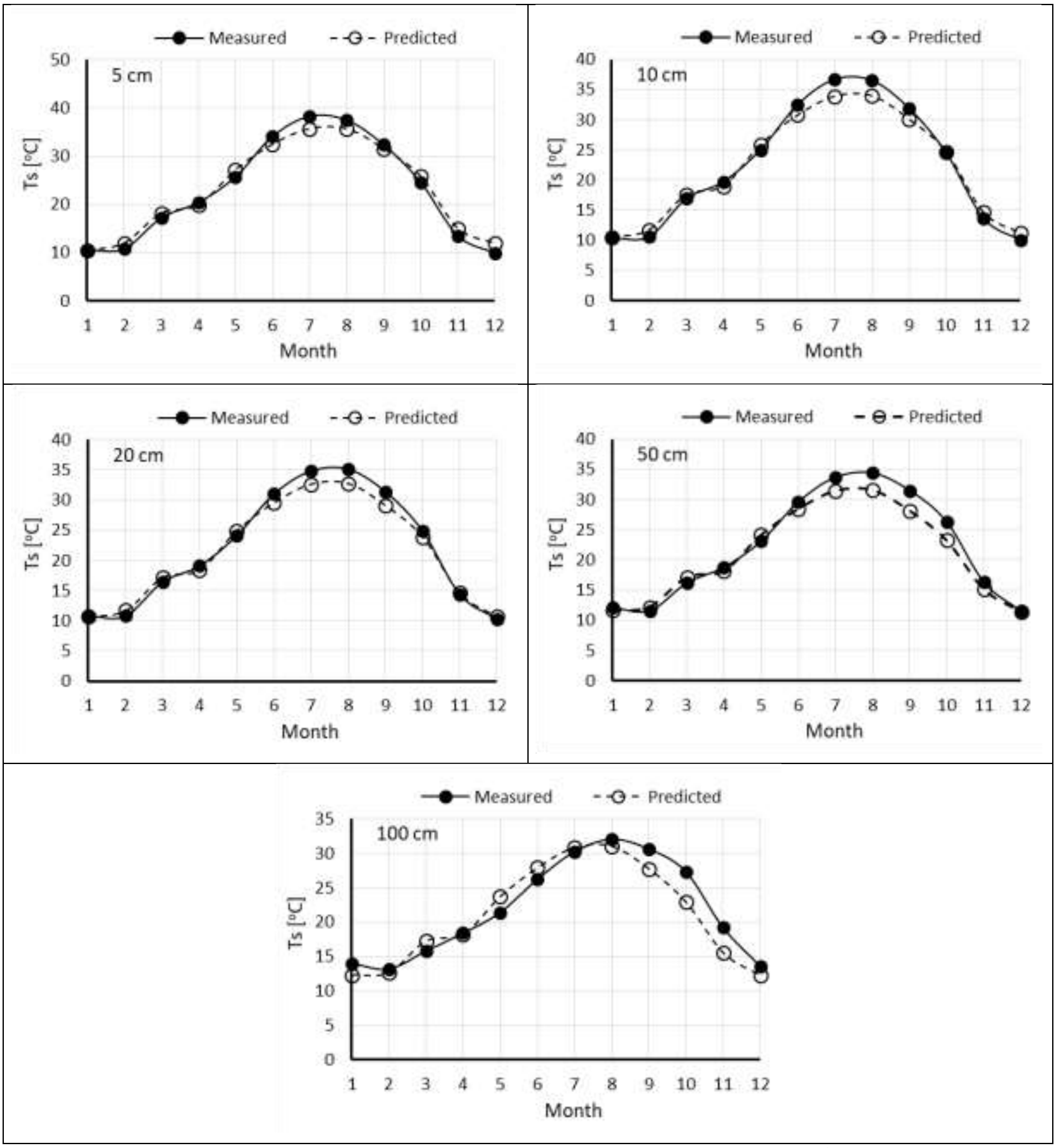



**Figure 6.** Comparison between GEP prediction and measured soil temperature for the Antalya

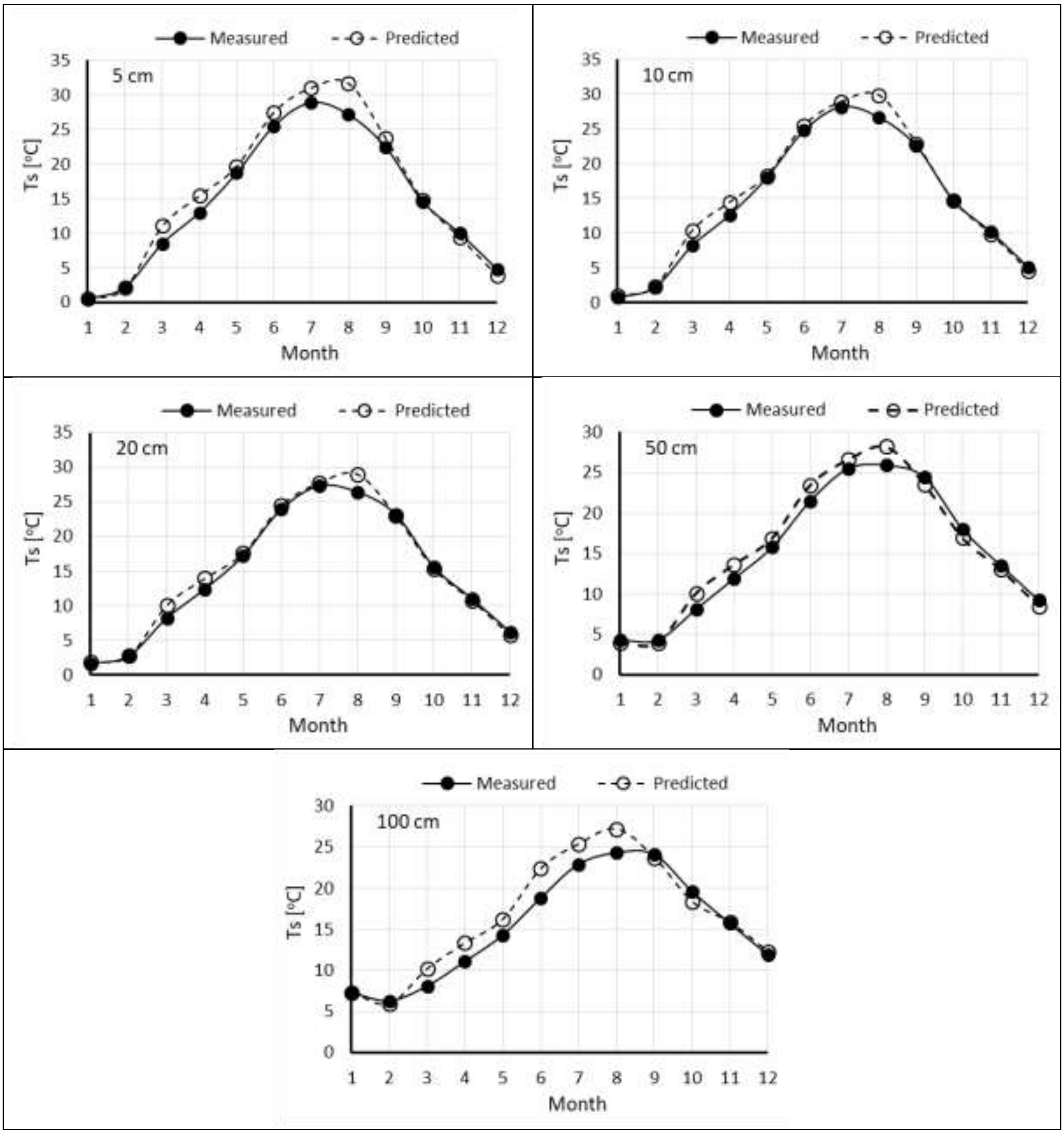



**Figure 7.** Comparison between GEP prediction and measured soil temperature for the Isparta

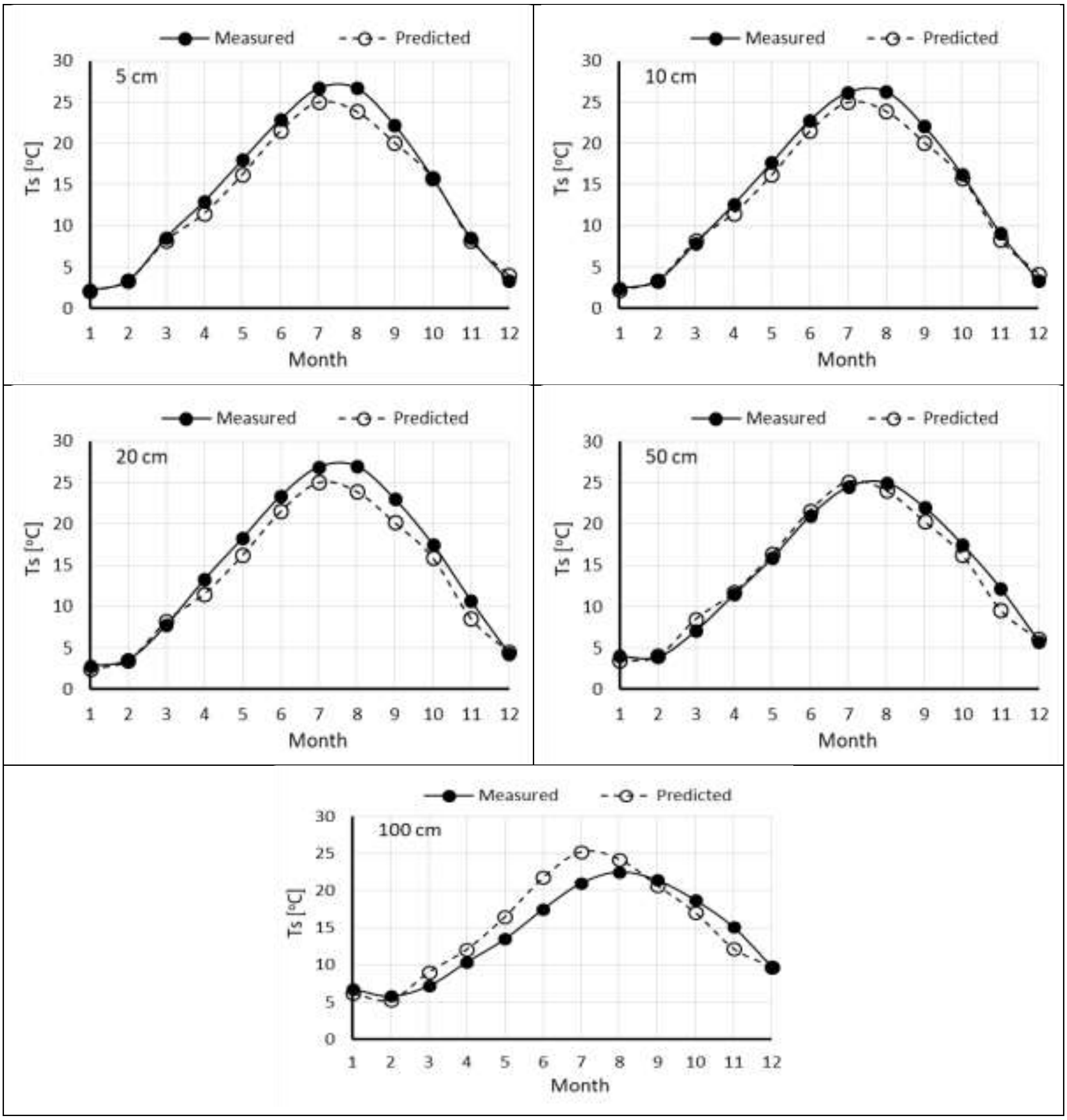



**Figure 8.** Comparison between GEP prediction and measured soil temperature for the Burdur

In addition, Tables 3-5 present a comparison of measured, GEP model, error and percentage difference for soil temperature at different depths of the Antalya, Isparta and Burdur. Obtained results from these tables, the error values for all stations are within acceptable limits.

**Table 3.** A comparison of measured, GEP prediction and error values for soil temperature of the Antalya.

| Antalya | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Month (M) | Ambient Temperature ($T_a$) | Relative Humidity ($R_h$) | Solar Radiation ($R_a$) | Depth (h) | Soil Temperature ($T_s$) Measured $T_s$ | Soil Temperature ($T_s$) Predicted $T_s$ | Error | Percentage difference |
| | (ºC) | (%) | (kcal/cm$^2$) | cm | (ºC) | (ºC) | | (%) |
| 1 | 11.4 | 67.5 | 6.0867 | 5 | 10.5 | 10.400 | 0.09918 | 0.944 |
| 2 | 11.5 | 59.8 | 8.1789 | 5 | 10.8 | 11.997 | -1.19726 | 11.085 |
| 3 | 15.9 | 66.6 | 11.8591 | 5 | 17.2 | 18.201 | -1.00196 | 5.822 |
| 4 | 16.8 | 67.8 | 15.3189 | 10 | 19.7 | 18.977 | 0.72281 | 3.669 |
| 5 | 21.7 | 61.0 | 17.2874 | 10 | 25.0 | 25.881 | -0.88145 | 3.525 |
| 6 | 25.6 | 63.4 | 20.1842 | 10 | 32.5 | 30.763 | 1.73602 | 5.341 |
| 7 | 28.5 | 69.1 | 19.1775 | 20 | 34.8 | 32.607 | 2.19258 | 6.300 |
| 8 | 28.7 | 68.6 | 17.0755 | 20 | 35.1 | 32.694 | 2.40554 | 6.853 |
| 9 | 25.6 | 67.7 | 14.5854 | 50 | 31.5 | 28.166 | 3.33358 | 10.582 |
| 10 | 21.0 | 55.5 | 10.7262 | 50 | 26.3 | 23.225 | 3.07470 | 11.690 |
| 11 | 14.2 | 67.9 | 5.3873 | 100 | 19.2 | 15.500 | 3.69915 | 19.266 |
| 12 | 11.1 | 71.6 | 4.3409 | 100 | 13.6 | 12.212 | 1.38738 | 10.201 |

**Table 4**. A comparison of measured, GEP prediction and error values for soil temperature of the Isparta.

| Isparta | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Month (M) | Ambient Temperature (Ta) | Relative Humidity (Rh) | Solar Radiation (Ra) | Depth (h) | Soil Temperature (Ts) Measured Ts | Soil Temperature (Ts) Predicted Ts | Error | Percentage difference |
| | (ºC) | (%) | (kcal/cm$^2$) | cm | (ºC) | (ºC) | | (%) |
| 1 | -0.1 | 64.6 | 6.9660 | 5 | 0.5 | 0.513 | -0.01395 | 2.791 |
| 2 | 1.3 | 65.3 | 8.5020 | 5 | 2.2 | 2.091 | 0.10848 | 4.931 |
| 3 | 8.9 | 64.1 | 10.8260 | 5 | 8.5 | 11.057 | -2.55775 | 30.091 |
| 4 | 12.5 | 57.9 | 13.3420 | 10 | 12.6 | 14.443 | -1.84371 | 14.632 |
| 5 | 15.9 | 49.2 | 18.5780 | 10 | 18.0 | 18.313 | -0.31376 | 1.743 |
| 6 | 22.3 | 39.9 | 20.1710 | 10 | 24.8 | 25.566 | -0.76695 | 3.092 |
| 7 | 25.1 | 35.3 | 20.8090 | 20 | 27.3 | 27.800 | -0.50086 | 1.834 |
| 8 | 25.7 | 38.3 | 18.4830 | 20 | 26.4 | 29.019 | -2.61932 | 9.921 |
| 9 | 19.7 | 53.4 | 13.5410 | 50 | 24.4 | 23.568 | 0.83179 | 3.408 |
| 10 | 12.6 | 67.8 | 10.6250 | 50 | 18.0 | 16.911 | 1.08837 | 6.046 |
| 11 | 8.8 | 72.7 | 7.1770 | 100 | 15.7 | 15.812 | -0.11205 | 0.713 |
| 12 | 3.7 | 69.2 | 5.9420 | 100 | 11.9 | 12.286 | -0.38688 | 3.251 |

**Table 5**. A comparison of measured, GEP prediction and error values for soil temperature of the Burdur.

| Burdur | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Month (M) | Ambient Temperature ($T_a$) | Relative Humidity ($R_h$) | Solar Radiation ($R_a$) | Depth (h) | Soil Temperature ($T_s$) | | Error | Percentage difference |
| | | | | | Measured $T_s$ | Predicted $T_s$ | | |
| | (°C) | (%) | (kcal/cm$^2$) | cm | (°C) | (°C) | | (%) |
| 1 | 1.5 | 74.9 | 4.9931 | 5 | 2.2 | 2.103 | 0.09622 | 4.373 |
| 2 | 3.3 | 66.9 | 7.9450 | 5 | 3.4 | 3.342 | 0.05705 | 1.678 |
| 3 | 8.2 | 52.6 | 12.7179 | 5 | 8.6 | 8.209 | 0.39032 | 4.538 |
| 4 | 11.5 | 57.6 | 14.2481 | 10 | 12.6 | 11.519 | 1.08029 | 8.573 |
| 5 | 16.2 | 58.6 | 18.8009 | 10 | 17.7 | 16.210 | 1.48978 | 8.416 |
| 6 | 21.5 | 53.1 | 19.3264 | 10 | 22.8 | 21.508 | 1.29155 | 5.664 |
| 7 | 25.0 | 41.1 | 20.7330 | 20 | 26.9 | 25.020 | 1.87984 | 6.988 |
| 8 | 23.9 | 49.0 | 18.4212 | 20 | 27.0 | 23.927 | 3.07238 | 11.379 |
| 9 | 20.1 | 50.2 | 15.2421 | 50 | 22.0 | 20.290 | 1.70964 | 7.771 |
| 10 | 15.8 | 57.9 | 10.9973 | 50 | 17.5 | 16.220 | 1.27930 | 7.310 |
| 11 | 8.2 | 65.6 | 6.4689 | 100 | 15.1 | 12.155 | 2.94474 | 19.501 |
| 12 | 4.0 | 73.5 | 5.63607 | 100 | 9.8 | 9.746 | 0.05335 | 0.544 |

## 4. Conclusion

The measurement of soil temperature is very important for various processes. The installation of thermometer correctly in the soil is very complicated. In this study, GEP model was used for estimation of monthly soil temperature using limited meteorological observations for the Antalya, Isparta, and Burdur in Turkey. The results obtained with GEP model were compared with the measured data. The values of MAPE, RMSE and $R^2$ for the soil temperature are 0.45248, 3.07284 and 0.9617 for the Antalya station, and 0.17994, 1.58698 and 0.9763 for the Isparta station, and 0.53234, 2.92649 and 0.9550 for the Burdur station, respectively. Errors obtained are within acceptable limits. The results show that GEP is an influential tool for estimating soil temperature. The new method does not require complex equations. The use of these equations will save the time as well as the finances for predicting soil temperature.

## Acknowledgments

Authors wish to thank the Turkish State Meteorological Service for supplying the meteorological data.

## References

[1] Gao Z., Horton R., Wang L., Liu H., Wen J. 2008. An improved force-restore method for soil temperature prediction. European Journal of Soil Science, 59 (5): 972–981.

[2] Citakoglu H. 2017. Comparison of artificial intelligence techniques for prediction of soil temperatures in Turkey. Theoretical and Applied Climatology, 130 (1-2): 545-556.

[3] Talaee P. H. 2014. Daily soil temperature modeling using neuro-fuzzy approach. Theoretical and Applied Climatology, 118 (3):481- 489.

[4] Behmanesh J., Mehdizadeh S. 2017. Estimation of soil temperature using gene expression programming and artificial neural networks in a semiarid region. Environmental Earth Sciences, 76 (2): 76.

[5] Kermani M. 2013. Hydrometeorological Parameters in Prediction of Soil Temperature by Means of Artificial Neural Network: Case Study in Wyoming. Journal of Hydrologic Engineering, 18 (6):707-718.

[6] Kim S., Singh V.P. 2014. Modeling daily soil temperature using data-driven models and spatial distribution. Theoretical and Applied Climatology, 118 (3): 465-479.

[7] Kisi O., Tombul M., Kermani M.Z. 2015. Modeling soil temperatures at different depths by using three different neural computing techniques. Theoretical and Applied Climatology, 121 (1-2): 377-387.

[8] Mihalakakou G. 2002. On estimating soil surface temperature profiles. Energy and Buildings, 34 (3): 251-259.

[9] Bilgili M. 2010. Prediction of soil temperature using regression and artificial neural network models. Meteorology and Atmospheric Physics, 110 (1-2): 59-70.

[10] Kisi O., Sanikhani H., Cobaner M. 2017. Soil temperature modeling at different depths using neuro-fuzzy, neural network, and genetic programming techniques. Theoretical and Applied Climatology, 129 (3-4): 833-848.

[11] Wu W., Tang X.P., Guo N.J., Yang C., Lui H.B., Shang Y.F. 2013. Spatiotemporal modeling of monthly soil temperature using artificial neural networks. Theoretical and Applied Climatology, 113 (3-4): 481-494.

[12] Ferreira C. 2006. Gene expression programming: mathematical modeling by an artificial intelligence (Vol. 21). Springer.

[13] Ferreira C. 2002. Combinatorial Optimization by Gene Expression Programming: Inversion Revisited. In: Santos J.M., Zapico A. eds., Proceedings of the Argentine Symposium on Artificial Intelligence,160-174, Santa Fe, Argentina.

[14] Ferreira C. 2001. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. Complex Systems, 13 (2): 87-129.

[15] Ferreira C. 2006. Designing Neural Networks Using Gene Expression Programming. In: Abraham A., de Baets B., Köppen M., Nickolay B. eds., Applied Soft Computing Technologies: The Challenge of Complexity, 517-536, Springer-Verlag.