https://doi.org/10.30910/turkjans.1101753



TURKISH JOURNAL of AGRICULTURAL and NATURAL SCIENCES

www.dergipark.gov.tr/turkjans

Research Article

Soil Temperature Prediction with Long Short Term Memory (LSTM)

Orhan İNİK1*, Özkan İNİK2, Taşkın ÖZTAŞ3, Alaaddin YÜKSEL1

^{1*}Bingöl University, Faculty of Agriculturel, Departmant of Soil Science and Plant Nutrition, Bingöl
²Tokat Gaziosmanpasa, Faculty of engineering and architecture, Departmant of Computer Engineering, Tokat
³Atatürk University, Faculty of Agriculturel, Departmant of Soil Science and Plant Nutrition, Erzurum

*Corresponding Author: oinik@bingol.edu.tr

Geliş Tarihi: 12.04.2022 Düzeltme Geliş Tarihi: 26.06.2022 Kabul Tarihi: 07.07.2022

Abstract

TÜRK

TARIM ve DOĞA BİLİMLERİ

DERGISI

Soil temperature not only affects many soil properties, but also has a significant effect on plant development. Knowing and correct estimation of soil temperature is important for both soil management and crop production. The accuracy of temperature forecasts is very important, especially for the countries that stand out with their agriculture-based economies. Therefore, in recent years, different artificial intelligence methods have been used in soil temperature predictions. Deep learning methods lead the way in achieving high prediction accuracy. In this study, a Long Short-Term Memory (LSTM) network, which is a deep learning (DL) sub-architecture, is proposed to create an effective model for soil temperature prediction. The data used in the study are the daily soil temperatures at a depth of 50 cm for the years 2013-2021 of Bingöl province. For the training of the proposed LSTM model, 89% of the data set within the scope of the study was used, and. The remaining 11% was estimated by the model for assessing model success. The RMSE value as a result of the estimation made by the trained LSTM model was obtained as 1.25. The high estimation accuracy of the proposed model showed that this model could be successfully applied in temperature data estimation studies.

Keywords: Soil Temperature, Soil Temperature Prediction, Deep Learning, LSTM

Uzun Kısa Süre Bellek (LSTM) ile Toprak Sıcaklığının Tahmini

Öz

Toprak sıcaklığı, toprağın birçok özelliğini etkilediği gibi bitki gelişimi süreçlerinde de önemli düzeyde etki yapmaktadır. Toprak sıcaklığının bilinmesi ve doğru tahmini hem toprak yönetimi hem de bitkisel üretim için önem arzetmektedir. Özelliklede tarıma dayalı ekonomileriyle öne çıkan ülkeler için sıcaklık tahminlerinin doğrululuğu çok önemlidir. Bu yüzden son yıllarda toprak sıcaklık tahminlerinde farklı yapay zeka yöntemleri kullanılmaya başlanmıştır. Derin öğrenme yöntemleri yüksek tahmin doğruluğu elde etmede bu konuda öncülük etmektedir. Bu çalışmada toprak sıcaklığı tahmininde etkin bir model oluşturmak için derin öğrenme (DL) alt mimarisi olan Uzun Kısa Süreli Bellek (LSTM) ağı önerilmiştir. Çalışmada kullanılan veriler Bingöl İline ait 2013-2021 yıllarına ait 50 cm derinlikteki günlük toprak sıcaklıklarıdır. Çalışma kapsamındaki veri setinin %80'ni önerilen LSTM modelinin eğitimi için kullanılmıştır. Geriye kalan %20'si ise model tarafından tahmin edilerek model başarısı ölçülmüştür. Eğitilen LSTM modelinin yapmış olduğu tahmin sonucundaki RMSE değeri 1.25 olarak elde edilmiştir. Önerilen modelin tahmin doğruluğunun yüksek olması, sıcaklık verileri tahmini çalışmalarında bu modelin başarılı bir şekilde uygulanabileceğini göstermiştir.

Anahtar kelimeler: Toprak sıcaklığı, Sıcaklık tahmini, Derin öğrenme, LSTM.

Introduction

Soil temperature significantly effects on physical, chemical and biological events that occur in the soil. Chemical and biological activities in soil almost cease in frozen soils, however, physical decomposition continues effectively. For example, at temperatures below 5°C, root development of most plants comes to a halt. Availability of soil and air temperature data is very important for understanding plant-soil relationships and for making comments on soil use (Dinc and Senol 1998). Because the temperature change in the soil has a significant effect on the formation of soil microclimate, soil properties and plant growth processes also change (Ekberli and Sarılar 2015). This effect directly influences microorganism activity of soil, germination, plant growth, prosthesis, physical decomposition, soil fertility, soil organic matter, soil moisture. For example, temperature fluctuations in the soil have a very significant effect on carbon (C) and nitrogen (N) mineralization of the soil and thus the vegetation period of the plants. The release of nitrogen or carbon dioxide depends on soil temperature, as well as growth and development of plants, nutrient mineralization of plants, the nutrient diffusion in soil, organic matter ratio in plants, water absorption and transport of water to stem, so the soil temperature has great effects on the functions of the plant root (Aslay and Ustun 2013; Bond-Lamberty et al. 2005; Buckman and Brady 1922; Demiralay 1993; Gao et al. 2007; Guntiñas et al. 2012; Guo et al. 2014; Li et al. 2013; Migała et al. 2014; Schütt et al. 2014; Seyfried et al. 2001; Tenge et al. 1998; Wang et al. 2006).

Soil temperature, an important meteorological parameter, is of great importance in many fields such as pedology, geology, agriculture, ecology and hydrology. Knowing the soil temperature is very important for the countries that have an important place in agriculture in their economies (Aslay and Ustun 2013). Therefore, knowing and estimating the soil temperature has become a necessity.

In recent years, studies have used today's technological possibilities to accurately predict meteorological data. In particular, it benefits from applications of numerous heuristic-driven models such as deep learning (DL), artificial neural networks (ANN), machine learning (ML) and statistical methods.

In a study, a model was developed that predicts monthly average soil temperatures of coming year by using monthly average meteorological data measured between 1970-2011 at 88 provincial stations in Turkey. Different ANN models were tested for soil temperatures at different depths. It has been determined that the prediction results in the ANN models were much better than the prediction results in the regression models, and the prediction results in the artificial neural network models were much closer to the measured real soil temperatures (Aslay and Üstün 2013).

Liu et al. (2021) proposed an adaptive deep learning prediction model for environmental data collection and accurate prediction of soil moisture and temperature. A data acquisition system is designed to collect long-term and equidistant time series data. A deep learning model based on LSTM was applied to predict soil moisture and temperature using the obtained temperature and humidity and soil moisture and temperature data. It has been proven that the proposed Model can obtain and manage the required data effectively and the prediction model can provide reliable prediction of soil moisture and temperature based on ambient temperature and humidity time series data (Liu et al. 2021). After that, Li et al. (2022) developed a new Long Short-Term Modeling (LSTM) model (ILSTM Soil) for soil temperature and moisture estimations, and reported that the ILSTM Soil model outperforms other existing prediction models in most cases (Li et al. 2022).

In another study, researchers developed a regional network of systems for soil moisture prediction based on the ERA5 climate reanalysis dataset, an open-source meteorological dataset published by the Copernicus Climate Change Service. Daily maximum and minimum air temperature, precipitation and vapor pressure values for the years 2011-2020 were used as data and they developed an iterative neural network to predict volumetric soil moisture after three days. A regional scale LSTM network was designed and trained using data from 2011 to 2016 at 28 locations covering four main soil types. The trained LSTM network was tested on 2017–2018 data. The obtained results were compared with a statistical algorithm and a classical machine learning approach. In the results obtained, it was seen that the LSTM network was more successful (Filipović et al. 2022).

Another research group estimated the average air temperature with the CNN model, using the monthly average vapor pressure and relative humidity of Antalya for the years 2000-2016, using the relevant month and year data. The CNN model performance was compared with statistical techniques. As a result, it has been observed that the predicted values in the CNN model were compatible with the real average air temperature values (Akyüz et al. 2020).

Suzen and Kayaalp (2018) made temperature predictions with the LSTM, using pressure, wind speed, humidity and temperature data obtained from the General Directorate of Meteorology for the city center of Isparta (Süzen and Kayaalp 2018). In the literature reviews, it has been understood that deep learning models are increasingly used in the estimation of meteorological data. It is known that deep learning models obtain different results on different data. For this reason, it is necessary to design new deep learning models for the prediction of soil temperature information in different locations.

The purpose of this study, for the first time, a new method has been proposed for the estimation of soil temperature values at 50 cm depth in Bingöl province. In the proposed method, firstly, the data obtained from the meteorological center was pre-processed and the data were made ready for the training of the LSTM model. Then, a unique LSTM network was designed and training and testing processes were carried out.

Material And Method

Study Area

The place where the values in the data set were obtained is the province of Bingöl, as shown in Figure 1. Bingöl has an average elevation of 1745 m. Mountainous areas, valleys, plateaus, plains and basins constitute the main landforms. (Avcı et al. 2018). Bingöl and its surroundings are open to humid-cool air masses from the north and due to the elevation factor, summers are hot and winters are cold. The annual average temperature is 12.1 ^oC. Annual rainfall is 873.7 mm. The number of days with snow is 24.5 days and the number of days with frost is 94.1 days.



Figure 1. World map (a), Map of Türkiye (b), City and central district map (c)

Deep Learning (DL)

Deep learning is a sub-branch of artificial intelligence and is the general name of architectures used in solving many different problems. There are many different architectures in deep learning these are convolutional neural networks, deep autoencoders, Long Short Term Memories (LSTM) and deep graph networks. These architectures achieve different successes in different data types (Inik and Ulker 2022). For example, convolutional neural networks generally provide high success on image data (Inik and Turan 2018). While deep autoencoders are used in segmentation problems, the LSTM architectures have been successfully implemented in time series. In the study, it has been understood that the best architectures that can be applied for temperature data are LSTM networks. For this reason, LSTM networks are used in this study.

Long-Short-Term Memory (LSTM)

Deep learning has a wide range of uses in artificial intelligence applications. For this reason, new deep learning architectures have been developed for different problems. One of them is the LSTM that is applied successfully in time series (Hochreiter and Schmidhuber 1997).

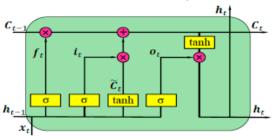


Figure 2. The LSTM architecture (Xiao and Yin 2019)

The LSTM architecture consists of sequential blocks that repeat each other as shown in Figure 2. The LSTM structure usually consists of 3 different layers: input, output, and forget. In this architecture, X_t and h_{t-1} information are used as inputs, and it is decided which information to delete. These operations are done using Equation 1 in the forget layer (f_t) . During this process, the sigmoid is used as the activation function.

$$f_t = \text{sigmoid} (W_{f,x} X_t + W_{f,h} h_{t-1} + b_f)$$
 1

In the second step, the input layer, where new information will be determined, comes into play. Here, the information is first updated with the sigmoid function (Equation 2).

$$i_t = \text{sigmoid} (W_{i,x} X_t + W_{i,h} h_{t-1} + b_i)$$
 2

In the next Equation, the candidate information that will form the new information is determined (Equation 3).

$$C_t = tanh (W_{c,x} * X_t + W_{c,h} * h_t - 1 + b_c)$$
 3

In Equation 4, *Ct* is obtained which denotes new information.

$$C_t = C_t - 1^* f_t + i_t * C_t$$

The result data is obtained by using the next equations from the last output layer (Equations 5 and 6)

$$O_t = \sigma \left(W_{o,x} * X_t + W_{o,h} * h_t - 1 + b_o \right)$$
 5

$$h_t = o_t * tanh\left(C_t\right) \tag{6}$$

By repeating the process outlined above, the Model weight (W) and bias (b) parameters are learned in a way that minimizes the difference between the LSTM output values and the actual training values.(Fischer and Krauss 2018; Kara 2019; Sagheer and Kotb 2019).

Proposed Method

The flow chart of the proposed method is given in Figure 3. Pre-processing was applied to make the received data suitable for the LSTM. The obtained data set is divided into two parts as training and testing. More than one LSTM model was designed and trained in experimental studies. Among these models, the model with the lowest error value was taken and applied on the test data. Thus, the model with the best performance in the proposed method was obtained.

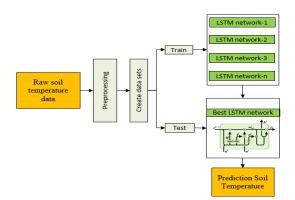


Figure 3. Flow chart of the proposed method

Data Preprocessing

The raw data of 3288-day soil temperature at 50 cm depth between 2013-2021 from the 17203-station affiliate to the Bingöl Provincial Meteorology Directorate are given in Figure 4. Analyzed data indicated that, the temperature in winter (December, January and February) is usually close to zero. In the studies carried out, it has been observed that the average temperature increases as you go deeper from the soil surface in winter, and the average temperature decreases as you go deeper from the surface in summer (Ekberli et al. 2017).

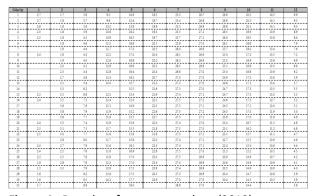


Figure 4. Raw data from meteorology (2013)

The data in Figure 4 has been converted into a time series suitable for training deep learning models. Since some days from the station were not recorded, it was recorded as blank. These blanks was filled with the data of the previous day. The edited data are shown in Figure 5.

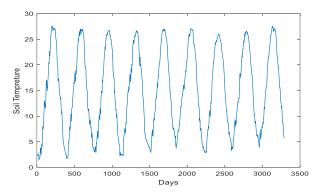
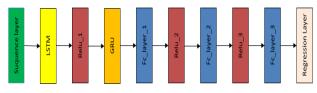


Figure 5. Graph of soil temperature data obtained as a result of data preprocessing

The LSTM Model Design

In the study, more than one LSTM network was designed for soil temperature prediction. The architecture of the model with the lowest regression error among these models is given in Figure 6. The properties of each layer of the model are given in Figure 7.



Figuru 6. Architecture of the Proposed LSTM Model

	Name	Туре	Activations	Learnables	Total Learnab	States
	sequenceinput Sequence input with 1 dimensions	Sequence input	1		0	
2	Istm LSTM with 200 hidden units	LSTM	288	InputWeights 800×1 RecurrentWe… 800×… Bias 800×1	161600	HiddenState 200×1 CellState 200×1
3	relu_1 ReLU	ReLU	200		0	
4	gru GRU with 200 hidden units	GRU	288	InputWeights 600× RecurrentWe 600× Bias 600×1	248588	HiddenState 200×1
5	fc_1 1000 fully connected layer	Fully Connected	1000	Weights 1000×200 Bias 1000×1	201000	
6	relu_2 ReLU	ReLU	1000		0	
7	fc_2 1000 fully connected layer	Fully Connected	1000	Weights 1000×1000 Bias 1000×1	1001000	
8	relu_3 ReLU	ReLU	1000		0	
9	fc_3 1 fully connected layer	Fully Connected	1	Weights 1×1000 Bias 1×1	1001	
10	regressionoutput mean-squared-error with response 'Response'	Regression Output	1		0	

Figure 7. Features in each layer of the proposed LSTM.

Performance Calculation Criteria

There are error criterion measures, which are special formulas that provide evidence about the working type and reliability of an algorithm or model. These criteria, which are widely used in the literature; Root Mean Square Error (RMSE), Mean Square Error (MSE), R Squared Score (R2) and Mean Absolute Error (MAE) (Sevinc and Kaya, 2021). In this study, the RMSE method given in Equation 7 was used to measure the performance of the proposed model. The RMSE is the root mean square error of the predicted values and is one of important performance the most criteria (Demirezen 2020)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}}$$
 7

Where, x_i is the predicted value, y_i is the actual value, and n is the total number of data.

Results And Discussion

The training processes of the proposed LSTM model were performed using the Deep Learning Library of the MATLAB R2021b software. The model was trained with an epoch number of 500 and a initial learning rate of 0.001. Eighty-nine percent of the data set was used for training and remaining 11% was used for testing. The RMSE and error convergence graph obtained by the model during the training phase are given in Figure 8.

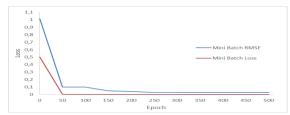
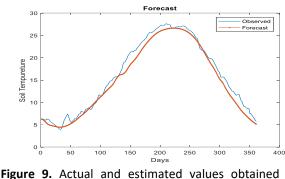


Figure 8. The RMSE and loss convergence graph obtained by the model during training

The regression graph obtained by the model is given in Figure 9. It is clearly seen that the estimated graph almost exactly coincides with the real graph.



with the proposed LSTM model

The error values obtained by the model at each point are given in Figure 10. The RMSE values of the errors in Figure 10 were obtained as 1.25.

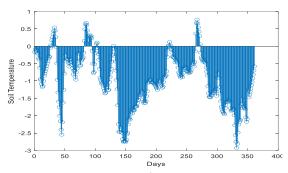


Figure 10. RMSE value of soil temperature data estimated by LSTM

In Figure 11, soil temperature data for a total of 3288 days between 2013-2021 are given. It is seen that in the summer months the soil temperature increases, while in the winter months the temperature decreases and periodically continues in this way. The soil temperature of the last 361 days, shown in red, was estimated by the proposed model. Looking at the graph, it is understood that the model has achieved a successful graph.

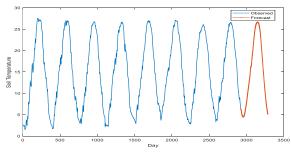


Figure 11. Estimated and actual values obtained with the proposed LSTM models

The LSTM model was used for the first time on the data set obtained within the scope of this study. However, it has been seen in the literature that different models are run on different data sets. Considering the LSTM accuracy rate of temperature prediction from Sevinç and Kaya deep learning methods, the RMSE values are seen as 1.859 (Sevinç and Kaya 2021). In another study, Kara found the RMSE value to be 19.39% (Kara 2019). Kreuzer et al. found the RSME value of the LSTM network accuracy ratio they proposed in the data obtained from different meteorological stations in Germany as 2.83 (Kreuzer et al. 2020).

Although the studies were conducted on the data in different regions, the performance of the proposed method was compared with these studies because the problem sought to be solved was the same and the method used was similar. When the proposed method was compared with similar studies, it was observed that the RMSE value obtained was much lower. It has been seen that the proposed LSTM architecture predicted soil temperature estimation with better performance. The most important reason why the proposed model exhibits better perfonmas than other studies is that a hybrid deep learning architecture is designed. In particular, the use of the LSTM model in the architecture as a hybrid with GRU has increased performance.

Conclusions And Recommendations

Soil temperature, one of the parameters of meteorology, is related to pedology, geology, agriculture, ecology and hydrology and many more. It is especially important in the field of soil science. Because quickly learning and estimating the soil temperature contributes to the healthy operation of both soil management and plant production plans. This situation is of great importance for countries whose economies consist mostly of agriculture. Today, the rapid development of technology provides the advantage of solving problems in an accurate and easy way. The advantage of technology is also felt in the agricultural field. In particular, deep learning (DL) architectures used in forecasting climate data provide great convenience. From this point of view, in this study, the LSTM architecture, which is the sub-architecture of DL, was used to predict soil temperature.

The data used in the study are the daily soil temperatures at a depth of 50 cm for the years 2013-2021 of Bingöl province. For the training of the proposed LSTM model, 89% of the data set within the scope of the study was used. The remaining 11% was estimated by the model and model success was measured. The RMSE value as a result of the estimation made by the trained LSTM model was obtained as 1.25. The high estimation accuracy of the proposed model has shown that this model can be successfully applied in temperature data estimation studies.

Acknowledge

The authors would like to thank the Bingöl Provincial Meteorology Directorate for the meteorological information provided for this study in order to support scientific studies.

References

- Akyüz AÖ, Kumaş K, Ayan M, Güngör A (2020) Antalya İli Meteorolojik Verileri Yardımıyla Hava Sıcaklığının Yapay Sinir Ağları Metodu ile Tahmini. Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi 10: 146-154.
- Aslay F, Üstün Ö (2013) Meteorolojik Parametreler Kullanılarak Yapay Sinir Ağları ile Toprak Sıcaklığının Tahmini. Politeknik Dergisi 16: 139-145.
- Avcı V, Esen F, Kıranşan K (2018) Bingöl İlinin Fiziki Çoğrafya Özellikleri. The Journal of Bingöl Studies 4.
- Bond-Lamberty B, Wang C, Gower ST (2005) Spatiotemporal measurement and modeling of stand-level boreal forest soil temperatures. Agricultural and Forest Meteorology 131: 27-40.
- Buckman HO, Brady NC (1922) The nature and properties of soils. Macmillan.
- Demiralay İ (1993) Toprak fiziksel analizleri. Atatürk Üniversitesi Ziraat Fakültesi Yayınları 143: 13-19.
- Demirezen S (2020) Türkiye'de Gün Öncesi Piyasası İçin Elektrik Fiyatlarının Tahmini.
- Dinç U, Şenol S (1998) Toprak etüd ve haritalama ders kitabı. Çukurova Üniversitesi Ziraat Fakültesi Genel Yayın.
- Ekberli İ, Gülser C, Özdemir N (2017) Farklı toprak derinliklerindeki sıcaklığın tahmininde parabolik fonksiyonun kullanımı. Toprak Bilimi ve Bitki Besleme Dergisi 5: 34-38.
- Ekberli İ, Sarılar Y (2015) Toprak sıcaklığı ve ısısal yayınımın belirlenmesi. Anadolu Tarım Bilimleri Dergisi 30: 74-85.
- Filipović N, Brdar S, Mimić G, Marko O, Crnojević V (2022) Regional soil moisture prediction system based on Long Short-Term Memory network. Biosystems Engineering 213: 30-38.
- Fischer T, Krauss C (2018) Deep learning with long short-term memory networks for financial

market predictions. European Journal of Operational Research 270: 654-669.

- Gao Z, Bian L, Hu Y, Wang L, Fan J (2007) Determination of soil temperature in an arid region. Journal of arid environments 71: 157-168.
- Guntiñas ME, Leirós M, Trasar-Cepeda C, Gil-Sotres F (2012) Effects of moisture and temperature on net soil nitrogen mineralization: A laboratory study. European Journal of Soil Biology 48: 73-80.
- Guo J, Yang Y, Chen G, Xie J, Yang Z (2014) Carbon mineralization of Chinese fir (Cunninghamia lanceolata) soils under different temperature and humidity conditions. Acta Ecologica Sinica 34: 66-71.
- Hochreiter S, Schmidhuber J (1997) Long shortterm memory. Neural computation 9: 1735-1780.
- Inik Ö, Turan B (2018) Classification of animals with different deep learning models. Journal of New Results in Science 7: 9-16.
- Inik Ö, Ülker E (2022) Optimization of deep learning based segmentation method. Soft Computing: 1-16.
- Kara A (2019) Global Solar Irradiance Time Series Prediction Using Long Short-Term Memory Network. Gazi Üniversitesi Fen Bilimleri Dergisi, Part C: Tasarım ve Teknoloji, vol 4: 7.
- Kreuzer D, Munz M, Schlüter S (2020) Short-term temperature forecasts using a convolutional neural network—An application to different weather stations in Germany. Machine Learning with Applications 2: 100007.
- Li L-J, You M-Y, Shi H-A, Ding X-L, Qiao Y-F, Han X-Z (2013) Soil CO2 emissions from a cultivated Mollisol: Effects of organic amendments, soil temperature, and moisture. European Journal of Soil Biology 55: 83-90.
- Li Q, Zhu Y, Shangguan W, Wang X, Li L, Yu F (2022) An attention-aware LSTM model for soil moisture and soil temperature prediction. Geoderma 409: 115651.
- Liu H, Yang Y, Wan X, Cui J, Zhang F, Cai T (2021) Prediction of soil moisture and temperature based on deep learning. 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA). IEEE.

- Migała K, Wojtuń B, Szymański W, Muskała P (2014) Soil moisture and temperature variation under different types of tundra vegetation during the growing season: A case study from the Fuglebekken catchment, SW Spitsbergen. Catena 116: 10-18.
- Sagheer A, Kotb M (2019) Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing 323: 203-213.
- Schütt M, Borken W, Spott O, Stange CF, Matzner E (2014) Temperature sensitivity of C and N mineralization in temperate forest soils at low temperatures. Soil Biology and Biochemistry 69: 320-327.
- Sevinç A, Kaya B (2021) Derin Öğrenme Yöntemleri ile Sıcaklık Tahmini: Diyarbakır İli Örneği. Computer Science: 217-225.
- Seyfried M, Flerchinger G, Murdock M, Hanson C, Van Vactor S (2001) Long-Term Soil Temperature Database, Reynolds Creek Experimental Watershed, Idaho, United States. Water Resources Research 37: 2843-2846.
- Süzen A, Kayaalp K (2018) Derin Öğrenme Yöntemleri İle Sıcaklık Tahmini: Isparta İli Örneği. International Academic Research Congress INES.
- Tenge AJ, Kaihura F, Lal R, Singh B (1998) Diurnal soil temperature fluctuations for different erosion classes of an oxisol at Mlingano, Tanzania. Soil and Tillage Research 49: 211-217.
- Wang C, Wan S, Xing X, Zhang L, Han X (2006) Temperature and soil moisture interactively affected soil net N mineralization in temperate grassland in Northern China. Soil Biology and Biochemistry 38: 1101-1110.
- Xiao Y, Yin Y (2019) Hybrid LSTM neural network for short-term traffic flow prediction. Information 10: 105.