JUNE 2021

M.Sc. in Civil Engineering

EBRU TARIM

HASAN KALYONCU UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

FORECASTING STREAMFLOW IN ASİ RIVER USING LONG-SHORT TERM MEMORY NETWORKS

M.Sc. THESIS

IN

CIVIL ENGINEERING

BY
EBRU TARIM

JUNE 2021

Forecasting Streamflow in Asi River Using Long-Short Term Memory Networks

M.Sc. Thesis

in

Civil Engineering

Hasan Kalyoncu University

Supervisor

Asst. Prof. Dr. Adem YURTSEVER

Co-Supervisor

Asst. Prof. Dr. Hüseyin Çağan KILINÇ

Ebru TARIM

June 2021

Civil Engineering Department, Civil Engineering M.Sc. (Master of Science) programme student **Ebru TARIM** prepared and submitted the thesis titled **Forecasting Streamflow In Asi River Using Long-Short Term Memory Networks** defened successfully on the date of 02/06/2021 and accepted by the jury as a M.Sc. thesis.

Position	**Title, Name and Surename** **Department/University**	**Signature**
Supervisor	Asst. Prof. Dr. Adem YURTSEVER Civil Engineering Department/ Hasan Kalyoncu University	
Jury Member	Prof. Dr. Mehmet KARPUZCU Civil Engineering Department/ Hasan Kalyoncu University	
Jury Member	Asst. Prof. Dr. Yunus ÖZTÜRK Civil Engineering Department/ Kilis 7 Aralık University	

This thesis is accepted by the jury members selected by the institute management board and approved by the institute management board.

Prof. Dr. İbrahim Halil GÜZELBEY
Director

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Ebru TARIM

ABSTRACT

FORECASTING STREAMFLOW IN ASİ RIVER USING LONG-SHORT TERM MEMORY NETWORKS

Tarim, Ebru
M.Sc. in Civil Engineering
Supervisor: Asst. Prof. Dr. Adem YURTSEVER
Co-Supervisor: Asst. Prof. Dr. Hüseyin Çağan KILINÇ
June 2021, 58 pages

River flow estimation is necessary in order to control water resources and to ensure sustainable use. Long-term flow forecasting is of great importance for water resources planning and management in both the short and long term. Traditional instantaneous flow measurements in river basins take a lot of time and require high costs. Instead of instantaneous measurements, artificial intelligence programs that can predict flow with today's technology reveal very important developments. Long-short term memory (LSTM) neural networks from deep learning neural networks, which support developments and are widespread in use, are adopted to predict the flow of river flows on daily and monthly scales. In this study, the LSTM model was created with daily flow data from Afrin flow measurement station (FMS) no D19A006, which is within the boundaries of the Asi river basin, and the predictive performance of the model was analyzed. The flow values used in the model between 2006-2018 were used as test data set. For the study, was tested with MSE, Huber and LogCosh loss functions using Adam and Adamax optimizer. When determining the best forecast model, statistical evaluation criterion were determined by comparing the observed data and forecast models. Considering the evaluation criterion, the LogCosh loss function of the Adamax optimizer was found to have better performance and R^2 0,9949 was achieved. Also, according to the model analysis, LSTM analysis of daily flow forecasting with different scenarios has shown that the performance impact can be expanded and help ensure the sustainability of water resources.

Keywords: LSTM, Deep Neural Network, River Streamflow, Flow Measurement Station

ÖZET

UZUN-KISA SÜRELİ BELLEK AĞLARI KULLANARAK ASİ NEHRİ'NDE AKIM TAHMİNİ

Tarim, Ebru
Yüksek Lisans Tezi, İnşaat Mühendisliği Anabilim Dalı
Tez Danışmanı: Dr. Öğr. Üyesi Adem YURTSEVER
Eş Danışman: Dr. Öğr. Üyesi Hüseyin Çağan KILINÇ
Haziran 2021, 58 sayfa

Su kaynaklarının kontrol altına alınabilmesi ve sürdürülebilir kullanımının sağlanabilmesi için nehir akış tahmini gereklidir. Uzun süreli akış tahmini, hem kısa hem de uzun vadede su kaynakları planlaması ve yönetimi için büyük önem taşımaktadır. Nehir havzalarında geleneksel olarak yapılan anlık akım ölçümleri çok zaman almakta ve yüksek maliyet gerektirmektedir. Anlık yapılan ölçümler yerine, günümüz teknolojisiyle akış tahmini yapabilen yapay zekâ programları oldukça önemli gelişmeler ortaya koymaktadır. Gelişmeleri destekleyen ve kullanımı yaygınlaşan derin öğrenme sinir ağlarından uzun kısa süreli bellek (LSTM) sinir ağı performansı, günlük ve aylık ölçeklerde nehir akımlarının akış tahmini için değerlendirilmektedir. Bu çalışmada, Asi nehir havzası sınırları içinde kalan D19A006 No’lu Afrin akım gözlem istasyonu’ndan (AGİ) alınan günlük akım verileri ile LSTM modeli oluşturularak, modelin tahmin performansı analiz edilmiştir. Modelde kullanılan 2006-2018 yılları arasını kapsayan akım değerleri, test veri seti olarak kullanılmıştır. Oluşturulan model, Adam ve Adamax iyileştiricileri kullanılarak MSE, Huber ve LogCosh kayıp fonksiyonları ile test edilmiştir. En iyi tahmin modeli belirlenirken, gözlemlenen veriler ve tahmin modelleri karşılaştırılarak istatistiksel değerlendirme kriterleri belirlenmiştir. Değerlendirme kriterleri göz önüne alındığında Adamax iyileştiricisinin LogCosh kayıp fonksiyonunun daha iyi performansa sahip olduğu ve R^2 0.9949 elde edildiği görülmüştür. Ayrıca, model analizine göre LSTM'nin günlük akış tahmininde farklı senaryolarla analizinde performans etkisinin genişletilebileceği ve su kaynaklarının sürdürülebilirliğini sağlamaya yardımcı olabileceği görülmüştür.

Anahtar Kelimeler: LSTM, Derin Sinir Ağı, Nehir Akımı, Akım Gözlem İstasyonu

ACKNOWLEDGEMENTS

I express my gratitude to my supervisor Asst. Prof. Dr. Adem YURTSEVER and Asst. Prof. Dr. Hüseyin Çağan KILINÇ for their endless support and guidance throughout the research. I would also like to thank the jurors and my family for supporting me throughout the work.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

SYMBOLS AND ABBREVIATIONS

ANFIS	: Adaptive Network Based Fuzzy Inference System
ANN	: Artificial Neural Networks
AS-LSTM	: Asset Sensitive Long Short Term Memory
CLSTM	: Convolutional Long Short Term Memory
CNN	: Convolutional Neural Networks
DL	: Deep Learning
FC-LSTM	: Fully Connected Long Short Term Memory
FCN	: Fully Connected Network
FFNN	: Feed-Forward Neural Networks
FFNN-ED	: Feed Forward Neural Network Based Encoder Decoder
FMS	: Flow Measurement Station
GCN	: Graph Convolutional Networks
GPR	: Gaussian Process Regression
GRU	: Gated Recurrent Unit
HA	: Historical Average
LASSO	: Least Absolute Shrinkage and Selection Operator
LSS-LSTM	: Local Spatial Sequential Long Short Term Memory
LSTM	: Long Short Term Memory
LSTM-ED	: Long Short Term Memory Encoder Decoder
MAE	: Mean Absolute Error
MAPE	: Mean Absolute Percentage Error
MI	: Mutual Information
ML	: Machine Learning

MLS	: Multi-Layer Perceptron
MODWT	: Maximal Overlap Discrete Wavelet Transform
MSE	: Mean Square Error
NARX	: Nonlinear Autoregressive External
NSE	: Nash-Sutcliffe Efficiency
RMSE	: Root Mean Square Error
RNN	: Recurrent Neural Network
SAC-SMA	: Sacramento Soil Moisture Calculation
SAE	: Sparse Autocoder
SGD	: Stochastic Gradient Descent
STA-LSTM	: Spatial-Temporal Attention Long Short Term Memory
STD	: Standart Deviation
SVM	: Support Vector Machine
SVR	: Support Vector Regression
SWAT	: Soil and Water Assessment Tool
WLSTM	: Wavelet Long Short Term Memory

CHAPTER 1

INTRODUCTION

1.1 General

Water is the main source and basis of both life and civilization. For this reason, the first people settled near freshwater sources, and the first civilizations developed in these areas. Over time, the rapid increase of the world population leads to a decrease in the amount of water available due to urbanization and industrial development, drying wetlands, expansion of agricultural land, intensively irrigated agriculture and water pollution. The water use and industrial requirements of the growing world population require effective and sustainable use of water resources (Tanrıverdi et al., 2011). Although it is a renewable resource, it is one of the main living materials with limited use. With the controlled and planned use of water, the operation, management and sustainability of water resources can be done effectively (Ağca et al., 2014).

In recent years, the use and sharing of water in transboundary waters and international waters facing water shortages sometimes brings regional and global crises. Transboundary streams are sometimes known as ' transboundary-forming streams' by forming borders between the two countries (Korkmaz and Karataş, 2009). On the other hand, in order for a stream that borders or draws borders to be defined as international water, first of all, a management agreement is required, which is accepted by all countries in the field of water drainage. Today, flow estimates about streams that are born from our country and pass to neighboring countries and enter our country from neighboring countries are of great importance for planning and management. In line with the forward planning of the transboundary streams, it prevents the formation of problems between countries depending on the water. In such streams, only coastal countries have the initiative and there is no use or intervention of third countries (Pazarcı, 1996).

In order to sustain vital activities, sufficient amounts and quality of water needed by living things must be obtained from the hydrological cycle. The most effective situation to support this hydrological cycle is the taking of sustainable working steps that will ensure that resources reach future generations (Yüksek, 2004). In order to ensure the intended sustainability, planning and management studies are required from the source of the water to the consumption stage. Otherwise, increasingly, drought, excessive consumption, lack of planning and management threaten the future sustainability of water. In the face of these threats, the continuity of life is also endangered. The basis for improving these problems is based on the planning and management of water by making forward-looking water estimation of river flows (Dedekayoğulları and Önal, 2009).

In order for the water structures to be designed correctly, the flow forecast must be done as accurately as possible. Different models are also used for these estimates. Since the factors affecting water in nature are uncertain and there is a nonlinear relationship between the parameters that will affect them, there are some difficulties during modeling and control. Traditional methods have been used for flow modeling in past years. Traditional methods cannot accurately model linear and non-static data like Artificial Intelligence (AI) methods. Therefore, AI methods have been used in place of traditional methods recently. AI methods are used quite widely because of their high accuracy compared to traditional methods (Nazimi, 2021). Since AI methods perform highly in stagnant and unstable data forecasting, there has been a lot of interest in determining water flow in the field of hydrology in recent years. The common purpose of AI methods is to use monthly flow data from past years taken from the Flow Measurement Station (FMS) as inputs and make predictions for the future. Widely used AI methods Artificial Neural Networks (ANN), Feed-Forward Neural Networks (FFNN), Adaptive Network Based Fuzzy Inference System (ANFIS), Machine Learning (ML), Deep Learning (DL) and Long Short Term Memory Networks (LSTM) (Altunparmak and Başakın, 2018).

The structure of ANNs consists of simple processor elements, and the elements created by ANN are connected between themselves with weights and connections that carry numerical data that can be expressed in different formats. The fact that the human brain that causes the progress of ANNs is mechanized has the features that artificial systems can make calculations that are difficult to understand (Yaşar, 2004).

The DL model emerges as a result of AI, technological advances in the hardware of computers and advances in ML techniques. Many real-world problems can be solved, especially with the emergence of DL algorithms that can be performed in parallel with multi-core computing devices such as GPUs. DL models show great success in applications such as classification, regression analysis and estimation in time series with the data they are trained in (Şişmanoğlu et al., 2020). DL enables the transfer of information between layers by enabling learning from large data sets with its layered architecture. Processor speeds and increased storage, gains in cheap computer hardware and ML are factors that drive DL development. Gradient descent-based optimization algorithms (Stochastic gradient descent (sgd), momentum, adam, adamax, adagrad, rmsprop and adadelta), which are used to minimize error in DL , are also part of this success (Seyyarer et al., 2020).

Recently, the use of the deep neural network, the LSTM network, for prediction in many areas of DL technique is emerging as a new situation. LSTM aims to maintain the error value from different times and layers in the backprop. It provides a more constant error value, allowing recurrent networks to continue learning steps. It does this by opening a new channel between cause and effect. The difference of the repeti repeated module in the LSTM structure is that it has 4 layers connected in a special way, rather than a single neural network layer. The cell can decide what to store, when to read, write or delete it. As with neurons, it stops the information coming by passing it on to its weight. These weights are calculated during the learning of the repetitive network and LSTM has different models according to different needs. Multiple optimization techniques, including data enhancement and batch normalization, are applied to further improve the performance of the method (Fang et al., 2020).

The aim of this innovative study is to reliably predict river flow in order to design correct planning and management in water resources by creating the LSTM model. The effects of the time gap between test set on the performance of the model were investigated. For the study, the estimation of river flows was examined by using the daily flow data from Afrin FMS located in the Asi river basin know as the transboundary basin. The estimation results are improved by using optimizers during the evaluation phase of flow estimation.

CHAPTER 2

LITERATURE REVIEW

2.1 General

Great efforts are being made to increase the efficiency of water resources in the world. In these cases, estimating water demand is a prerequisite for the successful implementation of water resources planning and management tools. However, forecasting approaches based on traditional neural networks often have difficulties in capturing the temporal dependencies that characterize water flow time series (Mouatadid et al., 2019). Daily river streamflow values are of great importance in forecasting studies, design and operation of water structures. Especially the development of a method that can make forward flow estimation can provide a suitable management opportunity for the producer, the user and the people of the region in the control of the water structures in the enterprise, hydroelectric production, environmental protection and flood control (Nacar et al., 2018).

Prediction of river flow as a fundamental source of hydrological information plays a crucial role in various fields of water projects. The flow in rivers is determined by measuring stations established by relevant organizations at certain points of the rivers. In the methods used in recent years, there are cases such as failure to receive data for various reasons and station malfunctions (Fathian et al., 2019). The need for accurate flow estimates in rivers encourages researchers to develop new methods and improve the predictive ability of traditional models. In such cases, AI techniques have been developed to complement the missing values and used in river flow estimation using various methods that have a wide range of uses in water resources engineering. AI methods developed in this context enable the completion of expected results in a shorter time and more efficiently, as well as long measurements and man-made studies (Rajaee et al., 2020).

Many ML models have been proposed for streaming forecasting in recent years. In order to define a model or MLsystem with ML techniques, the feature vector must first be removed. ML method, often called a time step, is used to reconstruct the time series of flow values into a standard ML input set and preprocess the data as a variable parameter (Dao et al., 2020). A new learning model is developed by conducting researches in studies where ML is not proficiency. Developed DL is making great progress. With the advent of DL algorithms, many problems can be solved by improving the problem of using very large values (Acet and Akkaya, 2019). DL makes great progress by eliminating this problem that ML field workers have been dealing with for many years. Because deep networks do the learning process on raw measurement values, unlike traditional machine learning and image processing techniques. While processing the raw measurement values, it obtains the necessary information with the representations it has created in different layers. Although the foundations of DL are based on the past, the most important reasons for its popularity in recent years are determined as having enough data for education and secondly, having a hardware infrastructure to process this data (İnik and Ülker, 2017). DL methods can automatically detect and analyze complex patterns in data and interactions between values in order to speed up the process and ensure accuracy in process steps (Sakarya and Yilmaz, 2019).

ANN in the field of ML are often used to solve many problems. It is known that it is convenient to use ANN models from AI methods very easily on river flows. Artificial intelligence methods are methods that are generally developed with programming languages and generate data through the learning algorithms of computer systems. With its developing technology, it increases its usage areas day by day and increases its use in daily life (Yıldıran, 2019). DL is artificial learning with multilayered models, increasing and improving the number of hidden layers in ANN. RNN is a kind of ANN model that provides the use of historical information by creating a loop. In this way, the information of the previous period can be used and it can be obtained and classifyed using new and old information. LSTM, a sub-branch of RNN, has been developed in order to eliminate the existing problems here (LeCun et al., 2015). LSTM model has been applied in many projects and studies for river flow estimations recently. As the studies are carried out, many innovations about LSTM emerge.

2.2 Overview Studies

Hu et al. (2018) examined the ANN and LSTM network models to simulate the precipitation-flow process in the Fen River basin. The input values to be used in the model are taken from 98 precipitation flow stations. Used as a training kit as a test set between 86 sedimentation flow stations and the rest. Considering the different delivery time models, it was seen that the LSTM model is more stable than the ANN model, which offers better simulation performance.

Nguyen and Bae (2020), proposes a framework that combines an advanced 1D/2D urban hydrological model with an estimating system to predict water levels and flooding events in an urban basin. When the LSTM model is trained and tested in the Gangnam urban basin in Seoul, South Korea, and a database is created, it shows the feasibleness of the proposed framework.

Yan et al. (2019) predicts flow using a time series model based on LSTM that takes into account historical streamflow measurement values, historical weather and weather forecast values of hydrological stations. In flow estimation, LSTM and Support Vector Machine (SVM) prediction models are compared. It appears that the LSTM estimation model has better prediction accuracy and the model gives satisfactory results in estimating river flows.

Lv et al. (2020) proposes a long-term cyclic hydrology prediction model with the help of improved LSTM, based on mutual information (MI) analysis of practical hydrological values of the Xixian Basin in 2011-2018 in China. Compared to some traditional and ML flood prediction schemes, it is seen that the LSTM model can accurately complete the long and short lead time hydrology prediction task.

Ni et al. (2020) developed two hybrid models based on LSTM for monthly flow and rainfall forecast in the basin. Wavelet-LSTM (WLSTM) implements three wavelet transformation algorithms for serial separation and implements a combined convolutional neural network to remove convolutional LSTM (CLSTM) timer properties. The prediction accuracy of WLSTM and CLSTM is compared to the multilayer sensor (MS) and LSTM. LSTM is found to be applicable to time series estimation.

Le et al. (2019) tried to demonstrate the capabilities of LSTM in flow predictions using 1, 2 and 3 days of flow data at Hoa Binh Station of Vietnam's Da River basin. Regarding the prediction accuracy in the model, it is seen that the prediction ability of the LSTM model is quite impressive.

Kratzert et al. (2018), proposed a novel data-driven approach, using the LSTM network, a special type of recurrent neural network. CCAMELS input values of 241 basins available free of charge are used to test the approach and also compare the results with the well-known Sacramento Soil Moisture Calculation Model (SAC-SMA) combined with the Snow-17 snow routine. SAC-SMA is shown to achieve better model performance.

Lee et al. (2018) developed an LSTM for daily flow simulations based on water level input values of 10 stations on the Mekong River, showing that the LSTM outperforms the Soil and Water Assessment Tool model (SWAT).

Li et al. (2020) revealed a data-oriented RR model using a sequential sequence LSTM network. The model is used to predict river discharge by designing various numerical tests using 10 years of rainfall from 153 rain gauges and river canal discharge data in a watershed in Houston, TX, known for severe flood events. The model results showed that the LSTM model was able to predict discharge efficiently and achieved good model performance.

Damavandi et al. (2019) presented a DL model in the Brazos basin, Texas, using the LSTM network for accurate prediction of flow flow. Previous rainfall observations and climate variability indices were used as potential predictors. Model performance was evaluated through the Nash-Sutcliffe model efficiency (NSE) coefficients between the Pearson correlation (ρ) and the predicted flow flow, giving achieved prediction results.

Hu et al. (2019) developed a new working model for flow prediction using the LSTM DL model that takes the time interval into account. Tunxi used runoff values from a hydrology station in China and rainfall from 11 nearby precipitation stations to estimate flow values from this hydrological station 6 hours in the future. The LSTM model outperforms Support Vector Regression (SVR) and multi-layer perceptron (MLP) models.

Fu et al. (2020) applied a DL method for daily flow simulation of the Kelantan River located in the northeast tropical region of the Malaysian peninsula. Using flow data from previous years for flow estimation, the model has been run from several perspectives. The LSTM model appears to be advantageous in processing constant flow data during the dry season and gives good results in capturing rapidly changing flow input characteristics during rainy seasons.

Xiang et al. (2020) presented an implementation of a prediction model based on LSTM and seq2seq structure to estimate hourly rainfall. Focusing on two Midwestern watersheds, namely Clear Creek and Upper Wapsipinicon River in Iowa, these models were used to predict hourly runoff for a 24 hour period using rainfall observation, rainfall forecast, runoff observation, and empirical monthly evapotranspiration data from all stations in these watersheds. It shows that the LSTM-seq2seq model outperforms linear regression, Lasso regression, Ridge regression, support vector regression, Gaussian process regression and LSTM at all stations in these two basins.

Shi et al. (2015) examines the convolutional LSTM (ConvLSTM) formed by expanding the fully connected LSTM (FC-LSTM) to predict future precipitation intensity in a local area. The ConvLSTM network performs better at capturing spatial-temporal correlations.

In addition to streamflow estimations in river basins, studies are carried out with LSTM model in many areas. These studies also provide good developments in their field. It cannot be denied that the LSTM model, which has become popular, will be seen more frequently in the future due to the curiosity in many areas of work.

Thapa et al. (2020) proposed to apply and compare LSTM, nonlinear autoregressive external (NARX), gaussian process regression (GPR) and support vector arrangement (SVR) models for melt snowwater flow prediction in Nepal's Himalayan region. Gamma testing was also performed to determine more suitable input combinations of models. It has been observed that the shallow LSTM method, which has one hidden layer, gives higher results than the LSTM method, which has a lot of hidden layers. Adamax, one of the seven optimizers used for LSTM, gives the most accurate result for melted snow water flow prediction.

Kratzert et al. (2019) examines the regional level Asset Sensitive LSTM (AS-LSTM) with 531 basins outperforming several hydrological models calibrated individually for each basin. LSTM networks offer unprecedented accuracy for prediction in ungauged basins. SAC-SMA was calibrated separately for each basin using 15 years of daily data. It is concluded that adding physical constraints to LSTM models can improve the simulations.

Karimi et al. (2019) designs first to compare the performance of three methods, namely ANN, LSTM, and Least Absolute Shrinkage and Selection Operator (LASSO), to predict flow. Machine performance and statistical learning techniques are tested using measurement values collected from Springfield City. While all three data-based methodologies provide acceptable predictive performance, it is observed that LSTM outperforms ANN due to internal memory integrated with a feedback structure.

Kao et al. (2020) examined the LSTM network-based Encoder Decoder (LSTM-ED) model for multi-stage flood prediction in the Shihmen Reservoir basin in Taiwan. It used a feed forward neural network based Encoder Decoder (FFNN-ED) model to compare model results. It shows that both models can generally provide forward-looking predictions and that the LSTM-ED model can only make a more reliable and accurate flood prediction with the long-term dependence between precipitation and runoff.

Fang et al. (2020), proposes a local spatial sequential LSTM neural network (LSS-LSTM) for flood susceptibility prediction in Shangyou County, China. Using the flood sensitivity model of the DL technique of LSTM for the prediction model, it applies two optimizations, input enhancement and batch normalization to further improve the performance of the proposed method. The LSS-LSTM method shows satisfactory predictive performance in terms of accuracy.

Ding et al. (2020) developed an interpretable Spatial-Temporal Attention LSTM model (STA-LSTM) based on LSTM and attention mechanism, focusing on the need for flood prediction. When creating the model, Adam optimizers are used to improve the model using the Max-Min method. Experiment results on three small and medium basins in China suggest that the proposed STA-LSTM model outperforms Historical Average (HA), Fully Connected Network (FCN), Convolutional Neural Networks (CNN), Graph Convolutional Networks (GCN) in most cases.

Zhang et al. (2018) developed a new time series model based on LSTM as an alternative to computationally expensive physical models. Input values are uses to estimate monthly water deflection, evaporation, precipitation, temperature and water table depth of Hetao Irrigation Area, a drought zone in China. With the LSTM model, it is seen that it can significantly prevent excessive seating and is an alternative approach that predicts the depth of the water table.

Li et al. (2018) tried a hybrid model based on the sparse autocoder (SAE) and LSTM network to improve the prediction accuracy of dissolved oxygen in aquaculture. The hidden layer data pre-trained by SAE contains deep latent features of water quality, and then input it into the LSTM to enhance the prediction accuracy. Experimental results show that SAE-LSTM surpasses LSTM through reducing MSE in forecast steps and surpasses SAE-BPNN.

Mouatadid et al. (2019), attempts to merge repetitive LSTM maximal overlap discrete wavelet transform (MODWT) and boot techniques for accurate and robust water flow forecast for the first time. For a daily delivery time belonging to Palos De La Frontera, an irrigation region in Huelva (Southern Spain), the autocorration function and partial autocorration function are used as inputs, resulting in a significant improvement in predicting water flow.

Dao et al. (2020) compared LSTM, GRU and ANN models to predict flow in the Yutan station control basin in Fujian Province, China, using hourly discharge measurements of a flow station and hourly rainfall measurements of the precipitation station. It is seen that the forecast accuracy of the LSTM and GRU models increases with the increased time step and eventually stabilizes. When the time step is optimized, LSTM and GRU models perform equally and better than ANN models.

CHAPTER 3

MATERIAL&METHOD

3.1 Study Area and Dataset

In this study, flow estimation for Afrin Stream of the Asi Basin, which is a transboundary basin, has been made (Figure 3.1). River flow data obtained from the FMS determined in the Asi basin were collected and used as input in the study model.

Figure 3.1. Location of Asi River Basin and Main Tributaries

Asi River, Nehr'ül Arabic name given Asi, the Mediterranean region is located in a stream, the water of Turkey, is shared by Syria and Lebanon. Asi River, Lebanon's Bekaa Valley in the east of the born here Syrian borders by moving in Homs (Hama) Lake flows, after the lake creates a time the Syria-Turkey border continued to flow and Turkey in the Samandağ (Hatay) Mediterranean ' e is poured. The total length of 448 km with 35 km of Lebanon Orontes River, 325 km of which 88 km of the territory of Turkey and flows through Syria (Scheumann et al., 2011) (Fig. 3.1). When the Asi River is compared with the Nil and Jordan Rivers in terms of geomorphological structure, it is known that the riverbed is not affected much by floods. In addition, it is stated in the literature that the geomorphological structure of the riverbed and the plants surrounding the riverbed act as a protective against floods (Pichon et al., 1975). Turkey is geographically and topographically divided into 25 hydrological basins, the total amount of the average annual flow of 186 billion m^3 of these basins "type. Asi Basin is one of these hydrological basins and the total rainfall area is 7,796 km^2. Since the precipitation amounts of these basins are not the same, their water regimes and potentials differ. Especially in the lower parts of the Asi basin in the Mediterranean Region, it has been exposed to pollution caused by industry, human and agricultural activities (Özcan et al., 2012).

The Flow Route of the Asi River rises from the eastern slopes of the El-Bekaa Valley, famous for its strong water resources, between the Lebanon and Anti-Lebanon Mountains, heading north, taking the springs of Hirmil, flowing 35 km from its source and crossing into Syrian lands. After the famous Kadesh, which witnessed the historical war between Hittite and Egypt, it expands and extends to the artificial Lake Hums (Kattina), which is associated with the ancient dam dated to the 3rd century (Bildirici, 1994). It passes through Hums (Emesa) and reaches one of the most important dam basins, elRastan, and then the fertile Hama (Epiphaneia) region. After Hama, which preserves its feature of "river city" and became famous as "City of Nauras", it is included in the Gâb Plain around Şeyzar (Larissa) where it enters a deep valley to the west. Here, the expanding valley floor creates wide plains. In the Gâb region, one of the most important agricultural areas of Syria, Asi divided into channels for irrigation in the direction of flow, leaving behind the ancient Apamea (Kal'at ul-Mudik), which it passes through, at the end of the Gâb, Kârkar (historical Tell Kârkûr) to. It moves along narrow and deep valleys in a mountainous terrain between the Ansariye (Emblonos) from the west and the Zaviye (Diokleion) Mountains from the east, which has been a crossroad and an important caravan stop since the Middle Ages. between steep rocky mountains, historic 'gateway town' Derküş forming the canyon river, then along the 22 km extends to form the Turkey-Syria border. The valley floor expands suddenly at the point of turning northwest in the provincial territory of Hatay. In the Amik Plain, Asi, who turns west from Demirköprü (Cisr'ül-Hadid), takes the two big arms from the north, Karasu and Afrin, and bends to the south. It passes through Antakya (Antiochia), bends in narrow valleys up to 40 km, and pours into the

Mediterranean Sea, in the area where the ancient al-Mina harbor ruins are located in Samandağ, with a total of 386 km (Fitchett and Deford, 1973).

FMSs are available at many points to make river flow measurements in the Asi basin. From these stations; Afrin FMS (D19A006) station is recorded by making measurements. Afrin FMSs is clearly visible as in Figure 3.2. Permission to collect the historical data was obtained from the General Directorate of Electrical Works Survey Administration. The flow of the river was measured by calculating the water level pressure with a sensor placed in the target area of the river. The daily stream flow values of the river were obtained for the sighted period of 2006-2018.

Figure 3.2. The Location of Afrin Stream measurement station

Time series includes 12 years of flow values survey from October 1, 2006 to September 31, 2018. In the study, flow values were used as test data.

3.2 Methodology

In this prediction study, Keras library of LSTM method was used to estimate the flow flow. Standardization of a dataset is a common requirement for many ML estimators. The flow data used are standardized by z-score normalization (Figure 3.3). Studies to evaluate the performance of the LSTM model, one of the AI methods proposed recently, in flow estimates are in demand. Everything that will and will be used before modeling is implemented we explain each of these repetitive units in this section.

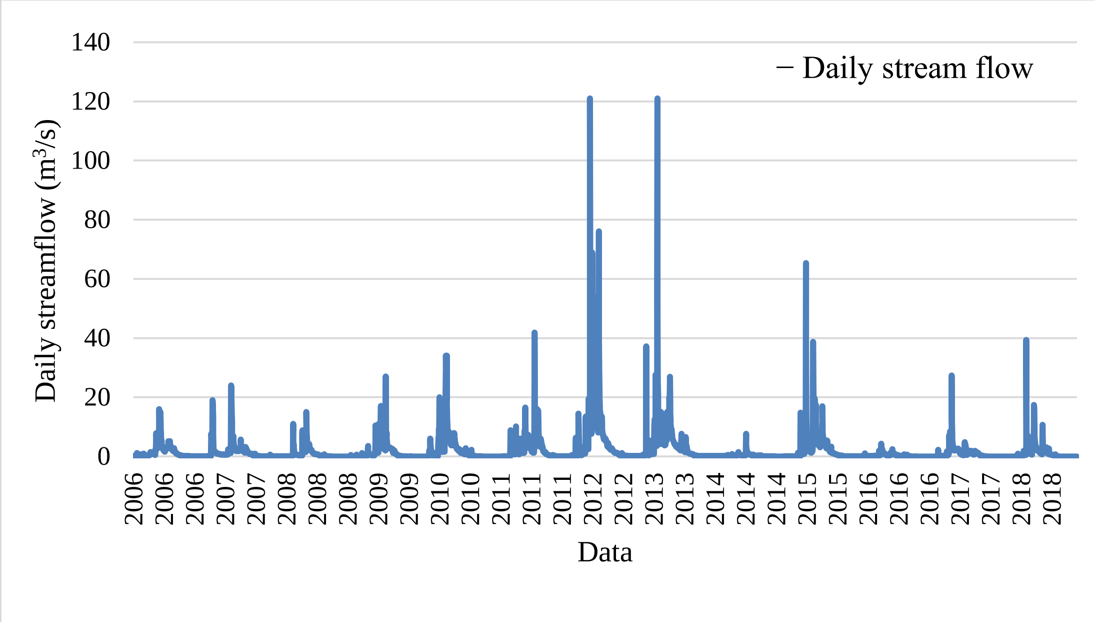

Figure 3.3. Shematic graph of Afrin Stream original dataset

3.2.1 Artificial Intelligence

AI actually started in the 1950s with intriguing personalities like Alan Turingn and John von Neumann (Tsang et al., 2017). To be even more precise; in the summer of 1956, John McCarthy, a professor of mathematics at The Dartmouth Summer Research Project (Dartmouth Conference) at Dartmouth College called this rapidly emerging technology artificial intelligence. The scientists and researchers who attended the conference were not always in total agreement on the problems or their possible solutions but they had a clear vision that they shared as they have stated in the proposal for the conference. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it (Moor, 2006).

AI is the behavior of the machine (intelligent behavior), which is called intelligence when done by human beings. It can also be said that AI is a theory that shows how the human mind works. The purpose of AI is to make machines smarter, understand what intelligence is, and make machines more useful. The aim of AI is to imitate human intelligence by means of computers, in this sense, to give computers the ability to learn to a certain extent. In this way, AI mostly consists of methods that try to model the human thinking ability, the working model of the brain, or the biological evolution of natüre (Tektaş et al., 2002).

ANN, one of the AI methods, have guided scientists on the superior features of the brain, and efforts have been made to create a mathematical model based on the neurophysical structure of the brain. In order to accurately model all the behaviors of the brain, various artificial cell and network models have been produced with the correct imaginary thought of its components. In ANN applications, applications and advanced networks are used. The ANN model created includes the input layer, hidden layer and output layer. In solving complex problems, it is generally sufficient to reach a single hidden buried solution (Cybenko, 1989). ANN method is used in many information flow estimates and successful results are obtained (Abdelwahab and Abdel-Aty, 2001).

3.2.2 Deep Learning and Algorithms

DL, ANNs have brought a new approach. This approach is called the Deep Convolution Neural Network (Hinton et al., 2006). Convolutional neural networks are known as multilayer neural networks. Important studies have been carried out with this neural network system and high performance results have been obtained. Deep convolutional neural network has achieved an important success by increasing these achievements to higher levels (Krizhevsky et al., 2012). Significant studies are emerging in many areas such as signal processing, video analysis, image analysis and detection, classification, medical image processing with convolutional neural network (Snoek et al., 2005). While using this neural network, some steps are taken. These are defined as pre-treatment, feature extraction and classification-detection. Special approaches are exhibited at each stage and studies are carried out to increase accuracy (Li et al., 2010). Many different approaches are presented especially for the feature extraction process. It is tried to reveal the specific points of the event desired to be detected by feature extraction. In the next process, neural networks are used to determine the class of properties determined using ANN (Scherer et al., 2010). With deep learning, many processes that have been done before are carried out together and the result is achieved. Here, structures such as pre-processing and feature extraction are ignored and these operations are performed automatically within the neural network. In the deep convolutional neural network, feature extraction is determined within the network and the properties of the structure desired to be detected within the layers are determined. There is an interconnected hierarchical structure between the lower layer and the upper layer. There is no special phase for feature extraction. In its structure within layers, the distinctive features of the object-event are determined and transferred to the next layer (Hinton and Salakhutdivot, 2006).

LeNet is the first basic DL network that enables the classification of numbers by learning handwriting numbers from 0 to 9 using gradient-based Convolutional neural networks. Although the DL structure was shown in 2006, DL became more popular with the ImageNet competition. Algorithms for image recognition have been put forward in the competition held by ImageNet. DL architectures have come to the fore in these algorithms. The results of DL architectures and object recognition have yielded very successful results. DL architectures are used in many areas such as pattern recognition, detection, classification, future prediction, pharmaceutical production, dictionary, signal processing, medical, finance, and defense industry. Looking at the results obtained, it was revealed that DL structures give much better results than many known methods (Hinton, 2006).

3.2.3 Long Short-Term Memory

LSTMs are different neural networks than the classical Multi-Layer Sensor (MLP). Like an MLP, LSTM is made up of layers made up of neurons. When predictions are made using LSTM, input data is spread over the network. Like the RNN, the LSTM also has repetitive connections so that the state from previous activations of the neuron in the previous time step is prepared an output and used as input with other inputs in the next step. In addition, LSTM has a perfect structure by preventing problems that prevent the training and scaling of other RNNs. This feature and its achieved and effective results cause this method to be popular (Brownlee, 2017).

The core idea of the LSTM is to replace every hidden unit in RNN with a cell of memory function (Hochreiter and Schmidhuber, 1997). Basically it is a variant of RNN model, and can be adapted to achieve excellent performance in processing time series data (Graves, 2013). In order to solve the problem of gradient isappearance, gradient expansion and error hindrance in the traditional LSTM, an extended memory block with a memory cell is applied to the LSTM (Gers et al., 1999). The memory cell can keep the historical data in the memory, and wake it up in any epoch when needed. Figure 3.4 illustrates the operating mechanism and internal memory cell of the LSTM.

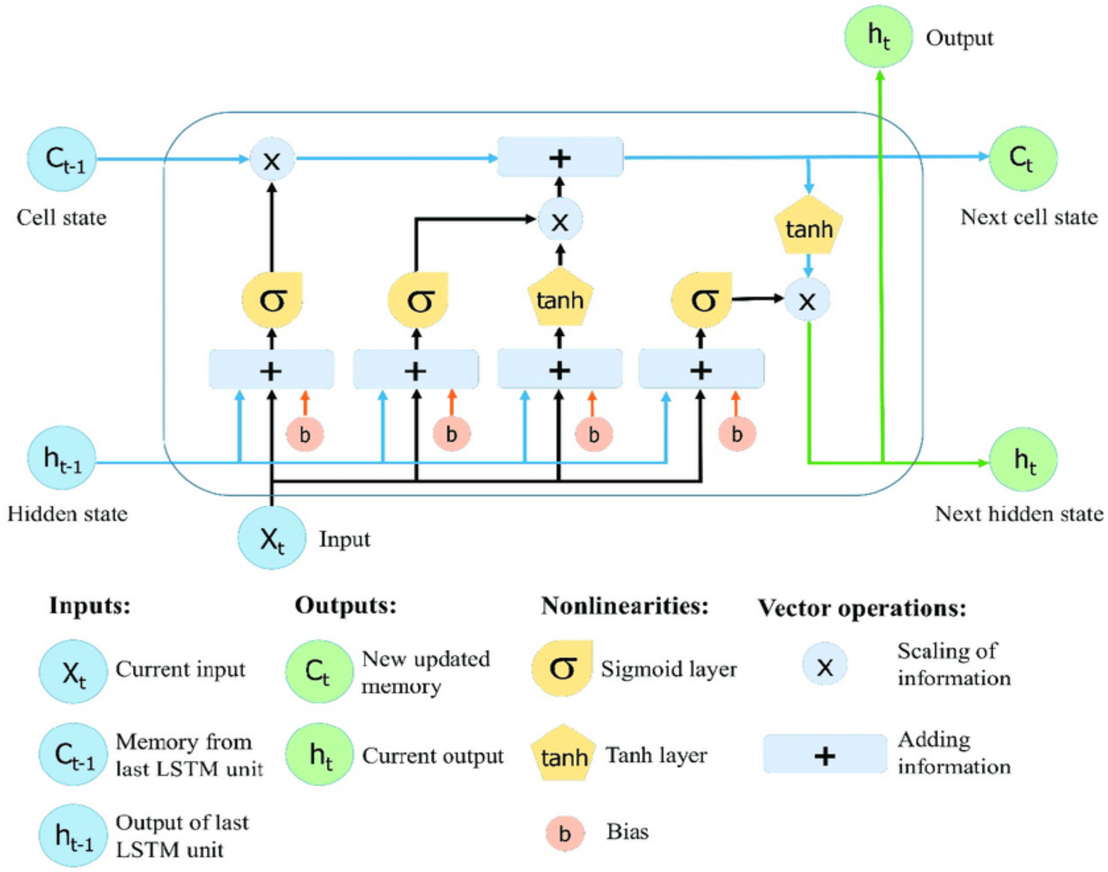

Figure 3.4. The internal structure of the LSTM memory cell

Forget gate, input gate, cell gate and output gate units are denoted by f_t, i_t C_t and O_t, respectively. The memory cell learns at the t time based on the input at the t time and the output at the (t-1) time in the network. The detail algorithm involved in the components of the memory cell is described as follows (Eqs. 1–6):

$$f_t = \sigma\ W(_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$it = \sigma\ W(i \cdot [\ h_t -1, xt] + bi \quad (2)$$

$$C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = ft * Ct-1 + it * Ct \quad (4)$$

$$O_t = softsign\ w\ h(\ _o [\ _{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = O_t * ReLU *(C_t) \quad (6)$$

3.2.4 LSTM Model Calibration and Parameters

As discussed in the LSTM model section, the parameters of the LSTM module/cell are Wu, Wf, Wo, Wc, bu, bf, bo, and bc. It use the parameters Wy and by to generate the predictions of ŷ from the hidden state of the LSTM. The dimension of the LSTM weight matrices depends on the dimension of the hidden state and the number of input features in x^4. At each timestep, we submit relevant variables x, such as marketing actions, customer behavior and seasonality indicators, in the form of a vector of dummy variables. In our illustration, the y variable is a vector of size one that indicates whether the customer has purchased during the following period, although the dependent variable can easily include multiple indicators. In the first empirical example, based on past donation and solicitation histories, the probability of an individual to donate to a charity over a period of time will be estimated. The calibration and prediction setup for the LSTM is shown in Figure 3.5. The analyst sets the number of hidden units in the LSTM exogenously. Settings that are used to control the learning of the algorithm, but are not parameters tuned by the learning algorithm itself, are referred to as hyperparameters (Kuhn and Johnson, 2013). Apart from the number of hidden units, the LSTM neural network has other hyperparameters such as the learning rate, batch size, or the relative contribution of norm penalties to the cost function. While training a model, the analyst aims at setting the parameters and hyperparameters such that the model reaches optimal capacity and therefore maximizes the chances that the model will generalize well to unseen data. Models with low capacity would underfit the training set and hence have a high bias. However, models with high capacity may overfit the training set and exhibit high variance. Representational capacity is the ability of the model to fit a wide range of functions. (Goodfellow et al., 2016).

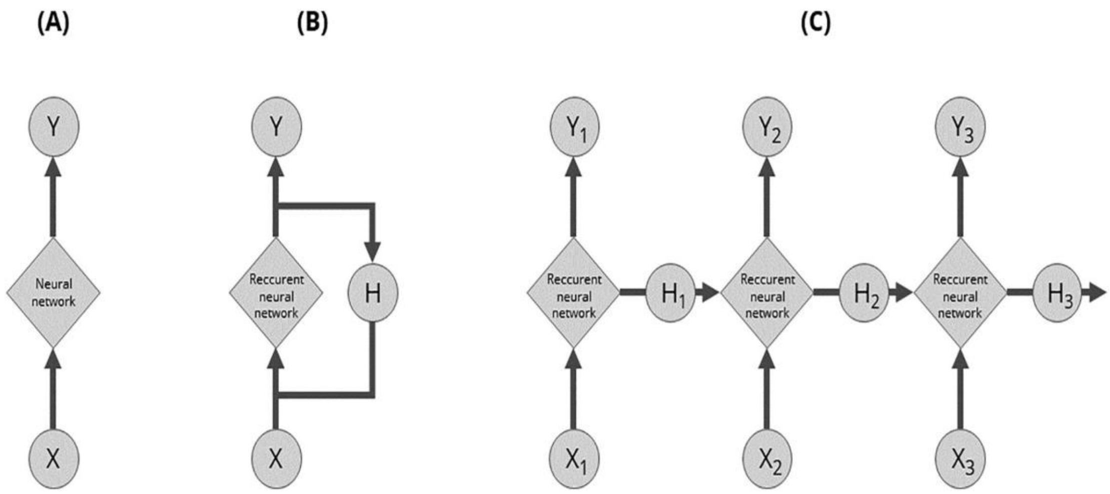

Figure 3.5. LSTM calibration and prediction setup scheme

The operating system used in this study was Microsoft Windows 10. TensorFlow 1.7.0, Cuda 9.0, and Python 3.6 were selected as the development environment. The learning rate is one of the key parameters of the developed LSTM model. The optimised value of the learning rate is set to 0.00001, as shown in Table 2. The other optimised values of the parameters of the developed LSTM model are as follows: the number of hidden layer units is 30; the input layer size is 2; the output layer size is 1; the learning rate is 0.00001; and the cycle index is 2000.

3.2.5 Python Keras Library and Programming Language

Keras is one of the most frequently used neural network libraries in open source artificial intelligence developed by François Chollet in Python in 2015 and also offers different modules together. Keras is one of the popular neural network model development APIs and can work with TensorFlow, R, Microsoft Cognitive Toolkit, Theano or PlaidML technologies. User-friendliness, simpler ANN models can be created, modular and open to development are some of the sub-reasons for being widely used in the field of ML. Since it is easy to create sequential deep network models using Keras, it is frequently used in DL studies. DL models are used to understand the structure in nonlinear and difficult to correlate data sets. It provides an environment where Keras offers great convenience to the researcher, the model architecture can be easily changed, the classification or regression problems can be easily identified, and the activation and cost functions can be personalized (Chollet, 2018).

The Python programming language started to be developed in Amsterdam in the early 90's by Guido Van Rossum. Python is an interpreted, interactive, object-oriented programming language. List and association arrays, dynamic writing and dynamic linking, modules, classes, exceptions, automatic memory management, etc. It provides high-level data structures such as. It has a very simple and elegant syntax and yet powerful. It was designed in 1990 by Guido van Rossum. Like many other scripting languages, it is free even for commercial purposes and can be run on practically any modern computer. A Python program is automatically compiled by the translator into platform independent bytecode and then interpreted (Sanner, 1999)

3.2.6 Optimizers

Adam optimization is an algorithm for first order gradient based optimization of stochastic functions based on predictions of lower order moments. The method can be easily used in non-stationary problems and has a minimum memory requirement. The success rate is much higher in efficient parameters. While creating the probability distribution of the outputs, it statistically optimizes the input-output map (Kingma and Ba, 2014).

Adam, it stores exponentially decaying average of past gradient (a_t) which represents the first moment (mean) and past squared gradient (u_t) which represents the second moment (variance) (Bera and Shrivastava, 2020). The at and ut are computed as follows:

$$a_t = \beta_1 a_{t-1} + (1 - \beta_1)\, d_t \tag{7}$$

$$a_t = \beta_2 u_{t-1} + (1 - \beta_2) {d_t}^2 \tag{8}$$

The at and ut are biased towards zero when decay rates are very small (i.e. β_1 and β_2 are near to zero). To overcome this situation, the first moment and second moment with biascorrected term are computed as follows:

$$\hat{a}_t = \frac{a_t}{1-{\beta_1}^t} \tag{9}$$

$$\hat{u}_t = \frac{u_t}{1-{\beta_2}^t} \tag{10}$$

Then, the Adam weight update rule is given by Equation (11).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\mathrm{u}_t + \epsilon}} \,.\, \hat{a}_t \tag{11}$$

Adamax is a variation of Adam dependent on the unendingness standard. It is a variant of Adam based on the infinity norm. Default parameters follow those provided in the paper. Adamax is sometimes superior to adam, specially in models with embeddings (Kingma and Ba, 2015).

AdaMax, weight update rule is defined by replacing the term $\sqrt{u_t + \epsilon}$ in Adam weight update rule as given in Equation (11) by exponentially weighted infinity norm (V_t) as shown below:

$$\theta_{t+1} = \theta_t - \frac{\eta}{V_t} . \hat{a}_t \tag{12}$$

$$\begin{aligned} V_t &= \beta_1 u_{t-1} + (1 - \beta_2)\, d_t \\ &= \max(\beta_2 u_{t-1}, d_t) \end{aligned} \tag{13}$$

3.2.7 Loss Functions

The loss function is the bread and butter of modern machine learning; it takes your algorithm from theoretical to practical and transforms neural networks from glorified matrix multiplication into deep learning. It's a method of evaluating how well algorithm models dataset. If predictions are totally off, loss function will output a higher number. If they're pretty good, it'll output a lower number. While changing parts of the algorithm to test and improve the model, the loss function can tell if it has arrived anywhere. Here, missing functions such as MSE, Huber and LogCosh are used:

MSE is often used in the measurement of the difference between the predicted value and the actual value in the system. It is a quadratic measure of the error magnitude. The estimate is also expressed as the standard deviation of residues. Its value can vary from zero to infinity. As in Mean Square Error, the closer to zero, the more successful. It can prevent the use of unwanted absolute values in many functions (Boyraz and Engin, 2018).

$$\text{MSE} = \frac{\sum_{i=1}^{n} (y_i - y_i^p)^2}{n} \tag{14}$$

Huber loss is used in robust regression to advantage from desirable properties of both I2 and I1 penalty. Noise may have forced the original matrix entries to deviate from their genuine quantization bounds. Thus, it intentionally use linearly growing (l1) penalty for violations from quantization bounds to be less sensitive to outliers as squared loss is. Thus, sharply varying behavior on feasible region is pointless. It also benefits the algorithm as it has non-zero learning rate for the values which hold in the problem constraints. (Esmaeili and Marvasti, 2019).

$$\mathrm{H}_{\Omega}(\mathrm{X})=\sum_{(i,j)\in\Omega} \widetilde{\mathrm{H}}\, i,j(X_{i,j}) \tag{15}$$

LogCosh is another function used in regression tasks that's smoother than L_2. LogCosh is the logarithm of the hyperbolic cosine of the prediction error. But LogCosh loss isn't perfect. It still suffers from the problem of gradient and hessian for very large off-target predictions being constant, therefore resulting in the absence of splits for XGBoost (Chen et al., 2018).

$$\mathrm{L}(\mathrm{y}, y^{p})= \sum_{i=1}^{n} \log(cosh(y_i^p - y_i)) \tag{16}$$

3.2.8 Model Evaluation Criteria

To evaluate the performance of estimation models in areas related to river flow, mean absolute error (MAE), root mean square error (RMSE), actual-predicted (MEAN) and standard deviation (STD), are statistical methods commonly used to compare predicted values with observed values.

MAE;

The measure of difference between two continuous variables is called MAE. MAE is the average vertical distance between each true value and the line that best fits the data. This value is also the average horizontal distance between all data points and the best fitting line. Since the MAE value can be easily interpreted, it is frequently used in regression and time series problems. The MAE value takes a value between 0 and ∞. The lower its value, the better it performs (Celestin et al., 2020).

$$\mathrm{MAE}= \frac{\sum_{i=1}^{n}(y_i - y_i^p)}{n} \tag{17}$$

MEAN;

In statistics, the actual value is the value that is obtained by observation or by measuring the available data. It is also called the observed value. The predicted value is the value of the variable predicted based on the regression analysis. The difference between the actual value or observed value and the predicted value is called the residual in regression analysis (Yao et al., 2009).

$$e = y - \hat{y} \quad (18)$$

RMSE;

RMSE represents the difference between the predicted value and the actual value. It is mainly used for calculating the values at a different time interval. The prediction is more misaligned when the value is larger (Chang et al., 2020).

$$\text{RMSE}=\sqrt{\frac{\sum_{i=0}^{n-1}(y_i - f_i)^2}{n}} \quad (19)$$

STD;

Most instruction on the standard deviation tends to emphasize teaching a formula, practice with performing calculations, and tying the standard deviation to the empirical rule of the normal distribution. This emphasis on calculations and procedures does not necessarily promote a conceptual understanding of standard deviation. A conceptual model of the standard deviation is needed to develop instruction that promotes the concept. It is through the coordination of distribution and deviation (as distance from the mean) that a dynamic conception of the standard deviation is derived as the relative density of values about the mean (Reading and Shaughnessy, 2004).

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Evaluation of Predicted and Observed Results Using Adam Optimizer and MSE Loss

Table 4.1 shows the results of the LSTM model using different loss functions and optimizers according to the performance evaluation criteria. When the test evaluation was made, multiple criteria were taken into consideration and comparison was made. According to the table, when looking at the RMSE criterion, the result for LSTM is 1,5934 from the model in which the Adam optimizer and the MSE loss function are used. According to the regression performance criteria, the result for LSTM is 0.9785. In the Adam MSE model, was found to be achieved the accuracy of MAE, RMSE, MEAN, STD and R^2 values for test.

Table 4.1 Criterion results for MSE loss with Adam optimizer

Statistics	Statistical Metrics
MAE	0,5614
RMSE	1,5934
MEAN	0,6231
STD	2,3530
R^2	0,9785

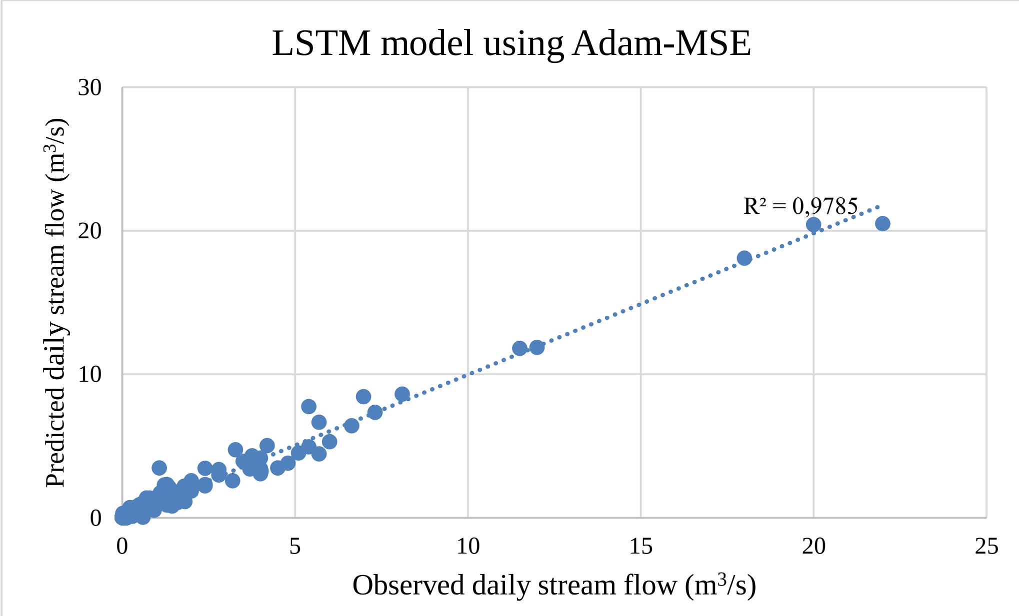

Figure 4.1 Comparison evaluation of Observed and Predicted values for Adam optimizer MSE loss

Fig. 4.1 indicate the distribution of the predicted and observed data of Adam optimizer MSE loss function. When looking at the scatter plots, the observed value of 20,49 and the predicted value of 22 are obtained, and the R^2 0.9785 value of the LSTM model can give close to realistic results.

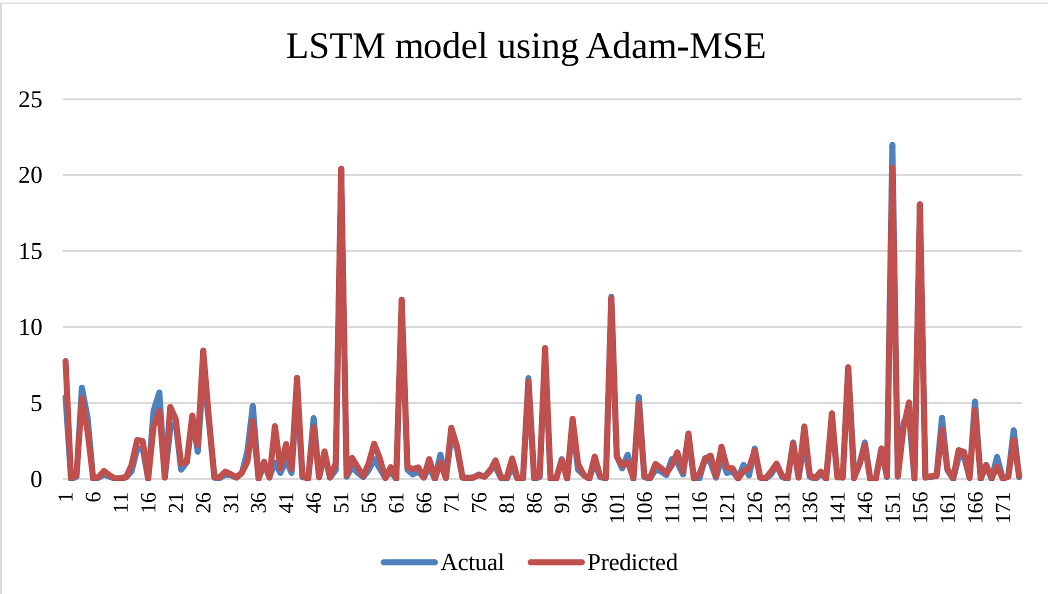

Figure 4.2 Evaluation of test accuracy for Adam optimizer and MSE loss

The comparison of and model values using Adam optimizer and MSE loss function together is shown in Fig. 4.2. As shown in the way, the Actual value is 22 and the line with the highest difference in Precdited is at the point "151".

4.2 Evaluation of Predicted and Observed Results Using Adam Optimizer and Huber Loss

Table 4.2 shows the results of the LSTM model using different loss functions and optimizers according to the performance evaluation criteria. When the test evaluation was made, multiple criteria were taken into consideration and comparison was made. According to the table, when looking at the RMSE criterion, the result for LSTM is 1,5692 from the model in which the Adam optimizer and the Huber loss function are used. According to the regression performance criteria, the result for LSTM is 0,9916. In the Adam Huber model, was found to be achieved the accuracy of MAE, RMSE, MEAN, STD and R^2 values for test.

Table 4.2 Criterion results for Huber loss with Adam optimizer

Statistics	**Statistical Metrics**
MAE	0,4241
RMSE	1,5692
MEAN	1,3068
Standart Deviation	2,9133
R^2	0,9916

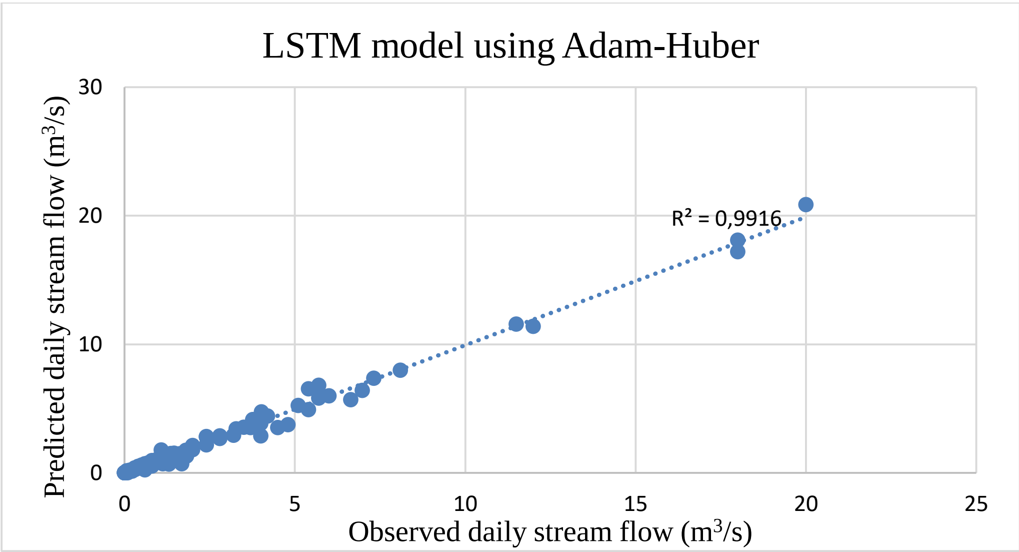

Figure 4.3 Comparison evaluation of Observed and Predicted values for Adam optimizer Huber loss

Fig. 4.3 indicate the distribution of the predicted and observed data of Adam optimizer Huber loss function. When looking at the scatter plots, the observed value of 20,48 and the predicted value of 20 are obtained, and the R² 0.9916 value of the LSTM model can give close to realistic results.

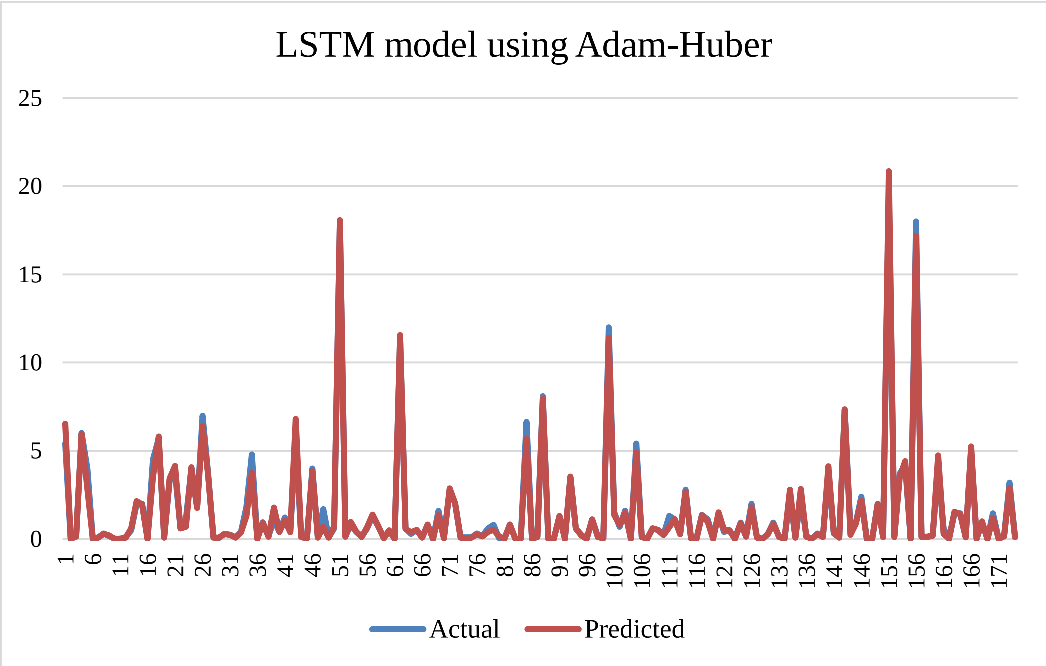

Figure 4.4 Evaluation of test accuracy for Adam optimizer and Huber loss

The comparison of test and model values using Adam optimizer and Huber loss function together is shown in Fig. 4.4. As shown in the way, the Actual value is 18 and the line with the highest difference in Precdited is at the point "156".

4.3 Evaluation of Predicted and Observed Results Using Adam Optimizer and LogCosh Loss

Table 4.3 shows the results of the LSTM model using different loss functions and optimizers according to the performance evaluation criteria. When the test evaluation was made, multiple criteria were taken into consideration and comparison was made. According to the table, when looking at the RMSE criterion, the result for LSTM is 1,7375 from the model in which the Adam optimizer and the LogCosh loss function are used. According to the regression performance criteria, the result for LSTM is

0,9882. In the Adam LogCosh model, was found to be achieved the accuracy of MAE, RMSE, MEAN, STD and R^2 values for test.

Table 4.3 Criterion results for LogCosh loss with Adam optimizer

Statistics	**Statistical Metrics**
MAE	0,4367
RMSE	1,7375
MEAN	0,9125
Standart Deviation	0,6151
R^2	0,9882

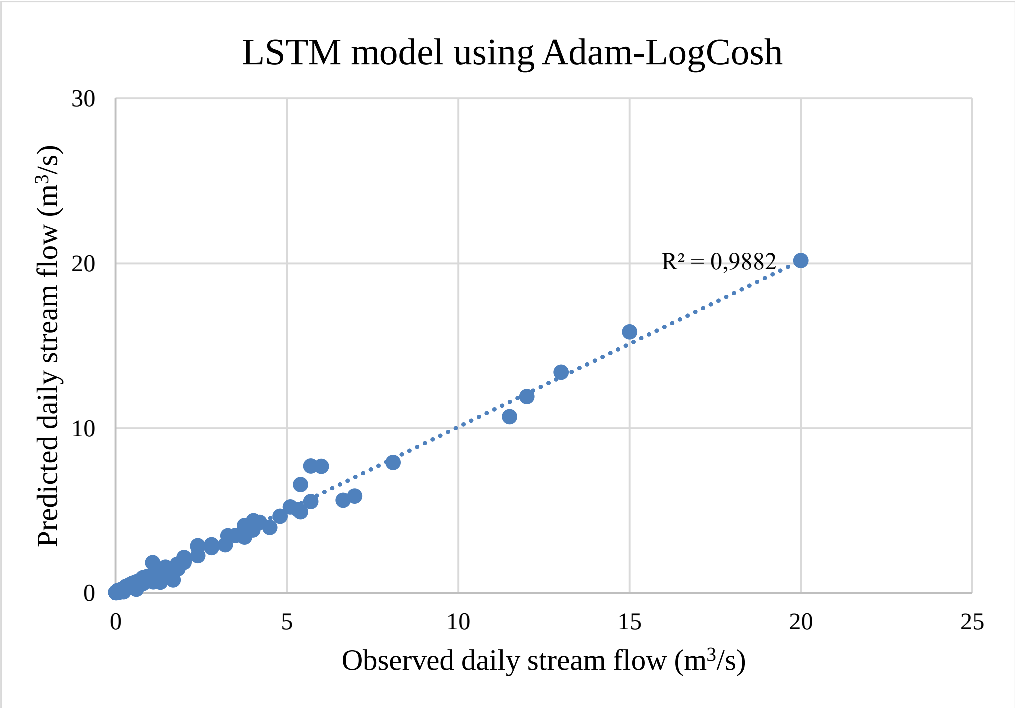

Figure 4.5 Comparison evaluation of Observed and Predicted values for Adam optimizer LogCosh loss

Fig. 4.5 indicate the distribution of the predicted and observed data of Adam optimizer LogCosh loss function. When looking at the scatter plots, the observed value of 20,17 and the predicted value of 20 are obtained, and the R^2 0.9882 value of the LSTM model can give close to realistic results.

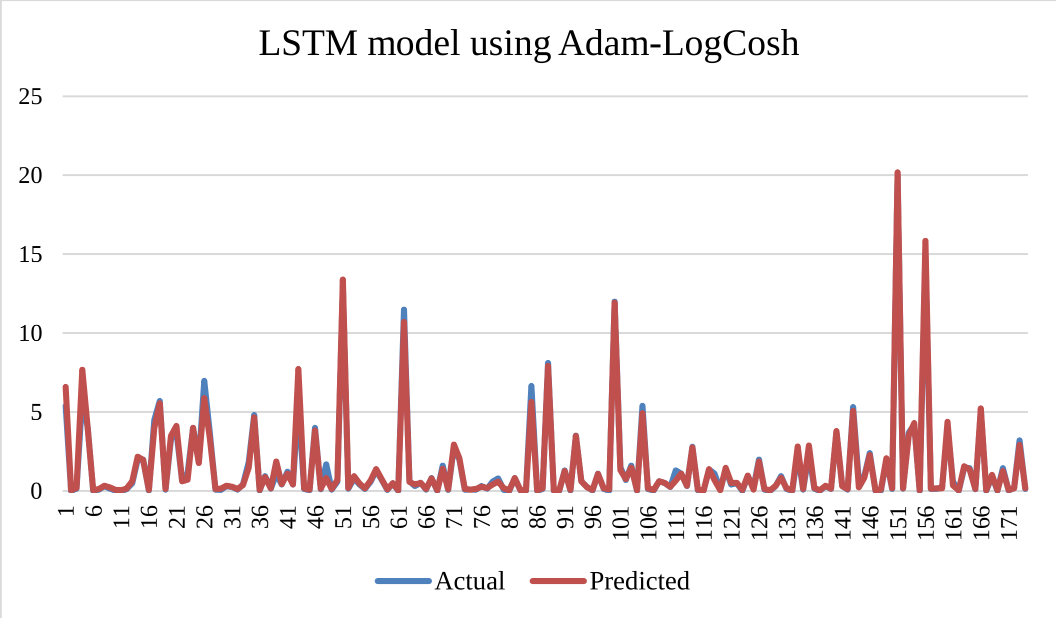

Figure 4.6 Evaluation of test accuracy for Adam optimizer and LogCosh loss

The comparison of test and model values using Adam optimizer and LogCosh loss function together is shown in Fig. 4.6. As shown in the way, the Actual value is 11,5 and the line with the highest difference in Precdited is at the point "62".

4.4 Evaluation of Predicted and Observed Results Using Adamax Optimizer and MSE Loss

Table 4.4 shows the results of the LSTM model using different loss functions and optimizers according to the performance evaluation criteria. When the test evaluation was made, multiple criteria were taken into consideration and comparison was made. According to the table, when looking at the RMSE criterion, the result for LSTM is 1,5232 from the model in which the Adamax optimizer and the MSE loss function are used. According to the regression performance criteria, the result for LSTM is 0,9854. In the Adamax MSE model, was found to be achieved the accuracy of MAE, RMSE, MEAN, STD and R^2 values for test.

Table 4.4 Criterion results for MSE loss with Adamax optimizer

Statistics	**Statistical Metrics**
MAE	0,4557
RMSE	1,5232
MEAN	0,6813
Standart Deviation	0,5277
R^2	0,9854

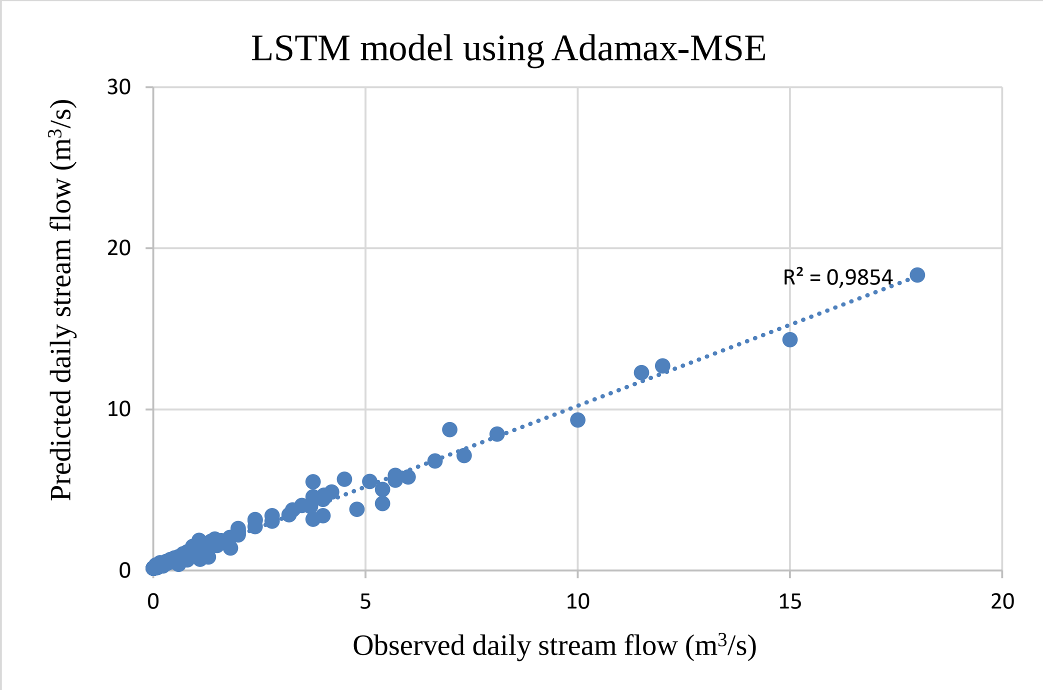

Figure 4.7 Comparison evaluation of Observed and Predicted values for Adamax optimizer MSE loss

Fig. 4.7 indicate the distribution of the predicted and observed data of Adamax optimizer MSE loss function. When looking at the scatter plots, the observed value of 18,33 and the predicted value of 18 are obtained, and the R^2 0.9854 value of the LSTM model can give close to realistic results.

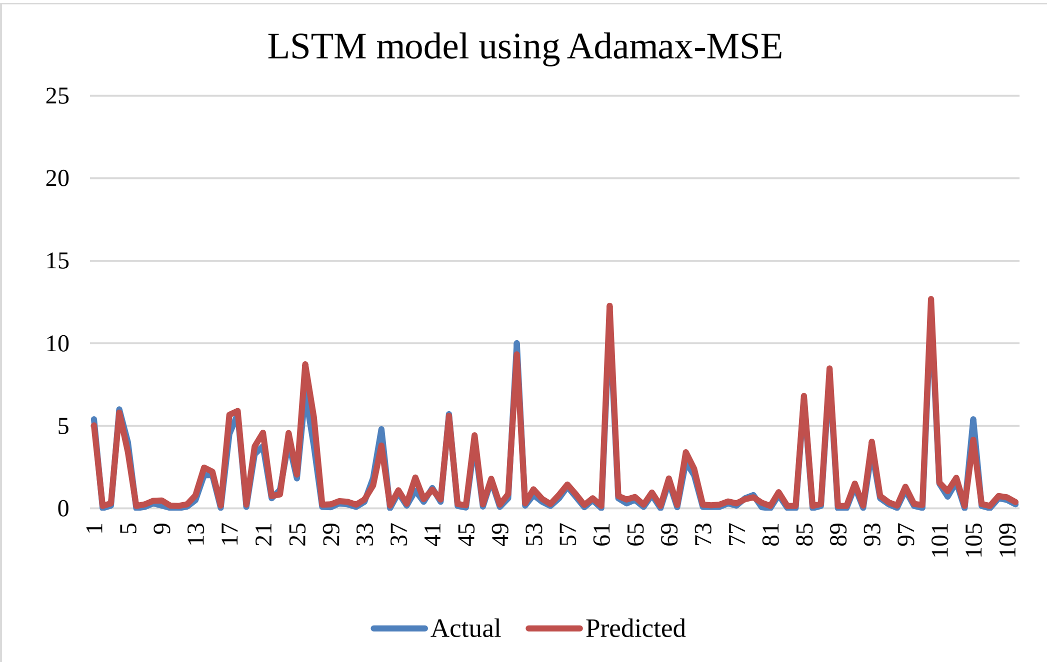

Figure 4.8 Evaluation of test accuracy for Adamax optimizer and MSE loss

The comparison of test model value using Adamax optimizer and MSE loss function together is shown in Fig. 4.8. As shown in the way, the Actual value is 10 and the line with the highest difference in Precdited is at the point "51".

4.5 Evaluation of Predicted and Observed Results Using Adamax Optimizer and Huber Loss

Table 4.5 shows the results of the LSTM model using different loss functions and optimizers according to the performance evaluation criteria. When the test evaluation was made, multiple criteria were taken into consideration and comparison was made. According to the table, when looking at the RMSE criterion, the result for LSTM is 1,2872 from the model in which the Adamax optimizer and the Huber loss function are used. According to the regression performance criteria, the result for LSTM is 0,9922. In the Adamax Huber model, was found to be achieved the accuracy of MAE, RMSE, MEAN, STD and R^2 values for test.

Table 4.5 Criterion results for Huber loss with Adamax optimizer

Statistics	**Statistical Metrics**
MAE	0,3385
RMSE	1,2872
MEAN	0,9496
Standart Deviation	0,5425
R^2	0,9922

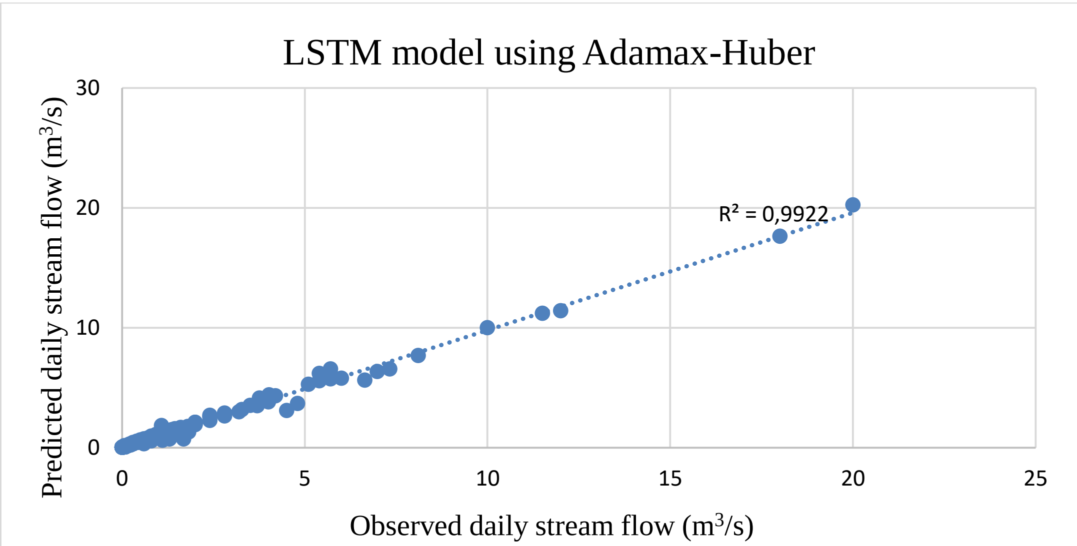

Figure 4.9 Comparison evaluation of Observed and Predicted values for Adamax optimizer Huber loss

Fig. 4.9 indicate the distribution of the predicted and observed data of Adamax optimizer Huber loss function. When looking at the scatter plots, the observed value of 20,37 and the predicted value of 20 are obtained, and the R^2 0.9922 value of the LSTM model can give close to realistic results.

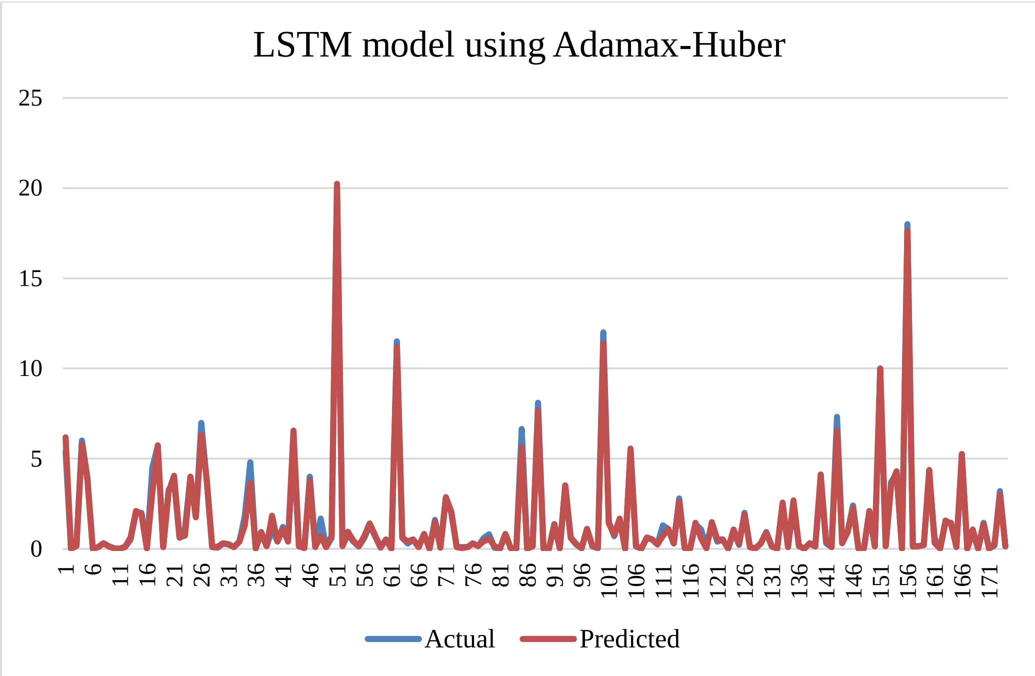

Figure 4.10 Evaluation of test accuracy for Adamax optimizer and Huber loss

The comparison of test model value using Adamax optimizer and Huber loss function together is shown in Fig. 4.10. As shown in the way, the Actual value is 18 and the line with the highest difference in Precdited is at the point "156".

4.6 Evaluation of Predicted and Observed Results Using Adamax Optimizer and LogCosh Loss

Table 4.6 shows the results of the LSTM model using different loss functions and optimizers according to the performance evaluation criteria. When the test evaluation was made, multiple criteria were taken into consideration and comparison was made. According to the table, when looking at the RMSE criterion, the result for LSTM is 1,3462 from the model in which the Adamax optimizer and the LogCosh loss function are used. According to the regression performance criteria, the result for LSTM is 0,9949. In the Adamax LogCosh model, was found to be achieved the accuracy of MAE, RMSE, MEAN, STD and R^2 values for test.

Table 4.6 Criterion results for LogCosh loss with Adamax optimizer

Statistics	Statistical Metrics
MAE	0,3551
RMSE	1,3462
MEAN	0,9309
Standart Deviation	0,5814
R^2	0,9949

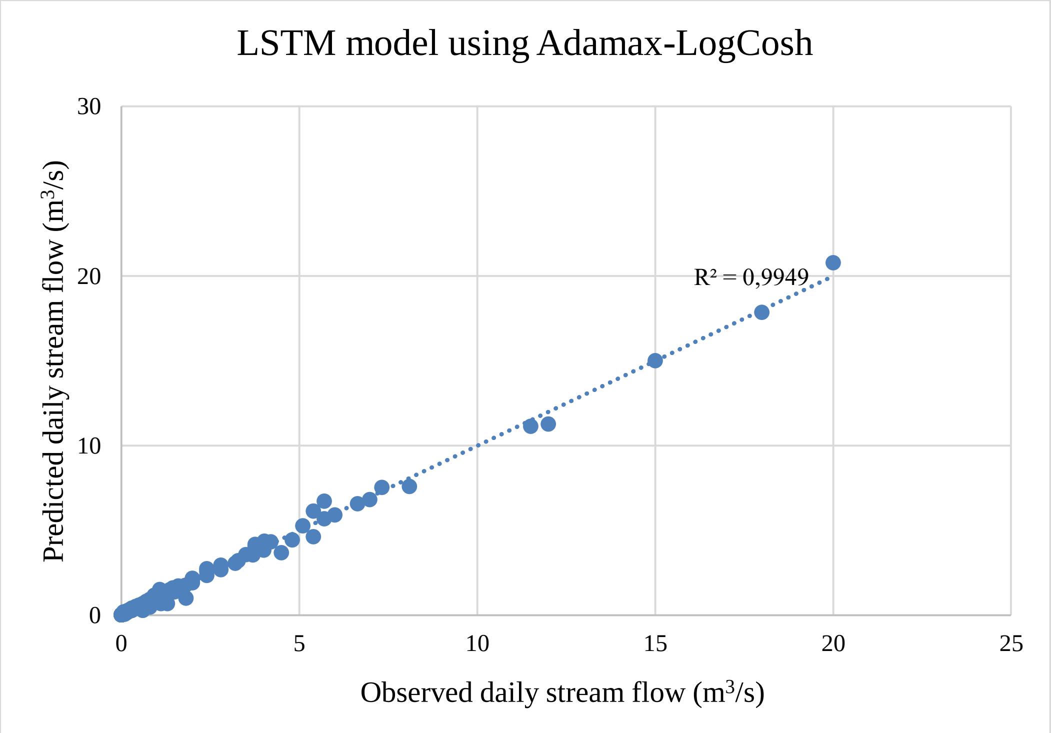

Figure 4.11 Comparison evaluation of Observed and Predicted values for Adamax optimizer LogCosh loss

Fig. 4.11 indicate the distribution of the predicted and observed data of Adamax optimizer LogCosh loss function. When looking at the scatter plots, the observed value of 20,77 and the predicted value of 20 are obtained, and the R^2 0.9949 value of the LSTM model can give close to realistic results.

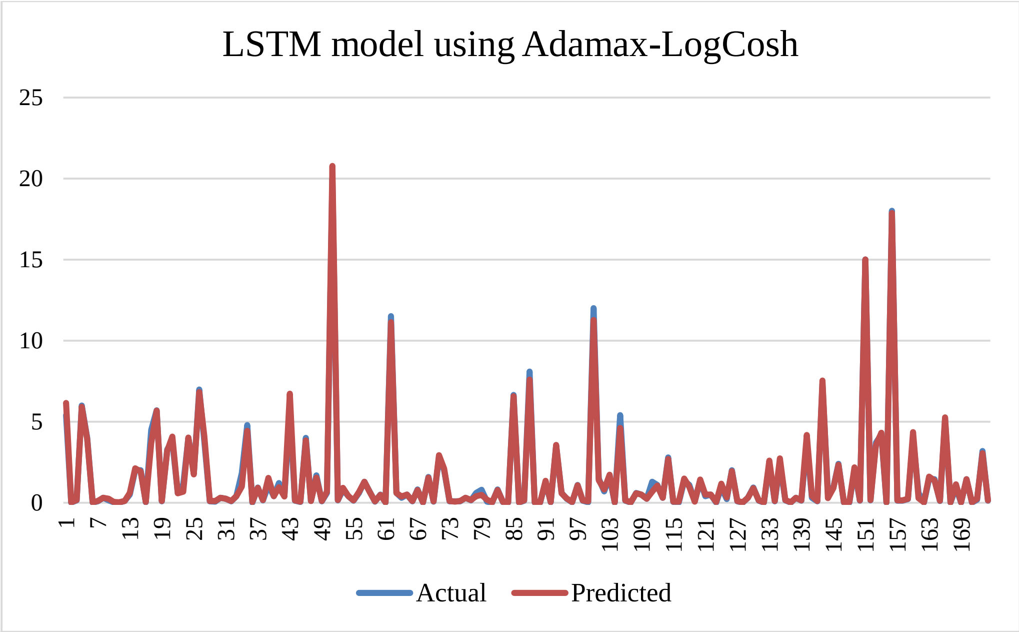

Figure 4.12 Evaluation of test accuracy for Adamax optimizer and LogCosh loss

The comparison of test model value using Adamax optimizer and LogCosh loss function together is shown in Fig. 4.12. As shown in the way, the Actual value is 12 and the line with the highest difference in Precdited is at the point "100".

Predictive models were validated using test data and evaluated using the MAE, RMSE, MEAN, STD and R^2 evaluation criteria while comparing observed and predicted data. With the selected evaluation parameters LSTM model was applied to forecast daily stream flows at Afrin river FMS. In the study, Adam and Adamax optimizers were tested and loss functions of MSE, Huber and LogCosh were analyzed.

Fig. 4.1, 4.3, 4.5, 4.7, 4.9, 4.11 indicate the distribution of predicted and observed data of LSTM. When looking at the scatter plots, the model in which the observed values and predicted values are closest to each other and the deviations are the least is the model that uses the LogCosh loss function and the Adamax optimizer function of Fig. 4.11.

Table 4.1, 4.2, 4.3, 4.4, 4.5, 4.6 show the results of LSTM model using different loss functions and optimizers according to performance evaluation criteria. When looking at the criterions in the tables, the best result for regression is from the model in table 4.6 which the Adamax optimizer and LogCosh is used is obtained.

According to MAE performance criterion, the best result for LSTM is 0.3385, from the model in which the Adamax optimizer and the Huber loss function are used; the worst result is from the model where Adam optimizer and MSE loss function are used with a value of 0.5614. According to this criterion, the most successful result belongs to the Adamax model, and the Huber loss performed better in both optimizers.

According to RMSE performance criterion, the best result for LSTM is 1,2872 from the model in which the Adamax optimizer and the Huber loss function are used; the worst result is from the model where Adam optimizer and LogCosh loss function are used with a value of 1.7375. According to this criterion, the most successful result belongs to the Adamax model, and the Huber loss performed better in both optimizers.

According to MEAN performance criterion, the best result for LSTM is 0.6231, from the model in which the Adam optimizer and the MSE loss function are used; the worst result is from the model where Adam optimizer and Huber loss function are used with a value of 1.3068. According to this criterion, the most successful result belongs to the Adam model, and the MSE loss performed better in both optimizers.

According to STD performance criterion, the best result for LSTM with a value of 0,5277 from the model in which the Adamax optimizer and the MSE lost function are used; The worst result is from the model using Adam optimizer and Huber loss function with a value of 2,9133. According to this criterion, the most successful result belongs to the Adamax model, and the MSE loss performed best.

Finally, according to the R^2 performance criterion, the best result for LSTM with 0.9949 value from the model where Adamax optimizer and LogCosh loss function are used; The worst result is from the model where Adam optimizer and MSE loss function are used with a value of 0.9785. According to this criterion, the most successful result belongs to the Adamax model, and the LogCosh loss performed best. When the best results of this criterion are compared, it is understood that the Adamax optimizer is more successful and the results with the worst performance for both optimizers belong to MSE. Considering all performance criteria; It is seen that the Adamax model is more successful than the Adam model.

CHAPTER 5

CONCLUSIONS

It is essential to use reliable flow estimates when planning regulations and applications in flows. Traditional flow estimation methods can be insufficient to make effective estimates with the uncertainties of the system and the nonlinear properties of the system. When the LSTM model, which are among the artificial intelligence methods that give effective results in the studies it was used in, was examined, it was found to be suitable as a solution method.

This paper proposed a method of a DL model based on LSTM was improved to estimate the streamflow of the Asi River. Optimizers were compared to prove which model is more advantageous in time series estimation. While the LSTM method was tested with the test data in the study, it was modeled using Adam and Adamax optimizers and using loss functions appropriate for regression (MSE, Huber, LogCosh). The optimizers and loss functions below R^2 value of 0.96 were ignored and were not included in the study. The results of the achieve of the LSTM model was evaluated by comparing the Adam and Adamax optimizers with R^2 value of 0.97 and above according to other statistical evaluation criteria. From the of view of performance assessment criteria, Adamax LogCosh model shows slightly better correctness in possessing the time serial of stream flow. The reason is probably that the this model decides what information to throw away and what new information to add. However, there is still room for improvement in the other stream-flow forecasting models: the results show errors in peak volume forecast which cannot be ignored. The models may be improved in the following ways: First, due to the limit of time and hardware capacity, the parameter choices of LSTM model are only based on simple tests and lack of thorough study. Also in the forecasting model phase, a combination of the Adamax LogCosh model and other models is an important consideration for future research.

REFERENCES

Abdelwahab, H. T., Abdel-Aty, M. A. (2001). Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record*, *1746*(1), 6-13.

Acet, A., Akkaya, A.E.A. (2019). Deep Learning Image Classification Using Tensorflow For Optical Aviation Systems. *The Journal of Cognitive Systems*, *5*(1), 1-4.

Ağca, N., Karanlık, S., Ödemiş, B. (2014). Assessment of Ammonium, Nitrate, Phosphate, and Heavy Metal Pollution in Groundwater from Amik Plain, Southern Turkey. *Environmental Monitoring and Assessment*, 186(9), 5921-5934.

Altunparmak, A., Basakin, E.B. (2018). River Flow Estimation Using Time Series and Comparison with Different Methods. *Journal of Science and Technology*, 11(1), 92-101.

Bera, S., Shrivastava, V. K. (2020). Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. *International Journal of Remote Sensing*, *41*(7), 2664-2683.

Bildirici, M. (1994). Historical Water Constructions-Konya, Karaman, Niğde, Aksaray, Yalvaç, Side, Mut, Silifke, DSI Technology Department, Printing and Photo-Film Branch Directorate, *T.C. Ministry of Public Works and Settlement, General Directorate of State Hydraulic Works*, 40th Establishment Year, Ankara.

Boyraz, C., Engin, Ş. N. (2018). Flow with Deep Learning, 2018 6th International Conference on Control Engineering and Information Technology, (1-5), *IEEE.*

Brownlee, J. (2017). Long short-term memory networks with python: develop sequence prediction models with deep learning. *Machine Learning Mastery (1)*, 1-229.

Celestine, M., Nadkarni, N. A., Garin, C. M., Bougacha, S., Dhenain, M. (2020). Sammba-MRI: A library for processing SmAll-MaMmal BrAin MRI data in Python. *Frontiers in neuroinformatics*, 14.

Chang, Y. S., Chiao, H. T., Abimannan, S., Huang, Y. P., Tsai, Y. T., Lin, K. M. (2020). An LSTM-based aggregated model for air pollution forecasting. *Atmospheric Pollution Research*, 11(8), 1451-1463.

Chen, P., Chen, G., Zhang, S. (2018). Log hyperbolic cosine loss improves variational auto-encoder, *International Conference on Learning Representations*, New Orleans, Louisiana, United States.

Chollet, F. (2018). Keras: The python deep learning library, *Astrophysics Source Code Library*, ascl-1806.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, *2*(4), 303-314.

Damavandi, H.R., Shah, R., Stampoulis, D., Wei, Y., Boscovic, D., Sabo, J (2019). Accurate Prediction of Streamflow Using Long Short-Term Memory Network, *International Journal of Environmental Science and Development*, 10(10), 294–300.

Dao, F. Y., Lv, H., Zhang, D., Zhang, Z. M., Liu, L., Lin, H. (2020). DeepYY1: a deep learning approach to identify YY1-mediated chromatin loops. *Briefings in Bioinformatics*.

Dao, J., Wang, J., Zhang, X. (2020). YNU-HPCC at SemEval-2020 Task 11: Lstm network for detection of propaganda techniques in news articles. *Cornell University Computer Science.*

Dedekayoğulları, H., Önal, A. (2009). The Importance of Water and Water Analysis in Terms of Environment-Human Health Relationship, *Journal of Istanbul Faculty of Medicine*, *72*(2), 65-70.

Ding, Y., Zhu, Y., Feng, J., Zhang, P., Cheng, Z. (2020). Interpretable spatio-temporal attention LSTM model for flood forecasting, *Neurocomputing*, *403*, 348-359.

Esmaeili, A., Marvasti, F. (2019). A novel approach to quantized matrix completion using huber loss measure. *IEEE Signal Processing Letters*, 26(2), 337-341.

Fang, Z., Wang, Y., Peng, L., Hong, H. (2020). Predicting flood susceptibility using long short-term memory (LSTM) neural network model, *Journal of Hydrology*, 125734.

Fang, Z., Wang, Y., Peng, L., Hong, H. (2020). Predicting flood susceptibility using long short-term memory (LSTM) neural network model, *Journal of Hydrology*, 125734.

Fathian, F., Mehdizadeh, S., Sales, A. K., Safari, M. J. S. (2019). Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models, *Journal of Hydrology*, *575*, 1200-1213.

Fitchett, J., Deford, M. (1973). A River Called Asi, *Saudi Aramco World Magazine*, Vol. 24, Number 3, 12-21.

Fu, M., Fan, T., Ding, Z. A., Salih, S. Q., Al-Ansari, N., Yaseen, Z. M. (2020). Deep learning data-intelligence model based on adjusted forecasting window scale: application in daily streamflow simulation, *IEEE Access*, 8, 32632-32651.

Gers, F.A., Schmidhuber, J., Cummins, F., (1999). Learning to forget: Continual prediction with LSTM, *9th International Conference on Artificial Neural Networks: ICANN '99*, 850–855.

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y. (2016). Deep learning (Vol 1, No. 2), Cambridge: *MIT press*, Healthc Inform Research.

Graves, A. (2013). Generating sequences with recurrent neural networks. Neural and Evolutionary Computing (cs.NE); Computation and Language (cs.CL),*Cornell University*, 1308.0850.

Hinton, G. E., Osindero, S., Teh, Y. W. (2006). A fast learning algorithm for deep belief nets, *Neural computation*, 18(7), 1527-1554.

Hinton, G. E., Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks, *Science*, 313(5786), 504-507.

Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. Neural Computation, *Massachusetts Institute of Technology*, 9 (8): 1735–1780.

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., Lou, Z. (2018). Deep learning with a long short-term memory networks approach for precipitation-flow simulation, *Water,* 10 (11), 1543.

Hu, Y., Yan, L., Hang, T. (2019). Stream-Flow Forecasting of Small Rivers Based on LSTM, *Proceedings of the 13th International Conference on Ubiquitous Information Management and Communication (IMCOM),* 1006-1014.

İnik, Ö., Ülker, E. (2017). Datasets and Software Libraries in Deep Learning. *1st International Symposium on Multidisciplinary Studies and Innovative Technologies Proceedings Book.*

Kao, I. F., Zhou, Y., Chang, L. C., Chang, F. J. (2020). Exploring a Long Short-Term Memory based Encoder-Decoder framework for multi-step-ahead flood forecasting, *Journal of Hydrology*, 124631.

Karimi, H. S., Natarajan, B., Ramsey, C. L., Henson, J., Tedder, J. L., Kemper, E. (2019). Comparison of learning-based wastewater flow prediction methodologies for smart sewer management, *Journal of Hydrology*, *577*, 123977.

Kingma, D. P., Ba, J. (2014). Adam: A method for stochastic optimization, *Cornell University Computer Science,* 1412.6980.

Kingman, D. P., Ba, J. (2015). Adam: A Method for Stochastic Optimization. Conference paper, *In 3rd International Conference for Learning Representations.*

Korkmaz, H., Çetin, B., Kuscu, V., Ege, I., Bom, A., Ozsahin, E., Karatas, A. (2012). Temporal changes of land use in Asi river delta (Hatay, Southern Turkey), *Journal of environmental biology*, *33*(2), 463.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M. (2018). Rainfall–runoff modelling using long short-term memory (LSTM) networks, *Hydrology and Earth System Sciences*, *22*(11), 6005-6022.

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, AK, Hochreiter, S., Nearing, GS (2019). Towards improved forecasts in non-empowered watersheds: harnessing the power of machine learning, *Water Resources Survey*, 55 (12), 11344-11354.

Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 25, (1097-1105).

Kuhn, M., Johnson, K. (2013). An introduction to feature selection. *Applied predictive modeling*, (487-519).

Le, X. H., Ho, H. V., Lee, G., Jung, S. (2019). Application of long short-term memory (LSTM) neural network for flood forecasting, *Water*, 11(7), 1387.

LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Springer nature*, *521*(7553), 436-444.

Lee, G., Jung, S., Lee, D. (2018). Comparison of physics-based and data-driven models for streamflow simulation of the Mekong river. *Journal of Korea Water Resources Association*, *51*(6), 503-514.

Li, T. L., Chan, A. B., Chun, A. (2010). Automatic musical pattern feature extraction using convolutional neural network, *Proc. Int. Conf. Data Mining and Applications*.

Li, W., Kiaghadi, A., Dawson, C. (2020). High temporal resolution rainfall–runoff modeling using long-short-term-memory (LSTM) networks, *Neural Computing and Applications*, 1-18.

Li, Z., Peng, F., Niu, B., Li, G., Wu, J., Miao, Z. (2018). Water quality prediction model combining sparse auto-encoder and LSTM network, *IFAC-PapersOnLine*, *51*(17), 831-836.

Lv, N., Liang, X., Chen, C., Zhou, Y., Li, J., Wei, H., Wang, H. (2020). Long Short Term Memory Cyclic Model with Mutual Information for Hydrology Prediction: A Case Study in the Xixian Basin, *Developments in Water Resources*, 103622.

Moor, J. (2006). The Dartmouth College artificial intelligence conference: The next fifty years. *Ai Magazine*, *27*(4), 87-87.

Mouatadid, S., Adamowski, J.F., Tiwari, M.K., Quilty, J.M. (2019). Combine maximum overlap discrete wavelet transform and long short term memory networks for irrigation flow estimation, *Agricultural Water Management*, *219*, 72-85.

Nacar, S., Kankal, M., Hınıs, M. A. (2018). Example of estimation-haldizen stream of daily stream flows with multivariate adaptive regression curves (MARC), *Gümüşhane University Journal of Science and Technology*, *8*(1), 38-47.

Nazimi, N., Ertugay, N. Monthly Flow Estimation of the Euphrates Basin Using FFNN, ANFIS and LSTM Models, *Erzincan Binali Yıldırım University Institute of Science*, Erzincan, Turkey, Master's Thesis, 1., 2020.

Nguyen, D. H., Bae, D. H. (2020). Correcting mean areal precipitation forecasts to improve urban flooding predictions by using long short-term memory network, *Journal of Hydrology*, *584*, 124710.

Ni, L., Wang, D., Singh, V. P., Wu, J., Wang, Y., Tao, Y., Zhang, J. (2020). Streamflow and rainfall forecasting by two long short-term memory-based models, *Journal of Hydrology*, *583*, 124296.

Özcan, T., Özcan, G., Erdoğan, H. (2012). Checklist of the freshwater decapods crustaceans from the Orontes River, *Arthropods*, 1(3), 118-120.

Pazarcı, H. (1996). The Role Of The Internatıonal Jurısdıctıon On The Bosnıa-Herzegovına Problem. *Ankara University SBF Journal*, 51 (01).

Pichon, F., Vallino, F., Albergoni, F., Albergoni, F.G. (1975). The Vegetation Cover Near Antioch, on the Orontes River, in the Classical Age (Forest and Plants Communities), *Arc – Bot – Biogeogr*, 51 (3): 123-133.

Rajaee, T., Khani, S., Ravansalar, M. (2020). Artificial intelligence-based single and hybrid models for prediction of water quality in rivers: A review, *Chemometrics and Intelligent Laboratory Systems*, 200, 103978.

Reading, C., Shaughnessy, J. M. (2004). Reasoning about variation. In *The challenge of developing statistical literacy, reasoning and thinking* (201-226), Springer, Dordrecht.

Sakarya, Ş., Yılmaz, Ü. (2019). Estimation of BIST30 Index Using Deep Learning Architecture, *European Journal of Educational and Social Sciences*, *4*(2), 106-121.

Sanner, M. F. (1999). Python: a programming language for software integration and development. *J Mol Graph Model*, *17*(1), 57-61.

Scherer, D., Müller, A., Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition, *In International conference on artificial neural networks* (92-101).

Scheumann, W., Sagsen, I., Tereci, E. (2011). Orontes River Basin: Downstream challenges and prospects for cooperation, In: Turkey's Water Policy. *Springer-Verlag Berlin Heidelberg*, 301-312.

Seyyarer E., Ayata F., Uçkan T. Karcı A. (2020). Application and Comparison of Optimization Algorithms Used in Deep Learning, *Anatolian Journal of Computer Sciences*, 5: 90-98.

Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, *28*.

Snoek, C. G., Worring, M., Smeulders, A. W. (2005). Early versus late fusion in semantic video analysis, *In Proceedings of the 13th annual ACM international conference on Multimedia* (399-402).

Şişmanoğlu, G., Koçer, F., Önde, M. A., Sahingoz, O. K. (2020). Price Prediction in the Stock Market with Deep Learning Methods, *Bitlis Eren University Journal of Science*, *9*(1), 434-445.

Tanrıverdi, T., Degirmenci, H., Sesveren, S. (2011). Assessment of irrigation schemes in Turkey based on management types. *African Journal of Biotechnology* Vol. 10(11), pp. 1997-2004.

Tektaş, M., Akbaş, A., Topuz, V. (2002). A review on the use of artificial intelligence techniques in traffic control. *Marmara University, Technical Sciences Vocational School*, 81040.

Thapa, D., Zebin, Z., Li, B., Lu, L., Fu, D., Shi, X., Tang, B., Qi, H. (2020). Snowmelt Driven Streamflow Prediction Using Machine Learning Techniques (LSTM, NARX, GPR, and SVR), *Water Journal*, 12(6), 1-18.

Tsang, L., Kracov, DA, Mulryne, J., Strom, L., Perkins, N., Dickinson, R., Jones, B. (2017). The impact of artificial intelligence on medical innovation in the European Union and the United States. *Intellect Prop Technol Law* J, 29 (8), 3-12.

Xiang, Z., Yan, J., Demir, I. (2020). A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water resources research*, *56*(1).

Yan, L., Feng, J., Hang, T. (2019). Small watershed stream-flow forecasting based on LSTM, *In International Conference on Ubiquitous Information Management and Communication* (1006-1014).

Yao, R., Li, B., Liu, J. (2009). A theoretical adaptive model of thermal comfort–Adaptive Predicted Mean Vote (aPMV), *Building and environment*, 44(10), 2089-2096.

Yaşar, M. (2004). Completing the gaps in daily flows by using artificial neural networks, Master's thesis, *Pamukkale University Institute of Science*.

Yıldıran, A. (2019). Estimation of river flow by artificial intelligence and trend analysis methods. Master's thesis, *Bilecik Şeyh Edebali University*, *Institute of Science*, Ref. No.: 10283422.

Yüksek, T. (2004). General evaluations regarding Turkey's water resources and basin planning. *Kafkas University Artvin Faculty of Forestry Journal*, 1-2 (71-83).

Zhang, J., Zhu, Y., Zhang, X., Ye, M., Yang, J. (2018). Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas, *Journal of hydrology*, *561*, 918-929.