Sigma J Eng Nat Sci, Vol. 42, No. 3, pp. 653−666, June, 2024

Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2024.00000

sigma Journal of Engineering and Natural Sciences Mühendislik ve Fen Bilimleri Dergisi

Research Article

Prediction of non-revenue water ratio in water distribution systems

Burak KIZILÖZ[1,*], Mehmet Emin BİRPINAR[2], Şükrü Ayhan GAZİOĞLU[3], Eyüp ŞİŞMAN[3]

[1]*Department of Environmental Protection and Control, Kocaeli Water and Sewerage Administration, Kocaeli, 41000, Türkiye*
[2]*Ministry of Environment and Urbanization, Ankara, 06570, Türkiye*
[3]*Department of Civil Engineering, Yıldız Technical University, İstanbul, 34220, Türkiye*

ARTICLE INFO

Article history
Received: 11 October 2022
Revised: 28 October 2022
Accepted: 12 April 2023

Keywords:
ANFIS; ANN; Modeling; Non Revenue Water; Water Distribution Systems

ABSTRACT

In the evaluations of water distribution systems (WDSs) in terms of water loss and performance, the Non-Revenue Water ratio (NRW) stands out as one of the most important parameters. Within the scope of this study, in order to predict the NRW ratio, a large number of models at different variable combinations were generated using the Adaptive Neuro-Fuzzy Inference System (ANFIS) and Artificial Neural Network (ANN) methods. The performance of the models formed has been evaluated by taking R^2, RMSE, MAE, SI, and Bias criteria as references. According to the study results, the model performances increase with the number of inputs in general, and the ANN models are more successful than ANFIS. Considering the modeling, the best-performing combination through the ANN method is WSQ-NJ-NL-NF, this one is the WSQ-NJ-NL-MPD combination in the ANFIS method which has three variables common. As a result, using variables common is significant for NRW predictions. On the other hand, NRW prediction performances need to improve by taking different variable combinations and methodological approaches into account, according to the ANFIS model results.

Cite this article as: Kızılöz B, Birpınar ME, Gazioğlu ŞA, Şişman E. Prediction of non-revenue water ratio in water distribution systems. Sigma J Eng Nat Sci 2024;42(3):653−666.

INTRODUCTION

Considering global warming and climate change, the management of water resources and especially water losses should be one of the most significant issues for the countries which are located in the Mediterranean basin, a weak basin in terms of water. The amount of water loss in the water distribution systems (WDSs) is a crucial issue to consider [1, 2]. The water losses and leakage activities are caused by many factors related to physical instances like water measurement errors, illegal water use, network age, and network pressure as well as the environmental reasons which depend on the topography and the state of the floor; the consumption habits and so on [3, 4]. To use water resources more efficiently and effectively, water utilities should calculate the Non-Revenue Water (NRW) quantities, a significant indicator for water leakages, and monitor them monthly by using water balance budget tables which were standardized by American Water Works Association (AWWA) and the International Water

***Corresponding author.**
*E-mail address: bkiziloz@isu.gov.tr

This paper was recommended for publication in revised form by Editor in Chief Ahmet Selim DAlkilic

Published by Yıldız Technical University Press, İstanbul, Turkey

Association (IWA). The high levels of the NRW ratio have negative impacts on the budgets and the future investment plans of the administrations.

Recently, studies on NRW have increased in the literature. Among these studies, there are many different methods, analyses, and methodologies to estimate and decrease the NRW, through calculations with the risk evaluation approaches, and define its components [5- 9]. The NRW ratio performances in different regions of Kocaeli have been evaluated through methodologies based on risk [10, 11].

[12] developed various models in order to predict the NRW ratio, using the parameters in the WDSs with the methods of artificial neural network (ANN) and the multiple regression analysis (MRA). When the obtained results of the used models are compared, it is seen that the ANN (R^2=0.6318) model has a higher prediction accuracy than the MRA (R^2=0.1906). Incheon (Republic of Korea) developed a model with the ANN method in order to predict the NRW ratio using some of the specific parameters that are influential on the leaks in the WDSs [13]. The models are developed by using 10, 20 and 30 neurons in the hidden layer, and it reached to its best model correlation as R^2=0.3973 with 20 neurons in the hidden layer. Various models have been developed to predict the NRW par [14] using ANN and ANFIS methods. According to one of the model results, the highest model accuracy is R^2=0.65 with ANN, and R^2=0.50 with ANFIS for the combination of WSQ/NL-ST/WM. In another study conducted on NRW ratio prediction, the ANN and Kriging methods have been preferred [15]. When the model results have been analyzed, the highest model performance has been obtained for the combinations with three variables, ST/NJ-NL-MAP (R^2=0.76) and NL-SCL-MPN (R^2=0.76). According to the Kriging model results, the combinations with two inputs, SIV-NJ/MPN (R^2=0.95) and SIV-NJ/MAP (R^2=0.94) have been quite successful. In another study on the NRW prediction conducted by [16] through ANN and Kriging methodologies, the model performance for two inputs of WSQ/SCL-MAP variables is R^2=0.397, and this one is R^2=0.89 for the Kriging model performance. [17] have used the Serial Triple Diagram Model (STDM) approach in another study and have obtained the model accuracy in the R^2=0.99 level for the NRW ratio model predictions.

In this study, the NRW ratios of WDS of Kocaeli's twelve regions are modeled through ANN and Adaptive Neuro-Fuzzy Inference System (ANFIS) approaches. Within the scope of this study, NRW predictions have been evaluated for the first time in our country with i) ANFIS models with three and four-input combinations, ii) ANN models with four-input combinations, iii) ANN models with NNF variables, and finally, iv) ANFIS models with four-input with WM variable.

Study Area and Data

Kocaeli is located between the 29° 22'-30° 21' east-west longitudes and the 40° 31'- 41° 13' north-south latitude. It is a metropolitan city with a total surface area of 3,623 km^2 (Figure 1).

Figure 1. Study area.

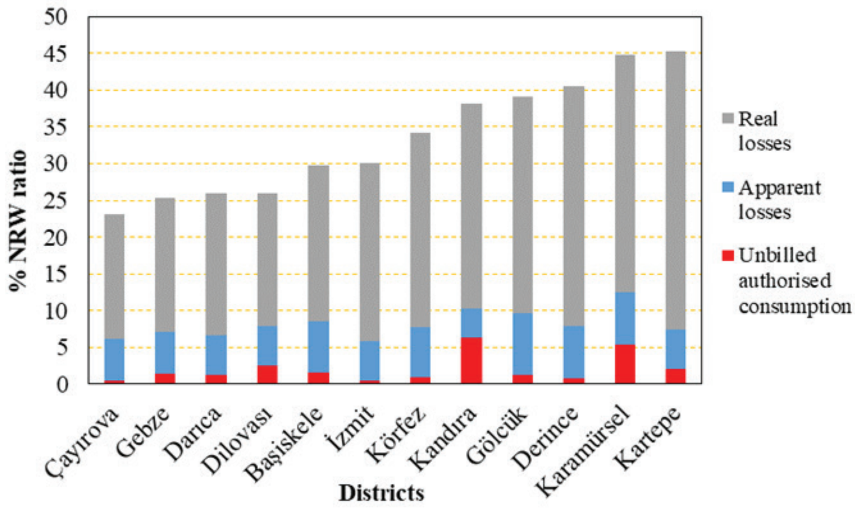

Figure 2. The NRW ratio changes of the regions.

It has 12 different regions in terms of water and sewerage administration. The water distribution system of the city has a total length of 8,936 km and the number of customers is 796,577 [15]. The total water consumption in the city is 163,627,918 m^3 for the year 2018 [18]. There are 195 drinking water storages, which are controlled via the SCADA system, 105 drinking water elevation centers and 11 potable water treatment facility in the Kocaeli [15]. NRW ratio was 28% at the end of 2022.

The physical losses are at a very high level of 24.79% and this is followed by the 6.03% of apparent losses and 1.49% of unbilled authorized consumption in Kocaeli. The 4.85% of authorized consumption errors component shows its impact on the NRW. In this study, the authorized consumption errors component, which has an impact on the NRW ratio, is included in the model study. The water balance components of 12 regions are also given separately in the Figure 2.

Figure 2 shows that Kartepe has the highest ratio of NRW, while Çayırova region has the lowest one. It is clear from the figure that the real loss component is much more efficient over the NRW ratio. Apparent loss component is another parameter that is influential over the NRW ratio.

MATERIALS AND METHODS

The NRW ratio is predicted for the 12 regions of Kocaeli by models using the ANFIS and the ANNs methods in this study. Models are developed with single input and single output (SISO), two inputs and single output (TISO), three inputs and single output (T3ISO), and finally, four inputs and single output (FISO).

As the input, the factors that represent the WDS are used such as the water supply quantity, the network length, the service connection length, the mean pipe diameter,

Table 1. Range of input-output parameters

Parameters	Abbreviation	Range	Unit
Water supply quantity [19]	WSQ	315,445-2,844,526	m^3
Domestic water storage tank [15, 19]	DWST	4,000-85,901	m^3
Number of failures [19]	NF	34-628	number
Failure Ratio [19]	FR	0.01–3.43	-
Number of service connection [15]	NSC	9,657–51,735	number
Service connection length [15]	SCL	130-527	km
Network length [15, 19]	NL	306-1600	km
Water meter [15, 19]	WM	15,124–160,135	number
Number of junctions [15, 19]	NJ	9,616–52,565	number
Mean pipe diameter [15, 19]	MPD	108-159	mm
Number of network failure	NNF	2-197	number
NRW ratio [19]	NRW	0.13–0.54	-

the number of service connection, the number of network failure, water meter, the number of junctions, the domestic water storage tank, the failure ratio, and the number of failures (Table 1).

In the study, the ANN NRW ratio models are developed as training data (55%), validation data (35%) and testing data (10%). The Levenberg Marquad (LM) back-propagation training algorithm is used. In all the models, sigmoid activation function is used in the hidden layer with a three-layer feed forward back propagation network (FFBP). The best model results are obtained when there are four neurons in the hidden layer.

Sugeno fuzzy logic (FL) System Inference principle is taken into account for the ANFIS models. In the models, grid partition method is used in order to generate the rules and categorize the input data. For all prediction models, the most appropriate MFs are selected. MFs of all models, linguistic terms such as "low", "medium" and "high" are considered. Hybrid learning algorithm has been used in order to get the best prediction in all models. In order to prevent over-learning, the number of epochs is selected as 20 in the models.

More detailed information about the ANN and ANIFS methodologies which are preferred frequently in the literature can be obtained from various resources [20, 21, 22].

The statistical operators below (correlation coefficient, R^2, root mean square error, RMSE, mean absolute error, MAE, and scatter index, SI, and bias) are used in order to compare the NRW ratio prediction results with actual results and also to evaluate the best model performances.

$$R^2 = \left[\frac{\sum_{i=1}^{n}(A_i - \bar{A})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n}(A_i - \bar{A})^2 \sum_{i=1}^{n}(P_i - \bar{P})^2}} \right]^2 \tag{1}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(A_i - P_i)^2}{n}} \tag{2}$$

$$MAE = \frac{\sum_{i=1}^{n}|A_i - P_i|}{n} \tag{3}$$

$$SI = \frac{RMSE}{\bar{A}} \tag{4}$$

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(A_i - P_i) \tag{5}$$

where, P_i is the predicted data, n is the total number of data points, and P_i is the actual data. Finally, $\bar{P}$ and $\bar{A}$ are the means of the prediction and actual data.

RESULTS AND DISCUSSION

ANN

In this study, of all the collected datasets, the 55% (80 data) is randomly divided as training datasets, the 35% (50 data) for validation datasets, and 10% (14 data) for testing datasets. ANN models' performance is given for each validation data set in Tables 2, 3, 4, and 5 for the prediction of the NRW ratio. The developed four-input models have generally higher model performance. In order to predict the NRW ratios a total of 146 ANN models are developed of which the 24 is single-input (SISO), 32 is two-input (TISO), 48 is three-input (T3ISO), and 42 is four-input (FISO) models.

When the SISO models in Table 2 are analyzed, it is seen that the components of WDSs, such as the domestic water storage tank, water supply quantity, water meter, number of junctions, number of service connection, mean pipe diameter, network length and number of failures are the variables that have an impact on the NRW ratio. The estimated NRW ratio scatter graph to the measured NRW ratio is also drawn in Figures 3 and 4. It is seen that the best model is the ANN-1.12.

Upon a close look at the 36 TISO models in Table 3, it is seen that the domestic water storage tank, number of junctions in water distribution networks and network length are variables that have an influence on the NRW ratio. The estimated NRW ratio scatter graph corresponding to the measured NRW ratio is drawn in Figures 5 and 6. It is seen that the best model is the ANN-2.27 Additionally, the number of inputs increases in the models, the performance of the models improves as well.

The T3ISO models in the Table 4 show that water supply quantity, number of junctions in water distribution networks, network length, water meter and number of failures are the effective on the NRW ratio. When the estimated NRW ratio scatter graph corresponding to the actual NRW ratio is drawn in Figures 7 and 8, one can see that the best two models are ANN-3.34 and ANN-3.37.

The FISO models in Table 5 show that water supply quantity, number of junctions in water distribution system, network length, mean pipe diameter and number of failures are the influential ones. The estimated NRW ratio scatter graph corresponding to the actual NRW ratio is drawn in Figures 9 and 10. Additionally, it is seen that the best two models are ANN-4.29 and ANN-4.34.

The WDS consist of different uncertain components and each component is effective on the NRW ratio. Several 146 ANN models are developed with different combinations in the study and the results show that the all component has

Table 2. ANN models performance for SISO

Model	Combinations	R2	RMSE	MAE	SI	Bias
ANN-1.1	WSQ/NSC	0.170	0.081	0.065	%24.13	-0.0001
ANN-1.2	WSQ/SCL	0.104	0.084	0.069	%25.09	0.0034
ANN-1.3	WSQ/NL	0.300	0.074	0.060	%22.28	0.0043
ANN-1.4	WSQ/DWST	0.129	0.083	0.069	%24.81	0.0076
ANN-1.5	WSQ/WM	0.128	0.083	0.069	%24.79	-0.0051
ANN-1.6	WSQ/NJ	0.061	0.085	0.071	%25.67	-0.0026
ANN-1.7	WSQ/MPD	0.168	0.081	0.065	%24.21	-0.0044
ANN-1.8	WSQ/NF	0.133	0.082	0.068	%24.68	-0.0016
ANN-1.9	DWST/NSC	0.524	0.061	0.047	%18.33	0.0002
ANN-1.10	DWST/SCL	0.621	0.054	0.043	%16.32	0.0025
ANN-1.11	DWST/NL	0.193	0.079	0.064	%23.82	0.0012
ANN-1.12	DWST/WM	0.662	0.051	0.041	%15.42	-0.0027
ANN-1.13	DWST/NJ	0.534	0.060	0.047	%18.13	0.0039
ANN-1.14	DWST/MPD	0.476	0.065	0.050	%19.42	0.0042
ANN.1.15	DWST/NF	0.035	0.087	0.073	%26.06	0.0052
ANN-1.16	NL/WM	0.556	0.059	0.047	%17.67	0.0020
ANN-1.17	NL/SCL	0.387	0.070	0.057	%20.96	-0.0060
ANN-1.18	NL/NSC	0.307	0.073	0.057	%22.04	-0.0006
ANN-1.19	MPD [15]	0.520	0.062	0.048	%18.53	0.0062
ANN-1.20	NL [15]	0.659	0.051	0.041	%15.46	0.0004
ANN-1.21	NNF	0.059	0.086	0.072	%25.79	-0.0069
ANN-1.22	NSC [15]	0.447	0.066	0.051	%19.80	0.0050
ANN-1.23	FR	0.114	0.083	0.069	%24.94	-0.0028
ANN-1.24	NJ	0.377	0.070	0.057	%20.94	0.0038
ANN-1.25	WM	0.274	0.075	0.060	%22.64	-0.0012
ANN-1.26	WSQ	0.263	0.076	0.063	%22.80	0.0019
ANN-1.27	NF	0.076	0.085	0.070	%25.48	-0.0020
ANN-1.28	DWST [15]	0.554	0.059	0.046	%17.74	0.0017

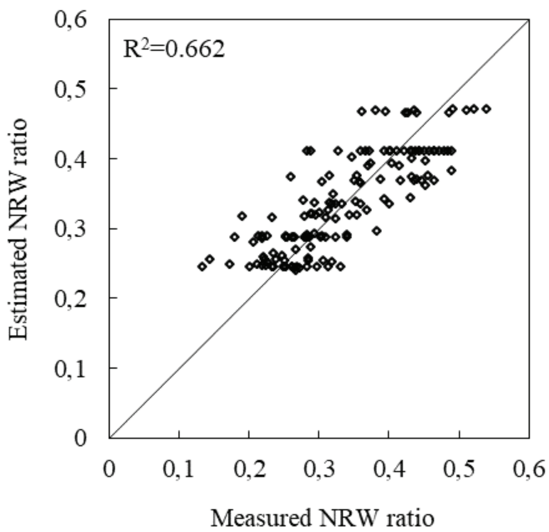

Figure 3. Scatter graph for DWST/WM model

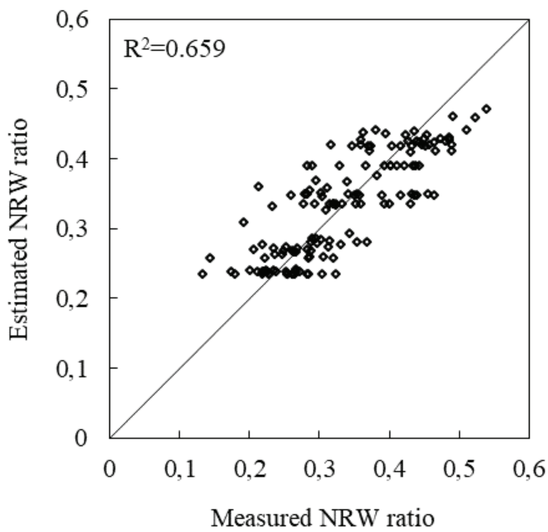

Figure 4. Scatter graph for NL model [15]

Table 3. ANN models performance for TISO

Model	Combinations	R^2	RMSE	MAE	SI	Bias
ANN-2.1	WSQ/NL–DWST/NSC	0.476	0.064	0.051	%19.19	-0.0030
ANN-2.2	WSQ/NL–DWST/SCL	0.684	0.049	0.039	%14.88	0.0003
ANN-2.3	WSQ/NL–DWST/WM	0.704	0.048	0.039	%14.57	0.0051
ANN-2.4	WSQ/NL–DWST/NJ	0.514	0.061	0.047	%18.48	0.0020
ANN-2.5	WSQ/NL–DWST/MPD	0.579	0.057	0.046	%17.23	-0.0036
ANN-2.6	WSQ/NL–NL/WM	0.550	0.059	0.046	%17.78	0.0027
ANN-2.7	WSQ/NL– NL/SCL	0.533	0.060	0.046	%18.13	0.0014
ANN-2.8	WSQ/NL– NL/NSC	0.514	0.063	0.051	%19.02	0.0135
ANN-2.9	WSQ/NL–MPD	0.640	0.053	0.043	%15.88	0.0014
ANN-2.10	WSQ/NL–NNF	0.443	0.066	0.054	%19.82	-0.0045
ANN-2.11	WSQ/NL–FR	0.458	0.065	0.052	%19.59	0.0065
ANN-2.12	WSQ/NL–WM [15]	0.666	0.051	0.041	%15.45	-0.0073
ANN-2.13	WSQ/NJ–WSQ/WM	0.181	0.080	0.065	%23.97	0.0009
ANN-2.14	WSQ/NJ–WSQ/MPD	0.282	0.075	0.059	%22.44	0.0002
ANN-2.15	WSQ/NJ–WSQ/NF	0.338	0.072	0.059	%21.70	-0.0073
ANN-2.16	WSQ/NJ–DWST/WM	0.673	0.050	0.040	%15.14	-0.0017
ANN-2.17	WSQ/NJ–DWST/MPD	0.422	0.067	0.052	%20.14	-0.0012
ANN-2.18	WSQ/NJ–DWST/NF	0.056	0.086	0.071	%25.75	0.0034
ANN-2.19	WSQ/NJ–MPD [15]	0.544	0.060	0.048	%18.15	0.0060
ANN-2.20	WSQ/NJ–FR	0.177	0.080	0.066	%24.18	0.0019
ANN-2.21	WSQ/NJ-NL	0.702	0.048	0.038	%14.48	0.0026
ANN-2.22	DWST/NJ–DWST/MPD	0.691	0.049	0.039	%14.72	0.0004
ANN-2.23	DWST/NJ–MPD	0.660	0.052	0.041	%15.56	0.0027
ANN-2.24	DWST/NJ–NL [15]	0.728	0.046	0.037	%13.81	-0.0018
ANN-2.25	DWST/NJ–FR [15]	0.633	0.054	0.042	%16.19	-0.0039
ANN-2.26	NL–MPD	0.683	0.050	0.040	%15.02	0.0015
ANN-2.27	NL–NJ	0.735	0.046	0.037	%13.73	0.0038
ANN-2.28	NL–FR	0.533	0.060	0.049	%18.13	0.048
ANN-2.29	NJ–MPD	0.694	0.049	0.040	%14.78	-0.0050
ANN-2.30	NJ-NSC	0.670	0.051	0.039	%15.22	-0.0001
ANN-2.31	WSQ-MPD	0.602	0.056	0.046	%16.77	-0.0040
ANN-2.32	WSQ-NJ	0.661	0.051	0.041	%15.45	-0.0010
ANN-2.33	WSQ-NL	0.687	0.049	0.039	%14.91	-0.0053
ANN-2.34	WSQ-WM	0.612	0.056	0.044	%16.76	0.0076
ANN-2.35	WSQ-NF	0.361	0.070	0.056	%21.18	0.0030
ANN-2.36	WSQ-FR	0.489	0.063	0.051	%19.00	0.0013

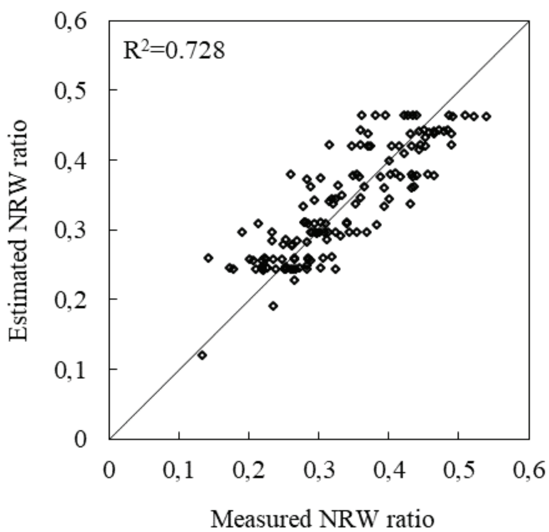

Figure 5. Scatter graph for DWST/NJ-NL model [15].

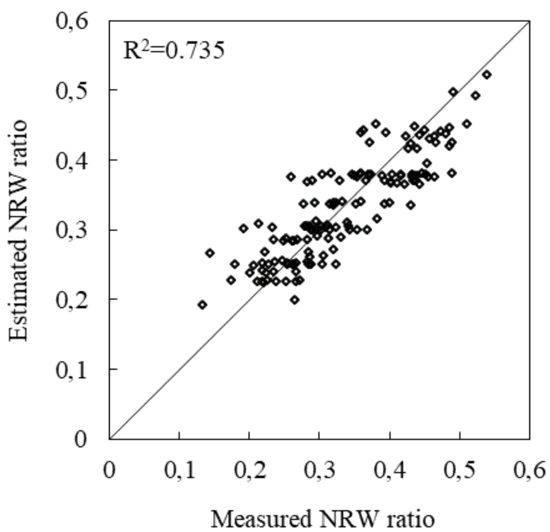

Figure 6. Scatter graph for NL-NJ model.

Table 4. ANN models performance for T3ISO

Model	Combinations	R^2	RMSE	MAE	SI	Bias
ANN-3.1	WSQ/NL–DWST/NSC-DWST/WM	0.690	0.050	0.040	%15.21	-0.0123
ANN-3.2	WSQ/NL–DWST/NS -DWST/NJ	0.646	0.052	0.042	%15.79	-0.0039
ANN-3.3	WSQ/NL–DWST/NSC-DWST/MPD	0.708	0.048	0.037	%14.33	0.0024
ANN-3.4	WSQ/NL–DWST/NSC-NL/WM	0.688	0.049	0.039	%14.86	0.0024
ANN-3.5	WSQ/NL–DWST/NSC-WM	0.723	0.046	0.036	%14.02	-0.0020
ANN-3.6	WSQ/NL–DWST/NSC-MPD	0.709	0.048	0.038	%14.48	0.0037
ANN-3.7	WSQ/NL–DWST/NSC-FR	0.558	0.059	0.045	%17.64	-0.0003
ANN-3.8	WSQ/NL–DWST/SCL-DWST/WM	0.708	0.048	0.038	%14.34	0.0035
ANN-3.9	WSQ/NL–DWST/SCL-DWST/NJ	0.634	0.053	0.042	%16.08	0.0008
ANN-3.10	WSQ/NL–DWST/SCL-DWST/MPD	0.699	0.048	0.039	%14.56	0.0036
ANN-3.11	WSQ/NL–DWST/SCL-MPD	0.694	0.049	0.039	%14.85	0.0046
ANN-3.12	WSQ/NL–DWST/SCL-NF	0.575	0.058	0.046	%17.55	-0.0074
ANN-3.13	WSQ/NL–DWST/SCL-WM	0.731	0.046	0.035	%13.81	0.0017
ANN-3.14	WSQ/NL–DWST/SCL-FR	0.705	0.048	0.038	%14.52	0.0071
ANN-3.15	WSQ/NL–DWST/WM-MPD	0.721	0.046	0.037	%13.99	0.0023
ANN-3.16	WSQ/NL–DWST/WM-NF	0.678	0.050	0.039	%15.15	0.0009
ANN-3.17	WSQ/NL–DWST/WM-FR	0.687	0.049	0.038	%14.91	0.0058
ANN-3.18	WSQ/NL–DWST/NJ-MPD	0.679	0.050	0.040	%15.04	0.0029
ANN-3.19	WSQ/NL–DWST/NJ-WM	0.696	0.049	0.039	%14.86	-0.0089
ANN-3.20	WSQ/NL–MPD-NF	0.554	0.059	0.046	%17.76	-0.0032
ANN-3.21	WSQ/NL–MPD-FR	0.629	0.054	0.042	%16.13	-0.0005
ANN-3.22	WSQ/NL–MPD-WM	0.666	0.051	0.040	%15.39	-0.0047
ANN-3.23	WSQ/NL–WM-FR	0.575	0.057	0.046	%17.27	0.0018
ANN-3.24	WSQ/NL–WM-NF	0.528	0.061	0.048	%18.28	0.0054
ANN-3.25	WSQ/NL–WSQ/WM-WSQ/MPD	0.646	0.053	0.042	%15.86	0.0061
ANN-3.26	WSQ/NL–WSQ/WM-WSQ/NF	0.673	0.050	0.040	%15.21	0.0049
ANN-3.27	WSQ/NJ–WSQ/WM-WSQ/MPD	0.700	0.048	0.038	%14.49	0.0089
ANN-3.28	WSQ/NJ–WSQ/WM-WSQ/NF	0.325	0.072	0.059	%21.75	0.0014
ANN-3.29	WSQ/NJ–DWST/WM-MPD	0.720	0.047	0.037	%14.31	0.0031
ANN-3.30	WSQ/NJ–DWST/MPD-WM	0.694	0.049	0.038	%14.66	0.0021
ANN-3.31	WSQ/NJ–MPD-WM	0.693	0.049	0.039	%14.77	0.0058
ANN-3.32	WSQ/NJ–MPD-FR	0.603	0.056	0.044	%16.89	0.0066
ANN-3.33	WSQ/NJ-NL-MPD	0.701	0.048	0.038	%14.52	-0.0040
ANN-3.34	WSQ/NJ-NL-WM	0.746	0.045	0.036	%13.59	0.0070
ANN-3.35	WSQ/NJ-NL-NF	0.724	0.047	0.037	%14.15	0.0024
ANN-3.36	WSQ-NJ-MPD	0.706	0.048	0.037	%14.39	-0.0001
ANN-3.37	WSQ-NJ-NL	0.737	0.046	0.036	%13.84	0.0011
ANN-3.38	WSQ-NJ-WM	0.714	0.047	0.036	%14.15	0.0009
ANN-3.39	WSQ-NJ-NF	0.515	0.062	0.049	%18.56	0.0008
ANN-3.40	WSQ-NL-MPD	0.709	0.047	0.039	%14.30	0.0022
ANN-3.41	WSQ-NL-WM	0.735	0.046	0.036	%13.73	0.0054
ANN-3.42	WSQ-NL-DWST	0.693	0.049	0.041	%14.74	0.0006
ANN-3.43	WSQ-NL-NF [19]	0.720	0.046	0.034	%13.02	-0.0024
ANN-3.44	WSQ –MPD-WM	0.693	0.049	0.038	%14.70	-0.0014
ANN-3.45	WSQ-MPD-NF [19]	0.650	0.052	0.040	%14.53	-0.0014
ANN-3.46	WSQ–NSC-WM	0.707	0.048	0.037	%14.36	0.0002
ANN-3.47	WSQ–NSC-MPD	0.708	0.048	0.039	%14.42	0.0061
ANN-3.48	WSQ–NSC–NF	0.708	0.048	0.036	%14.34	-0.0031
ANN-3.49	WSQ–WM-MPD	0.699	0.048	0.038	%14.59	0.0038
ANN-3.50	WSQ–WM-NF	0.682	0.050	0.039	%15.11	0.0071

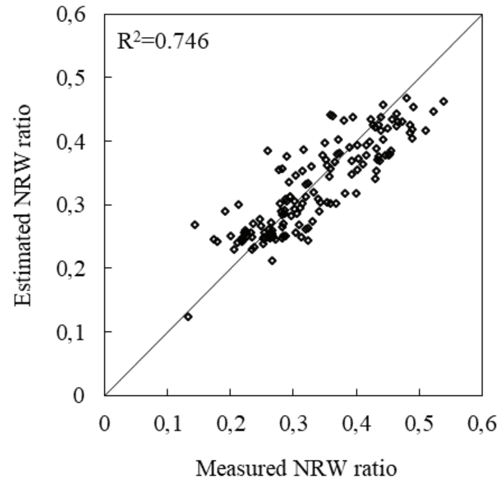

Figure 7. Scatter graph for WSQ/NJ-NL-WM model.

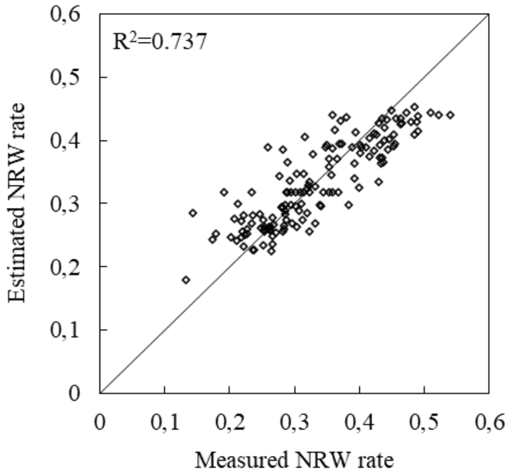

Figure 8. Scatter graph for WSQ-NL-NF model.

Table 5. ANN models performance for FISO

Model	Combinations	R^2	RMSE	MAE	SI	Bias
ANN-4.1	WSQ/NL–DWST/NSC-DWST/WM-DWST/NJ	0.676	0.050	0.040	%15.08	0.0003
ANN-4.2	WSQ/NL–DWST/NSC-DWST/MPD-DWST/NJ	0.730	0.046	0.035	%13.75	-0.0005
ANN-4.3	WSQ/NL–DWST/NSC-DWST/MPD-DWST/WM	0.721	0.046	0.038	%14.02	0.0040
ANN-4.4	WSQ/NL–DWST/NSC-NL/WM-MPD	0.720	0.047	0.036	%14.04	-0.0022
ANN-4.5	WSQ/NL–DWST/NSC-WM-MPD	0.717	0.047	0.038	%14.13	0.0033
ANN-4.6	WSQ/NL–DWST/NSC-MPD-FR	0.732	0.046	0.036	%13.77	0.0041
ANN-4.7	WSQ/NL–DWST/NSC-WM-FR	0.709	0.048	0.037	%14.33	-0.0035
ANN-4.8	WSQ/NL–DWST/SCL-DWST/WM-MPD	0.720	0.047	0.037	%14.04	0.0033
ANN-4.9	WSQ/NL–DWST/SCL-DWST/NJ-WM	0.726	0.046	0.037	%13.90	0.0042
ANN-4.10	WSQ/NL–DWST/SCL-DWST/MPD-WM	0.709	0.047	0.037	%14.29	0.0016
ANN-4.11	WSQ/NL –DWST/SCL-WM-MPD	0.717	0.047	0.037	%14.14	0.0037
ANN-4.12	WSQ/NL–DWST/SCL-WM-FR	0.654	0.052	0.040	%15.59	0.0025
ANN-4.13	WSQ/NL–DWST/SCL-MPD-FR	0.711	0.047	0.037	%14.32	-0.0053
ANN-4.14	WSQ/NL–DWST/WM-MPD-FR	0.723	0.047	0.037	%14.11	-0.0014
ANN-4.15	WSQ/NL–DWST/NJ MPD-WM	0.710	0.048	0.037	%14.42	0.0072
ANN-4.16	WSQ/NL–DWST/NJ-MPD-FR	0.669	0.051	0.039	%15.34	0.0045
ANN-4.17	WSQ/NL–MPD-WM-FR	0.684	0.049	0.039	%14.92	-0.0034
ANN-4.18	WSQ/NL–WSQ/WM-WSQ/MPD-WSQ/NF	0.736	0.045	0.035	%13.70	-0.0002
ANN-4.19	WSQ/NJ–WSQ/WM -WSQ/MPD-WSQ/NF	0.708	0.048	0.036	%14.35	-0.0035
ANN-4.20	WSQ/NJ–DWST WM-MPD-FR	0.708	0.048	0.037	%14.39	0.0056
ANN-4.21	WSQ/NJ–DWST/MPD-WM-FR	0.683	0.049	0.039	%14.91	0.0008
ANN-4.22	WSQ/NJ–MPD -WM-FR	0.719	0.047	0.035	%14.06	-0.0018
ANN-4.23	WSQ/NJ -NL-MPD-WM	0.740	0.045	0.035	%13.55	0.0026
ANN-4.24	WSQ/NJ -NL-WM-FR	0.715	0.048	0.037	%14.43	-0.0095
ANN-4.25	WSQ/NJ -NL-MPD-FR	0.712	0.047	0.036	%14.27	0.0022
ANN-4.26	WSQ/NJ-MPD-WM	0.716	0.047	0.038	%14.33	-0.0083
ANN-4.27	WSQ-NJ- NL-WM	0.751	0.044	0.034	%13.28	0.0027
ANN-4.28	WSQ-NJ-NL-MPD	0.749	0.044	0.034	%13.28	0.0012
ANN-4.29	WSQ-NJ-NL-NF	0.794	0.040	0.031	%12.02	0.0013
ANN-4.30	WSQ/NJ -NL-FR	0.754	0.044	0.034	%13.16	-0.0022
ANN-4.31	WSQ/NJ-WM-NF	0.761	0.043	0.033	%12.97	-0.0005
ANN-4.32	WSQ/NJ-MPD-NF	0.744	0.045	0.034	%13.48	0.0033
ANN-4.33	WSQ-NL-MPD-WM	0.744	0.045	0.035	%13.49	0.0009
ANN-4.34	WSQ-NL-MPD-NF	0.794	0.040	0.031	%12.07	-0.0041
ANN-4.35	WSQ-NL-WM-NF	0.774	0.042	0.033	%12.61	-0.0027
ANN-4.36	WSQ-NL-DWST-WM	0.717	0.047	0.037	%14.14	-0.0042
ANN-4.37	WSQ-NL-DWST-MPD	0.715	0.047	0.038	%14.22	0.0026
ANN-4.38	WSQ-MPD-WM-DWST	0.724	0.046	0.036	%14.02	0.0057
ANN-4.39	WSQ-MPD-WM-NF	0.750	0.044	0.035	%13.29	0.0029
ANN-4.40	WSQ–NSC-WM-MPD	0.707	0.048	0.038	%14.34	0.0017
ANN-4.41	WSQ–NSC-MPD-NF	0.767	0.043	0.034	%12.88	0.0050
ANN-4.42	WSQ–NSC–WM-NF	0.735	0.046	0.036	%13.86	-0.0037

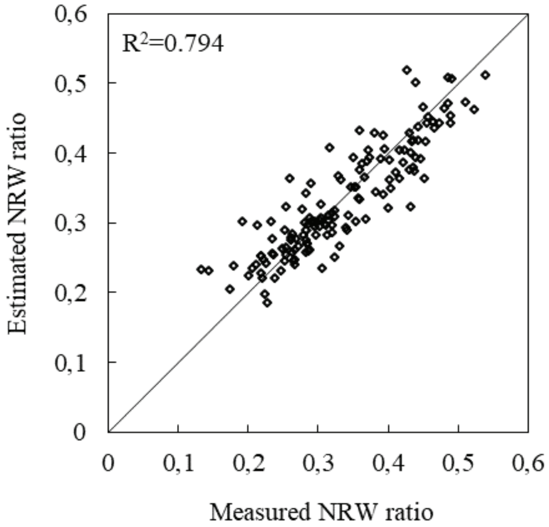

Figure 9. Scatter graph for WSQ-NJ-NL-NF model.

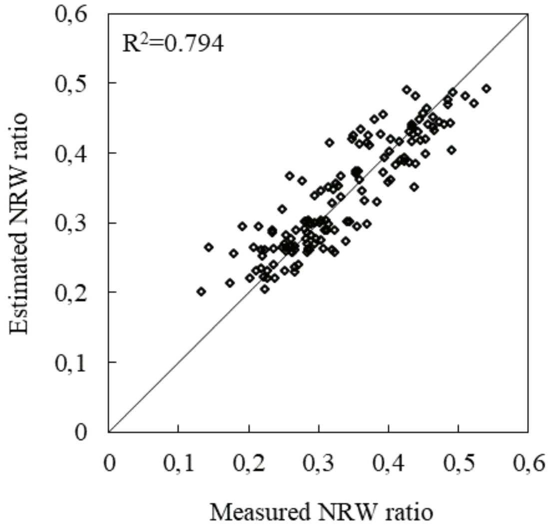

Figure 10. Scatter graph for WSQ-NL-MPD-NF model

an impact on the NRW. Each component in the water distribution networks is evaluated by separate modeling, and its impact on the NRW ratio is studied. With the developed plenty of ANN models, the NRW ratios can be predicted at the partly acceptable levels.

ANFIS

In order to predict the NRW ratio, ANFIS method is also used in this investigation and the results compare with the ANN model. The data used in the ANN method are also preferred for the ANFIS. Here, 66% of the total data is separated for the training, while the 34% testing for the validation. Starting from one input, and increasing the number of inputs one by one; the models are developed with the ANFIS until the four-input (see Tables 6, 7, 8, 9). In this part of the research, the most meaningful models are selected among all the ANN models, and total 57 ANFIS model including 8 single-input, 15 two-input, 23 three-input and 11 four-input have been developed.

The SISO ANFIS models and their performance in Table 6 has been presented. The estimated NRW ratio scatter graph corresponding to the actual NRW ratio is drawn in Figures 11 and 12. It is seen that the best model is ANFIS-1.3. ANN models with single input are better than ANFIS model performances.

The TISO ANFIS models in Table 7 show that the WSQ, NL, DWST, WM, NJ and FR are the variables, which have an impact on the NRW ratio. When the estimated NRW ratio scatter graph corresponding to the actual NRW ratio is drawn in Figures 13 and 14, one can see that the best model is ANFIS-2.8. It is seen that as the number of input increases as the correlation coefficient increases, while the scatterings decrease. The ANN method is better than the ANFIS method for model performances with two-input model combinations.

Table 6. ANFIS models performance for SISO

Model	Combinations	R^2	RMSE	MAE	SI	Bias
ANFIS-1.1	WSQ/ NL [14]	0.329	0.087	0.073	%26.40	-0.0062
ANFIS-1.2	DWST / SCL	0.016	0.102	0.087	%30.89	-0.0029
ANFIS-1.3	DWST/ WM [14]	0.386	0.081	0.067	%24.45	-0.0031
ANFIS-1.4	DWST / NJ	0.212	0.091	0.075	%27.61	-0.0031
ANFIS-1.5	NL / WM	0.150	0.096	0.080	%29.07	-0.0036
ANFIS-1.6	MPD	0.367	0.083	0.070	%25.22	-0.0029
ANFIS-1.7	NL	0.057	0.100	0.085	%30.25	-0.0035
ANFIS-1.8	NSC	0.262	0.092	0.076	%27.76	-0.0031
ANFIS-1.9	NJ	0.104	0.099	0.084	%29.82	-0.0029
ANFIS-1.10	DWST	0.029	0.102	0.086	%30.77	-0.0030

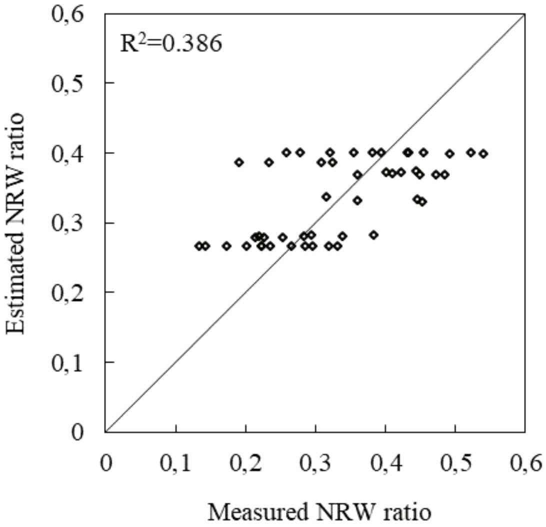

Figure 11. Scatter graph for DWST/WM model [14]

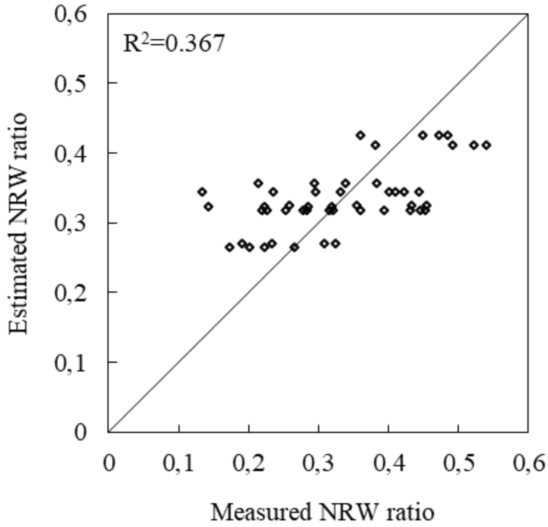

Figure 12. Scatter graph for MPD model

The T3ISO ANFIS models in Table 8 show that WSQ, NL, DWST, WM, NJ, SCL and FR are the variables, which have an impact on the NRW ratio. When the estimated NRW ratio scatter graphs corresponding to the actual NRW ratio are drawn in Figures 15 and 16; the best two models are ANFIS-3.2 and ANFIS-3.7.

When the ANFIS model performances with TISO given in Table 7 are compared with the model performances with T3ISO given in Table 8, it is seen that there is no significant improvement.

In conclusion, when the ANFIS models with FISO are analyzed, the FISO ANFIS model in Table 9 indicates that the WSQ, NL, WM, NJ, and MPD are the variables that are most effective on the NRW ratio. When the estimated NRW ratio scatter graphs corresponding to the actual NRW ratios are drawn in Figures 17 and 18; them the best two models are ANFIS-4.2 and the ANFIS-4.7. As the number of inputs in the models increases, the correlation coefficient little increases, while the scatterings decrease. The best performance among the ANFIS models has been obtained for the model combination with four-input (WSQ-NJ-NL-MPD).

Out of all several ANN and ANFIS models that have been developed within the study ANN-4.29 and ANFIS-4.2 have the highest correlation coefficient and the smallest root mean square errors for the ANN and ANFIS methodologies

Table 7. ANFIS models performance for TISO

Model	Combinations	R2	RMSE	MAE	SI	Bias
ANFIS-2.1	WSQ/NL–DWST/SCL	0.496	0.076	0.066	%23.07	-0.0077
ANFIS-2.2	WSQ/NL–DWST/WM [14]	0.500	0.076	0.062	%22.86	-0.0049
ANFIS-2.3	WSQ/NL–MPD	0.463	0.077	0.066	%23.21	-0.0040
ANFIS-2.4	WSQ/NL–WM	0.442	0.080	0.066	%24.29	-0.0045
ANFIS-2.5	WSQ/NJ–DWST/WM	0.276	0.088	0.071	%26.45	0.0031
ANFIS-2.6	DWST/NJ–DWST/MPD	0.281	0.089	0.075	%26.98	0.0121
ANFIS-2.7	DWST/NJ–NL	0.425	0.079	0.066	%23.77	-0.0112
ANFIS-2.8	DWST/NJ–FR	0.586	0.066	0.055	%20.08	-0.0044
ANFIS-2.9	NL–MPD	0.480	0.075	0.062	%22.81	0.0068
ANFIS-2.10	NL–NJ	0.378	0.081	0.066	%24.53	0.0044
ANFIS-2.11	NJ–MPD	0.463	0.076	0.063	%23.05	-0.0025
ANFIS-2.12	NJ-NSC	0.448	0.077	0.065	%23.14	-0.0018
ANFIS-2.13	WSQ-MPD	0.447	0.078	0.064	%23.59	-0.0045
ANFIS-2.14	WSQ-NJ	0.231	0.091	0.079	%27.64	-0.0021
ANFIS-2.15	WSQ–NL	0.285	0.089	0.077	%26.76	-0.0128
ANFIS-2.16	WSQ-WM	0.175	0.095	0.079	%28.57	-0.0040

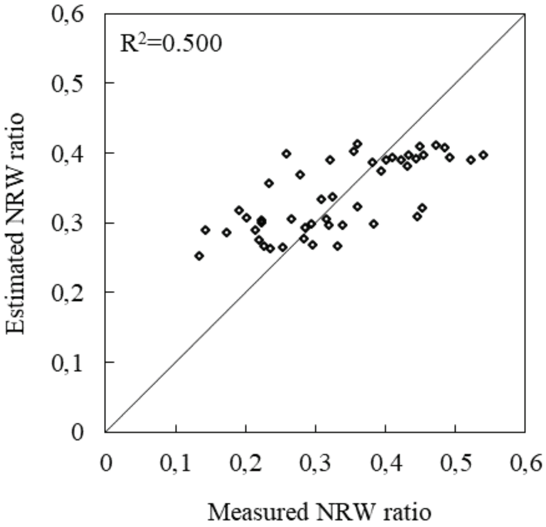

Figure 13. Scatter graph for WSQ/NL-DQST/WM model [14].

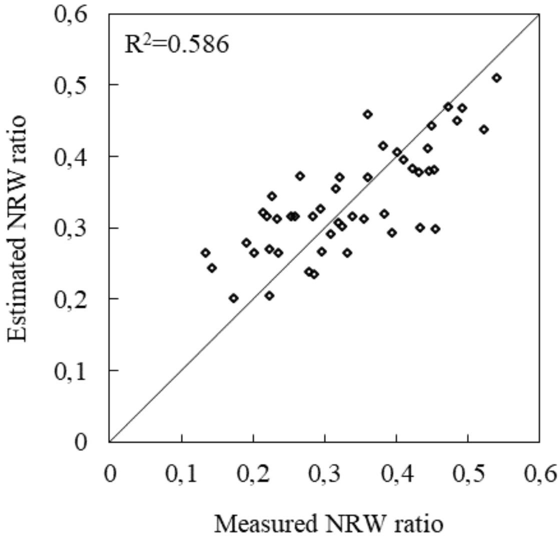

Figure 14. Scatter graph for DWST/NJ-FR model.

Table 8. ANFIS models performance for T3ISO

Model	Combinations	R^2	RMSE	MAE	SI	Bias
ANFIS-3.1	WSQ/NL–DWST/NSC-DWST/MPD	0.078	0.143	0.109	%43.17	-0.0311
ANFIS-3.2	WSQ/NL–DWST/NSC-WM	0.522	0.072	0.062	%21.86	-0.0110
ANFIS-3.3	WSQ/NL–DWST/NSC-MPD	0.448	0.077	0.067	%23.33	-0.0019
ANFIS-3.4	WSQ/NL–DWST/SCL-DWST/WM	0.120	0.118	0.080	%35.71	0.0149
ANFIS-3.5	WSQ/NL–DWST/SC -DWST/MPD	0.358	0.084	0.069	%25.51	-0.0177
ANFIS-3.6	WSQ/NL–DWST/SCL-WM	0.437	0.077	0.066	%23.43	-0.0013
ANFIS-3.7	WSQ/NL–DWST/SCL-FR	0.560	0.070	0.056	%21.31	-0.0168
ANFIS-3.8	WSQ/N –DWST/WM-MPD	0.423	0.078	0.061	%23.59	0.0037
ANFIS-3.9	WSQ/NL–DWST/JUC-WM	0.457	0.077	0.062	%23.33	-0.0137
ANFIS-3.10	WSQ/NJ–WSQ/WM-SIV/MPD	0.384	0.081	0.067	%24.37	-0.0008
ANFIS-3.11	WSQ/NJ–DWST/WM-MPD	0.301	0.089	0.068	%26.81	0.0083
ANFIS-3.12	WSQ/NJ-NL-MPD	0.142	0.098	0.070	%29.69	0.0023
ANFIS-3.13	WSQ/NJ-MPD-WM	0.366	0.084	0.071	%25.30	-0.0096
ANFIS-3.14	WSQ/NJ-MPD-NF	0.042	0.106	0.083	%32.05	-0.0115
ANFIS-3.15	WSQ-NJ-MPD	0.477	0.075	0.060	%22.56	-0.0038
ANFIS-3.16	WSQ-NJ-MPD	0.107	0.108	0.083	%32.72	-0.0098
ANFIS-3.17	WSQ-NJ-WM	0.411	0.081	0.069	%24.48	-0.0090
ANFIS-3.18	WSQ-MPD-MPD	0.113	0.103	0.077	%30.95	-0.0106
ANFIS-3.19	WSQ-MPD-WM	0.382	0.082	0.070	%24.83	-0.0071
ANFIS-3.20	WSQ–NSC-WM	0.379	0.083	0.071	%25.02	-0.0071
ANFIS-3.21	WSQ–NSC-MPD	0.504	0.073	0.059	%21.99	-0.0030
ANFIS-3.22	WSQ–NSC–NF	0.240	0.094	0.074	%28.32	-0.0179
ANFIS-3.23	WSQ–WM-MPD	0.359	0.082	0.065	%24.90	-0.0002

respectively. These models are evaluated together in Figure 19. It should be noted that there is an amount of difference in the method among the model scatterings in which model results are compatible with each other.

The results of this study are also compared with the earlier studies in Table 10. It is seen in Table 10 that the ANN, Kriging, STDM, and ANFIS methodologies have been used in the literature to predict NRW ratios. As a result, the best ANN and ANFIS models have been determined by forming NRW ratio prediction models over many model input variables combinations.

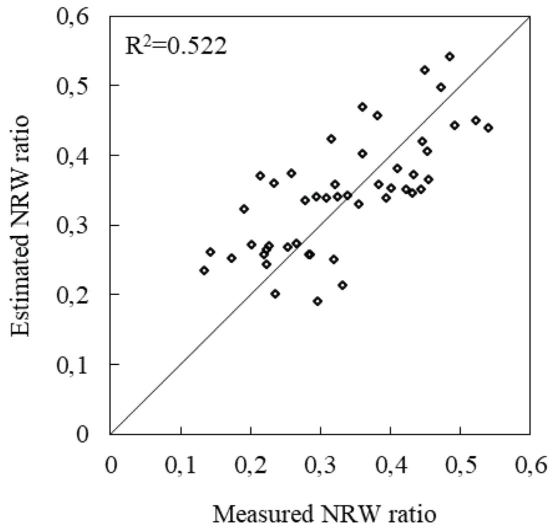

Figure 15. Scatter graph for WSQ/NL-DWST/NSC-WM model

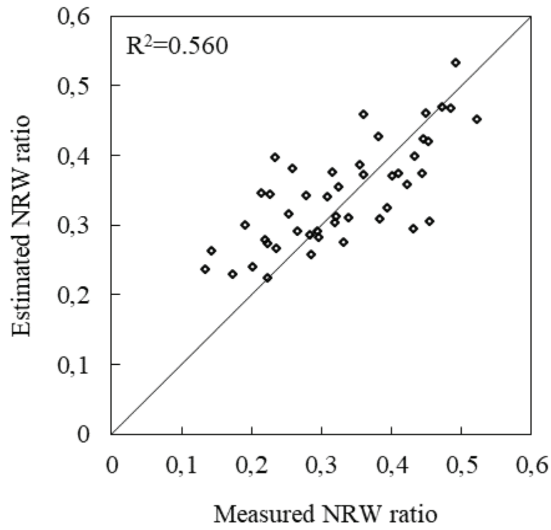

Figure 16. Scatter graph for WSQ/NL-DWST/SCL-FR model

Table 9. ANFIS models performance for FISO

Model	Combinations	R^2	RMSE	MAE	SI	Bias
ANFIS-4.1	WSQ-NJ-NL-WM	0.439	0.078	0.065	%23.68	-0.0142
ANFIS-4.2	WSQ-NJ-NL-MPD	0.606	0.066	0.053	%19.92	-0.0117
ANFIS-4.3	WSQ-NJ-NL-NF	0.291	0.094	0.073	%28.41	-0.0210
ANFIS-4.4	WSQ-NJ-NL-FR	0.313	0.093	0.074	%28.05	0.0013
ANFIS-4.5	WSQ-NJ-WM-NF	0.276	0.091	0.075	%27.49	-0.0218
ANFIS-4.6	WSQ-NJ-MPD-NF	0.332	0.091	0.072	%27.44	-0.0150
ANFIS-4.7	WSQ-NL-MPD-WM	0.603	0.068	0.053	%20.69	-0.0197
ANFIS-4.8	WSQ-NL-MPD -NF	0.092	0.108	0.084	%32.48	-0.0172
ANFIS-4.9	WSQ-NL-WM-NF	0.154	0.108	0.079	%32.59	-0.0320
ANFIS-4.10	WSQ-MPD-WM-NF	0.102	0.109	0.084	%32.76	-0.0176
ANFIS-4.11	WSQ–NSC-MPD-NF	0.319	0.093	0.073	%28.19	-0.0079

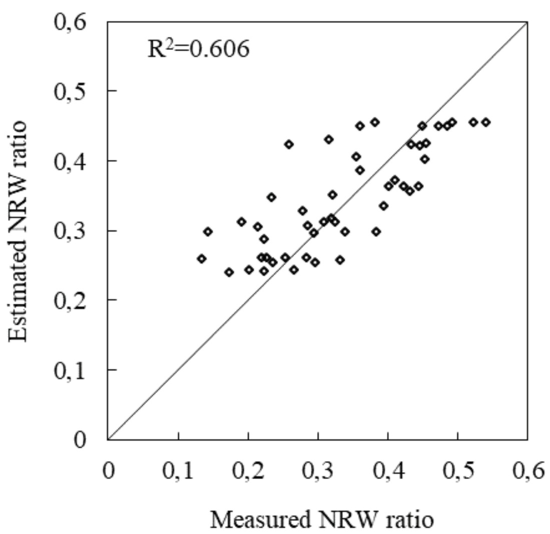

Figure 17. Scatter graph for WSQ-NJ-NL-MPD model.

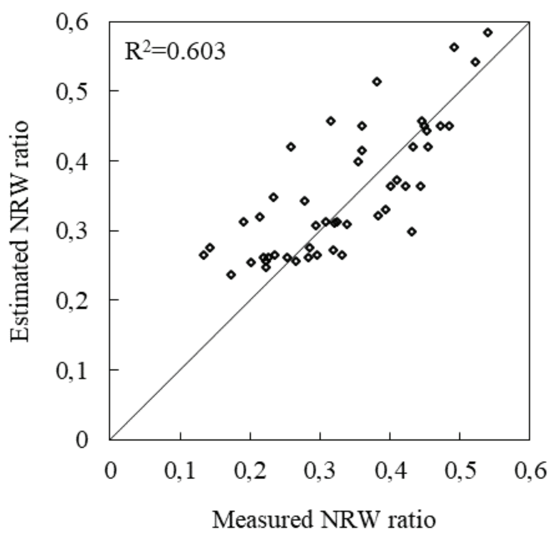

Figure 18. Scatter graph for WSQ-NL-MPD-WM model.

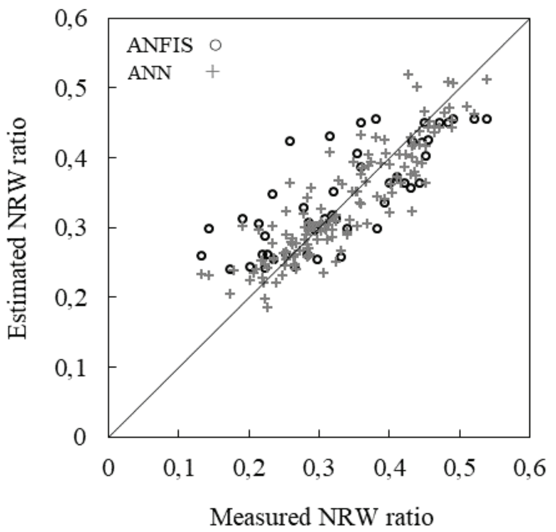

Figure 19. Compare of the best ANN and ANFIS models for this study.

Table 10. Comparison of the previous studies

No	Study	Model	Input number	R2
1	[12]	ANN	6	0.631
2	[13]	ANN	3	0.397
3	[14]	ANN	2	0.650
	[14]	ANFIS	2	0.500
4	[15]	ANN	3	0.760
	[15]	Kriging	2	0.950
5	[16]	ANN	2	0.397
	[16]	Kriging	2	0.890
6	[17]	STDM	4	0.990
7	**This study**	**ANN**	**4**	**0.794**
8	**This study**	**ANFIS**	**4**	**0.606**

CONCLUSION

Within this study, research is operated on the NRW ratios in the WDSs through ANN and ANFIS models. A total of new 216 models are developed as SISO, TISO, T3ISO and FISO. The best prediction model with a maximum accuracy by both methods is researched. Among the methods, the best ones are obtained using the FISO variables. When all models are evaluated; it is seen that the WSQ, NL, DWST, NJ, MPD, NF, SCL and WM variables have an impact on the NRW ratio and that they should be taken into account in prediction models.

The best ANN model is obtained by using the WSQ-NJ-NL-NF (R^2=0.794) input parameters. Any prediction accuracy on ANN methodology could not be achieved with ANFIS models which are formed with various combinations. The best model accuracy is obtained by the (R^2=0.606) combination of the ANFIS approach. Both models have average correlation and a small error to predict the NRW ratio. There is some amount of slight difference between the models due to a difference in method.

Predicting and interpreting the NRW ratio is so difficult due to uncertainties. One of the leading causes of uncertainties is the inability to measure and monitor WDS parameters and variables with sufficient accuracy. Technological investments are needed to reduce NRW and to monitor accurately. These investments are only possible with an additional budget. It is expected that this study and its developed models can guide NRW ratio predictions for water utilities with relatively limited resources and prioritize their future investments accordingly. In conclusion, it is also expected that the developed models can serve as a model for the future studies of experts and researchers studying this subject.

AUTHORSHIP CONTRIBUTIONS

All authors contributed to the study conception and design. The data analysis was performed by B.K. initiated the research. The first draft of the manuscript was written by B.K., E.Ş., and all authors commented on and added to the following versions of the manuscript. All authors read, review, editing and approved the final manuscript.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

REFERENCES

[1] Kanakoudis V, Tsitsifli S. Water volume vs. revenues-oriented water balance calculation for urban water networks: The "Minimum Charge Difference" component makes a difference! Available at: https://www.researchgate.net/profile/Vasilis-Kanakoudis/publication/270645306_Water_volume_vs_revenues_oriented_water_balance_calculation_for_urban_water_networks_the_Minimum_Charge_Difference_component_makes_a_difference/links/54d3e9180cf246475804046c/

Water-volume-vs-revenues-oriented-water-balance-calculation-for-urban-water-networks-the-Minimum-Charge-Difference-component-makes-a-difference.pdf. Accessed on May 15, 2024.
[2] Van Den Berg C. Drivers of non-revenue water: A cross-national analysis. Util Policy 2015;36:71–78. [CrossRef]
[3] Rajani B, Kleiner Y. Comprehensive review of structural deterioration of water mains: Physically based models. Urban Water 2001;3:151–164. [CrossRef]
[4] González-Gómez F, Martínez-Espiñeira R, García-Valiñas MA, García-Rubio M. A. Explanatory factors of urban water leakage rates in Southern Spain. Utilities Policy 2012;22:22–30. [CrossRef]
[5] Tabesh M, Asadiyami Yekta AH, Burrows R. An integrated model to evaluate losses in water distribution systems. Water Resour Manag 2009;23:477–492. [CrossRef]
[6] Tabesh M, Roozbahani A, Roghani B, Faghihi NR, Heydarzadeh R. Risk assessment of factors influencing non-revenue water using bayesian networks and fuzzy logic. Water Resour Manage 2018;32:3647–3670. [CrossRef]
[7] González-Gómez F, García-Rubio MA, Guardiola J. Why is non-revenue water so high in so many cities? Int J Water Resour Dev 2011;27:345–360. [CrossRef]
[8] Güngör-Demirci G, Lee J, Keck J, Guzzetta R, Yang P. Determinants of non-revenue water for a water utility in California. J Water Supply Res Technol AQUA 2018;67:270–278. [CrossRef]
[9] Şişman E, Kızılöz B. Trend-risk model for predicting non-revenue water: An application in Turkey. Util Policy 2020;67:101137. [CrossRef]
[10] Kizilöz B, Şişman E. A new performance analysis model for urban water supply systems evaluation. Desalin Water Treat 2021;235:177–192. [CrossRef]
[11] Kizilöz B, Şişman E. Exceedance probabilities of non-revenue water and performance analysis. Int J Environ Sci Technol 2021;18:2559–2570. [CrossRef]
[12] Jang D, Choi G. Estimation of non-revenue water ratio using MRA and ANN in water distribution networks. Water 2018;10:2. [CrossRef]
[13] Jang D, Choi G. Estimation of non-revenue water ratio for sustainable management using artificial neural network and Z-score in Incheon, Republic of Korea. Sustainability 2017;9:1933. [CrossRef]
[14] Kiziloz B, Şişman E. Estimation of non-revenue water rate using artificial neural networks and adaptive neuro fuzzy inference systems. In Proceedings of the 4. Eurasian Conference on Civil and Environmental Engineering; 2019 Jun 17-18; İstanbul, Türkiye. ECCOCE; 2019. p. 1175–1186.
[15] Şişman E, Kizilöz B. Artificial neural network system analysis and Kriging methodology for estimation of non-revenue water ratio. Water Sci Technol Water Supply 2020;20:1871–1883. [CrossRef]
[16] Kızılöz B, Şişman E. Prediction models for non-revenue water ratio. Ömer Halisdemir Univ J Eng Sci 2021;10:276–283.
[17] Kızılöz B, Şişman E. Non-revenue water ratio prediction with serial triple diagram model. Water Supply 2021;21:4263–4275. [CrossRef]
[18] ISU. 2018 faaliyet raporu. Available at: https://www.isu.gov.tr/component/GetFiles.ashx?t=2&FileName=7fcb77ca-f3d2-4cf8-b1bd-1e930bdde327.pdf. Accessed on May 15, 2024.
[19] Şen Z, Şişman E, Kizilöz B. A new innovative method for model efficiency performance. Water Supply 2022;22:589–601. [CrossRef]
[20] Şen Z. Fuzzy Logic and Hydrological Modeling. 1st ed. Boca Raton: CRC Press; 2009. [CrossRef]
[21] Kizilöz B. Prediction model for the leakage rate in a water distribution system. Water Supply 2021;21:4481–4492. [CrossRef]
[22] Kizilöz, B. Prediction of daily failure rate using the serial triple diagram model and artificial neural network. Water Supply 2022;22:7040–7058. [CrossRef]