Water Supply

© 2021 The Authors

Water Supply Vol 21 No 8, 4263 doi: 10.2166/ws.2021.173

Non-revenue water ratio prediction with serial triple diagram model

Burak Kızılöz[a] and Eyüp Şişman[b,c,*]
[a] Kocaeli Metropolitan Municipality Water and Sewerage Administration, İzmit, Kocaeli, Turkey
[b] School of Engineering and Natural Sciences, Civil Engineering Department, Istanbul Medipol University, Kavacık, 34181 Istanbul, Turkey
[c] Climate Change Researches Application and Research Center, Istanbul Medipol University, (IKLIMER) Kavacık, 34181 Istanbul, Turkey
*Corresponding author. E-mail: esisman@medipol.edu.tr

EŞ, 0000-0003-3696-9967

ABSTRACT

Water administrations attempt to control the Non-Revenue Water Ratio (NRWR) values in sustainable and well-performing water distribution infrastructures. In this respect, the NRWR value prediction through appropriate models over a small number of controllable variables is significant. The collection, monitoring, and predictions of data on variables that are used in the NRWR calculations are not practical and required significant time besides financial resources. In this study, the NRWR predictions have been made through the suggested method over three parameters. The model prediction accuracies, in the literature obtained by using the Triple Diagram Model (TDM) over two parameters, have been increased through the Serial Triple Diagram Model (STDM) suggested in this study. This method shows that better predictions are possible in the NRWR modeling. Thanks to the model applications developed in this study, water administrations can make predictions with the least error (less than 10% relative error) and certain variables, according to the characteristics of each water distribution network.

Key words: Kriging, NRW Ratio, Serial Triple Diagram Model, water distribution network, water management

HIGHLIGHTS

- The variables and combinations which have an impact on the prediction of NRW ratios have been researched by using the Serial Triple Diagram Models.
- Kriging method has been used in this study for NRW ratio prediction.
- This model product is the best in the literature for NRW ratio predictions.

INTRODUCTION

The International Water Association (IWA) makes important contributions to the non-revenue water (NRW) issue through numerous conferences, workshops, and publications every year. Despite all these efforts, the water loss rate is high and especially the awareness level of the water utilities in most developing countries has not reached the expected level yet. It is, therefore, necessary to provide much support on NRW studies. In order to meet increasing water demand around the world, the water utilities should focus on after the present and future climate change impacts, NRW calculation, prediction, monitoring, and reducing issues. Therefore, the water infrastructure is very important for continuous water supply and quality to the consumers (Shuang *et al.* 2017). Analysis of water distribution systems infrastructure components shows that they are composed of treatment units, transmission and distributions lines, tanks, pumping stations, valves, drains and air release valve. Any failure in the distribution components can cause significant water losses (WL). The production and operating costs increase due to these losses. Such problems have different consequences, from service disruption to water quality deterioration. For this reason, the studies on WL in water distribution systems (WDSs) are quite important (Kanakoudis & Tsitsifli 2010; Kanakoudis *et al.* 2015; Kanakoudis *et al.* 2016).

The distribution system losses can be separated into components as the real and apparent losses (Alegre *et al.* 2016). The water loss quantity and the unbilled authorized consumption contribute to the NRW loss (Kingdom *et al.* 2006; Kanakoudis & Muhammetoglu 2014; Alegre *et al.* 2016; Boztaş *et al.* 2019; Şişman & Kizilöz 2020a, 2020b). The NRW amount is predicted as about 126 billion m^3 around the world and this amount corresponds to 39 billion dollars (Liemberger & Wyatt 2019). The water authorities must keep the NRW at low levels (under 10%) as in developed countries to continue their

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

investments and to properly operate water distribution infrastructures. In case of insufficient feasibility, design and project application studies, the NRW amount reaches high levels in a short time and causes significant budget waste. Recently, Turkey has paid more attention to the legal regulations and practices including water losses reduction in WDSs in accordance with the above-mentioned issues. In this context, the regulation on control of water losses in water supply and distribution systems (Şişman & Kizilöz 2020b), most recently revised and published in the Official Gazette dated on 8 May 2014, has entered into force. According to the amendment in the existing regulation, the metropolitan and provincial municipalities are responsible for reducing water losses to the level of 30% until 2023, and 25% in 2028; as for the other municipalities; this level should be at most 35% until 2023, 30% until 2028 and 25% until 2033 (RTMAF 2019). In parallel with these developments, NRW applications have gathered speed in Turkey. In addition, for water resources based on climate change impacts, water quantity reduction has had negative effects, especially in the last 20 years in the country geography, which has brought the NRW studies to the forefront. Improvement of both the current methodologies has been tried with the support of the new approaches. Güngör-Demirci *et al.* (2018) have used the panel regression analysis method to identify the effective components on NRW. Tabesh *et al.* (2018) have considered the fuzzy logic risk evaluation and Bayesian network methods on three principal components of WDS including apparent losses, real losses and non-revenue water authorized consumptions. The NRW amount prediction has been made through alternative Artificial Neural Networks (ANNs) models by using physical and operational variables such as water supply quantity per demand junction, pipe ratio deterioration, demand energy ratio, average pipe diameter, pipe length per demand junction and number of leaks (Jang & Choi 2017, 2018). The network performances have been analyzed by a new risk approach in Kocaeli (Kızılöz & Şişman 2021). However, the first model prediction results are also obtained through the TDM and Kriging methods for NRW rate (Şişman & Kizilöz 2020a).

In the literature, there are various studies in which TDM and Kriging methods are successfully applied (Sen 2009). The contour diagrams are obtained through the Kriging method, where the relationship between the two independent variables and the single dependent variable has been predicted through the most appropriate interpolation approach (Özger & Sen 2007). The method flexibility is high in this relationship to generate the accurate model for the data scatter point. Despite the intensively scattered points, the TDM charts help understanding of variable behaviors and interpreting them easily.

In this paper, the NRW ratio (NRWR) prediction is achieved through the Serial Triple Diagram Modeling (STDM) method for the first time. In this approach, firstly, the two independent variables, which are effective on NRWR, are taken into consideration and then the initial TDM with two inputs is used for NRWR prediction. Thereafter, the second and last TDM is considered by the TDM errors with an additional variable. This sequential model structure for NRWR prediction is called STDM.

STUDY AREA AND DATA DESCRIPTION

The suggested model application is carried out in Kocaeli, one of the biggest industrial cities of Turkey, in accordance with the data obtained from the water distribution infrastructure (Figure 1).

A total of 163,627,918 m^3 water was supplied for the consumers in Kocaeli City in 2018 (ISU 2018), the total water loss quantity was 50,422,227 m^3 and the NRW was 52,865,847 m^3. The water loss rate has been reduced respectively between 2014 and 2019 in sequence to 39.3%, 37%, 35.6%, 33%, 31%, and 30.2%. The investments after the year 2015 have a significant role in these reductions. Dividing a WDS into district metered areas (DMAs) is an important step and task that usually troubles water utility managers towards pressure management for reducing leakages (Chondronasios *et al.* 2017; Creaco & Haidar 2019). Certain DMAs and pressure metered areas (PMAs) were installed in WDSs by taking hydraulic model results as a reference, especially starting from the districts with high water supply and water losses, in order to reduce water losses within the framework of the investments in question. The studies on reducing leakages at ideally operating pressure were carried out by controlling the network pressures (Creaco *et al.* 2019; Güngör *et al.* 2019). At the end of 2020, the total network length was 10,656 km in Kocaeli and the isolated network length was 1,488 km. In order to reduce the water losses to 25% in 2023, DMA and PMA isolated region studies are continuing. In addition, the water meters have expired their economic life or are about to expire and therefore they are being replaced with smart ones to reduce apparent losses in NRW components. As a result of ongoing studies, the total water loss of Kocaeli was measured as 27.95% at the end of 2020, and the NRW ratio was 29.86%. In the last five years, the studies have focused mainly on reducing leakages in fittings, building service connections and pressure management system applications in water distribution networks.

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

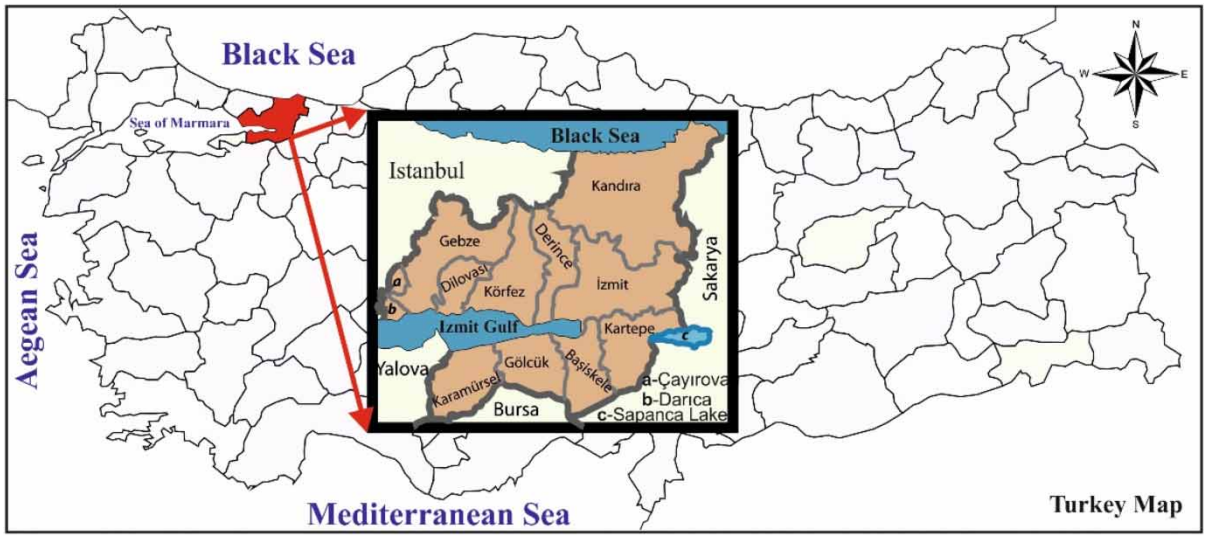


Figure 1 | General view of Kocaeli.

Each district in Kocaeli (12 districts) has been planned as a separate principal DMA region. In these regions, the system input volume (SIV), network length (NL), number of failures (NF), pipe diameters, pipe ages, number of water meters and NRWR amounts are followed in the most detailed way. The significant numerical characteristics of Kocaeli water distribution system can be seen in Table 1 for the year 2018.

The NRWR model inputs are determined through certain variables that are effective on NRWR and these model inputs are studied by Şişman & Kızılöz (2020a), considering network pressure, network age, network length, water meter, pipe diameter, failure ratio, system input volume, number of junctions and service connection length. In addition, the statistical information on all parameters for the model is presented in Table 2.

METHODOLOGY

The practical benefit of this method is that it treats multiparameter variables to reach a better solution, which cannot be easily obtained by a set mathematical formulation such as multiple regression, which provides a smoothened surface with overall error minimization. In the solely mathematical models, there are a set of restrictive assumptions. However, in our method, there is not such restrictive assumption.

Table 1 | Significant numerical characteristics of Kocaeli water distribution system (Şişman & Kizilöz 2020a)

Districts	SIV (m^3)	NL (km)	NF	NRW ratio%
DMA 1	30,840,477	1,114	1,106	29.74
DMA 2	6,647,316	1,600	900	34.05
DMA 3	14,160,461	918	1,114	44.36
DMA 4	14,702,104	680	1,153	32.28
DMA 5	29,810,207	1,050	672	25.40
DMA 6	8,941,011	854	484	28.50
DMA 7	14,216,123	803	560	37.02
DMA 8	5,169,076	443	269	41.10
DMA 9	11,181,395	447	563	40.11
DMA 10	6,284,492	310	439	26.16
DMA 11	9,474,883	339	173	23.01
DMA 12	12,200,373	378	48	26.01

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

Table 2 | STDM model parameters

Measured variables	Proposed	Min. and max. values	Unit
System input volume (SIV)	IWA (Alegre *et al.* 2016)	315,445–2,844,526	m^3
Failure ratio (FR)	IWA (Alegre *et al.* 2016)	0.01–0.64	–
Service connection FR (SCFR)	IWA (Alegre *et al.* 2016)	0.15–2.23	–
Service connection length (SCL)	IWA (Alegre *et al.* 2016)	130–527	km
Network length (NL)	IWA (Alegre *et al.* 2016)	310–1,600	km
Water meter (WM)	IWA (Alegre *et al.* 2016)	15,471–160,135	Number
Number of junctions (NJ)	Jang & Choi (2018)	9,660–52,565	Number
Mean pressure of network (MPN)	IWA (Alegre *et al.* 2016)	43–64	m
Mean age of pipe (MAP)	EPA (2006)	8–30	year
Mean diameter of pipe (MDP)	EPA (2006)	108–159	mm
Non- revenue water ratio (NRWR)	IWA (Alegre *et al.* 2016)	0.13–0.54	–

On the other hand, the monitoring of 170 Performance indicators (Pls) that are suggested by the IWA is carried out using the IWA Water Balance tables over 232 variables (Kanakoudis *et al.* 2013). Kanakoudis *et al.* (2013) have explained in their studies the formula of 75 IWA Pls and the 98 IWA variables required for calculations such as system input volume (SIV) (m^3), water losses (m^3), real losses (m^3), apparent losses (m^3), non-revenue water (NRW) (m^3), main (network) lengths (NL-km), service connections, average service connection length (SCL-m), meter replacement (WM), and average operating pressure (kPa). They have also indicated that the 170 Pls suggested by the IWA do not include all issues that water administrations encounter and there are various parameters in water losses that cannot be explained by the existing IWA Pls. In the international literature, it is clearly stated that various factors and parameters have an effect upon pipe failures and leakages; for example, pipe characteristics (diameter, age, and etc.), environmental/climatic conditions (soil temperature, traffic load, etc.), and operational and maintenance factors (water quality, maintenance characteristics, etc.). Jang & Choi (2018) has developed the NRW ratio prediction models through the multiple regression analysis (MRA) and artificial neural network (ANN) methods using operational parameters (demand energy ratio and no of leaks) and physical parameters (average pipe diameter, pipe length/demand junction, water supply quantity/demand junction, and deteriorated pipe ratio). Şişman & Kızılöz (2020a) have identified the parameters affecting the NRW ratio and have predicted this ratio by using the ANN and Kriging methods in their first study. The model inputs selected in this study are the same as the independent variables suggested by the IWA, EPA (2006), Kanakoudis *et al.* (2013), Alegre *et al.* (2016), Jang & Choi (2018), and Şişman & Kızılöz (2020a).

As for modeling, while the change of two variables with each other can be seen easily in the cartesian coordinate system, it can be seen only through the contour plots if three variables are available (Sırdaş & Şen 2003). The X and Y axes indicate the coordinate in classical geostatistics applications. Each axis is defined in terms of the variables used in the study. The NRWR values are predicted through the selected variables. The Surfer program is employed for the preparation of TDM prediction charts. In this study, the TDM charts are established by means of the point Kriging interpolation method. The methods used for establishing model charts through the Kriging method are explained in a detailed manner in the studies carried out by Matheron (1963) and Journal & Huijbrets (1978).

It is important for problem solution to prepare the variograms revealing the structural forms and to generate the Kriging model for prediction problem. One of the most important geostatistical methods is the Kriging for the prediction based on stochastic regional change (Matheron 1963; Krige 1966). The geostatistical method provides information concerning the quantity of the examined or investigated value against the possible variable values defined in the XY coordinate system for prediction. The geostatistical method benefits from variograms explaining the change based on independent variables to use the current information in the best way. According to this method, the current data distances in the XY coordinate system and the changeable weighted means are taken as an average to predict the dependent variable value against the certain values of independent variables. The prediction model is given below in a simple way:

$$Z_i = f(X_i,\ Y_i) + \varepsilon_i \quad i = 1,\ 2,\ 3 \ldots..,\ n \tag{1}$$

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

Z indicates non- revenue water ratio; X represents various variables in each model such as SIV/NJ, SIV/SCL, SIV/NL, NL/MPN, NL/MAP, SCL/MPN, ERROR; Y shows different variables for each model such as MAP, MPN, SCFR, MAP, FR, SIW/WM,

In Equation (1), n means the number of measuring points and ε is the error term in accordance with the model chart.

According to the cartesian coordinate system, if the random variables, which are close to each other and positioned x and $x+h$ have $Z(x)$ and $Z(x+h)$ values to indicate h distance, this h distance vector actually means the geometric distance between the two input variables. NRW counter maps can be drawn by using the Kriging method with two input variables. The variance of $Z(x)$ can be defined as the quadratic mean of $m(x)$ (mean value) deviations.

$$Var[Z(x)] = E[Z(x)] - m(x)]^2 \tag{2}$$

where, E is the expected value (mean) operator. The variance of $Z(x+h)$ can be explained similarly.

On the other hand, the covariance indicates the change between $Z(x)$ and $Z(x+h)$, which is explained as follows:

$$C(x,\ x+h) = E\{[Z(x) - m(x)][Z(x+h) - (m(x+h)]\} \tag{3}$$

The correlation structure of the semi-variogram is expressed as follows (Matheron 1963).

$$\gamma(h) = \frac{1}{2}Var[Z(x) - Z(x+h)] \tag{4}$$

The relationship between $\gamma(h)$ and $C(h)$ is defined as,

$$\gamma(|h|) = \sigma^2 - C(h) \tag{5}$$

where σ^2 is the constant variance in this equation.

Before prediction, the semi-variogram (SV) is obtained from the current data. For this purpose, the data values matched according to their positions should be divided into groups by taking the distance between them into consideration. The mathematical statement of experimental semi-variogram is given for each group distance as,

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z(x_i) - Z(x_i+h)]^2 \tag{6}$$

In this equation, $\gamma(h)$, is the semi-variogram for h distance; h is the mean distance (lag-size) between the remaining data at intervals grouped by distance; $N(h)$ is the total data matching number remaining in these intervals; $Z(x_i)$ is the value in x_i point; and finally, $Z(x_i+h)$ gives the measurement value in x_i+h point.

This semi-variogram function should be calculated in each direction and for each pair of measurements. Then, it is defined through the functions such as linear, gaussian, exponential, and spherical. The mathematical model selected for the experimental semi-variogram provides the required input parameters (weight coefficients) for Kriging interpolation. Furthermore, the Kriging interpolation technique can be applied to calculate the unknown point values. The expected value of the Z variable (at x_0 position) can be predicted as the weighted average of measured values $[Z(x_1),\ Z(x_2),\ Z(x_3),\ \ldots\ldots,\ Z(x_n)]$. The basic relation of the Kriging method is available below in the most general sense:

$$Z^{\omega}(x_0) = \sum_{i=1}^{n}\theta_i z(x_i) \tag{7}$$

where $Z^{\omega}(x_0)$ is the appraisal value of Z in x_0; θ_i is the weight values corresponding to each $Z(x_i)$, and n is the point number remaining inside the influence area for prediction. In this modeling, the counter maps have been formed by using the Kriging interpolation technique and the intermediate values have been estimated.

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

Serial Triple Diagram Model (STDM)

In this study, STDM, a structure developed for the first time in this research, is used in order to predict the NRWR. The Serial Triple Diagram Model (STDM) has a scientific philosophical basis, although there may not be many mathematical equations. There is no flaw in the methodology because any model will need numerical data for application, whether conceptual, logical, statistical, or mathematical. In the literature, there are plenty of models that cannot be applied in the absence of measurements. The STDM procedure provides every detail in spatial mapping better than regression models, which provide a smoothening effect through a crisp mathematical model.

The first NRWR prediction is made by the TDM charts for the two selected independent variables. In this initial TDM chart, a model chart with two inputs and one output is established by considering the correlation between the selected variables. Thus, the initially established model can be evaluated according to the certain performance criteria. If this model structure cannot predict the dependent variable sufficiently, a new TDM model chart is established to improve the performance by reducing the errors. On the other hand, the error terms (difference between the prediction and the calculated value) after the initial model and another independent variable, effectively in NRWR change, is taken into consideration for the next prediction charts; the NRWR predictions are repeated in this way by the new TDM charts. A model structure with four inputs in total including three different independent input variables and error terms in accordance with the model result in the initial stage is established for the NRWR variable model prediction. In this way, it is possible to continue with sequential models until the desired prediction accuracy is achieved. This structure is called the Serial Triple Diagram Model for the prediction of Non-Revenue Water Ratio (NRWR) in Figure 2.

The equations used for the STDM and the sample model charts are given below:

$$P_{1j} = f\,[m_{1j},\ m_{2j}] \tag{8}$$

$$P_{2j} = f\,[m_{3j},\ e_{1j}] \tag{9}$$

$$P_{(1j)} = a_{1j}.m_{1j} + b_{1j}.m_{2j} \tag{10}$$

$$e_{1j} = M_j - P_{1j} \tag{11}$$

$$P_{2j} = c_{1j}.m_{3j} + d_{1j}.e_{1j} \tag{12}$$

$m_{ij} = ij.$ independent measured input data of $i.$ series, $(i = 1, 2, 3)$; $j = 1, 2, \ldots\ldots, n$ where n is the record length, $M_j =$ measured dependent variable, $j = 1, 2, \ldots\ldots, n$ where n is the total number of data,

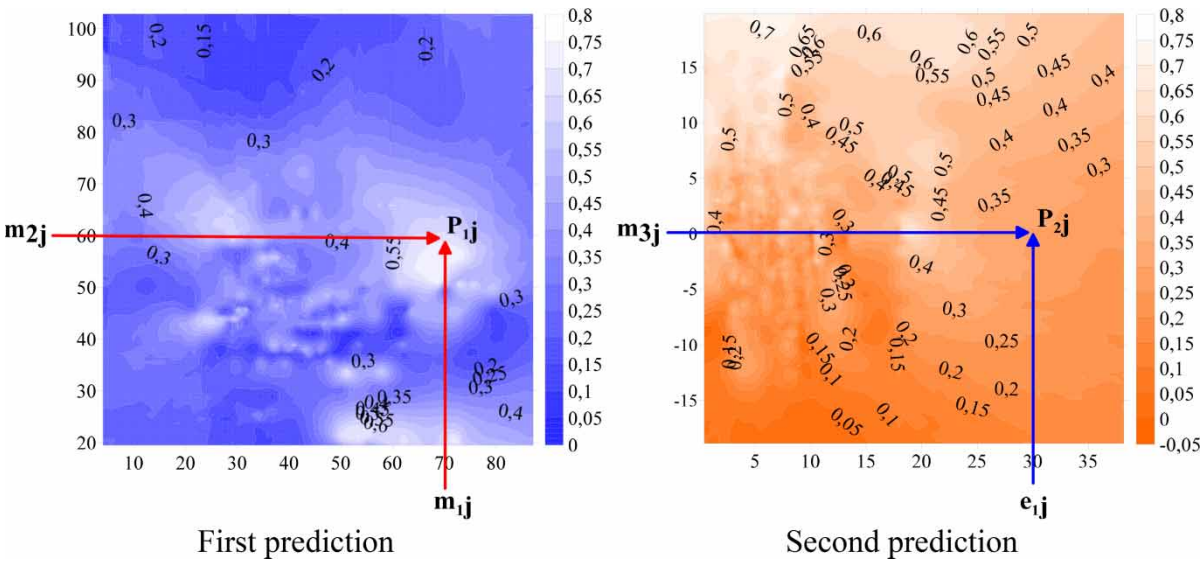


Figure 2 | Serial Triple Diagram Models for the prediction of NRWR.

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

$P_{ij} = ij$. model output as prediction, $(i = 1, 2)$,
$e_{1j} = 1j$. 1. model error terms,
$e_{2j} = 2j$. 2. model result errors, and
a_{1j}, b_{1j}, c_{1j} and d_{1j} = model constant.

Evaluation of STDM model performance

The performances of models in this research are evaluated according to the mean square error (MSE) and coefficient of determination, R^2 parameters.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - M_i)^2 \tag{13}$$

$$R^2 = \frac{\sum_{i=1}^{n}(P_i - \bar{M})^2 - \sum_{i=1}^{n}(P_i - M_i)^2}{\sum_{i=1}^{n}(P_i - \bar{M})^2} \tag{14}$$

In these equations, M_i is the measured data, P_i is the estimation data and n is the data number. $\bar{M}$ and $\bar{P}$ are the averages of measured and estimation data.

RESULTS AND DISCUSSION

When the suggested model prediction charts are established, the independent variables are identified on the x and y axes and the dependent variable appears on the z axis. The NRWR prediction charts are constructed through the Surfer software by using linear semi variograms.

The NRWR variable value prediction for Kocaeli City is made by Şişman & Kizilöz (2020a) through the TDM prediction charts for the first time with performance results given in Table 3. When the performance indicators are analyzed, it is observed that the R^2 values vary between 0.74 and 0.95 and the MSE values between 0.0004 and 0.0021. In this study, for better model performance, the next TDM chart is obtained by considering the model errors of the initial model setup and of a third variable that is effective in NRWR prediction with higher performance. Thus, the TDM structure with two inputs and one output is converted by the proposed approach into a new model with four inputs in total coupled with the error terms in the initial model. The model performance predictions from the TDM and STDM models are given in Table 3. The analysis of the models' performance shows that the STDM predictions are better than the TDM among the ten serial models.

Table 3 | STDM prediction results

	First TDM Şişman & Kizilöz (2020a)				Second TDM (In this study)			
No.	Inputs		R^2	MSE	Inputs		R^2	MSE
1	SIV/NJ	MPN	0.95	0.0004	Error	MDP	0.97	0.0003
2	SIV/NJ	MAP	0.94	0.0005	Error	MPN	0.99	0.0001
3	SIV/SCL	MAP	0.93	0.0006	Error	SCFR	0.97	0.0003
4	SIV/SCL	MPN	0.93	0.0006	Error	MAP	0.98	0.0002
5	SIV/NL	MPN	0.90	0.0008	Error	MAP	0.97	0.0002
6	SIV/NL	MAP	0.90	0.0008	Error	MDP	0.95	0.0004
7	NL/MPN	MAP	0.79	0.0017	Error	FR	0.92	0.0008
8	NL/MAP	MPN	0.78	0.0017	Error	SIV/WM	0.97	0.0003
9	NL/MPN	MDP	0.78	0.0017	Error	MAP	0.99	0.0001
10	SCL/MPN	MAP	0.74	0.0021	Error	MDP	0.94	0.0005

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

The R^2 performance criteria, for which is the minimum as 0.74 for SCL/MPN-MAP input parameters in TDM structure is 0.94 in terms of SCL/MPN-MAP and initial TDM errors-MDP inputs in STDM structure. The STDM has significant improvements rather than the TDM model predictions in each of the ten combinations taken into consideration.

The TDM and STDM model prediction charts are given in Figure 3. SCL/MPN-MAP input combination model prediction results for the TDM against the calculated NRWR variable values are exposed in Figure 3(b). The 1:1 straight line (45°) in red represents the comparison of calculations and model predictions. Comparison of Figure 3(b) and 3(d) according to the reference line shows clearly that the STDM predictions in Figure 3(d) are better than the TDM model predictions in Figure 3(b). According to the chart in Figure 3(b), the TDM model prediction results are generally below or above the calculations. On the other hand, the predictions overlap mainly with the STDM chart in Figure 3(d). In addition, it is possible to evaluate the prediction charts given in Figure 3(a) and 3(c) for the models numbered as ten. It is understood from the Figure 4(a) model chart that the NRWR increases when the service connection length per unit average pressure and the network age increase. The model prediction performances are improved by obtaining the NRWR behavior and change for the current sample application according to the model chart, depending on the errors and the average pipe diameter as in Figure 4(c).

The second worst model performance is the NL/MPN-MPD combination for the TDM and its performance scatter plot is in Figure 4(b). According to the model chart, it is observed that the NWRW increases in general when the network length

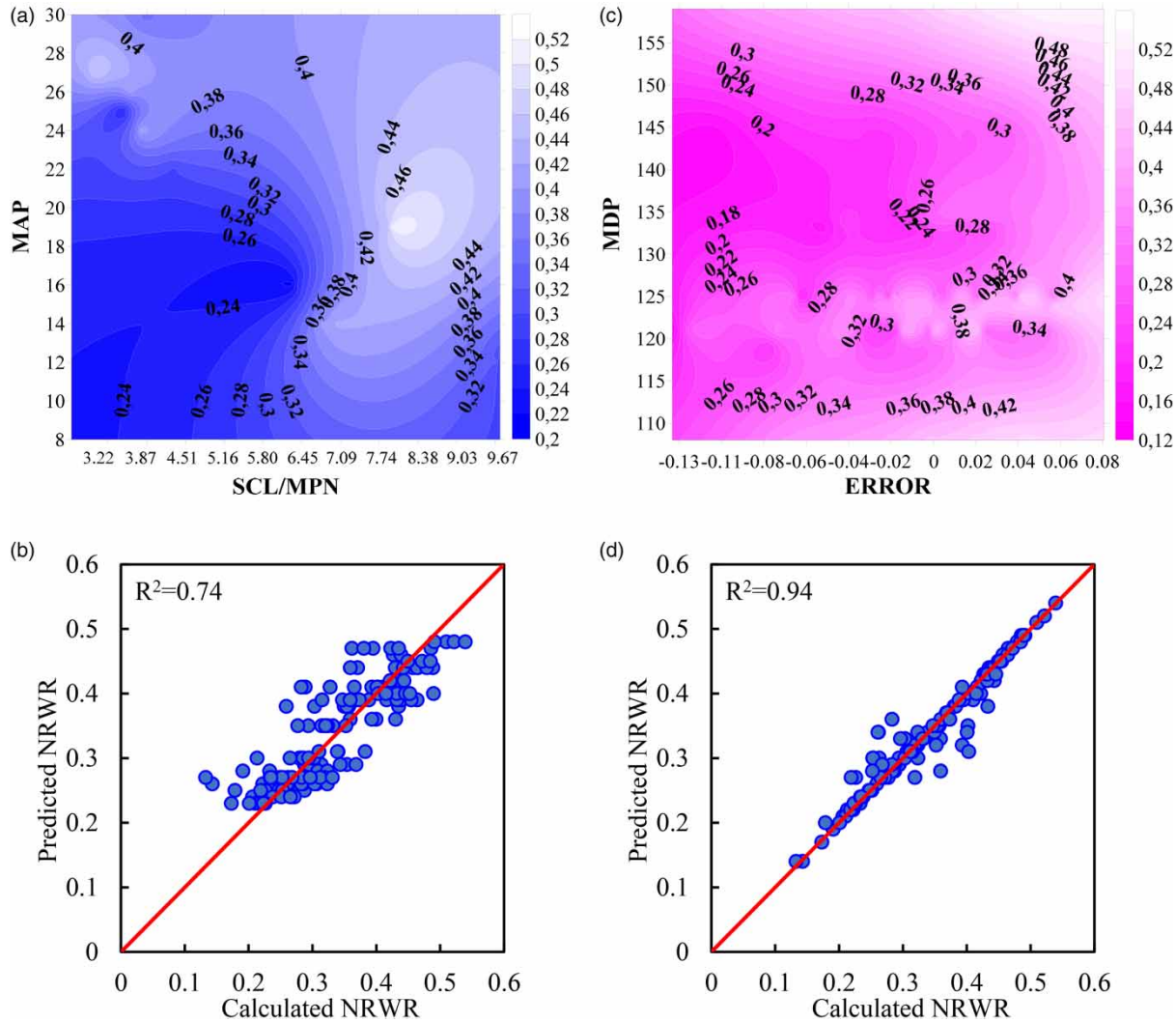


Figure 3 | STDM, (a) first TDM for SCL/MPN-MAP, (b) first comparison of predicted and calculated NRWR, (c) second TDM for ERROR-MDP and (d) second comparison of predicted and calculated NRWR.

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

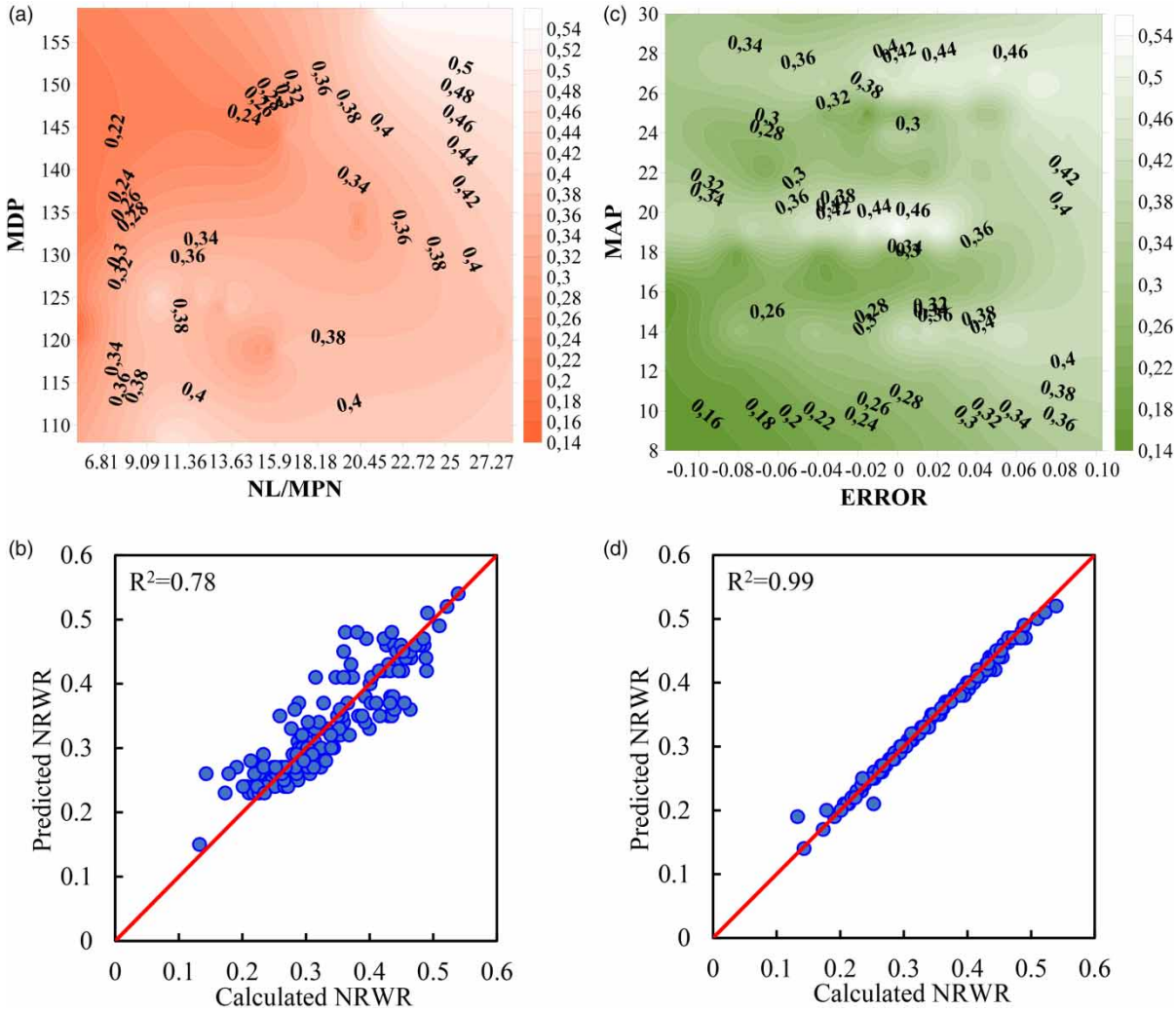


Figure 4 | STDM, (a) first TDM for NL/MPN-MDP, (b) first comparison of predicted and calculated NRWR, (c) second TDM for ERROR-MAP and (d) second comparison of predicted and calculated NRWR.

increases per unit average operating pressure. On the other hand, there seems to be a much more complex relationship to the average pipe diameter and the NRWR change (Figure 4(a)). Figure 4(b) indicates that it is not possible to explain the NRWR behavior sufficiently by referring to these parameters only. The points above the red line in Figure 4(b) show that the model predictions are high and the points below this line show that the predictions are low. In this study, the next prediction model chart in Figure 4(c) is established by using the error value of the initial prediction model and the MAP variable. This chart reflects that the NRWR increases depending on the age increase. After the establishment of the next prediction model chart, the performance has increased and the R^2 value has risen to 0.99 (Figure 4(d)). According to the measurement results, the NRWR and STDM model predictions fall mostly on the 45° slope line. The STDM structure helps the NRWR prediction with three variables with a high accuracy.

The best TDM chart, according to the initial model results for NRWR prediction, are available in Figure 5(a) (Şişman & Kizilöz 2020a). The relevant chart shows that the NRWR ratios fluctuate by the average pressure change in the network. The NRWR change from the system input volume per network junction at average 44–64 bar network pressures can be interpreted through the related figure. If the average pressure value is at 50–54 mSS level, in general, for network operation, it is possible to control the NRWR below 30%. The model performance in Figure 5(b) shows clearly that the prediction performance of the TDM is quite good. This performance improves the predictions slightly and the R^2 value of model performance

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

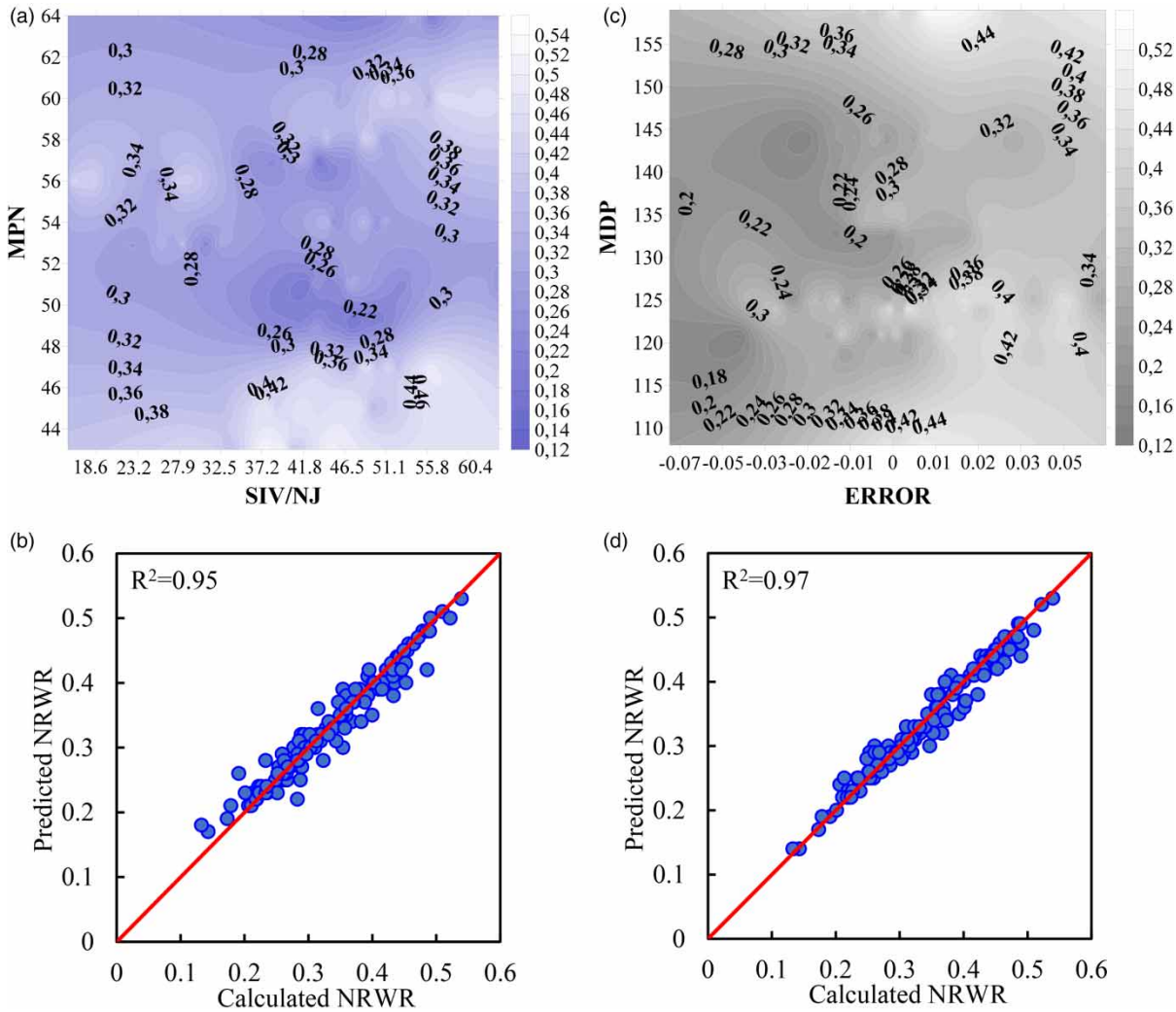


Figure 5 | STDM, (a) first TDM for SIV/NJ-MPN, (b) first comparison of predicted and calculated NRWR, (c) second TDM for ERROR-MDP and (d) second comparison of predicted and calculated NRWR.

criterion increases from 0.95 to 0.97 when the application is performed for the same NRWR output data set with the STDM structure as suggested in this study. Comparison of the STDM chart given in Figure 5(b) with the TDM chart in Figure 5(b) indicates the possibility of evaluating the differences between the model results. The STDM predictions are much more successful in low NRWR values. On the other hand, higher prediction performance in high NRWR values is achieved by the initial TDM chart in Figure 6(a) for the model application and data combination.

The tendency for increase in the NRWR value depending on increase in age is clearly seen in Figure 6(a), by the initial prediction chart regarding the other combination. There is an age increase in the model, and therefore a little decrease is observed in NRWR in some small areas. This situation indicates that the network is in good condition despite network aging or the network operating efficiency (pressure management system applications) is sufficient. The performance chart regarding the model efficiency is given in Figure 6(b). According to the model predictions and the 1:1 scattering line of the calculated NRWR values, the calculations and the model predictions are generally compatible, and the R^2 value performance criterion is 0.94. The next model prediction chart in Figure 6(c) is established through the STDM approach by using the errors of the initial model predictions and the average pressure variable in the network. The model performance is improved visibly in the final model chart, and the performance value for R^2 has increased from 0.94 to 0.99. In addition, the compatibility between the calculations and the model predictions can be seen from the scattering plot in Figure 6(d).

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

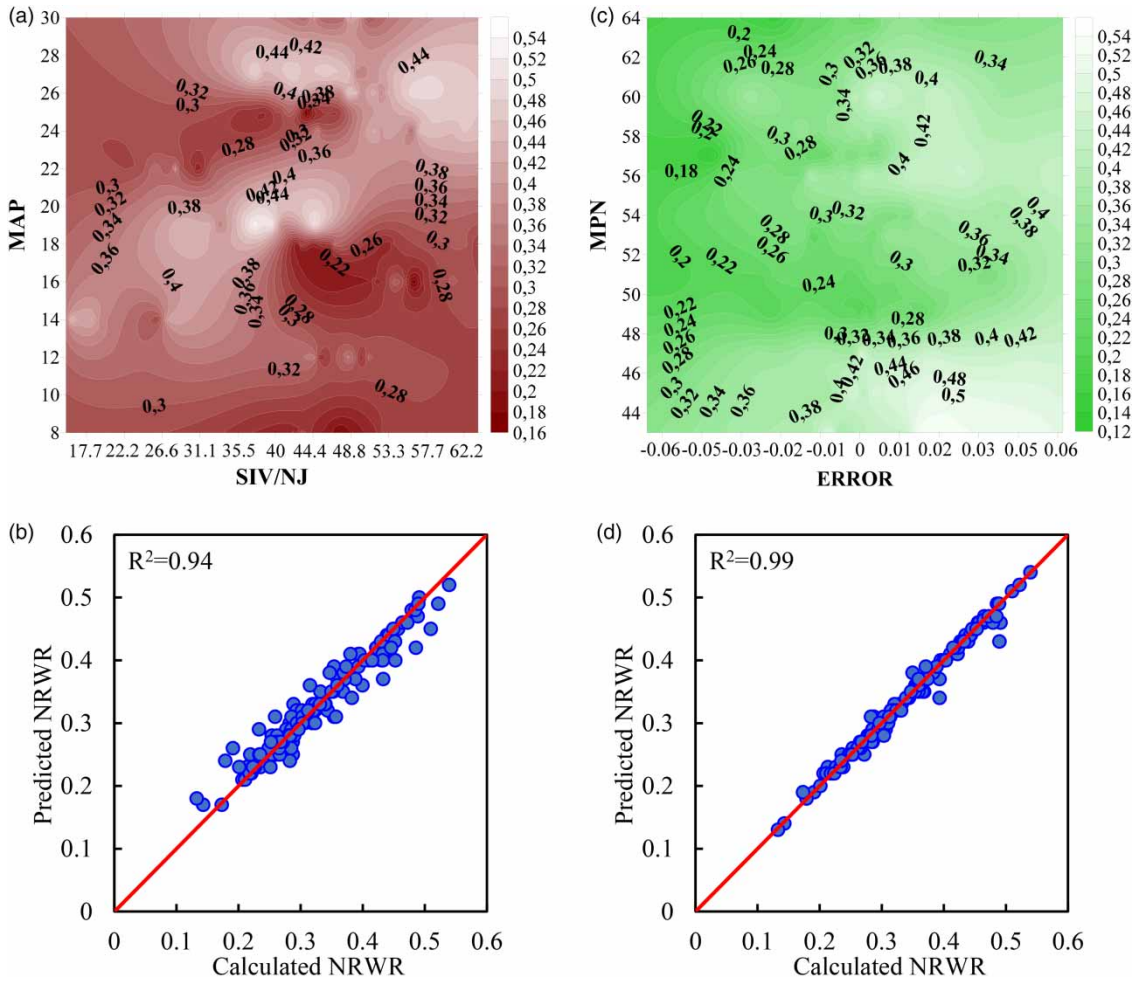


Figure 6 | STDM, (a) first TDM for SIV/NJ-MAP, (b) first comparison of predicted and calculated NRWR, (c) second TDM for ERROR-MPN and (d) second comparison of predicted and calculated NRWR.

With the STDM structure suggested among the ten model applications, the predictions are better than the TDM results. Additionally, the previous model studies and the highest performances in this model study are presented in summary in Table 4.

Finally, the comparison of the best prediction results regarding the TDM and STDM by Şişman & Kizilöz (2020a) is given in Figure 7. The significant improvement in model performance can be seen clearly through the scattering by red points when model predictions are interpreted in accordance with 1:1 straight line.

Table 4 | Comparison of the model results

No.	Study	Model	The number of input	Variables	R^2	MSE
1	Jang & Choi (2017)	ANN	3	WSQ*/DJ*, DPR, DER	0.40	97.18
2	Jang & Choi (2018)	ANN	6	WSQ*/DJ*, NOL, DPR, DER, APD, PL/DJ	0.63	63.70
3	Şişman & Kizilöz (2020a)	ANN	3	ST/NJ, NL, MAP	0.76	0.0019
4	Şişman & Kizilöz (2020a)	TDM	2	SIV*/NJ*, MPN	0.95	0.0004
5	This study	STDM	4	SIV*/NJ*, MPN, MAP, ERROR	0.99	0.0001

*System input volume (SIV) = water supply quantity (WSQ) and demand junction (DJ) = number junction (NJ).

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

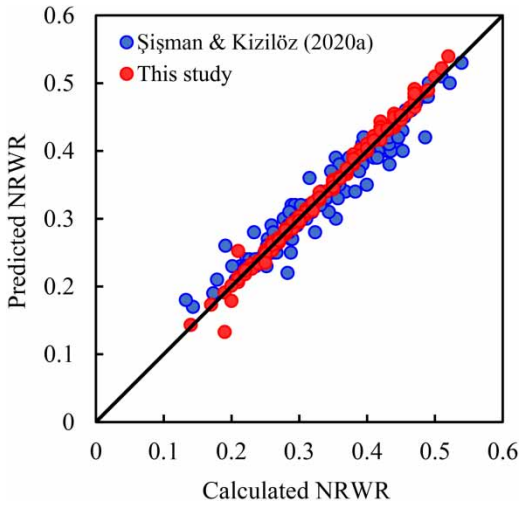


Figure 7 | Comparison of the TDM and STDM.

CONCLUSIONS

In this study, the geostatistical approach is used in order to model the prediction uncertainties and to reveal the correlation structure between the selected dependent variable and the independent variables. The non-linear regression between the input and output data is expressed geometrically through the charts. In this study, the NRWR variable is predicted using the parameters suggested by IWA identifying the WDS. The models with two inputs and one output are developed through the TDM for NRWR predictions. It is not always possible with variables depending on only two inputs to find a high-accuracy solution to complicated problems such as NRWR prediction. It is observed that much more successful results can be achieved through the STDM structure, as suggested in this research, with the addition of a third variable to the TDM prediction models.

The prediction charts reflect close agreement between the measurement and calculation results. The STDM procedure provides every detail in spatial mapping better than regression models, which provide a smoothening effect through a crisp mathematical model.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Alegre, H., Baptista, J. M., Cabrera, E., Cubillo, F., Duarte, P., Hirner, W., Merkel, W. & Parena, R. 2016 *Performance Indicators for Water Supply Services*, 3rd edn. IWA Publishing, London, UK.

Boztaş, F., Özdemir, Ö., Durmuşçelebi, F. M. & Firat, M. 2019 Analyzing the effect of the unreported leakages in service connections of water distribution networks on non-revenue water. *International Journal of Environmental Science and Technology* **16** (8), 4393–4406.

Chondronasios, A., Gonelas, K., Kanakoudis, V., Patelis, M. & Korkana, P. 2017 Optimizing DMAs' formation in a water pipe network: the water aging and the operating pressure factors. *Journal of Hydroinformatics* **19** (6), 890–899.

Creaco, E. & Haidar, H. 2019 Multi objective optimization of control valve installation and DMA creation for reducing leakage in water distribution networks. *Journal of Water Resources Planning and Management* **145** (10), 04019046.

Creaco, E., Campisano, A., Fontana, N., Marini, G., Page, P. R. & Walski, T. 2019 Real time control of water distribution networks: a state-of-the-art review. *Water Research* **161**, 517–530.

Güngör, M., Yarar, U., Cantürk, Ü. & Fırat, M. 2019 Increasing performance of water distribution network by using pressure management and database integration. *Journal of Pipeline Systems Engineering and Practice* **10** (2), 04019003.

Güngör-Demirci, G., Lee, J., Keck, J., Guzzetta, R. & Yang, P. 2018 Determinants of non-revenue water for a water utility in california. *Journal of Water Supply: Research and Technology – AQUA* **67** (3), 270–278.

ISU (Kocaeli Water and Sewage Administrative General Directorate). 2018 *Kocaeli Water and Sewage Water Administrative: Annual Activity Report for Water Management*. ISU, Kocaeli, Turkey.

Jang, D. & Choi, G. 2017 Estimation of non-revenue water ratio for sustainable management using artificial neural network and Z-score in Incheon, Republic of Korea. *Sustainability* **9** (11), 1933.

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest

Jang, D. & Choi, G. 2018 Estimation of non-revenue water ratio using MRA and ANN in water distribution networks. *Water* **10** (1), 2.

Journel, A. & Huijbregts, A. 1978 *Mining Geostatistics*. Academic Press, London, UK.

Kanakoudis, V. & Muhammetoglu, H. 2014 Urban water pipe networks management towards non-revenue water reduction: two case studies from Greece and Turkey. *Clean – Soil, Air, Water* **42** (7), 880–892.

Kanakoudis, V. & Tsitsifli, S. 2010 Water volume vs. revenues oriented water balance calculation for urban water networks: The 'Minimum charge difference' component makes a difference!. In: *Proceedings of the IWA Int Special Conf Water Loss*, pp. 27–30.

Kanakoudis, V., Tsitsifli, S., Samaras, P., Zouboulis, A. & Banovec, P. 2013 A new set of water losses-related performance indicators focused on areas facing water scarcity conditions. *Desalination and Water Treatment* **51** (13–15), 2994–3010.

Kanakoudis, V., Tsitsifli, S. & Zouboulis, A. I. 2015 WATERLOSS project: developing from theory to practice an integrated approach towards NRW reduction in urban water systems. *Desalination and Water Treatment* **54** (8), 2147–2157.

Kanakoudis, V., Tsitsifli, S. & Demetriou, G. 2016 Applying an integrated methodology toward non-revenue water reduction: the case of Nicosia, Cyprus. *Desalination and Water Treatment* **57** (25), 11447–11461.

Kingdom, B., Liemberger, R. & Marin, P. 2006 The challenge of reducing non-revenue water (NRW) in developing countries – how the private sector can help : a look at performance-based service contracting. In: *Water Supply and Sanitation Sector Board Discussion Paper Series*, pp. 1–52. World Bank, Washington, DC. https://siteresources.worldbank.org/INTWSS/Resources/WSS8fin4.pdf.

Kizilöz, B. & Şişman, E. 2021 Exceedance probabilities of non-revenue water and performance analysis. *International Journal of Environmental Science and Technology*. https://doi.org/10.1007/s13762-020-03018-y.

Krige, D. G. 1966 Two-dimensional weighted moving average trend surfaces for ore evaluation. *Journal of the Southern African Institute of Mining and Metallurgy* **66**, 13–38.

Liemberger, R. & Wyatt, A. 2019 Quantifying the global non-revenue water problem. *Water Science and Technology: Water Supply* **19** (3), 831–837.

Matheron, G. 1963 Principles of geostatistics. *Economical Geology* **58**, 1246–1266.

Özger, M. & Sen, Z. 2007 Triple diagram method for the prediction of wave height and period. *Ocean Engineering* **34**, 1060–1068.

Republic of Turkey Ministry of Agriculture and Forestry (RTMAF) 2019 *Regulation on the Control of Water Losses in Water Supply and Distribution Systems (in Turkish)*.

Sen, Z. 2009 *Spatial Modeling Principles in Earth Sciences*, 1st edn. Springer, Dordrecht, pp. 203–269. https://doi.org/10.1007/978-1-4020-9672-3_1.

Shuang, Q., Liu, Y., Tang, Y., Liu, J. & Shuang, K. 2017 System reliability evaluation in water distribution networks with the impact of valves experiencing cascading failures. *Water* **9** (6), 413.

Sırdaş, S. & Şen, Z. 2003 Spatio-temporal drought analysis in the Trakya region, Turkey. *Hydrological Sciences Journal* **48** (5), 809–820.

Şişman, E. & Kizilöz, B. 2020a Artificial neural network system analysis and Kriging methodology for estimation of non-revenue water ratio. *Water Supply* **20**, 1871–1883.

Şişman, E. & Kizilöz, B. 2020b Trend-risk model for predicting non-revenue water: an application in Turkey. *Utilities Policy* **67**, 101137.

Tabesh, M., Roozbahani, A., Roghani, B., Faghihi, N. R. & Heydarzadeh, R. 2018 Risk assessment of factors influencing non-revenue water using Bayesian networks and fuzzy logic. *Water Resources Management* **32** (11), 3647–3670.

US Environmental Protection Agency 2006 *Decision-Support Tools for Predicting the Performance of Water Distribution and Wastewater Collection Systems*. Vol. EPA/600/R-02/029. US EPA, Washington, DC, USA.

First received 1 February 2021; accepted in revised form 27 May 2021. Available online 8 June 2021

Downloaded from http://iwaponline.com/ws/article-pdf/21/8/4263/969705/ws021084263.pdf
by guest