Condition Assessment and Rehabilitation for Trunk Sewer Deterioration based on Semi-Markov Model

Dr. Awatif Soaded Alsaqqar
Assistant professor
Engineering College - Baghdad University
d.alsaqqar@yahoo.com

Dr. Basim Hussein Khudair
Assistant professor
Engineering College - Baghdad University
basim22003@yahoo.com

Rehab Karim Jbbar
M.Sc. Student
Engineering College - Baghdad University
rehabkarim818@gmail.com

Abstract

An accurate assessment of the pipes' conditions is required for effective management of the trunk sewers. In this paper the semi-Markov model was developed and tested using the sewer dataset from the Zublin trunk sewer in Baghdad, Iraq, in order to evaluate the future performance of the sewer. For the development of this model the cumulative waiting time distribution of sewers was used in each condition that was derived directly from the sewer condition class and age data. Results showed that the semi-Markov model was inconsistent with the data by adopting (χ^2 test) and also, showed that the error in prediction is due to lack of data on the sewer waiting times at each condition state which can be solved by using successive condition inspection data for measuring the waiting times of the pipes at each condition class.

Keywords: semi-Markov process, deterioration model, Zublin trunk sewer, transition probability matrix, sewer conditions.

1. Introduction

The ageing of the sewers adversely affects their performance. There are two categories of these adverse effects: structural deterioration (e.g. breakage or deformation of pipes) and loss of serviceability (e.g. blockages or tree intrusion that reduces pipe's hydraulic efficiency) [13]. Thus ageing and deterioration cause cities to be susceptible to their effect of flooding and collapse. The solution of such problems was the ability of predicting the future performance of the infrastructure [2]. Condition ratings were used to estimate the structural deterioration of sewers in the form of discrete classes for reducing the computational complexity related with the system of continuous condition ratings [12].

A Markov chain process was developed for modeling the sanitary sewers deterioration and divided the asset life into four phases [1]. A stationary transition matrix was used upon expert opinions to compile these transition matrices and to characterize the deterioration in each phase. Result showed that the developed deterioration curves closely resembled the curves acquired based on expert opinions. They recommended that the work in the future should focus on integrating the automated models of sewer condition assessment, with deterioration models, as well as with management systems of infrastructure.

A good and significant results were found when modeling the likelihood of a sewer being in a deficient class using a multinomial log model considering age, effluent transported, material type, depth and diameter categories. Then, an ascending order of likelihood was used to rank the sewers, to provide list of priority for inspection [4].

The evaluation of sewer deterioration using a semi-Markov process was advised partitioning sewers into groups involving relatively homogenous characteristics, since different sewers could deteriorate at different rates [8]. The probability distribution of the waiting times in each class condition was derived based on continuous updating over time for expert opinion. Monte-Carlo simulation was used to calculate the cumulative waiting time distributions. The non-homogenous transition probability matrices in different classes were compiled using conditional survival probabilities.

This research, aims to develop a proactive methodology for evaluating Zublin sewer existing condition by considering the sewer condition class and age data. Based on these data, structural condition assessment model can be developed using the semi-Markov technique. This model is utilized to predict the probability mass function for the sewer over time and to predict it's future performance.

2. Material and Method

2.1 Case study description

The Zublin trunk sewer was chosen as a case study which established in 1983 with a total design capacity about 8.8 m^3/s. It is one of the main lines that collect sewage from Al-Rusafa side in Baghdad city, Iraq with an estimated total length around 25.4 km as shown in Fig. 1. This line starts at Al-Shaab municipality up to Safi al-Din Alhalli Street with diameters of 1800 and 2400 mm at depths of 3-7 m. Then, this line enters Al-Sadr city with a diameter of 3000 mm and continues with the same diameter until reaching Al-Habibia station at depths 7-9 m. After leaving Al-Habibia station, this line continues its flow until reaching Al-Rustamiya sewage treatment plant (3rd expansion) south of Baghdad with diameter 3000 mm at depths 6-10 m. This line is designed to serve about one million people in eastern Baghdad but with growing populations to nearly four million and not to conduct periodic maintenance of the line since the nineties this led to the deterioration of the line.

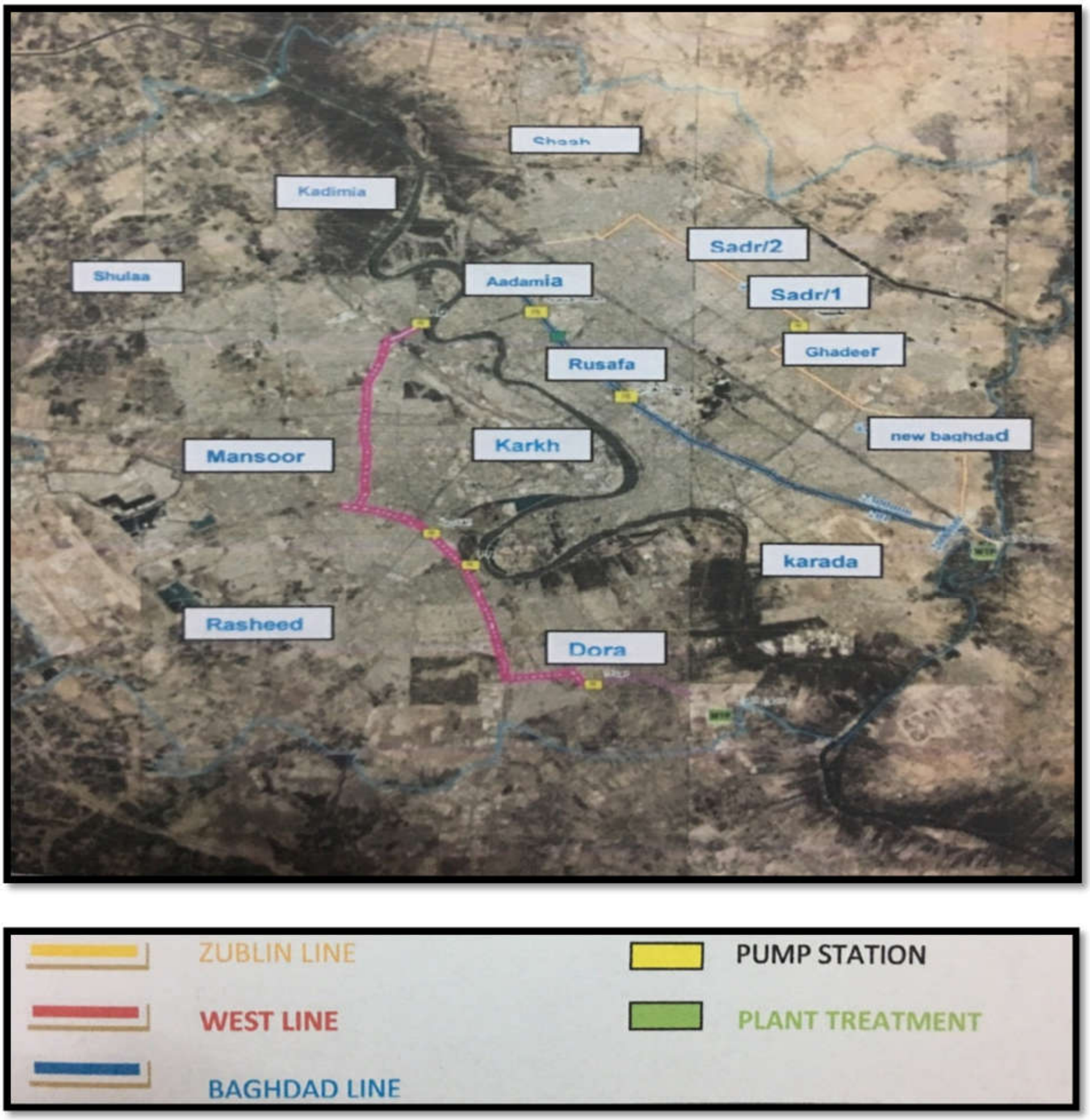

Fig. 1 Zublin trunk sewer layout

2.2 **Data collection and analysis**

In this paper, the data was obtained from distributing questionnaires on sewer management experts in Baghdad Mayoralty and to different municipalities of Al-Rusafa where Zublin line serves them. The data included: sewer condition and age. The entire dataset represented 99 samples, 79 were set aside for calibration and 20 were for validation as shown in Table 1. A condition rating system with five classes was used to describe sewers structural conditions as shown in Table 2. There are no sewers in condition class 1 and few samples in condition class 2 so they combined with condition class 3.

Table. 1 Calibration and test datasets Details for the semi-Markov model

Calibration Dataset					Test dataset				
Age year	Structural Condition			Total	Age year	Structural Condition			Total
	3	4	5			3	4	5	
26	1	0	0	1	28	1	0	0	1
27	2	0	0	2	29	2	0	0	2
28	5	0	0	5	30	1	0	0	1
29	4	0	0	4	33	2	0	0	2
30	1	0	0	1	34	0	2	0	2
31	4	0	0	4	35	0	2	0	2
32	1	0	0	1	36	1	1	0	2
33	1	3	0	4	37	1	0	0	1
34	2	7	0	9	38	1	0	0	1
35	1	6	0	7	40	1	0	3	4
36	1	9	0	10	43	0	0	2	2
37	0	1	1	2	Total	10	5	5	20
38	1	0	2	3					
39	0	0	1	1					
40	1	3	10	14					
43	0	0	7	7					
45	0	0	4	4					
Total	25	29	25	79					

Table. 2 Structural condition ratings description [18]

Case	Structural conditions	Physical description
1	Excellent condition	Only planned maintenance is required; no defect is detected in the system.
2	Very good	Minor maintenance is required as well as maintenance programmed and rehabilitation can be scheduled for long term construction, for example simple deformation of plastic pipes and simple crack.
3	Good	Significant maintenance is required; rehabilitation is necessary for medium-term within 3-5 year.
4	Poor	Significant renewal/upgrade is required, rehabilitation is keen and has must be completed in (one-two years), necessary emergency has

		to be check, construction damages with insufficient static safety and hydraulic or tightness.
5	Very poor	The network is unserviceable i.e. complete collapse of the sewer, rehabilitation stringent and short term maintenance to prevent collapse necessary, the sewer will be shortly be impermeable.

2.3 The semi-Markov model

2.3.1 Fundamentals of Markov processes

Markov-chains are a stochastic process that is used to describe the behavior of a system that passes through a countable number of possible condition classes [9]. It is a memory less random process in which the prediction of the future condition depends on the current condition only regardless of its history [16]. The current condition state of the system may be changed to another worse condition or remain in the same condition at each time step, depending on a given probability distribution [3]. Associated with this change a transition term which expresses the change in condition class and a term transition probability expresses the probability are used. The transition probabilities can be either time dependent or time independent, which means a non-homogenous and homogenous Markov model respectively [2]. For sewer deterioration modeling the first one is usually used since transition probability depends on pipe age and older pipes may deteriorate faster [8] and the condition changes were predicted for pipe population, but not individual pipes because of the lack of longitudinal data [17].

2.3.2 Model description

Mathematically, a matrix P with size (m x m) is used to express the transition probabilities with (m) being the condition classes number and i = m being the worst condition class. In this matrix, the row elements summation is always 1 and the sewer cannot improve its condition class without rehabilitation and intervention activities [9]. For simplicity in calculation, it can be assumed that each condition class can only drop one level (i.e. the transition can be only to the next worse condition class) during small time interval in one year, so most parts of the matrix are equal to zero [10].

$$P = \begin{bmatrix} p_{11}^{t,t+1} & p_{12}^{t,t+1} & 0 & \cdots & 0 \\ 0 & p_{22}^{t,t+1} & p_{23}^{t,t+1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & p_{m-1,m-1}^{t,t+1} & p_{m-1,m}^{t,t+!} \\ 0 & 0 & \cdots & 0 & p_{mm}^{t,t+1}=1 \end{bmatrix} \quad (1)$$

- $p_{11}^{t,t+1}$ is the probability that the sewer stays in condition i between time t and t+1
- $p_{12}^{t,t+1}$ is the probability that the sewer transits in the next worse condition between time t and t+1
- $p_{mm}^{t,t+1}$ =1 since class m is a failed class this means when a sewer enters this class it cannot leave it without intervention.

2.3.3 The semi-Markov models

In these models, the transition probabilities do not only depend on the sewer's current condition class, but also on the time already spent in this class [5]. It was assumed that the time spent in each condition class (waiting time) is a randomly distributed variable and are denoted as T1, T2,…, Tm-1 where their corresponding probability density function (pdf), a survival function (sf) and a cumulative density function (cdf) are denoted as $f_i(t)$, $S_i(t)$, $F_i(t)$ where i={1, 2, …, m-1} [8]. The time that process take it to go from class i to class k is denoted as T_{i-k}, which is a random variable representing the sum of waiting times in classes {i, i+1,…, k-1}. The mathematical expression of it is $T_{i-k}=\sum_{j=i}^{k-1} T_{j,j+1}$; i = {1, 2, ..., n-1}, k = {2, 3, ..., n}. In addition, $f_{i-k}(T_{i-k})$, $S_{i-k}(T_{i-k})$, $F_{i-k}(T_{i-k})$ are the pdf, sf and cdf of T_{i-k}, respectively. The conditional probability of the process to transit from class i at time t to the next class in the next time step (assumed to be 1 year) is given by [8]:

$$P_{i,i+1}(t)=\frac{f_{1\to i}(T_{1\to i})}{S_{1\to i}(T_{1\to i})-S_{1\to(i-1)}(T_{1\to(i-1)})} \; ; \quad i = \{1, 2, \ldots, m-1\} \qquad (2)$$

To populate the transition probability matrix of the semi-Markov model, the conditional probability in equation (2) is used. This process is non-stationary since the above transition probabilities are time-dependent. The infrastructure assets with the non-stationary deterioration property has been noticed by others, e.g., [7, 11, 12, and 6].

2.3.4 Prediction

After establishing the transition probability matrix P, the following equations can be used to model the deterioration process:

$A(t)= \{a_1^t, a_2^t, \ldots, a_m^t\};$

$$\sum_{i=1}^{m} a_i^t = 1 \quad (3)$$

Where: the m-dimensional vector A (t) represents the probability mass function (pmf) that is used to define the class of the process at any time t, a_i represents the probability of the process is in class i at time t while the pmf of the process at time (t +1) can be expressed as follows:

$$A\ (t+1)=A\ (t)P^{t,t+1} \quad (4)$$

To find the pmf of the process at time (t+s), the above equation can be extended as follows:

$$A(t+s)= A\ (t)P^{t,t+1}P^{t+1,t+2} \ldots P^{t+s-1,t+s} \quad (5)$$

2.3.5 Calibration

The determination of the cumulative waiting times' distributions is considered the key in the semi-Markov process application since the transition probability matrix Eq. 1 can easily be populated using Eq. 2, given the cumulative waiting times' pdfs and sfs. It was assumed that the waiting time Ti in any class i could be modeled as a random variable following Weibull distribution of a two-parameter [8]:

$$F_i(t)=P[T_i \leq t]= 1-e^{-(\frac{t}{\alpha_i})^{\beta_i}} \quad (6)$$

$$S_i(t)=1-F_i(t)= e^{-(\frac{t}{\alpha_i})^{\beta_i}} \quad (7)$$

$$f_i(t)=\frac{\partial F_i(t)}{\partial t} = \frac{\beta_i}{\alpha_i^{\beta}}\ t^{(\beta_i-1)}\ e^{-(\frac{t}{\alpha_i})^{\beta_i}} \quad (8)$$

Where α & β > 0 are the scale and shape parameters of the Weibull distribution respectively. The closed circuit television (CCTV) inspection technique that is used for gathering data of the sewer current condition normally just records the condition class of the pipe being inspected with its age at inspection without the waiting time in each condition class. This is greatly hampered the above modeling technique's application to sewer deterioration.

In this work, the (pdfs) & (sfs) of the sewers cumulative waiting times were derived directly from the sewer condition class and age data. Intuitively, the sewers' age in condition class 3 can be interpreted as the cumulative waiting times from their previous classes (i.e. condition classes 1 and 2) up to present classes. Hence the sewers' age in condition class 4 can also be interpreted as the cumulative waiting times from their previous classes (i.e. condition classes 1, 2 and 3) up to present classes, and so on. Based on the above assumption, the pdfs, cdfs and sfs of the cumulative waiting times can directly be derived from the sewer condition and age data. Here it was hypothesized that the data follow the Weibull distribution [8].

An empirical technique (regression method) was used for calculating the parameters of the Weibull distribution (α and β). The following steps are used based on the principles of least squares method [15]:

1. Linearize the (cdf) of the Weibull distribution (Equation 6) into the following form:

$$Y = a + bX \quad (9)$$

$$a = -\beta \ln(\alpha) \quad (10)$$

$$b = \beta \quad (11)$$

$$X = \ln(t) \quad (12)$$

$$Y = \ln(\ln(\frac{1}{1-F(t)})) \quad (13)$$

Where t the waiting times.

2. Sort in ascending order the N number of waiting times, t_i (i=1 to N) and assign a rank d_i to them.

3. Estimate the median rank of t_i as:

$$F(t_i) = \frac{d_i - 0.3}{N + 0.4} \quad (14)$$

4. The parameters a & b of Eq. 9 (i.e. intercept and slope respectively) were estimated using the principles of least squares according to the following equations:

$$a = \frac{\sum_{i=1}^{N} Y_i}{N} - b\frac{\sum_{i=1}^{N} X_i}{N} \quad (15)$$

And

$$b = \frac{\sum_{i=1}^{N} X_i Y_i - \frac{\sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i}{N}}{\sum_{i=1}^{N} X_i^2 - \frac{(\sum_{i=1}^{N} X_i)^2}{N}} \quad (16)$$

Where X_i and Y_i are calculated using Eq.12 and 13 respectively for each t_i.

5. Once a & b are determined, α & β can easily be determined from Eq. 10 and 11.
6. The correlation coefficient ρ is used to fit the distribution calculated as:

$$\rho = \frac{\sum_{i=1}^{N}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i-\bar{X})^2 \bullet \sum_{i=1}^{N}(Y_i-\bar{Y})^2}} \quad (17)$$

Where $\bar{X}$ is the mean of X_i and $\bar{Y}$ is the mean of Y_i and this procedure was implemented in Microsoft Excel.

3. RESULTS AND DISCUSSION

3.1 Parameters of Weibull distribution

Table 3 shows the estimated Weibull distribution's parameters of the sewers cumulative waiting times according to condition classes with their corresponding correlation coefficients.

Table. 3 The estimated Weibull distribution's parameters

condition classes	Weibull parameters		Correlation coefficient, ρ
	α [years]	B	
3	32.477	9.533	0.92
4	36.4	19.078	0.90

3.2 Sample prediction

To illustrate the use of the semi-Markov model, consider the sewers being in condition class 3 at time t = 26 year. These sewers will have a pmf of A (26) = {1,0,0}, since there is only one sewer which is in condition 3 (i.e. 100% is in condition class 3 and 0% are in condition classes 4 or 5 at t = 0). Fig. 2 shows the predicted pmf of the sewers beginning from time t = 27 until the age = 39 years, assuming no rehabilitation is achieved on them. It appears from this figure that sewers in condition 3 have high probability values at the beginning, but these values will decrease overtime, coupled with increasing in the probability values of sewers being in conditions 4 and 5. At age 35 year, for example, the sewers will have a pmf of A (35) = {0.0646, 0.4694, 0.466}. The algorithm of the semi-Markov model was implemented in MATLAB, to facilitate the prediction of the pmf and condition of the sewers.

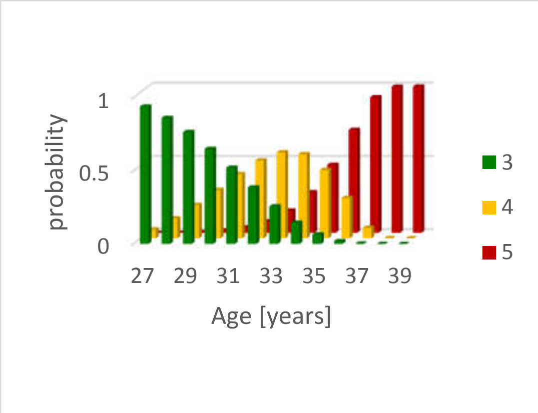

Fig. 2 Progression of sewers pmf over time.

3.3 Model performance evaluation

3.3.1 Goodness-of-fit test

This test showed how confidently the semi-Markov model fit with the observations. The Pearson chi-squared (χ^2) goodness-of-fit test was used with a null hypothesis ($H_\circ$) that the observed pipes number in condition m (O_m) is matched with the predicted pipes number in condition m (P_m) [13]. To conclude the fitness of the model a confidence level of 95% or 99% is often used [14]. The test statistic (χ_L^2) for this model can be calculated as follows:

$$\chi_L^2 = \sum_{m=1}^{3} \frac{(O_m - P_m)^2}{P_m} \quad (18)$$

The null hypothesis is rejected If the test statistic χ_L^2 is larger than the critical value of $\chi_{0.05,2}^2$ (the confidence level is 95% with 2 degree of freedom). To ensure the accuracy of χ_L^2 the predicted sewers number in any condition m must be at least 5 [14]. The expected number of pipes in each condition class over a time interval and the observed frequencies at the same time interval can be used to compute the test statistic (χ_L^2). The predicted number of pipes at a given time interval is determined by multiplying the pmf with the sewer stock.

Tables 4 and 5 show the results of this test on the calibration and validation datasets respectively. The χ^2 values provided unsatisfactory prediction by the semi-Markov model. From these tables it can be seen that the model underestimates the frequencies in conditions 3 and 4, while it overestimated the sewers frequencies in the more deteriorated condition class 5. Therefore, the deterioration of the sewers was overestimated by the model, which could be attributed to the assumption of estimating the cumulative waiting times from the sewers age at inspection and this result refers to the difficulty in applying this model.

Table. 4 Summary of χ^2 test for the calibration dataset

Condition class	No. of pipes		Chi-square χ_L^2
	Observed	Predicted	
3	25	16.3608	19.706
4	29	18.096	
5	25	44.5431	

* Statistically significant at 5% level if $\chi^2 > \chi^2_{0.05;2}$= 5.99

Table. 5 Summary of χ^2 test for the validation dataset

Condition class	No. of pipes		Chi-square χ_L^2
	Observed	Predicted	
3	10	4.0177	12.312
4	5	4.8255	
5	5	11.1567	

* Statistically significant at 5% level if $\chi^2 > \chi^2_{0.05;2}$= 5.99

4. CONCLUSIONS AND RECOMMENDATIONS

A non-homogeneous semi-Markov model has been presented in this paper for the structural deterioration of Zublin trunk sewer and provided a framework for predicting the sewer's future condition at the group pipe level. This type of information is extremely important in projecting budgetary requirements for sewer system maintenance and rehabilitation, and the success of proactive pipe maintenance strategies. In this model, sewers deterioration is represented as processes of transitions from a better condition class, (condition 1), to the next worst condition class, ending in the failed condition class, (condition 5), which illustrates the form of the age-dependent transition probabilities matrix of the model. The prediction ability of the semi-Markov model during calibration and validation using χ^2 goodness-of-fit test showed to be inappropriate for Zublin trunk sewer deterioration. The error of this model prediction seems to be rooted in the underestimation of the sewers cumulative waiting times at different condition classes, which can be solved by using successive condition inspection data for measuring the waiting times of the pipes at every condition class.

References

1. Abraham, D.M. and Wirahadikusumah, R., 1999. Development of prediction model for sewer deterioration. Proceedings of the 8th conference Durability of Building Materials and Components, Edited by M.A. Lacasse and D.J. Vanier, IRC, NRC, pp. 1257-1267, Vancouver.
2. Ana, E. V. and Bauwens, W., 2010. Modeling the structural deterioration of urban drainage pipes: The state-of-the-art in statistical methods. Urban Water Journal 7: 47-59.
3. Ana, E.V., 2009. Sewer asset management - sewer structural deterioration modeling and multicriteria decision making in sewer rehabilitation projects prioritization. PhD thesis, Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussel.
4. Ariaratnam, S.T., El-Assaly, A. and Yang, Y., 1999. Sewer Infrastructure assessment using logit statistical models. *Proceedings, Annual Conference of the Canadian Society for Civil Engineering*, Regina, pp.330-338, June.
5. Dirksen, J. and Clemens, F. H. L. R., 2008. Probabilistic modeling of sewer deterioration using inspection data. Water Science and Technology 57 (10): 1635-41.
6. Guignier, F. and Madanat, S. M., 1999. Optimisation of infrastructure systems maintenance and improvement policies. Journal of Infrastructure Systems, ASCE, 5(4), Dec.
7. Jiang, Y., Saito, M. and Sinha, K. C., 1989. Bridge performance prediction model using Markov chain. Transp. Res. Rec. 1180, Transportation Research Board, Washington DC, pp. 25-32.
8. Kleiner, Y., 2001. Scheduling inspection and renewal of large infrastructure assets. Journal of Infrastructure Systems 7(4): 136-143.
9. Kley, G. and Caradot, N., 2013. D 1.2 REVIEW OF SEWER DETERIORATION MODELS. KWB report, project SEMA, 2013 Berlin, Germany.
10. Le Gat, Y., 2008. Modeling the deterioration process of drainage pipelines. Urban Water Journal 5: 97-106.
11. Madanat, S. M., Karlaftis, M. G. and McCarthy, P. S., 1997. Probabilistic infrastructure deterioration models with panel data. Journal of Infrastructure Systems, ASCE, 3(1),
12. Madanat, S.M., Mishalani, R. and Ibrahim, W.H.W., 1995. Estimation

of infrastructure transition probabilities from condition rating data. Journal of Infrastructure Systems, ASCE, 1(2): 120-125.
13. Micevski, T., Kuszera, G. and Coombes, P., 2002. Markov model for storm water pipe deterioration. Journal of Infrastructure Systems 8(2): 49-56.
14. Montgomery, D. C., Runger, G. C. and Hubele, N. F., 2004. Engineering Statistics. John Wiley & Sons, Florida.
15. ReliaSoft, 2005. Life Data Analysis and Reliability Engineering Theory & Principles Reference. Tucson, Arizona: ReliaSoft Publishing.
16. Ross, S.M., 2000. Introduction to probability models. 7th edition. San Diego: Academic Press.
17. Tran, D.H., 2007. Investigation of deterioration models for stormwater pipe systems. Doctoral dissertation, School of Architectural, Civil and Mechanical Engineering, Faculty of Health, Engineering and Science, Victoria University, Australia.
18. Wagga Wagga city of council, 2010. Asset management plan.

تقييم اوضاع واعادة تأهيل لتدهور انابيب المجاري الناقلة بالاعتماد على نموذج شبه ماركوف

د.عواطف سؤدد عبدالحميد
استاذ مساعد
كلية الهندسة-جامعة بغداد
rehabkarim818@gmail.com

د.باسم حسين خضير
استاذ مساعد
كلية الهندسة-جامعة بغداد
basim22003@yahoo.com

رحاب كريم جبار
طالبة ماجستير
كلية الهندسة-جامعة بغداد
d.alsaqqar@yahoo.com

الخلاصة:

إن التقييم الدقيق لاوضاع الانابيب مطلوب من اجل الادارة الفعالة لانابيب المجاري الناقلة. في هذا البحث تم بناء نموذج شبه ماركوف واختباره باستخدام بيانات الصرف الصحي لخط زبلن الناقل في بغداد، العراق من أجل تقييم أداءه في المستقبل. لبناء هذا النموذج تم استخدام توزيع وقت الانتظار التراكمي لشبكات الصرف الصحي في كل حالة والتي كانت مستمدة مباشرة من بيانات حالة الصرف الصحي والعمر. وأظهرت النتائج أن النموذج شبه ماركوف لا يتفق مع البيانات من خلال اعتماد (اختبار $\chi2$) وكذلك، أظهرت أن الخطأ في التنبؤ يرجع إلى عدم وجود بيانات حول فترات الانتظار في كل حالة والتي يمكن حلها باستخدام بيانات التفتيش المتعاقبة لقياس فترات الانتظار في الأنابيب في كل حالة.