Modeling state of waste water system in Dakar, Senegal

Jacob Schennings
Rasmus Pettersson

Institutionen för informationsteknologi
Department of Information Technology
Abstract

Modeling state of waste water system in Dakar, Senegal

Jacob Schennings & Rasmus Pettersson

This report contains a method to mathematically determine the condition of pipes in a sewer system in Dakar, Senegal. First a survival analysis model is established based on probability density functions. The functions are then fitted to a data set consisting of the age of the sewer pipes and their inspected status. The status is ranked on a scale from 0 to 5 where 0 is a practically new pipe and 5 being a pipe with severe deterioration. The data fitting is done using the maximum likelihood function. Based on the established functions a prediction of the deterioration of the system is done, modeling how the proportion of pipes in different states develops.

It is found that a model describing the deterioration can be established. It is also found that a few hundred inspected pipes is sufficient to model large quantities of similar pipes with good accuracy as long as consistent classification of the pipes is applied.
1 Introduction

Between 1990 and 2015 12% of the population in Senegal gained access to improved facilities for sanitation (WHO/UNICEF 2015). Thus the system is under a very rapid development and methods to maintain the structural state of the pipes are necessary.

A major problem regarding pipes is the difficulties of inspections. In Dakar, sewers are almost never inspected without a reported default with non existing precautionary interventions as a result. This leads to unnecessary large and expensive interventions. If a preventive maintenance was established the system would be both more cost efficient and reliable.

In larger cities around the world, studies have been conducted to investigate the possibility to determine the status, or state, of sewer and waste water systems with the goal to simplify the preventive work and keep the sewer network in a good condition. For the past fifteen years the number of studies regarding the condition of urban sewer networks and the aging process of single sewer pipes have increased. The reason is the increased demand for knowledge and information about the condition or state of sewer network systems (Ana 2010). This information is crucial for governments and organizations that handle the maintenance of the networks. This process also involves preventive work on the network to minimize failure caused by pipe breakage. The importance of the preventive and maintenance work is described from three main perspectives.

The first is a societal perspective where a more well conditioned sewer network system means less discomfort and inconvenience for the public when a sewer system fails.

The second perspective regards the public health since an obstructed sewer pipe can cause a flood resulting in the creation of large basins of stagnated water. These basins provide a growth place for a variety of bacteria which is both a direct and indirect threat to the public health through ingestion or pollution of drinking water sources. In warmer climate they also provide a growth place for malaria mosquitoes which is a severe problem in most third world countries. A functioning water distribution and waste system drastically reduces the risk of disease spreading and contributes to a better public health.

It is also important from the economic perspective since a lot of resources are invested in these systems and the maintenance of them (Wiradikusumah et al 2001). Resources can be spared by reducing the emergency inspection and repairs caused when pipes suddenly fail.

This preventive maintenance can be planned based on statistical and probabilistic methods describing the conditions and development of the pipes.

One of the most important factors when considering the deterioration of a pipe is the age of the pipe and an older age has a significant effect of the probability of a pipe being in a deteriorated state. (Ana et al 2009). Based on age we can thus analyse the present pipes and use it to predict the future of the network. This prediction gives helpful insight when planning the maintenance and the economy.

In this paper we use the WRc standard from 2001 to classify the degree of the default of the pipes, on a scale from 0 to 5, and based on the classification it is possible to analyse the entire system and its development over time.

2 Background

2.1 The Millenium Development Goals

The Millenium Development Goals is a declaration of a vision to improve the situation for the most exposed people in the world. The program was established in 1990 and the objective is to reach the goals within 2015.
One of the goals is to halve the proportion of the population without sustainable access to safe drinking water and basic sanitation. The goal had already been achieved in 2010 but in 2012 there were still 748 million people relying on unsafe drinking water sources.

2.2 Senegal

Senegal is the most western African country, situated on the coastline between Mauretania and Guinea Bissau. The country was established out of a French colony in 1960 and has since been one of the most stable democracies in Africa (CIA 2015).

However, the economic development has been struggling. In 2011 46.7 percent of the population lived below the national poverty line, showing almost no decrease since 2005 (IMF 2013). The fresh water and sanitation sector is one of the most developed in sub-Saharan Africa and 98 percent of the urban population had access to fresh water in 2011. On the other side only 47 percent of the rural population had access to sanitation at the same time. One of the main challenges for the sector, in order to maintain the advancement, is to consolidate the financial viability (IMF 2013).

The climate in the country is tropical with a hot and humid environment and the uneven rainfall is stressing the waste water and sewer system. The rain season is occurring between May to November. Flooding occurs repeatedly in the capital Dakar during this season.

2.3 ONAS

The sanitation system in Senegal is maintained and developed by a public sanitation company ONAS, Office National de l’Assainissement du Senegal, which has been operating since 1996.

Unfortunately, the company is poorly funded due to the economic state of the government and most of the funding is provided by sanitation surcharge included in the water bills and foreign funding. These tariffs cover most of the operating costs but the room for development and expansion is small. As a consequence ONAS priority actions consist of mobilizing funding to replace aging sewer networks and achieve financial stability (UN WSP 2011).

2.4 Previous research

The focus of the previous studies conducted around the world have firstly been to determine which factors that are of highest significance in the aging process of single pipes and secondly how this knowledge can be implemented in a mathematical model which would combine the states of the single pipes to create a prediction of the overall state of the sewer network system.

According to previous studies there are mainly three approaches used (Ana 2010), (i) physical models, (ii) artificial intelligence-based models, and (iii) statistical models. As discussed by (Ana 2010) the three models have advantages and disadvantages regarding the precision of the prediction, the amount of data they require to create the prediction and the computation resources needed to execute the model.

2.4.1 The physical model

The physical model is constructed by the analysis of the physical properties of and around the pipe such as pipe material and the type of soil that surrounds the pipe and the different loads that affect the pipe. Together with the material deterioration caused by e.g. internal and external corrosion the aging process should then be determined. The problem with the physical method is that the generated models are often to simplistic to reflect the actual aging process since it is very complex in reality and not completely understood (Kleiner 2001).
2.4.2 The artificial intelligence-based model

The artificial intelligence-based model is contrary to the physical model much more capable of handling complex problems that can not be described by analytical and exact models. There are several methods that are artificial intelligent-based as described by (Ana 2010). However, these models are looked upon as 'black boxes' which means that it is hard to explicitly define causal relations of important parameters. This method also requires substantially more computing resources and a large amount of data to be calibrated which makes the method even less attractive since the availability of diverse accurate data is rare, even in developed countries.

2.4.3 The statistical model

The third and last model is the statistical model. This model relates the knowledge of the pipe condition, from e.g. visual inspections, to significant factors such as pipe material, age of the pipe and type of soil e.g. As discussed by (Ana 2010) there are two kinds of statistical models, the pipe group models and the pipe level models. The difference between these two is the formation of subgroups or cohorts in the pipe group model. Pipes sharing similar properties such as material, length, diameter etc. are placed in cohorts since they are expected to deteriorate similarly (Baur and Herz 2002). For the pipe level method each individual property of each pipe is taken into consideration when determining the deterioration of the pipe.

3 Method

3.1 Limitations

In order to simulate the deterioration of the Senegalese sewer system, there is a need to identify relevant and measurable variables. These variables consist of specific properties of the pipe on one hand and the corresponding deterioration of the pipe on the other hand. One natural assumption of a depending variable is the age of the pipe, where the deterioration increases as age increases. Other types of variables could include the material of the pipe, amount of water flowing or geographic locations. Due to the limit of this report and the difficulties of finding data describing the system this report handles a model based purely on how a sewer pipe system deteriorates with age. However, when suitable data is used the model can easily be extended. This is primarily done by performing the age analysis for different categories of pipes and identifying differences in the behavior of the system.

3.2 Choice of method

The chosen method is described in A Survival Analysis Model for Sewer Pipe Structural Deterioration (Duchesne et al). However the model has been extended to include all five possible stages in the WRCs standard of classifying pipes.

3.3 Survival model

The survival analysis model is based on probability density functions, PDFs. The PDF describes the probability of a pipe at a certain age staying in a specific state. From the PDFs, survival functions describing the probability that a pipe of a certain age will stay at least time $t$ in the specific state can be created. These functions are necessary due to the inspection of the pipes. The inspection only gives the current state of the pipe, without information of when transition
will occur. Hence, the deterioration will take place in some unknown future. The survival function \( S_i(t) \) is

\[
S_i(t) = \int_t^\infty f_i(u) du ,
\]

where, \( i \) corresponds to the state of the pipe.

As previously discussed, at inspection we only find information about the state and the age of the pipe. If for instance a pipe at stage 2 is found it is not possible to conclude anything about when the deterioration between the stages occurred. The age of the pipe corresponds to the time the pipe has been in the three stages 0, 1 and 2 in total and as a consequence we need to create cumulative staying times.

The cumulative PDF for a pipe staying in either state 0 or 1 is

\[
f_{01} = f_0 * f_1(t) = \int_0^t f_0(\tau) f_1(t - \tau) d\tau
\]

The surviving function corresponding to \( f_{01} \) describing the probability that a pipe of age \( t \) will be in state 0 or 1 can therefore be calculated as

\[
S_{01}(t) = \int_t^\infty f_{01}(u) du = 1 - \int_0^t f_{01}(u) du
\]

The model is developed further in the same manner

\[
f_{012} = f_{01} * f_2(t)
\]

for the PDF to be in state 0123 and 01234.

### 3.4 Choice of Probability Density Function

In order to fit data to the developed survival model a probability density function is necessary. Any distribution function is possible but in this paper the exponential is used giving the distribution function below

\[
f_i = a_i e^{-a_i t} \quad i = 0, 1, ..., 5
\]

where \( i \) represents the state.

### 3.5 Data fitting

In order to fit the measured data containing age and state of the pipes to the survival model the maximum likelihood method is used. The state of the pipes can be assumed to be independent of each other, meaning that the data is uncorrelated. A consequence of this is that the combined likelihood for all pipes can be calculated as the product of the probabilities for each single pipe. The likelihood for the survival function \( S_i \) is shown below where the data \( t_j \) is the age of the pipe \( j \), \( rc \) are the pipes in state \( i \), \( lc \) are the pipes outside of state \( i \) and \( n_{rc}, n_{lc} \) are the number of pipes in the corresponding group.

\[
L_i = \prod_{i \in rc} S_i(t_j) \prod_{j \in lc}(1 - S_i(t_j))
\]

The parameters to the corresponding survival functions are found where the maximum likelihood occurs. By successively maximizing \( L_j \) for the corresponding survival functions the parameters to each PDF can be determined and the entire model can be developed.
Practically the problem of finding the maximum likelihood can be treated as a minimization problem when inverted. This results in a linear problem which can be solved by e.g. the function \texttt{fminsearch} in Matlab which determines the parameter minimizing the inverted maximum likelihood function. However, before calculating the maximum likelihood the equation 6 was logarithmized to get the sum of the survival functions. This was performed since the product sum quickly reduces to a very small number that is too small for computers to handle. The logarithm gives

$$\log L_i = \sum_{i \in rc} \log S_i(t_j) + \sum_{j \in lc} \log(1 - S_i(t_j)) \quad j = 0, 012, 0123, 01234$$  \hspace{1cm} (7)

which then was inverted, by changing the sign, to work with \texttt{fminsearch} in Matlab since the function finds the minimum of the input function.

### 3.6 Modeling

Finally, when survival functions are estimated, the entire system can be modeled. Planning of maintenance and economy can then be done based on the model. As previously discussed the survival functions give the probability that a pipe of a certain age will stay in a specific state. Based on this probability we can estimate the proportion of pipes, of the entire system, that are in the specified state. The number of pipes of age \( y \) at state \( i \) is found by multiplying the number of pipes with the survival function \( S_i \). By summing over all ages \( y \) the total number of pipes in the state can be found and the proportion of pipes as a function of \( t \) can be calculated as below

$$P_0(t) = \frac{\sum_{y=1}^{n_{age}} (n_y(t)S_0(y))}{n_{tot}}$$  \hspace{1cm} (8)

where \( n_y \) denotes the number of pipes which are at age \( y \) at time \( t \) and \( n_{tot} \) denotes the total number of pipes in the database.

The proportion of pipes in the other five states can be calculated in the same way.

$$P_1(t) = \frac{\sum_{y=1}^{n_{age}} (n_y(t)S_{01}(y))}{n_{tot}} - P_0(t)$$  \hspace{1cm} (9)

$$P_2(t) = \frac{\sum_{y=1}^{n_{age}} (n_y(t)S_{012}(y))}{n_{tot}} - P_0(t) - P_1(t)$$  \hspace{1cm} (10)

$$P_3(t) = \frac{\sum_{y=1}^{n_{age}} (n_y(t)S_{0123}(y))}{n_{tot}} - P_0(t) - P_1(t) - P_2(t)$$  \hspace{1cm} (11)

$$P_4(t) = \frac{\sum_{y=1}^{n_{age}} (n_y(t)S_{01234}(y))}{n_{tot}} - P_0(t) - P_1(t) - P_2(t) - P_3(t)$$  \hspace{1cm} (12)

$$P_5(t) = 1 - P_0(t) - P_1(t) - P_2(t) - P_3(t) - P_4(t)$$  \hspace{1cm} (13)

### 3.7 Sources of error

The model has generally two sources of error. One cause of error in the modeling of the system is how the randomly inspected pipes are drawn. Naturally a larger set of pipes gives a better model describing the system but on the other hand a large number of inspections generates a greater cost. A well estimated sample size regarding these factors is desirable in order to efficiently model the system.
Another source of error is the classification of the pipes in different states. The classification is done by a human inspector and the result can be inconsistent depending on for instance measurement technique and education.

3.8 Stability analysis regarding sample size

The robustness of the model is determined by performing a stability analysis by altering the initial data the model uses to determine all parameters in the survival functions. The first part of the stability analysis is aimed to estimate the necessary sample size of pipes. This is done by investigating if the size of the data set used for model calibration from the database has any significance and if so by how much. This is done by randomly extracting \( m \) pipes from the database and calculating the parameter values for the survival functions.

3.9 Stability regarding inspection

The second stability test of the model is to study how the results alter with a consistent over- or undergrading of the state of the pipes. This will result in a weighting of the data sets. There are three cases of this misclassification, (1) consistent underestimation of state, (2) consistent overestimating of state and (3) random over- and underestimation of state. This can in reality represent an operator that conducts inspection of the sewer system with wrong measurement technique, thus constantly over- or underestimating the state of the pipe.

4 Results

4.1 Data fitting on the entire data base

The probability density functions are chosen based on the exponential distribution in (5).

<table>
<thead>
<tr>
<th>Function</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_0 )</td>
<td>( a_0 e^{-a_0 t} )</td>
</tr>
<tr>
<td>( f_1 )</td>
<td>( a_1 e^{-a_1 t} )</td>
</tr>
<tr>
<td>( f_2 )</td>
<td>( a_2 e^{-a_2 t} )</td>
</tr>
<tr>
<td>( f_3 )</td>
<td>( a_3 e^{-a_3 t} )</td>
</tr>
<tr>
<td>( f_4 )</td>
<td>( a_4 e^{-a_4 t} )</td>
</tr>
</tbody>
</table>

From the cumulative distribution function provided in (2) we can calculate the cumulative probability density functions which are shown below.

<table>
<thead>
<tr>
<th>Function</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{01} )</td>
<td>( \int_{\tau=0}^{\tau} f_0(t) f_1(t) d\tau )</td>
</tr>
<tr>
<td>( f_{012} )</td>
<td>( \int_{\tau=0}^{\tau} f_0(t) f_1(t) f_2(t) d\tau )</td>
</tr>
<tr>
<td>( f_{0123} )</td>
<td>( \int_{\tau=0}^{\tau} f_0(t) f_1(t) f_2(t) f_3(t) d\tau )</td>
</tr>
<tr>
<td>( f_{01234} )</td>
<td>( \int_{\tau=0}^{\tau} f_0(t) f_1(t) f_2(t) f_3(t) f_4(t) d\tau )</td>
</tr>
</tbody>
</table>

The cumulative density functions are as previously discussed used to calculate the cumulative survival functions and these are shown below.
In order to fit the necessary parameters $a_0$, $a_1$, $a_2$, $a_3$, $a_4$ to provided data it can be seen that they can be found step by step. First $a_0$ was found by finding the maximum likelihood for the function $L_0$ defined in (6). The found $a_0$ was then used when maximizing $L_01$, giving $a_1$ and so forth.

All function parameters are given in the table below and the corresponding survival functions are shown in Figure 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>0.0836</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.0887</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.0867</td>
</tr>
<tr>
<td>$a_3$</td>
<td>0.0732</td>
</tr>
<tr>
<td>$a_4$</td>
<td>0.0489</td>
</tr>
</tbody>
</table>

Figure 1 shows the probability of being in state group $i$ at age $t$. Naturally the probability of being in for instance state 0, 1, 2, 3 or 4 ($S_{01234}$) is higher than the probability of being in state 0, 1, 2 or 3 ($S_{0123}$) for all time.
4.2 Simulation of the system using the full data base

After establishing the survival functions, the entire system was modeled using the equations for calculating the proportions of pipes, (8 - 13). This gives an overview of how large proportions of the pipes that will be in the specific states in the future as a function of time. The result is shown in Figure 2.

It can be seen how the proportion of pipes in the low states, with small defects, decreases rapidly in the beginning. The proportion of pipes in higher states is increasing as pipes deteriorate in time. Since the system is modeled without replacement all pipes will deteriorate in the worse state 5, represented by the bright blue line.

![Proportion of pipes in state 0,1,2,3 and 4 without replacement](image)

*Figure 2: Proportion of pipes in the different states as a function of time*

4.3 Stability analysis regarding sample size

In this section a sample of pipes was drawn from the database. The previously discussed model was then used on the sample and the system was modeled. In Figure 3, the survival functions for different sample sizes are shown. Every sample size was randomly drawn 100 times out of the database containing 26 425 pipes and modeled.
Figure 3: Survival functions with different sample sizes

As can be seen in Figure 3 the survival functions seem to deviate more when a smaller sample size is used, giving a more insecure model of the system. This can also be noted from the standard deviation of the parameters derived by the maximum likelihood as shown in Table 1 decreases with an increased sample size.

Table 1: The standard deviation of the PDF parameters corresponding to number of pipes in sample.

<table>
<thead>
<tr>
<th>Data set size</th>
<th>50</th>
<th>100</th>
<th>500</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>0.0168</td>
<td>0.0095</td>
<td>0.0043</td>
<td>0.0034</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.0177</td>
<td>0.0154</td>
<td>0.0053</td>
<td>0.0037</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.0230</td>
<td>0.0153</td>
<td>0.0062</td>
<td>0.0044</td>
</tr>
<tr>
<td>$a_3$</td>
<td>0.0269</td>
<td>0.0137</td>
<td>0.0069</td>
<td>0.0048</td>
</tr>
<tr>
<td>$a_4$</td>
<td>0.0194</td>
<td>0.0143</td>
<td>0.0059</td>
<td>0.0045</td>
</tr>
</tbody>
</table>

The standard deviation in combination with the number of pipes is presented graphically in Figure 4.
It can be seen that the standard deviation does not decrease rapidly when more pipes are added in the data set. The very small increment of accuracy does not motivate the negative effects of large samples such as expensive inspections.

4.4 Stability with perturbation in data

In this section the case of incorrect inspection is investigated. A percentage of the pipes is perturbed by raising the state one step. The entire database is then modeled. This scenario corresponds to an inspector who systematically underestimates the state of the pipe. In Figure 5 the resulting parameters from the data fitting are shown as a function of how many percent of the pipes that are elevated to the higher state. As we can see all the parameters are increasing when the database is modified.

An natural approach is to investigate how the changed parameters affect the final simulation showing the proportion of pipes which are in a certain state. In Figure 6, a part of the pipes are elevated to a higher state and the corresponding proportions of the system are shown as a function of time. Naturally an underestimation of the system at inspection will result in a
model where the system deteriorates faster. This can be seen when comparing Figure 6a with Figure 6b, Figure 6c and Figure 6d. The latter shows a higher proportion of pipes in the worst state.

\[ |e| = \left| \frac{f(t, x + \delta x) - f(t, x)}{f(t, x)} \right| \leq \frac{x + \delta x - x}{x} = \frac{\delta x}{x} \]

The relationship guarantees that a perturbation in the initial data does not blow up in the final solution. Due to the structure of the model, where all pipes are moving towards the final state, a perturbed result will converge in time towards the unperturbed result. However, for small \( t \) the results deviate. If the model is investigated for results 20 years after inspection the following table can be built from Figure 6.

Figure 6: Proportion of pipes with underestimated inspection
Table 2: Proportion of pipes in states at age = 20 after perturbation.

<table>
<thead>
<tr>
<th>state</th>
<th>0% perturbation</th>
<th>20% perturbation</th>
<th>40% perturbation</th>
<th>90% perturbation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.026</td>
<td>0.002</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>1</td>
<td>0.070</td>
<td>0.059</td>
<td>0.051</td>
<td>0.004</td>
</tr>
<tr>
<td>2</td>
<td>0.118</td>
<td>0.108</td>
<td>0.100</td>
<td>0.009</td>
</tr>
<tr>
<td>3</td>
<td>0.172</td>
<td>0.162</td>
<td>0.154</td>
<td>0.143</td>
</tr>
<tr>
<td>4</td>
<td>0.234</td>
<td>0.221</td>
<td>0.213</td>
<td>0.193</td>
</tr>
<tr>
<td>5</td>
<td>0.380</td>
<td>0.431</td>
<td>0.498</td>
<td>0.563</td>
</tr>
</tbody>
</table>

From Table 2 the relative error can be calculated and plotted as a function of perturbation for all states. This is shown in Figure 7. It can be seen that the lowest states, 0, has a relative error larger than the perturbation for almost all changes in the database. It can also be seen that the relative error for the proportions of pipes that are in state 1 is above the critical area when the database is perturbed by a very high portion. However, the proportions of state 2, 3, 4 and 5 are satisfying the previously mentioned relationship for all sizes of perturbation.

5 Discussion

5.1 Implementation and usage of the model in Dakar

The ability to predict the future deterioration of the sewer pipe system is necessary to create a more sustainable system, both structurally and economically. This ability can be enhanced by using this model.
The implementation of this model in the ONAS organization in Senegal will not be instantaneous. First the current database of the pipes in Dakar must be completed with the respect to at least the age of the pipes. In the second stage visual inspection of pipes must be conducted to create the connection between age and the inspected state. After these procedures are completed the use of the model can be initiated.

The properties of this model result in a reduction in the number of inspected pipes that need to be carried out to be able to calibrate the model. There is no need to inspect all pipes in the system to be able to use the model. As we can see in Section 4.3 more investigated pipes will only give a small contribution to the accuracy of the results after a few hundred inspections.

The education of the inspectors can not be unnoticed and is necessary for providing accurate results. As we can see in Section 4.2 wrong measures affect the final results. If for instance there are five inspectors, inspecting 20% of the pipes each, and one inspector constantly underrates the state it will give large effects on the modeled proportions in the near future.

5.2 Other possible sources of error
There are a few sources of error that cause the model to produce faulty results. A few of them are discussed in the method section above.

Another important error that needs to be taken in consideration is the evolutionary evolution of the system. This results in an overestimation of the state of the system. Pipes in worse states, naturally, are replaced in a higher frequency compared to pipes in lower states. During this process, history about pipes is most often not preserved which means that as soon as a pipe is replaced the data about the state and age is reset in the database. The pipes that are still in the database, and are used in the modelling, are therefore better than the average pipe (Scheidegger 2010). By resetting the data on replaced pipes the knowledge of pipes with a short lifespan is lost and not included in the calibration process. The result of this missing data is that the model tends to overestimate the lifespan of pipes in the model since it is calibrated with data from pipes with a longer lifespan.

The implementation of this model comes with some initial costs. The most evident cost, and perhaps also the largest, regards the process of inspecting pipes. The conclusion that can be drawn from the simulation where 100 sets of sample data were drawn from the database is that the calibration process is sensitive regarding the number of pipes in each data set, corresponding to the number of pipes inspected. However, the standard deviation decreases exponentially as the number of pipes in each data set is increased.

5.3 Further development
The next step in the model development stage is to group pipes with similar properties in cohorts and calculate the survival functions for these subgroups. This increases the accuracy of the model since pipes with different properties e.g. material, diameter, length, type of soil etc. could age at a different rate. This is accomplished by introducing more variables in the data extraction procedure to group pipes in the initial calibration process. An implementation of cohorts would be a swift procedure but is not covered in this report since no correct data was available to use.
6 Acknowledgments

We would like to thank the Swedish International Development Agency and the International Science Programme at Uppsala University whose funding, through the Minor Field Study-program, has been necessary to carry through this project.

Thanks also to our supervisors in Senegal, Dr Babacar Touboun and Dr Mamadou Bousso, for helping us with technical aspects of this report. We would also like to thank them for all practical aid and the great hospitality during our stay. We would also like to thank our team colleagues Assane Gning and Awa Gaye for their participation in the project and their hospitality in showing us around in Senegal and the local culture.

We would also like to thank Pr. Per Lötstedt at Uppsala University for his support in the completion of this project.
References


