FIELD DATA TO ESTIMATE INFILTRATION INTO SEWER SYSTEMS

Torsten Franz, Mandy Rutsch, and Peter Krebs, Institute for urban water management, Dresden University of Technology

Abstract:

Within the European project APUSS (Assessing infiltration and exfiltration on the performance of urban sewer systems) a model is developed which allows the estimation of infiltration and exfiltration of an urban catchment and which is supported and accompanied by software development and application. Thus, a data driven model ought to be developed using the approach "similar pipe characteristics lead to similar infiltration rates", i.e. similarities between groups of pipes or sub-catchments are used to transfer available information on infiltration. In this paper, general rules for data acquisition (as first step of modelling) as well as results of a parameter reduction are introduced.

Keywords
ANOVA, data acquisition, data sources, infiltration, multidimensional scaling

1. Background

Parasite water is detrimental to the efficiency of sewers and wastewater treatment plants and thus an important cost factor. The infiltration is driven by the water head and the leakage of the pipe. However, the knowledge about cause-effect relationships between (i) pipe characteristics, pipe environment, and structural state of pipes and (ii) the structural state and infiltration rates is fairly limited [Rutsch, 2003]. The analysis shows that a new approach is necessary. With the assumption "similar pipe conditions lead to similar infiltration and exfiltration (I/E) rates" it is possible to look for and work with "similarities" within a sewer system (similarity approach). With a two-step-procedure containing classification and generalisation homogeneous areas and homogeneous groups of pipes, respectively, can be identified. Between these groups information should be assignable. The first step of the implementation of the similarity approach was the verification of the basic assumption "similar pipe conditions lead to similar I/E rates". By means of

- typical data sets of the Dresden sewer system
- methods of classical statistics (analysis of variance, least significant difference test)
- methods of exploratory data analysis (multidimensional scaling MDS)

similarities and dissimilarities of sub-catchments with different infiltration rates were studied. The investigation area was divided in sub-catchments, which fulfil a common set of boundary conditions. By using a variety of conventional (measuring) methods as basis, a more detailed view of the parasite water situation was expected.

2. Compiling of data sets for modelling

2.1. Data Acquisition

In general the data situation for large sewer systems is fair but not satisfying. Based on experiences in several German cities the sources listed in Table 1 are relevant. A successful acquisition depends mainly on a good co-operation with the data owners.

General rules are:

- The data owners - esp. the operator - should have a serious interest in the investigation.
- It must be clearly identified, which data are needed for which purpose in order to reach which goal.
- Public or public-like data owners are more generous with their data.

At the beginning, there are some general ideas about data requirements. Still, one does not know, if and where the data are available. A certain time factor must be considered. The data collection for a city requires several months, for a river shed more than half a year. In order to receive a minimum data quality the acquisition should focus on data quantity. Anyway, the strategy "take-what-you-can-get" is recommended.

Table 1: Data sources

data donors	**data formats**
Operator	graphic based
- sewer system	- Geographic information system
- water supply	- Network Information system
city authorities	- CAD
- environmental data	numeric based
- socio-economic data	- sewer register
regional authorities	- process control systems
- environmental data, esp. ground water	- mercantile software
	- measurements, simulations

2.2.Data pre-processing:

The pre-processing depends on the individual situation. The necessary steps are:

- basic part: digitalisation of analogue data, conversion into the formats needed, validation
- main part: linking of data, implementation into main software, adaptation to a consistent spatial-temporal environment
- additional part: simplification, time series analyses, distribution transformation, etc.

The parameters mentioned in Table 2 represent a typical data set. It is estimated that acquisition and pre-processing consume at least half of any project working time.

For the city of Dresden a data set was compiled containing reaches and sub-catchments characterised by independent parameters as well as measured infiltration rates. As a first step to link infiltration and parameters three questions were discussed.

3. Field data and infiltration

3.1.Are there significant differences within the subcatchment?

The answer is to reject the hypothesis that reach populations of sub-catchments do not belong to one basic population (reach population – mean sample of characteristics of single reaches, basic population – mean value of characteristics of the whole catchment). An "all-in-one" test was not available, because the considered parameters belong to different scales. Therefore, every parameter was analysed individually with one-way ANOVA (metric scale), Kruskal-Wallis-ANOVA (ordinal scale), and contingency tables (nominal scale), respectively. These methods compare mean values of groups to verify a significant inter-group difference. It could be shown that reach populations of sub-catchments with different infiltration rates differ significantly from each other in their characteristics. It can be concluded, that in general, there are significant dissimilarities between the reach populations of the sub-catchments.

Beside this general awareness, the relationships between certain groups are of interest. Therefore, least-significance-tests for metrical and multiple comparisons of mean ranks for ordinal parameters were conducted. These post-hoc tests compare groups with each other. The number of not rejected tests (i.e. two areas were defined as similar) was summarised and standardized. These results represent a ranking of parameters responsible for similarities among sub-catchments. A low number of rejected tests stand for a low importance for similarities. The condensed results are shown in Table 2. High percentages signify a high frequency of significant similarities. The result indicates the parameters most unimportant for dissimilarities: distance to buildings, distance to streets, reach length, and slope. This result is obvious, because these parameters have to be similar due to technical reasons and the purposes of sewerage.

Table 2: Frequency of significant similarities over all sub-catchments

parameter	**Dresden**
dist. to river Elbe	5%
population density	5%
dist. to drainage	6%
pop.-spec. length	6%
thickness cohesive layers	6%
dist. to storm sewers	9%
dist. to rivers	9%
groundwater level	11%
coverage	15%
date of construction	15%
profile circumference	16%
dist. to buildings	30%
dist. to streets	32%
length	39%
traffic load[1]	69%
Slope	85%

[1] linked to street type

3.2. Are the parameters sufficient to describe infiltration?

The sub-catchments can be seen as objects in an *n*-dimensional space with *n* as number of parameters. Multi-dimensional-scaling (MDS) transforms efficiently *n*-dimensional arrangements of objects (e.g. sub-catchments defined by *n* parameters) to low-dimensional configurations with optimal approximation of observed distances. The distances are an abstract measure for similarity/correlation [Borg and Groenen, 1997] that has, however, no physical meaning.

In using MDS the *n*-dimensional data transformed to one dimension were related to the corresponding infiltration rates. Good correlation between pipe-surface-specific infiltration rate q_f and groundwater-level weighted parameters was found (Figure 1), although only 6 subcatchments had the required groundwater information. Due to the reasonable correlation it can be concluded that the selected parameters - or a part of them - are sufficient to describe infiltration.

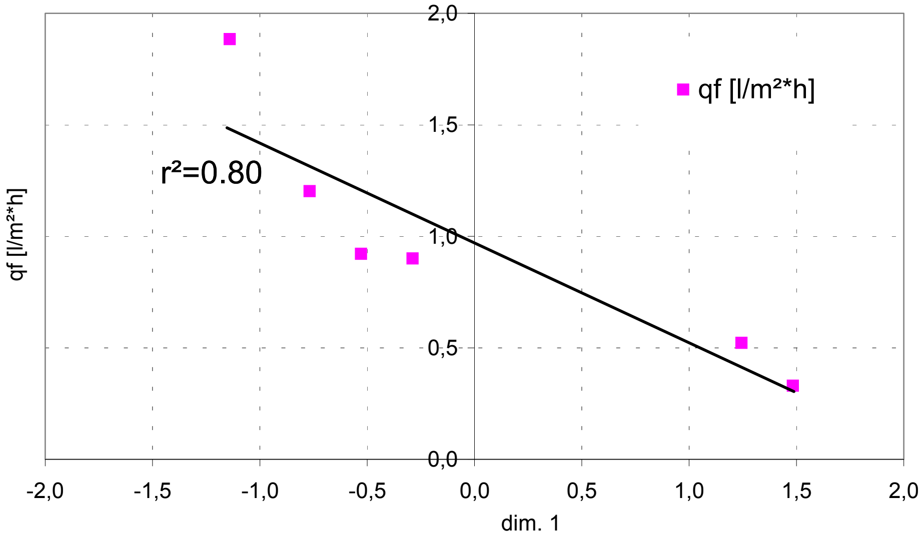

Figure 1: MDS result for sub-catchments vs. infiltration rate

3.3. Which parameters are relevant to describe dissimilarities between the sub-catchments – and so infiltration?

First results were obtained in analysing the similarity of the subcatchments not considering the infiltration rates (Table 2). Furthermore, MDS was used to investigate the parameters directly. The result of MDS applied to the parameters listed in Table 2 is shown in Figure 2. The more or less circular pattern of parameters (dashed line) indicates that a general merging for reduction purposes is not reasonable [Gifi, 1990]. The observed cluster effects (full lines) on structural data do not stand inevitably for a possible simplification potential. Profile type and circumference are correlated because the amount of discrete circumferences depends on the shape. Whether material, sewer system and date of construction can be simplified to a factor "construction period" must be investigated in future.

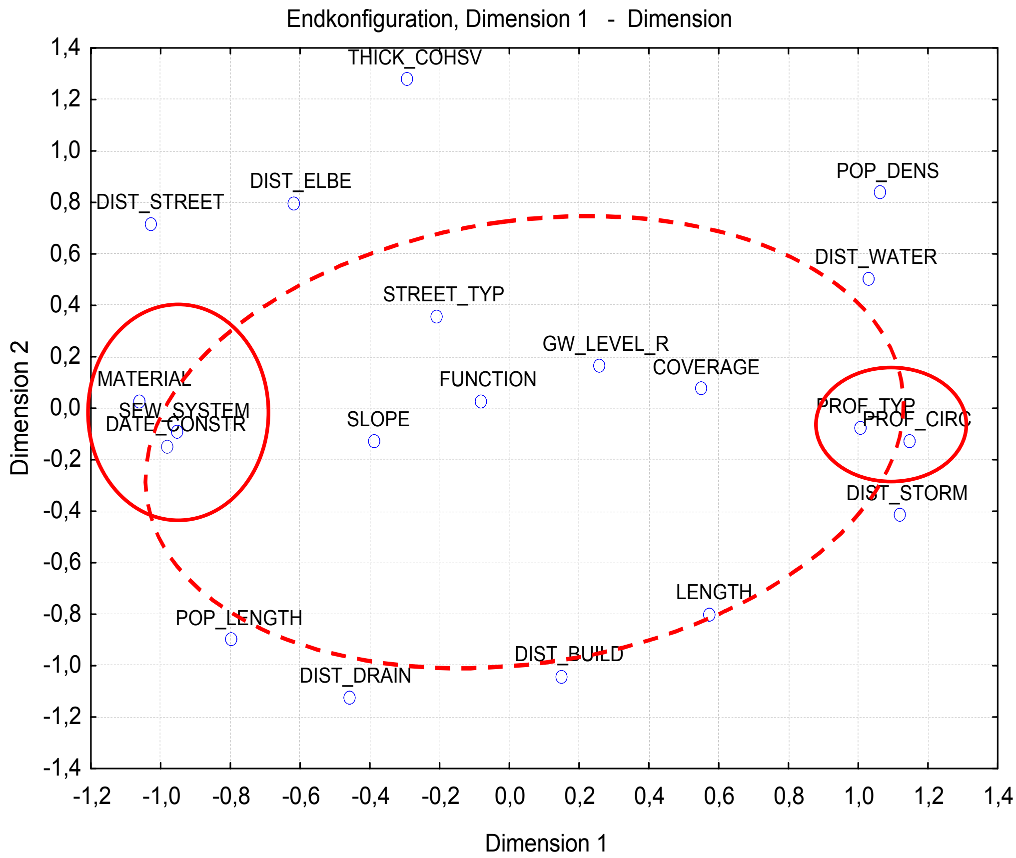

Figure 2: MDS results for parameters

An analysis of single pipes was possible by means of neural networks (NN) [Sarle, 1997]. Several training runs were made with leakage factors for single pipes [Karpf and Krebs, subm.] as output. Input parameters were varied in number and composition. The error between actual [see Karpf and Krebs, subm.) and modeled output averages to 9. The analysis of networks produces a weighted parameter ranking shown in Table 3. The ranking was performed through systematic non-consideration of single parameters. The greater the resulting error calculated from the NN the more significant is the left out parameter for the description of the leakage factor. Distance to buildings and distance to surface water are supposed to have the least influence on the leakage factor.

Table 3: Weighted ranking according to NNs

parameter	rank
COVERAGE	1.0
DATE_CONST	3.0
DIST_ELBE	3.2
DIST_STREET	3.5
POP_DENS	5.3
STREET_TYP	5.5
SURFACE[1]	5.6
DIST_BUILD	6.4
DIST_WATER	7.0

[1] Parameters length and circumference were combined to surface for simplification purposes.

In applying NN to analyse damage numbers [Rutsch and Uibrig, 2003] a reasonable result could not be obtained. Hence, a rank variance analysis was performed with the Mann-Whitney-Test or Wilcoxon-Test. This test is based on assigning ranks to the sample values. To detect differences between two populations the null hypothesis that two populations are identical is to be rejected. The null hypothesis can be rejected if the ranks associated with one sample tend to be larger than those of the other sample.

Approximately 7.2 % of the Dresden sewer system surveyed via closed-circuit-television (CCTV) was used for the analysis. They were attributed to their structural data/pipe characteristics. This sample is used to identify a possible relevance of structural data on the condition class calculated from CCTV records. The independent variables are material, construction period and dimension (Table 4).

Table 4: Compilation of criteria for the relevance analysis

material		dimension		construction period	
specification	number	specification	number	specification	number
concrete	1	< 300 mm	1	until 1899	1
clay	2	300 - 600 mm	2	1900 - 1916	2
		600 - 900 mm	3	1917 - 1928	3
		> 900 mm	4	1929 - 1945	4
				1946 - 1989	5
				since 1990	6

The obtained levels of significance are shown class-wise as cumulative function in Figure 3. The bisecting line represents the cumulative function of the significance random distribution. The further down and the further right the location of the cumulative function the more relevant is the criterion [Müller, 2002]. As an example the comparison of different dimensions is shown in Figure 3.

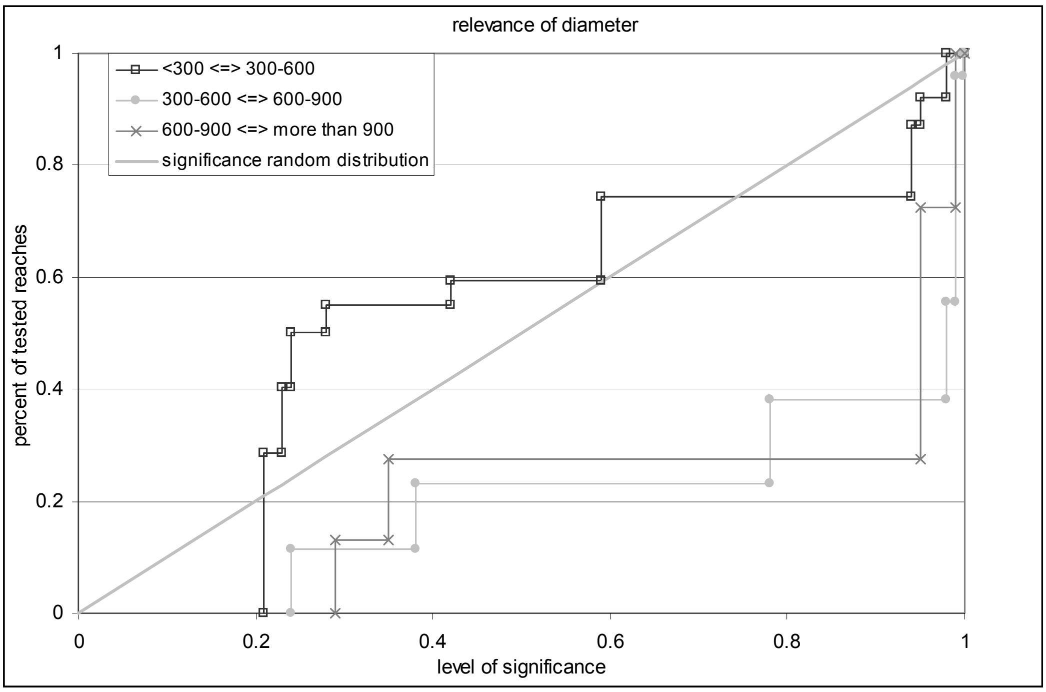

Figure 3: Relevance of pipe diameter for condition classes

The cumulative function of the diameters smaller than 600 mm does not show any relevance on the condition class, in contrary to the larger diameters. There is no difference in the defect frequency between small sewers (< 300 mm) and medium sewers (300-600 mm). However, the defect frequency of reaches with diameters of more than 600 mm differs significantly. The comparison of concrete and clay (materials with a sufficient high amount of data) revealed a difference in the defect frequency. Concerning the construction period, relevance on the defect frequency was expected also due to the construction periods chosen by historical development. But only two cumulative functions indicated a significant difference (construction period of 1946 to 1989 and of 1990 to today and tendentially the periods 1929 to 1945 and 1946 to 1989). The data availability is improving and with a larger dataset and finer resolution the content of information will increase.

4. Conclusions

In general, the quality of the various analyses is satisfactory but still not very good due to the relatively low number of available objects. It was not possible to pinpoint one parameter or a small group of parameters as being of overriding importance. The parameter groups external load (e.g. coverage, population density), groundwater situation, and construction circumstances (e.g. date of construction) are of high importance. Pipe location (e.g. distance to streets, buildings) and characteristics (e.g. profile type, circumference, length) are less important. Dependencies of sewer condition and parameters were demonstrated. Definition of construction periods would be very valuable.

By means of limited but typical data sets similarities and dissimilarities of sub-catchments with different infiltration rates as well as correlations between independent parameters and infiltration rates could be detected. Thus, the assumption “similar pipe conditions lead to similar I/E rates” can be the fundament for further research.

Acknowledgements:

This study has been carried out within the framework of the European research project APUSS (Assessing Infiltration and Exfiltration on the Performance of Urban Sewer Systems) which partners are INSA de LYON (FR), EAWAG (CH), Technical University of Dresden (DE), Faculty of Civil Engineering at University of Prague (CZ), DHI Hydroinform a.s. (CZ), Hydroprojekt a.s. (CZ), Middlesex University (UK), LNEC (PT), Emschergenossenschaft (DE) and IRSA-CNR (IT). APUSS is supported by the European Commission under the 5th Framework Programme and contributes to the implementation of the Key Action “Sustainable Management and Quality of Water” within the Energy, Environment and Sustainable Development Contract n° EVK1-CT-2000-00072.

References

Borg I. and Groenen P. [1997]: Modern Multidimensional Scaling. Springer New York Inc. ISBN 0-387-94845-7.

Gifi A. [1990]: Nonlinear multivariate analysis. John Wiley and Sons Ltd. Chichester. ISBN 0-471-92620-5.

Karpf C. and Krebs, P. [subm.]. Sewers as drainage systems – quantification of groundwater infiltration. Submitted for Novatech 2004, Lyon

Müller, K. [2002]. Entwicklung eines allgemein anwendbaren Verfahrens zur selektiven Erstinspektion von Abwasserkanälen und Anschlussleitungen. http://www.isa.rwth-aachen.de/content_templates/ctpl_list/ctpl_sitemap.asp

Rutsch M. [2003]: Factors expected to influence the leakage of sewers. Internal report for the APUSS project.

Rutsch M. and Uibrig H. [2003]: Classification System to Estimate the Leakage of Sewers. In: Baur R., Kropp I. and Herz R. [eds.]: Rehabilitation Management of Urban Infrastructure Networks. Proceedings of the 17th European Junior Scientist Workshop. pp. 153 – 158.