Valuing information for sewer replacement decisions
Wouter van Riel, Jeroen Langeveld, Paulien Herder and François Clemens

ABSTRACT
Decision-making for sewer asset management is partially based on intuition and often lacks explicit argumentation, hampering decision transparency and reproducibility. This is not to be preferred in light of public accountability and cost-effectiveness. It is unknown to what extent each decision criterion is appreciated by decision-makers. Further insight into this relative importance improves understanding of decision-making of sewer system managers. As such, a digital questionnaire (response ratio 43%), containing pairwise comparisons between 10 relevant information sources, was sent to all 407 municipalities in the Netherlands to analyse the relative importance and assess whether a shared frame of reasoning is present. Thurstone’s law of comparative judgment was used for analysis, combined with several consistency tests. Results show that camera inspections were valued highest, while pipe age was considered least important. The respondents were pretty consistent per individual and also showed consistency as a group. This indicated a common framework of reasoning among the group. The feedback of the group showed, however, the respondents found it difficult to make general comparisons without having a context. This indicates decision-making in practice is more likely to be steered by other mechanisms than purely combining information sources.

Key words | consistency, decision-making, information, paired comparisons, sewer asset management

INTRODUCTION
Sewer systems are critical infrastructures, because they provide essential services to society (Murray & Grubesic 2007): protecting public health and preventing urban flooding. Appropriate management of sewer systems, termed sewer asset management, is required to continue their service provision. In the Netherlands, approximately 1.5 billion Euro per year is currently spent for management of the 120,000 km of sewers (RIONED Foundation 2013), which equals 88 Euro/inhabitant/year. The largest portion of this budget, approximately 60%, is deployed for sewer replacement.

Sewer asset management typically is a public responsibility. As such, public accountability applies, implying preconditions could be specified for the entities responsible for sewer asset management. These are: the obligation to explain and justify a course of action, and efficient and effective management at lowest public costs. Especially this latter precondition, defined as cost-effectiveness by Katz & Kahn (1978), has received increased attention over the last decade. Decision-making transparency is required in order to meet these preconditions. This is important because it could enhance the integrity of public governance, could improve performance, and provides managers and citizens with input for judging the fairness, effectiveness and efficiency of governance (Bovens 2009). Decision-making for sewer asset management is, however, not transparent enough to specify its level of cost-effectiveness or to assess whether, where and how it could be improved.

Increasing transparency for sewer asset management, for example for sewer replacement decisions, is difficult. This has two causes: system and process complexity. System complexity refers to the many interactions between the objects and their direct surroundings, resulting in difficulties in determining the current structural condition, the deterioration process and the effects of management actions. Process complexity refers to the network of involved stakeholders, resulting in difficulties in assessing the interactions between relevant stakeholders, their

doi: 10.2166/wst.2016.253
interests, mutual negotiations and effects on the sewer asset manager. These complexities decrease the transparency of decision-making, which is not preferred in light of public accountability and cost-effectiveness.

Multiple decision support systems have been developed to assist sewer asset managers in optimising their maintenance planning, where ‘optimal’ may refer to, for example, lowest life cycle costs. These systems generally contain a mathematical optimisation procedure (single or multi-objective), a deterioration process and maintenance strategies. These normative decision support tools propose maintenance strategies over time to help the managers with their decision-making (e.g. Liu & Frangopol 2005; Sægrov et al. 2006; Marzouk & Omar 2012; Egger et al. 2015; Lounis & Daigle 2015; Tscheikner-Gratl et al. 2015). A general drawback of these tools is that they fail to meet decision-making in reality. First, the measures the tools propose are based on a relatively small set of data input: pipe age, CCTV inspection data and parameters about sewer system dimensions. Second, support tools that do include additional decision-making criteria, for example criticality of the sewer pipes or road works, use arbitrary and static weights for each decision criterion. Decision-making in reality, however, is based on a relatively large set of interrelated decision criteria, where the importance of each criterion varies per replacement project, depending on unique local circumstances and personal preferences of a sewer asset manager (Van Riel et al. 2014, 2016). Figure 1 illustrates this by showing the relations between a wide variety of decision-making criteria for 150 executed replacement projects, obtained through interviews. Each node represents a unique decision criterion and each connection represents a combination of criteria as mentioned by the interviewees. The degree of a node is defined as the number of connections to this node (Boccaletti et al. 2006).

Although the graph in Figure 1 suggests some decision criteria seem more important than others, no information is available yet about the actual relative importance decision criteria so far.

The objective of the presented study is to assess the weight of individual decision criteria for sewer replacement decisions, i.e. the extent to which each decision criterion is

![Figure 1](https://iwaponline.com/wst/article-pdf/74/4/796/459081/wst074040796.pdf)

**Figure 1** | Network graph of all information combinations. Node size represents degree. Edge thickness and colour intensity represent relative weight (frequency) of connection (Van Riel et al. (2016), reprinted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com/)).
appreciated or valued by decision-makers. Two aspects are considered: first, the perceived importance of decision criteria for hypothetical sewer replacement decisions, and second, the presence or absence of a shared frame of reference for judging about this relative importance. To this end, specific contextual information is excluded from the study. Further insight into these relative importances creates better understanding of the decision-making behaviour of sewer system managers. This understanding is required to increase decision transparency and cost-effectiveness in sewer asset management.

METHODS AND MATERIALS

Data collection instrument

A digital questionnaire was set up in ‘Survalyzer’ (software for online surveys), containing pairwise comparisons between relevant decision criteria. These criteria, the variables, were selected from Van Riel et al. (2016), in which decision argumentation of 150 sewer replacement projects in the Netherlands was analysed through interviews. Table 1 lists the mentioned decision criteria from these interviews mentioned most often. For each mentioned criterion, the underlying source and the information that is obtained from it are described. The right column shows how often the decision criterion was mentioned, normalised with respect to the total number of replacement projects. A project is a defined set of activities to replace sewer pipes at a predefined location, with a given budget and time limit. The length of pipe replacement per project in the Netherlands generally ranges from 10 to 500 metres. The retrieved decision criteria reflected a broad spectrum of argumentations, creating a representative data set, evaluated in a feedback session with experts from the urban drainage sector (Van Riel et al. 2016).

Citizens’ complaint calls concern feedback of the public about blocked gully pots, occurrence of stench or blocked household connection. These aspects may indicate insufficient hydraulic performance. Information about the occurrence of potholes or gaps in the road’s surface indicate ingress of soil in the sewer, influencing the stability of the road on top. This information is usually obtained through complaint calls or through feedback from maintenance activities. Hydraulic performance concerns the system’s transport capacity to minimise flooding. Environmental performance relates to the system’s storage capacity in order to minimise overflow of wastewater on surface water. Both aspects are obtained from a hydraulic model. Pipe age is data retrieved from the sewer system management database. It typically serves as an indicator for pipe quality for operational replacement decisions. Pipe age is important for strategic decisions as well, since it is the basis for long-term budget allocations. The planning of road or urban development works is communicated internally at the municipality. This may be through ad-hoc face to face discussions or through collective planning procedures. Uneven soil subsidence rates may cause open joints, fractures, loss of storage capacity

Table 1: Indicated decision argumentation mentioned more than five times (Van Riel et al. 2016, reprinted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com/))

<table>
<thead>
<tr>
<th>Source</th>
<th>Obtained information</th>
<th>Relative occurrence frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCTV inspection images</td>
<td>Pipe quality</td>
<td>0.57</td>
</tr>
<tr>
<td>Citizen’s complaints calls</td>
<td>Insufficient hydraulic performance (e.g. stench, gully pot blockages)</td>
<td>0.29</td>
</tr>
<tr>
<td>Complaint calls, feedback from maintenance activities</td>
<td>Gaps in the road</td>
<td>0.26</td>
</tr>
<tr>
<td>Hydraulic model</td>
<td>Hydraulic performance</td>
<td>0.19</td>
</tr>
<tr>
<td>Hydraulic model</td>
<td>Environmental performance</td>
<td>0.17</td>
</tr>
<tr>
<td>Sewer system management database</td>
<td>Pipe age</td>
<td>0.13</td>
</tr>
<tr>
<td>Road manager</td>
<td>Planning of road works</td>
<td>0.13</td>
</tr>
<tr>
<td>Urban planner</td>
<td>Planning of urban development</td>
<td>0.13</td>
</tr>
<tr>
<td>Soil subsidence measurements</td>
<td>Soil subsidence differences</td>
<td>0.08</td>
</tr>
<tr>
<td>Storm water policies</td>
<td>Preference to change system type or layout</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*aMentioned by the interviewees as such and included in the questionnaire in this wording.*
and fouling for sewer sections without pile foundations. Soil subsidence differences on a network level are measured by measuring the sewer invert level at the location of a manhole. Subsidence differences on pipe level are measured by analysing the sewer’s slope profile with the CCTV tractor (Dirksen et al. 2014). The obtained measurement data are converted typically by means of intuitive reasoning and rules of thumb into judgements about the severity of the subsidence rates, and consequently, need for sewer pipe replacement. Storm water policies relate to the organisational preferences about urban water strategies concerning system type and layout. This decision criterion concerns the long-term urban drainage strategy, including measures to cope with climate change (Kleidorfer et al. 2013), and is usually not a direct reason to replace sewer pipes but supplementary to other decision criteria.

The questionnaire was tested and adjusted twice before it was completed. Feedback of the first test session showed the wording of some variables were to be changed and the number of included variables were to be reduced from 15 to 10. A maximum of 10 variables was selected to minimise fatigue effects when filling in the survey. A second test was initiated to create a dataset that was used to evaluate the data analysis procedures for inconsistencies. The wording in the introduction and variable names were slightly changed for clarity. The final questionnaire started with an introduction of the research and an example of how to weigh and fill in the paired comparisons. Second, the respondents were asked about their gender, age, years of working experience, whether they work at a municipality and whether they work in an area prone to soil subsidence. Third, the 10 variables were randomly offered in a complete design (Street & Burgess 2007) in 45 pairs, asking respondents for a preference for one variable per pair. At the end, the respondents were thanked and asked for feedback.

Sample selection

The target population is Dutch sewer system managers. As such, judgmental sampling was applied. RIONED Foundation (centre of expertise in urban drainage in the Netherlands) were asked to distribute the survey, because they have contact data of all urban drainage departments at Dutch municipalities. On 25 November 2013, they e-mailed an invitation for participation in the survey to all 407 municipalities in the Netherlands (one e-mail per municipality). A reminder was sent on 3 December to increase the response.

Data analysis

The questionnaire was anonymous, implying no data about the respondents, their organisation and their sewer system was collected. Such metadata are not relevant here, because specific context is explicitly omitted from the questionnaire in order to assess whether a shared frame of reasoning is present.

The intangible property, weight or importance of information, was assessed by applying Thurstone’s law of comparative judgment, case V (Thurstone 1927b). Another common approach for comparative judgement is to use the analytic hierarchy process (AHP), developed for multi-criteria decision-making (Saaty 1987). The AHP is essentially an expansion of the pairwise comparison approach, creating a hierarchical structure of the decision criteria that are evaluated by paired comparisons. The main difference between the comparison procedure is that comparing pair through Thurstone’s approach requires the respondents to express a preference only, while the AHP requires the respondents to rate each preferred variable, usually on a 1–9 scale. The reciprocal of the given rate is associated with the variable in the same pair that is not rated, i.e. \( a_{ij} = 1/a_{ji} \). Hence, the AHP method puts an additional cognitive load on the respondents.

Thurstone’s model assumes that a variable’s quality is normally distributed on a psychological scale. It describes that different people may have different opinions on the quality of a variable. Each variable’s \( T \) quality score (the perceived value) is taken to be the mean quality of the corresponding normal distribution. Each respondent \( N \) is presented with every \( \frac{1}{2} \cdot (T^2 - T) \) possible pair of \( T \) items, and is asked which of two items is more favourable to the issue in question. An individual chooses the alternative with the highest perceived utility, which the respondent realises from the quality distributions of the two variables in the pair under consideration. For each pair of items the proportion is obtained (the empirical probability) of the times one variable was judged to be more favourable than the other. From the empirical probabilities of each pair, the mean quality score of each variable can be calculated using the normal cumulative density function.

A respondent is not always consistent in making a comparative judgment from one occasion to the next. An inconsistency occurs whenever a circular triad is present in the judgments (Kendall & Babington Smith 1940). A circular triad is illustrated as follows: item A is preferred over B, B over C, and C over A. The greater the number of circular triads in the data, the more
inconsistent a respondent is said to be (Thurstone 1927a). Consistency in the AHP approach is assessed with the ‘consistency ratio’ (Saaty 1987). It is interesting to assess whether individual consistency is correlated to the working experience of the respondents, because working experience is typically regarded as important to make sound decisions in sewer asset management. Next to internal consistency, validity of the results may also be analysed by determining concordance between judges by applying statistics described by Kendall (1938) and Kendall & Babington Smith (1940).

The following aspects were analysed:

1. The questionnaire results from Survalyzer were converted to a $T \times T$ comparison matrix for all respondents.
2. The mean quality scores per variable were calculated from the empirical probabilities in the comparison matrix.
3. The coefficient of consistence, zeta, per respondent was calculated. Zeta is the ratio of the number of circular triads each respondent makes and the maximum possible number of triads. A zeta of 0 equals complete inconsistency and 1 equals complete consistency.
4. Kendall’s tau test for every $\frac{1}{2}(N^2 - N)$ possible pair of respondents was applied to assess concordance between respondents. This non-parametric test computes the correlation between ranked data, with the test result ranging between $-1$ for complete disagreement and 1 for complete agreement; 0 equals no correlation.
5. The coefficient of agreement was calculated to assess concordance for the entire sample. This statistic, $u$, represents the extent of concordance for all judges together, where $u = 1$ equals complete agreement (Cohen 1960).
6. The relation between the number of years of working experience and individual consistency was assessed. Kendall’s tau test was used to test whether both variables are correlated.

The assumptions underlying the law of comparative judgment, case V are debatable (see Sjöberg 1962), especially equal and independent variance for all variables and between respondents. Yet, it is a reproducible approach to analyse intangible properties of information that provides plausible results. Next to that, the participants make choices in hypothetical situations, which can differ from their choice behaviour in reality. The goal of this study is, however, not to mimic reality, but to identify a general framework of reasoning.

### RESULTS AND DISCUSSION

#### Sample characteristics

The final response rate was 43%, yielding 177 completed responses from 407 invitations; 106 respondents (26%) left the questionnaire before finishing it, resulting in a non-response of 31%. The average completion time was 10.4 minutes. Table 2 shows several sample characteristics of the 177 completed responses.

The respondents not working at a municipality were excluded from the data for further analysis, resulting in a final sample size of 174.

The average age of the respondents is 46 years and the average working experience is 17 years, which indicates the urban drainage sector is relatively aged and experienced.

#### Variables’ quality scores

The 10 selected decision criteria were put onto a relative psychological scale, shown in Figure 2. The scale unit is expressed in the number of standard deviations from the mean quality score. The numbers do not have intrinsic meaning: they may be shifted by choosing another zero point or scale size and, thus, only indicate the relative distance between the points. Here, the least important variable was chosen as zero.

Figure 2 shows that sewer system managers perceive camera inspection images as the most important information source from these 10 variables. This can be explained by the fact that performing and evaluating inspections is normalised and often used in practice as the primary source of information, despite the drawbacks of the method (Dirksen et al. 2013). Information about road or urban development works was found to be important considering the

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Sample characteristics ($n = 177$) with rounded percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td>Male</td>
</tr>
<tr>
<td>Male</td>
<td>92%</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>&lt;30</td>
</tr>
<tr>
<td>3%</td>
<td>25%</td>
</tr>
<tr>
<td><strong>Years of working experience</strong></td>
<td>&lt;10</td>
</tr>
<tr>
<td>25%</td>
<td>33%</td>
</tr>
<tr>
<td><strong>Work at municipality?</strong></td>
<td>Yes</td>
</tr>
<tr>
<td>98%</td>
<td>2%</td>
</tr>
<tr>
<td><strong>Municipality in settlement prone area?</strong></td>
<td>Yes</td>
</tr>
<tr>
<td>31%</td>
<td>68%</td>
</tr>
</tbody>
</table>
initiation of sewer replacement projects (Van Riel et al. 2016). Apparently the respondents value camera inspections higher than information about the planning of other public works. It is surprising that citizens’ complaint calls were valued relatively high. The municipal complaint registration and solving procedures are usually found to be ineffective in practice. Gaps in the road were valued relatively high, although this type of information requires immediate action. Yet, such gaps typically do not occur along the length of the sewer section, implying local repair works are preferred over replacement of pipes. Environmental performance is considered less important than hydraulic performance. Loss of storage capacity may be compensated for by an increased cleaning frequency. Yet, replacing individual pipes may only partly compensate for loss of storage capacity, because environmental performance relates to the scale of the catchment while the replacement project relates to object scale. This may be the explanation for its positioning on the scale. Storm water policies are positioned plausibly on the scale since this criterion is supplementary to other decision criteria. Pipe age is considered least important for operational sewer replacement decisions, although it was mentioned as a criterion in 13% of the executed replacement projects (see Table 1). Overall, information about pipe quality obtained by CCTV data is valued higher than any other type of information, including information about the planning of other public works or measures for climate adaptation.

Several differences can be observed between the respondents working in areas with and without the influence of soil subsidence. First, information about soil subsidence was considered least important for areas that were not prone to it. This information source is considered significantly, and logically, more important in areas prone to subsidence. Second, camera inspections were considered more important in areas that are not prone to subsidence. A possible reason is that an important failure mechanism in stable soils (pipe degradation) is more easily detectable by camera inspection than an important failure mechanism in settling soils (change of storage capacity and hydraulic performance). Therefore, the usefulness of the information source might be perceived higher, depending on local soil conditions. Third, hydraulic models to assess hydraulic and environmental performance were also considered more important in areas that are not prone to subsidence. A possible explanation is that hydraulic models have a higher chance of producing useful results, at least for dry weather conditions, because of a lack of data about uneven changes in sewer pipe gradients and effect on hydraulic performance.

**Consistency and concordance**

Are the respondents’ answers consistent? As indicated in the section ‘Data analysis’, the spread between the variables is an indication of the perceived quality difference. It is also an indication of the respondents’ capability to discriminate between the variables’ qualities. Figure 3 shows the zeta statistics in a cumulative probability distribution. Figure 3 shows that the majority of the respondents, approximately 70%, have a zeta value of at least 0.8. This
means that the group is fairly consistent in their judgments, implying that they are capable of discriminating between the variables. Thus, it is concluded that most of the respondents are trustworthy judges. It also implies that small differences between variable qualities are probably caused by the fact that the quality differences are small, i.e. almost equally important information.

Figure 4 shows the relation between the years of working experience and individual coefficient of consistence. Figure 4 shows no clear relation between the coefficient of consistence and the number of years of working experience. The result from Kendall’s tau test ($\tau = 0.76 \times 10^{-2}$) shows both variables are approximately uncorrelated.

These results suggest that the assumption of more working experience equals higher consistency does not hold.

Do the respondents agree with each other, regardless of their consistency? Figure 5 shows the results in a cumulative distribution function. Agreement means that respondents agree both in their consistencies and their inconsistencies.

Figure 5 shows that the data are approximately normally distributed, implying the sample mean is the best estimator to judge respondents’ overall concordance. The sample means are higher than zero (approximately 0.2), indicating reasonable concordance between the judges. This suggests the presence of a shared frame of reference for judging the relative value of decision criteria. The coefficient of agreement supports this suggestion.

The coefficient of agreement $u$ for the total sample equals 0.21. For settlement prone areas: $u = 0.19$. For areas not prone to settlement: $u = 0.24$. These results show that sewer system managers in areas not prone to subsidence agree slightly more with each other compared to the other groups. Figure 2 supports this result by a larger spread of the variables. All three $u$ values are statistically significant at the 95% confidence interval ($p \ll 0.001$). These results mean that the respondents show significant agreement in their judgments, i.e. the judging is not done at random and a common line of thinking is apparent.

The applied Thurstone’s model of paired comparison may be used to produce weights in multi-criteria decision support models. Yet, given the importance of specific local circumstances, the paired comparison procedure would have to be repeated for every setting the decision support tool is applied to.
Relation to decision-making in reality

The unique circumstances of a real sewer replacement project were omitted in this study. This could decrease the agreement between respondents, because they judge their preferences from different perspectives, i.e. they use different frames of reference for their judgments. Several respondents mentioned this in the feedback section at the end of the questionnaire. They found it difficult to make a preference judgment at each pair, because they missed context. For example, in replacement project 1, they would prefer variable A over B, but would choose variable B over A in replacement project 2 depending on local circumstances. The content of these comments shows that deciding about sewer replacement is an art of fine tuning, combining and negotiating about available information and interests of other actors, due to a variety of local circumstances. This does not mean, however, that a common frame of reference is absent.

CONCLUSIONS

This study was aimed at analysing the perceived importance of decision criteria in hypothetical sewer replacement decisions and the presence of a shared frame of reasoning among Dutch sewer system managers. It is concluded that conventional CCTV images are valued most and that the majority of decision criteria are supplementary to this decision criterion. It suggest that the theoretical replacement strategy is primarily system oriented. This conclusion supports the analysis of the executed replacement projects.

A shared frame of reasoning about the relative value of decision criteria is indeed present. Despite the important influence of specific local circumstances, sewer asset managers appear to be comparable in their manner to judge the importance of information.

Working experience is not correlated with consistency in judging the value of information. The described results allow taking a peek into the way sewer system managers weigh or value sources of information relevant for initiating replacement decisions. Although this shared frame is present, the respondents’ feedback implies that purely combining information sources cannot drive the decision process for sewer replacement, although this is essentially how current decision models portray sewer replacement decisions. The trade-off of interests, values and information other than conventional camera inspection images plays a prominent role, which is neglected in current decision models for sewer asset management. Therefore, it is recommended to introduce relevant intangible decision factors into current decision models. To do so, decision processes in sewer asset management should be analysed in relation to their context due to the important influence of specific local circumstances. In order to close the gap between support tools and reality, model developers could pay more attention to multi-criteria decision support tools that can incorporate tacit next to explicit knowledge. Of course, the incorporated tacit knowledge needs to be properly motivated and evaluated frequently, because it should remain transparent for every user and insights may evolve over time.

The presented results illustrate how Dutch sewer asset managers make their replacement decisions. Despite specific local circumstance influencing sewer replacement decisions in reality, a common frame of reasoning about the relative value of decision criteria is present. This might be explained by the sewer asset management education in the Netherlands. As such, it may be interesting to study the existence of a common frame of reasoning in other countries, having differences in, for example, education provision, organisational setup and culture, sewer asset management challenges, perception of citizens and available budget.

ACKNOWLEDGEMENTS

The research is performed within the Dutch ‘Kennisprogramma Urban Drainage’ (Knowledge Programme Urban Drainage). The involved parties are: ARCADIS, Deltares, Evides, Gemeente Almere, Gemeente Arnhem, Gemeente Breda, Gemeente ’s-Gravenhage, Gemeentewerken Rotterdam, Gemeente Utrecht, GMB Rioleringstechniek, Grontmij, KWR Watercycle Research Institute, Royal HaskoningDHV, Stichting RIONED, STOWA, Tauw, vandervalk + degroot, Waterboard De Dommel, Waternet and Witteveen + Bos.

REFERENCES


First received 12 January 2016; accepted in revised form 17 May 2016. Available online 1 June 2016