Data in Urban Drainage Studies

Ian Garside MSc MIWEM Senior Engineer
Montgomery Watson
St James Court
Wilderspool Causeway
Warrington
WA4 6PS

1. Use of Data in Urban Drainage Studies

The data gathered to describe assets are the most important input to urban drainage studies.

Studies cannot be undertaken without this data and studies cannot be effectively and economically undertaken if the data is of dubious quality.

When considering the subject of data and data quality one must ensure that the data is fit for its end use and all intermediate tasks required of it.

The end use of this data is principally to evaluate the present and future performance of the asset either individually or part of a network. The principal inputs to this process are twofold.

i) The raw data and
ii) The manipulation techniques applied to the data

The manipulation techniques used will not be considered in detail here. This paper will concentrate on the raw data.

The raw data is initially used to describe the physical attributes of the sewer system. After manipulation the performance of the system is evaluated. Following system assessment the data is used to evaluate the optimum solution to any problems which exist in the system.

In developing solutions to system deficiencies the largest Capital Cost Savings can be made in this planning phase. It is therefore crucial that this data and its interpretation be as correct as possible. The cost reduction potential for each phase of a capital project is illustrated in Figure 1.

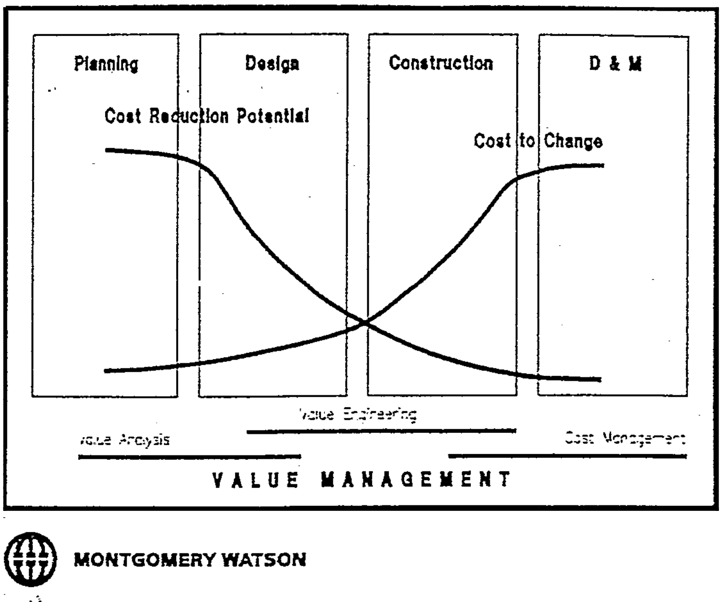

Figure 1 : Value Management

2. Review of Data Sources

There are a large number of data sources involved in the development of Urban Drainage Studies. These are reviewed below;

(i) Manhole Location and Survey Data

This data is measured on site and entered into a proprietary database such as SUS25, or STC25, etc. These databases have validation procedures which ensure that the data entered is consistent. A third party check is normally undertaken on site to validate measurements. This data is usually collected and validated in accordance with the WRc / WAA Model Contract Document with amendments made by clients as deemed appropriate. This data refers to the location and attributes of manholes and links between the manholes.

(ii) CCTV Survey Data

Results from CCTV survey are usually in two forms. The first form is video tape recordings of the survey. The second is a series of information entered by the operative of the location of type and extent of existing or potential structural defects. The data relates to the same entities described by the manhole location survey. However, it is often difficult to cross reference the manhole and CCTV data.

The data is gathered in accordance with the appropriate WRc / WAA Model Contract Documentation. The data is often entered on PRISM or Examiner databases.

(iii) Contributory Area Data

Results from contributory area surveys either undertaken by means of a field survey or a desk study are often kept separately from the network model until the two data sets are incorporated at model test stage. These surveys whether undertaken in the field or at the desk are undertaken by a sampling exercise and often on maps which are out of date. Such surveys are frequently undertaken independently from surveys undertaken to determine property numbers and densities for dry weather flow analysis.

(iv) Flow Data

Flow data is collected in accordance with WRc / WAA Model Contract Documentation, again amended to satisfy individual client requirements. This data is often gathered using quality systems. Third party software is available for the client to perform an independent check.

(v) Silt and Sediment Data

Silt and Sediment data is collected in an often 'ad-hoc' manner and no distinction is made on the type of sediment fractions encountered. Quantitative data concerning the occurrence of silt and sediment is of upmost importance when concerning the quality of effluent in the sewer system.

(vi) 'Point' Feature Data

Collection of data concerning the operation of CSOs, detention tanks, pumping station and in sewer control devices is often undertaken by the modelling engineer or technician. There are no published standards to which the data must conform.

The data collection requirement is large and the quality of data collected varies.

3. Sources of Data Errors

Data errors are often introduced to the process for numerous reasons. Some errors are systematic; these are the most important type of error. Errors which result in an apparent random fluctuation of data are less important as the effect is to produce data around a mean figure which is often the correct figure.

Examples of systematic data errors are as follows;

- Errors of interpretation in Internal Condition Grade Assessment
- Interpretation errors made during measurement of node features
- Incorrect specification of data for use with hydraulic model. Errors often relate to the use of the WALLRUS runoff model(s).

Examples of errors of a random nature are as follows;

- Incorrect measurement of manhole levels and depths. (Errors may be included whilst the measurement is within the specified tolerance). The occurrence of such errors often revolve around a mean, that value being the correct value. The errors will typically form a 'normal' distribution around this mean figure.
- Raingauge readings. Raingauges can either over read or under read depending upon

their location relative to macro - or micro environmental factors such as wind or shelter.
- Flow measurement. As with raingauges flow meters tend to over read or under read depending on environmental conditions in the sewer. However, errors will not be 'normally' distributed, as proportionally more readings will under read than over read.
- Sampling of contributory area and the consequential factoring up of data can introduce errors. As with raingauge data these errors will be normally distributed around the mean or correct value. The amount of skewness around this value will depend on the type of development being sampled, the sampling rate and the rate of development transience.

The random errors described above are acceptable to the industry as they reflect economic methods of data collection. The resultant data is considered as 'fit for purpose'.

The systematic errors are often unacceptable.

4. Finding Data Errors

Since the data gathered in such exercises cover a wide variety of sources and the resultant datasets can be huge, the potential for the introduction of both 'Systematic' and 'Random' errors in large.

Even when good quality systems are in place the presence of errors can have serious consequences.

The most common method of identifying errors is model verification. Experienced practitioners undertaking the verification exercise can easily identify errors using this technique. Finding data errors in this way is an iterative procedure, however. It is often better to establish systems which can identify and thereby allow for error elimination before the costly exercise of verification is embarked upon.

The process of finding data errors can be made more comprehensive and quicker if the large data set is organised properly and is easily cross-referenced. Furthermore, it is crucial that the data which inevitably refers to the attributes on performance of pipes is related to that pipe. Reference to that pipe can then be easily made.

5. Consequences of Data Errors

In looking at the whole question of data collection and data errors it is relevant to discuss the potential consequences of allowing data errors into the capital scheme design process. It should be stressed that any suitable system of data validation should be capable of identifying the errors highlighted here.

Case 1: Application to the Prediction of Hydraulic Inadequacy

The first example is that of a small sewer system whose correct dimensions have been coded into a WALLRUS model. A parallel model has been developed to represent a scenario where each successive manhole working downstream are at the maximum tolerances of measurement, ie., the levels are + or - 25 mm of the true value. Figure 2 is a schematic of the network and Figure 3 illustrates how the levels may have been measured.

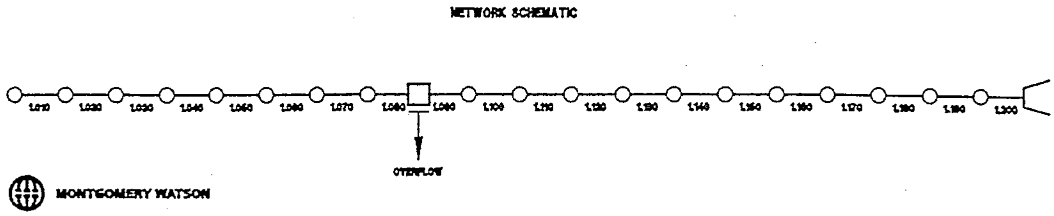

Figure 2 :

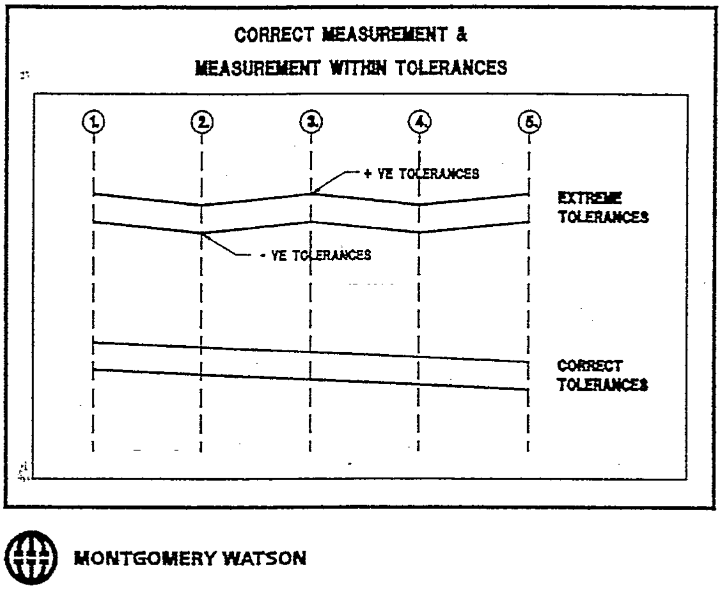

Figure 3 :

This second model probably represents the worst case scenario in terms of adherence with tolerances, but not the worst case had more, systematic errors, been made.

The effect of applying the extreme tolerance on the system is produce a different shape of hydrograph of the outfall characterised by less attenuation and a more pronounced peak. This is illustrated in Figure 4. More seriously is the fact that in the catchment, at pipe location 1.150, the surcharge from the extreme tolerance model is way in excess of that for the correct model. (See Figure 5).

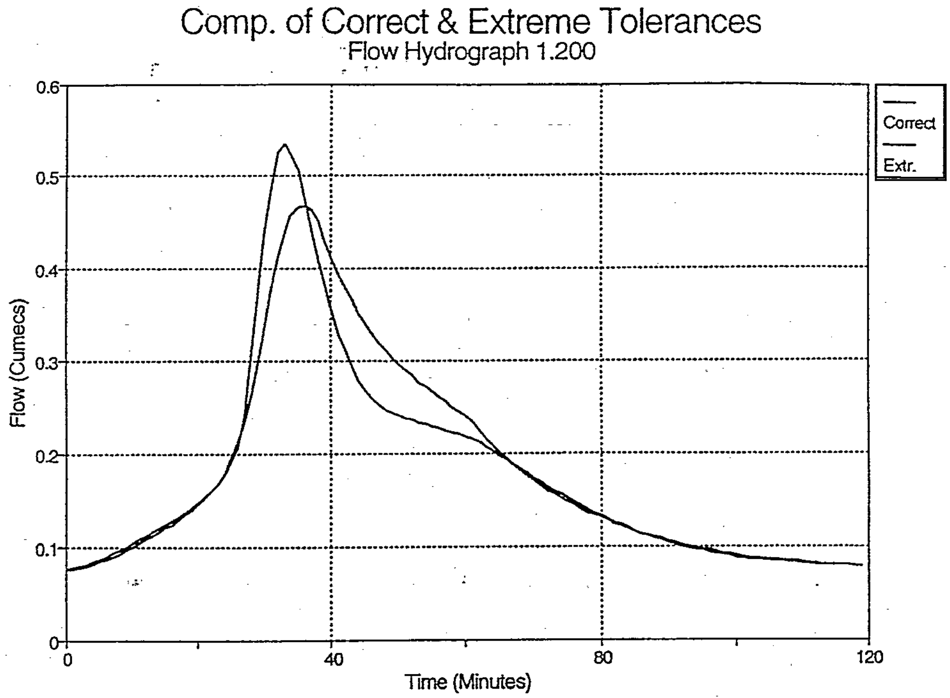

Figure 4 : Outfall Hydrograph

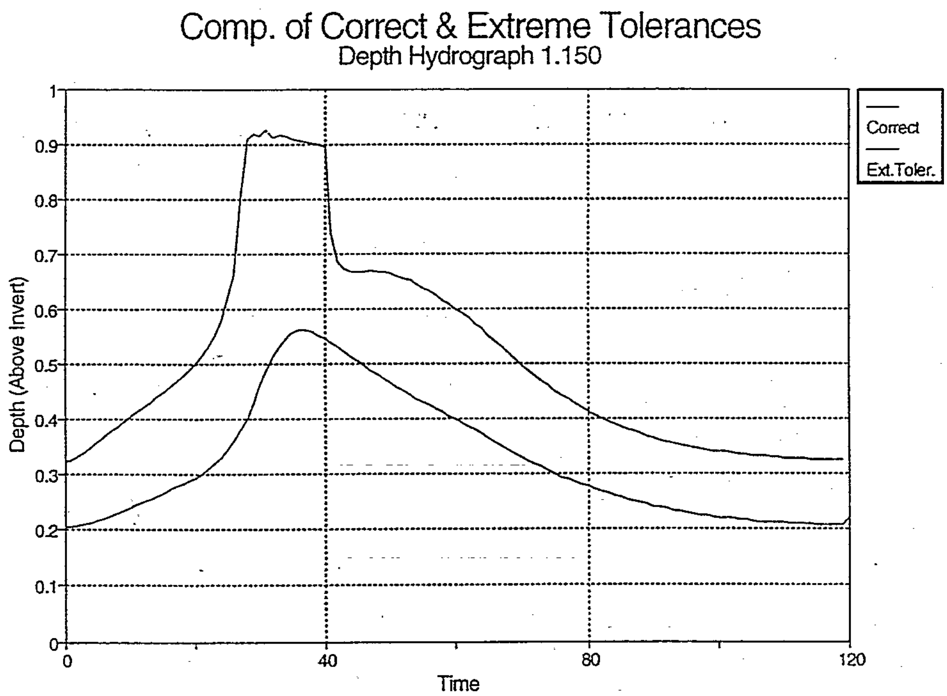

Figure 5 : Surcharge in the Network

The consequences of this are two fold:-

a) The effect may 'trigger' a perceived hydraulic inadequacy on the extreme tolerance case. The correct situation does not warrant such attention. This difference can be the difference between spending capital and not spending capital. Any capital scheme is likely to be in excess of the study cost by a considerable margin.

b) Should both models indicate a 'trigger' for rehabilitation then the extreme tolerance case would require more comprehensive capital works to achieve the required 'target' level of service.

Case 2: Pollution Assessment

The same models as used above were used with a Water Quality sewer model and the models predictions of spill over an overflow were observed. The model indicated a significant overflow operation in terms of flow. But what is more interesting is that the pollutant peak in the case of the extreme model is a very high load in terms of BOD. Such acute occurrences can be harmful to rivers and in particular fish and the associated river impact assessment may reflect this. Figure 6 illustrates this.

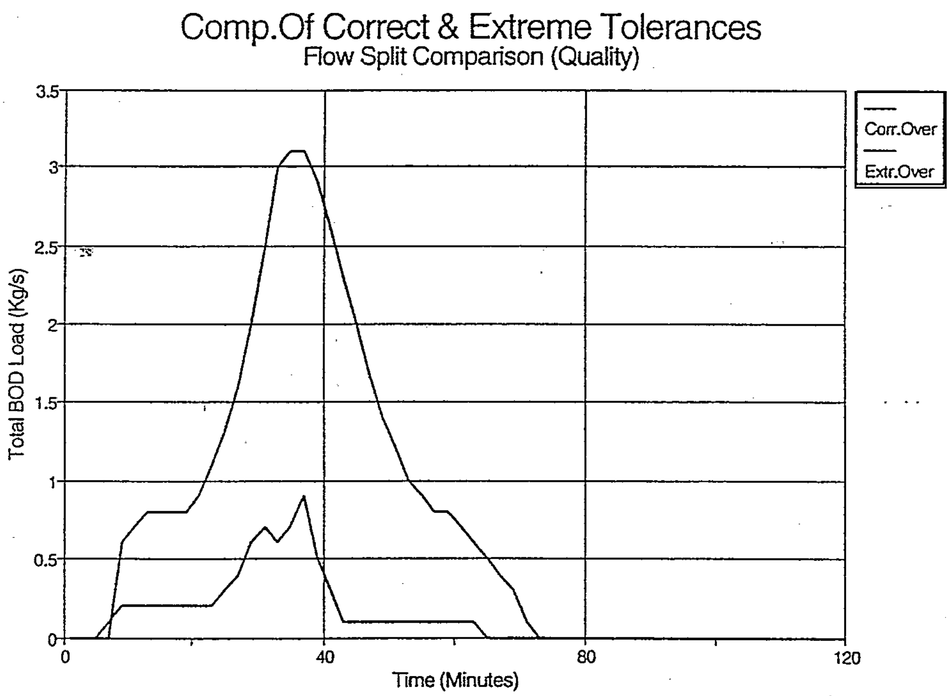

Figure 6 : Overflow Spill Quality

The consequences of this again are twofold:-

(i) The rehabilitation works may be 'triggered'.

(ii) Should the solution advocate storage then storage within the system may be required to be up to twice the size advocated by the correct model.

The ambient sediment deposition in the two models was also different. In the extreme tolerance case where sediment was deposited, as theory would suggest, thereby increasing the total sediment load available for spill.

6. Conclusions

- The use of incorrect data can have adverse effects on the capital spending of water companies.
- The capital spending can be reduced by a higher standard of data assisted by properly implemented Quality Systems.
- The interpretation of data can be assisted by the interpretation of data using structured retrieval systems.
- If resources are tight the areas in which problems are known to occur should be concentrated upon.