Paper 3 - Estimating the Frequency of Operation of Storm Outfalls and Overflows - Development of Method 2 Procedures

(Summary of Paper Only)

David W Page BSc CEng MICE MCIWEM and Gian P Illari BEng, Hyder Consulting Limited

INTRODUCTION

The accurate prediction of the spill volume and frequency of storm discharges from combined sewer overflows and storm outfalls is essential to determine compliance with environmental water quality standards. Ever larger storm tanks, involving significant capital expenditure, are being designed and constructed to limit spill frequency and volume, and it is obviously important that these are sized as accurately as possible.

The accuracy of the predictions is dependent not only on the adequacy of the hydraulic model, but as importantly on the rainfall storms that are selected for the analysis and design. The normal requirement is for a minimum ten year historical or stochastic rainfall data set to be used involving many hundreds of discrete rainfall events. Total confidence can only be achieved if every storm from the complete data set is run through a detailed verified hydraulic model, such as Hydroworks. Even with the increasing computational power available today, this procedure is very expensive and time consuming.

A number of existing methodologies have been developed to avoid analysing the complete rainfall data set. These generally involve ranking the rainfall data set in terms of the severity of spill volumes. A much smaller number of selected rainfall events are then assessed. However, the existing methodologies can be subject to significant inaccuracy with all the most severe events for a particular sewerage system not always identified and included in the analysis.

A more accurate, catchment specific method of selecting rainfall events is required to identify and rank the most severe storms, and the purpose of this paper is to investigate potential improvements to the existing procedure. The recommended procedure basically builds on the methodologies presented in the Coastal Sewerage Research Reports produced by WRc in the early 1990s [(1) (2) (3) (4)].

EXISTING METHODS

Method 1

This was a fairly simplistic approach that relied on producing a regression equation from a relatively small number of spill results and only generally included two or three rainfall parameters. The method also ignored the effect of antecedent conditions. The accuracy of the method was found to be in the region of ± 20 % [(1)].

Method 2

This method is inherently more accurate than Method 1 in as much as the actual events are accurately simulated and antecedent conditions are taken into account. However, the ranking and choice of the rainfall events to be included in the data set are critical. Ranking used to be undertaken in relation to total rainfall depth only, but this can be inaccurate. An aggregate ranking is better, using gross rainfall characteristics (total rainfall depth, peak intensity, mean intensity and duration), but this also can be inaccurate. The true ranking is catchment specific and dependent on a whole range of factors such as inflow/outflow relationship, percentage run-off and time of concentration.

Simpol

SIMPOL is a simplified spreadsheet model to conceptualise the overall sewerage system, developed to assist with UPM procedures(5). It can be used to rank the individual rainfall events in terms of severity. However, in our experience, it has not proved suitable for representing the hydraulic performance of typical sewerage systems with the complexity of many combined sewer overflows and storage tanks, particularly where backwater conditions occur. Consequently the accuracy of the results and the predicted rainfall event ranking is questionable. Calibration of the SIMPOL model can also be a very time consuming process for the more complex systems, and in many cases, it can be as quick to run the entire rainfall data set through the detailed hydraulic model.

Simplified Hydroworks Model

Another approach is to produce a simplification of the detailed hydraulic model. The aim is to significantly prune the size of the model to greatly reduce the run times whilst maintaining an acceptable level of accuracy. The complete rainfall data set is then analysed and the results used to obtain the event ranking. Similar to the SIMPOL methodology, the ranked events are then run with the detailed un-simplified model to produce the final results.

DEVELOPMENT OF REVISED METHODOLOGY

The revised methodology aims to combine the advantages of the previous methods. The proposed methodology uses regression analysis to rank the complete rainfall data set according to the spill volumes predicted. A design set, selected from the ranked data set can then be run with the verified hydraulic model.
The basic methodology proposed is listed below;

- select a sample of events from the rainfall record and run them with the verified hydraulic model;
- obtain predicted spill volumes;
- derive a relationship (using regression analysis) between spill volume and event characteristics;
- apply relationship to the complete rainfall record and rank by spill volume;
- select a design set of appropriate size;
- run the design set with the hydraulic model (some simulations would have already been performed during the regression analysis);
- calculate spill frequencies.

An improvement is demonstrated if the regression equation improves the severity ranking of discrete events, when compared to others such as total depth or aggregated rankings. The accuracy of the regression equation is dependant on the sample originally selected and the event characteristics chosen as independent (x) values. How do we sample the rainfall record? What rainfall characteristics are to be used? Are some more significant than others? Data from previous studies was used to investigate a number of sampling techniques, using various rainfall characteristics, and quantify their accuracy. The data from two studies is presented as examples.

Select a sample of events from the rainfall record and run them with the verified hydraulic model.
Obtain predicted spill volumes.

Initially, an aggregated ranking was applied to the data sets based on total rainfall depth, peak intensity and mean intensity. The data sets were sampled in various ways, selected events were run with the verified hydraulic model and the predicted spill volumes were obtained.

Derive a relationship between spill volume and event characteristics.

The various sampling techniques are listed below together with results from the linear regression analysis:

Table I - Sampling techniques - correlation between known and predicted ‘y’ (spill) values

STUDY DATA SET A

Sample	Nr of events run with model	Correlation between input and predicted ‘y’ values (1 highest, 6 lowest)
A. Top 10 storms + a sample of storms spread throughout the data set	28	2
B. Top 20 storms + a sample of storms spread throughout the data set	28	3
C. Top 5 storms + every 5th storm up to No. 100	24	5
D. Top 30 storms + a sample of storms spread throughout the data set	49	4
E. Top 30 storms	30	6
F. A sample of storms spread throughout the data set	20	1

STUDY DATA SET B

Sample	Nr of events run with model	Correlation between input and predicted ‘y’ values (1 highest, 6 lowest)
A. Top 10 storms + a sample of storms spread throughout the data set	24	3
B. Top 20 storms + a sample of storms spread throughout the data set	34	5
C. Top 5 storms + every 5th storm up to No. 100	24	2
D. Top 30 storms + a sample of storms spread throughout the data set	44	4
E. Top 30 storms	30	6
F. A sample of storms spread throughout the data set	17	1

In order of significance, the 5 gross storm characteristics used for Study Data set A were total depth, event duration, UCWI, maximum intensity, and mean intensity. Similarly, for Study Data set B they were total depth, UCWI, event duration, mean intensity, and maximum intensity. It is suggested that all available characteristics should be included, as characteristics that are later found to be insignificant do not reduce the overall regression accuracy. Spreadsheet packages available today can perform the data manipulation and regression analysis in a matter of minutes. Noticeably, antecedent conditions (represented by the UCWI value) were a significant characteristic for both catchments. A poorer correlation was obtained when UCWI was not included. The results demonstrate that an event with a high UCWI value should be used in the sample set which generates the regression equation.

Any characteristic which further distinguishes the discrete rainfall events from each other could be considered. Statistical techniques are available to measure the departure from symmetry of distributions of grouped data, e.g. skewness and kurtosis. Their inclusion could further increase regression accuracy. The start time of the event may also be a significant characteristic if historical rainfall data is being used together with a DWF profile for model simulations. However, the true test of the accuracy of the regression is not whether it achieved a satisfactory correlation with its input ‘y’ set, but whether it produced an improved severity ranking upon applying it to the whole data set. The ranking predicted by sample technique F was investigated further.

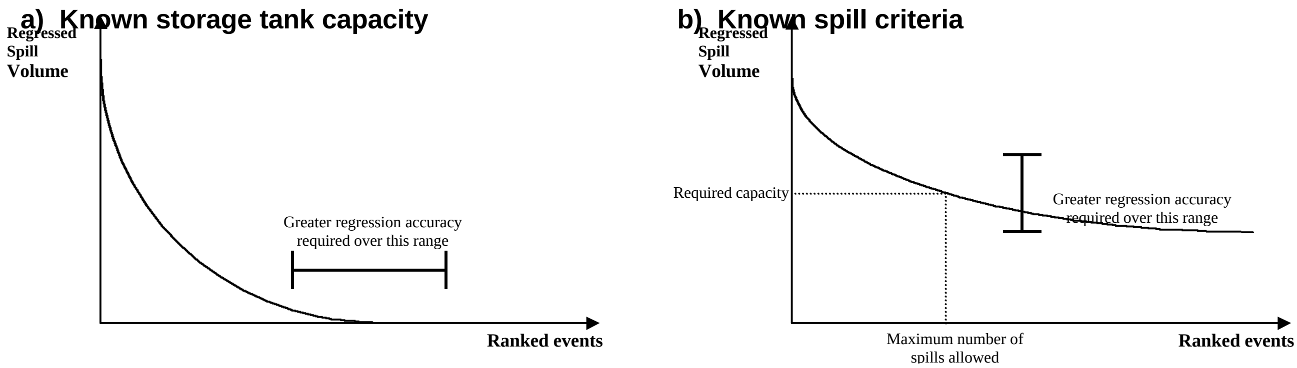
When performing spill frequency analysis, we either,

a) have a known overflow arrangement and require an assessment of frequency of operation, or
b) know the spill criteria to be achieved and require the minimum storage to achieve compliance.

Regression analysis performed on a sample of the data set will give a spill result for every event. Referring to the two conditions above, either,

a) the event at which spill first occurs, or
b) the volume at which the spill criteria is first exceeded, can be obtained.

Figure I - Events ranked by regressed spill volume



Selecting a range of storms either side of this critical point, a new regression equation can be produced, improving accuracy in the chosen range. With reference to Figure I, predictions for storms outside this range may reduce in accuracy, but the larger predicted spills will still be incorporated into the final design set. Therefore, a re-iteration of regression was performed on the study data sets. A band of 20 storms was selected, centered on the 30th ranked storm, in an attempt to improve the ranking of the top 40 storms. This gave a suitably large final set of results with which to compare accuracy against traditional approaches.

Apply relationship to the complete rainfall record and rank by spill volume. Select a design set of appropriate size.

The regression equation was applied to the complete rainfall record. Events were then ranked on predicted storm volume and the top 40 events chosen to form the design set. A number of the model simulations required had already been performed as part of the re-ranking exercise, as detailed below.

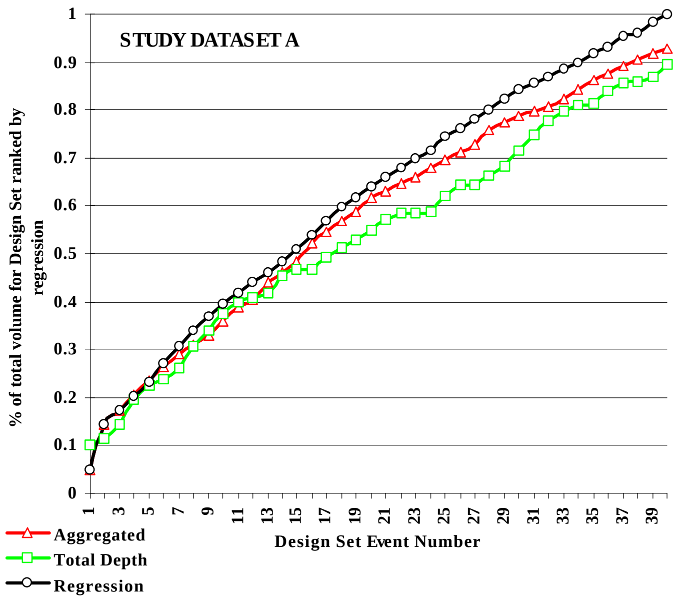
Table II - Regression ranking - total number of model simulations performed

	Number of simulations performed prior to determination of design set	Number present in design set	Total number of simulations performed
STUDY DATASET A	40	15	65
STUDY DATASET B	35	12	63

Run the design set with the verified hydraulic model. Calculate spill frequencies.

To enable a comparative analysis to be performed, simulations from design sets based on a total depth and an aggregated ranking were assessed. Results are shown below:

Figure II - Cumulative spill volumes



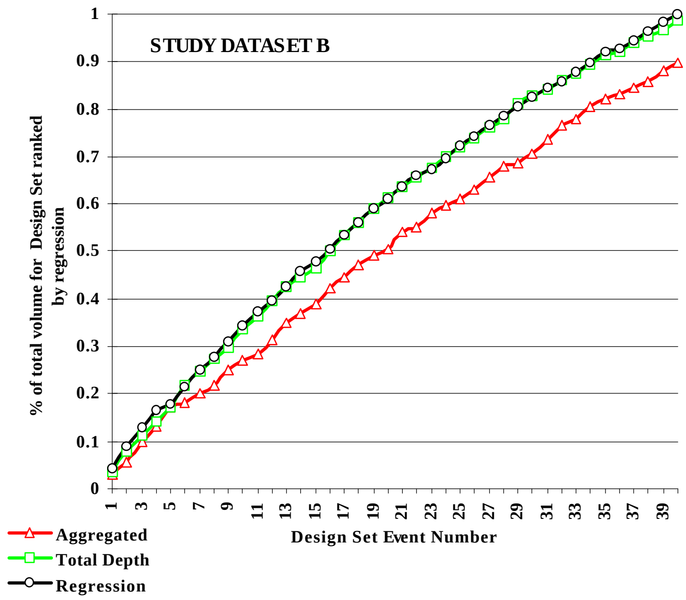


Table III - Spill volumes obtained from design sets

Design Set (Top 40)	STUDY DATA SET A Spill volume obtained from Hydroworks simulation (m^3)				STUDY DATA SET B			
	>400	>600	>800	>1000	>800	>1000	>1200	>1400
Total depth	30	21	12	1	32	21	16	8
Aggregated	36	23	12	1	26	18	15	7
Regression	38	29	15	1	35	22	16	8

The following observations can be drawn from the results:

- Table III shows that the regression ranking incorporated events causing spill into the design set which were omitted from the others.
- The design set determined by regression consistently produced the highest spill frequencies. For example, it predicted approximately 25% more spills over 400 m^3 compared to the total depth design set for study data set A.
- Considering study data set B, a tank sized at 800 m^3 would fail a 3 spill per bathing season criteria if assessed using the regression ranked design set, but would be satisfactory using the aggregate ranked design set.
- The regression ranked design set also predicted the largest total spill volume. Interestingly, the design ranked by total depth resulted in a comparative 10% error for data set A, while there was very little difference for data set B. For this data set, a 10% comparative error results if the aggregated ranking was used. The validity of the total depth ranking has been proven for catchment B. However, the need to vary the ranking to suit the catchment in question has been demonstrated.

Reading down the rankings obtained, the number of simulations that would have to be performed in order to obtain similar results can be determined. This is shown in the Table below:

Table IV - Simulation run numbers required to achieve similar results

STUDY DATA SET A

Spill volume (m^3)	Number of simulations required		
	Total depth Ranking	**Aggregated Ranking**	**Regression Ranking**
>400 (38 nr)	43	50	40
>600 (29 nr)	59	82	39
>800 (15 nr)	70	190	25

STUDY DATA SET B

Spill volume (m^3)	Number of simulations required		
	Total depth Ranking	**Aggregated Ranking**	**Regression Ranking**
>800 (35 nr)	102	63	40
>1000 (22 nr)	73	57	26
>1200 (16 nr)	46	29	20
>1400 (8 nr)	46	29	10

Table IV above shows that the revised ranking reduces simulation times considerably, if a comparative result is to be achieved. However, the larger simulation numbers are generally caused by only one ‘rogue’ event being ranked much lower down the data set than model simulations predict.

SUMMARY AND CONCLUSIONS

- Method 2 proposed that the rainfall record was ranked on the basis of total rainfall depth and a suitable number of storms were to be chosen for the design set. However it was recognised that the selection of storms can be a “subjective choice and is based on the need to include those storms most likely to cause the highest spills”(3). The revised methodology aims to remove subjectivity in the choice of storms, reducing the risk of not including ‘rogue’ events in the final design set.

- Analysis performed on study data sets has shown that a regression ranking can be developed which improves upon rankings traditionally used for spill frequency analysis, i.e. aggregated and total depth rankings. However, the original total depth ranking may still be valid for some catchments. Regression analysis can be used to determine the significance of this characteristic, in order to apply a suitable ranking.

- Further investigation is required to accurately define a procedure to be adopted, as the proposed methodology has only been applied to two catchments. Future analysis may consider the addition of further gross storm characteristics, other ways to sample the original data set and the use of non-linear regression. This could produce an improved original ranking, decreasing the time taken to arrive at a satisfactory design set.

REFERENCES

(1) Gent, R. (1991): An Investigation of Error Bands Associated With Different Spill Prediction Methods, Coastal Research Report Nr 14, WRc, Swindon.

(2) Henderson, R.J., Dempsey, P. (1990): Estimating the Frequency of Operation of Storm Outfalls and Overflows - Method 1, Coastal Research Report Nr 5, WRc, Swindon.

(3) Therlfall, J., R.J., Dempsey, P. (1990): Estimating the Frequency of Operation of Storm Outfalls and Overflows - Method 2, Coastal Research Report Nr 4, WRc, Swindon.

(4) Therlfall, J., R.J., Dempsey, P., Hutchings, C.J. and Murrell, K.N. (1991): Recommended Method for Estimating Storm Discharge Volumes and Frequencies, Coastal Research Report Nr 12, WRc, Swindon.

(5) Foundation for Water Research (1994), Urban Pollution Management Manual, Report Nr. FR/CL0002, Foundation for Water Research, Marlow.

DISCUSSION

Question George Heywood Tynemarch

It seems that the regression analysis was dependent on the event characteristics and their choice what were these and their relative importance ?
Did you look at average intensity over time within the events ?

Answer

Event parameters were, total depth, peak intensity, average intensity, duration and UCWI (Urban Catchment Wetness Index). The key parameters in order were: total depth, event duration, closely followed by UCWI, then maximum intensity. Mean intensity was the least significant parameter, which was expected as this is simply peak intensity divided by duration. If you look at average intensity over certain periods within the event you are attempting to distinguish, between differing storm "shapes" .We are investigating this as it does improve accuracy.

Question Elliot Gill WRc

In many systems with large tanks then you see small events that come after large events cause spill because they occur when the storage is still partly full. Have you looked at this ?

Answer

Yes this is an area of concern and we are looking at sequencing of storms and the inter storm dry period. We are developing regression analysis methods that piggy back these type of events.