Stochastic Generation of Rainfall Time Series: A Case Study in Sewerage Rehabilitation

Paul S.P. Cowpertwait, BSc, PhD, FSS.
Model Solutions, Northumbrian Water Group, Manor Walks, Cramlington, Northumberland, NE23 6UP.

Abstract

The Stochastic Rainfall Generator (SRG) is used to study the overflow performance of a sewerage system in the North of England. One hundred years of *hourly* rainfall time series are simulated using the SRG, and a *typical* year extracted from the 100-year record by a method described in this paper. The 100 most severe rainfall events are then extracted from the simulated typical year, and ranked in order of severity. The events are disaggregated and converted into a form suitable for WALLRUS simulation. Overflow performance is assessed by considering the number, volume, and duration of spills for the typical year.

1. Introduction

Rainfall time series provide an effective means of assessing the hydraulic performance and pollution impact of sewer systems by quantifying the frequency, rates, volumes and durations of discharges (*Henderson*, 1986). However, the available WRc annual series do not accurately represent all geographical regions nor do they contain extreme events (*Henderson*, 1986). The Stochastic Rainfall Generator (SRG), which is based on the work by *Cowpertwait* (1991a, 1991b) and *Cowpertwait et al* (1991, 1992), is a regionalised stochastic rainfall model and improves upon the available WRc annual series.

2. Background

The SRG is based on a *Clustered Poisson Process* which takes into account the *dependence* inherent in rainfall events (see, for example, *Rodriguez-Iturbe et al, 1987, or Cowpertwait, 1991a,b*). The model has five parameters (which are estimated for each month): 1) the mean waiting time between the beginning of the storms (in hours), 2) the mean number of rain cells per storm, 3) the mean duration of each rain cell (in hours), 4) the mean intensity of each rain cell (in mm/hour), and 5) the mean waiting time for each rain cell after the beginning of the storm (in hours).

The parameters of the model were estimated for each month from rainfall statistics that were extracted from historical records of rainfall data (see figure 1). These parameter estimates were then regressed on site characteristics known to influence rainfall (e.g. altitude, coastal distance, etc) so that the parameters of the model could be estimated at sites lacking data. This is known as the *regionalised* version of the model and enables hourly rainfall time series to be generated at any location in the UK by inputting site variables.

On average, the regionalised version of the model has an accuracy equivalent to using about 20 years of daily data to estimate the parameters of the model at the site under investigation. However, this average value of 20 years has quite a high standard deviation (about 10 years) associated with it so that the accuracy of the regionalised version of the model is likely to vary from site to site. An alternative to using the regionalised version of the model is to use the *nearest neighbour* version, which involves fitting the model to the nearest daily rainfall station of a similar altitude to the site under investigation. The nearest station may already be available in the database of rainfall statistics (shown in figure 1). Alternatively, daily rainfall data could be purchased from the Meteorological Office, Bracknell (at about £400 for 30 years) and used to fit the stochastic model.

In summary, there are *two* versions of the SRG: 1) a regionalised version, and 2) a nearest neighbour version, both of which can be used with confidence to generate *hourly* rainfall time series for any location in the UK. The regionalised version can be used as a tool for *interpolating* between the available data shown in figure 1, and the nearest neighbour version used when the site under investigation is *close* to one of the available rainfall stations. The regionalised version of the model has been tested by *Gill et al* (1992), and found to predict spill volumes to within ±10%, although a correction factor of 5% has to be applied to the predicted volumes.

The two versions of the SRG can be used to simulate *hourly* rainfall time series. However, for sewerage rehabilitation rainfall data are usually required at a finer resolution. A *disaggregation* model was developed to meet this requirement (*Cowpertwait, 1991b*). This model assumes that rainfall occurs in discrete pulses, which are randomly distributed over the wet hours (according to a probability density function that has a shape dependent on the depth of rain within the hour, and the depths in the preceding and following hours). The disaggregation program converts the hourly rainfall events to 5 minutely rainfall profiles in a form suitable for WALLRUS or WASSP simulation. The disaggregation model has also been tested by *Gill et al* (1992) and found to predict spill volumes to within ±5%, when compared to volumes predicted when using historical rainfall time series.

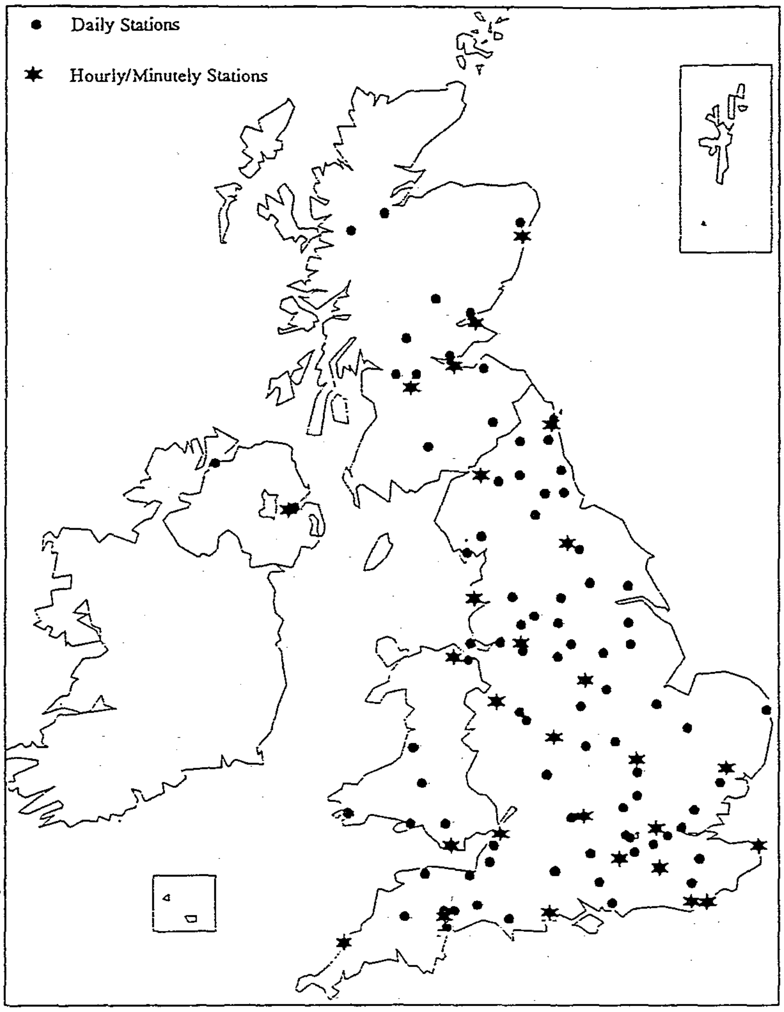

Figure 1: Available rainfall data for nearest neighbour method

3. Generating a rainfall time series for Esh Winning, Durham

3.1 Estimating the parameters of the SRG

A rainfall time series was required by Northumbrian Water Ltd for the Esh Winning sewerage system, which was known to have serious overflow problems. It was evident (in figure 1) that a neighbouring site had been used in the study by *Cowpertwait* et al (1992), so that the *nearest neighbour* method could be adopted. The neighbouring site was Durham City, which is about 7km from Esh Winning and has a similar altitude (about 100m above sea level).

The daily rainfall statistics for Durham City were extracted from the data base, and the parameters of the SRG estimated (table 1 gives the estimates for January and July as examples).

Table 1: Example parameter estimates for the SRG

Month	Mean time between storms in hours	Mean waiting time for rain cells in hours	Mean cell duration in hours	Mean number of cells per storm	Mean cell intensity in mm / hour
January	43.0	6.0	1.4	3.1	0.72
July	64.4	6.8	1.1	3.2	1.6

3.2 Extracting a typical year from the simulated rainfall time series

A 100-year record of hourly rainfall time series was generated using the SRG, and a *typical* year extracted by the following method.

The historical rainfall statistics* used to estimate the parameters of the SRG for each month were stored in a database. These statistics, a total of 16 per month, were identified as summarising the properties of a rainfall time series by *Cowpertwait* (1991b) and include the mean, variance, probability of rain, and wet/dry spell transition probabilities for various time steps. The statistics were then found for each month of each year for the simulated 100-year record, and the following calculated:-

$$SS = \sum (1 - S_f / S_g)^2,$$

where S_f and S_g are the fitted (historical) and generated statistics respectively, and the summation is taken over all 16 rainfall statistics. The month-year with the smallest summation was regarded as the most typical month in the 100-year record (because the generated statistics were closest to the fitted values). The typical months were then concatenated to form a typical year.

3.3 How representative is the typical year?

The accuracy of the typical year can be assessed by summing SS for each month to obtain a total sum of squares TSS (which is the summation of $16 \times 12 = 192$ values). The value obtained for TSS for the nearest neighbour method was 1.73, which amounts to an

* The rainfall statistics used to fit the SRG to each month are:- the daily mean, the variance of the rainfall total in a range of time steps, the probability of no rain in a range of time steps, the probability of rain in a time step given the previous time step had rain, where the time steps were 1, 3, 6, 12, and 24 hours. This amounts to using 16 rainfall statistics to fit the model to each month of the rainfall record.

average error of about ±10% when spread over the 192 values. A visual assessment of this error was made by considering monthly means, variances, and proportion of dry days (figures 3, 4, 5, and 6). The figures also show the simulated statistics over the 100-year period, which verifies the computer program for the model (i.e. shows that the program reproduces the statistics used to fit the model). From the figures it is evident that the statistics for the typical year closely follow those for the historical and simulated rainfall time series, with no consistent over- or under-estimation being evident. Hence, the typical year was regarded as representative and could be used with confidence in assessing the number, duration, and volume of spills per year for the Esh Winning sewerage system.

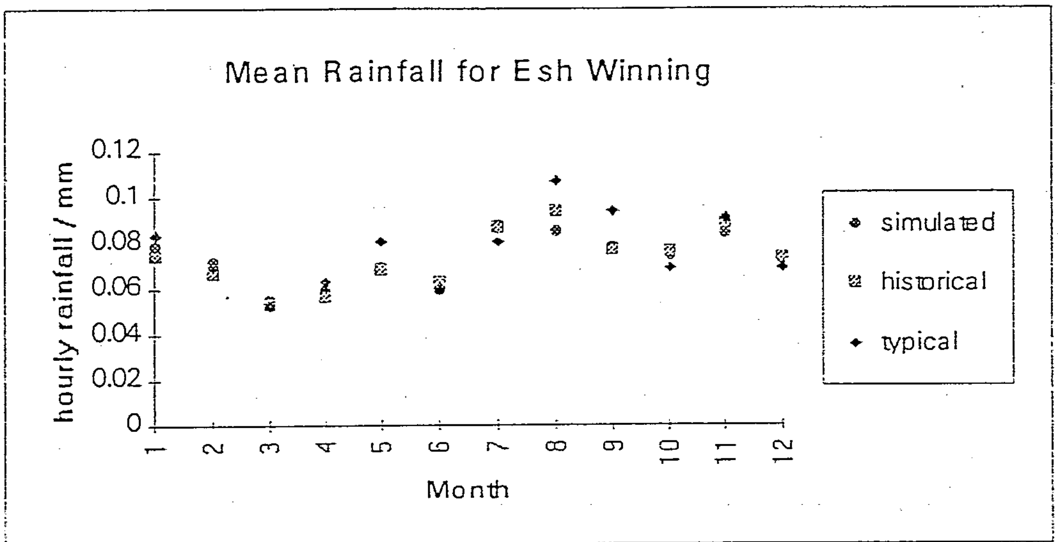

Figure 3: Comparison of mean monthly rainfalls

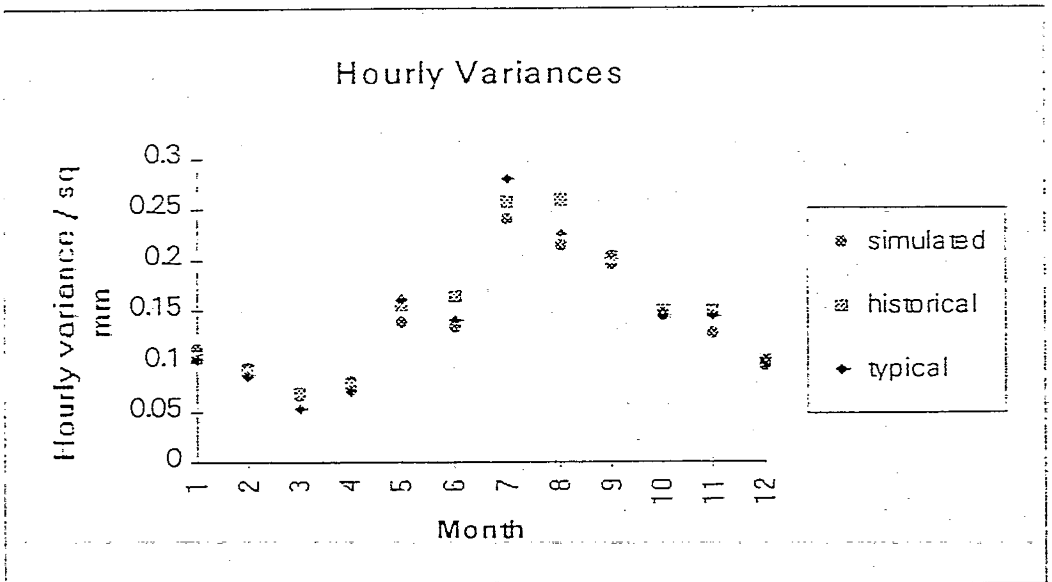

Figure 4: Comparison of monthly hourly variances

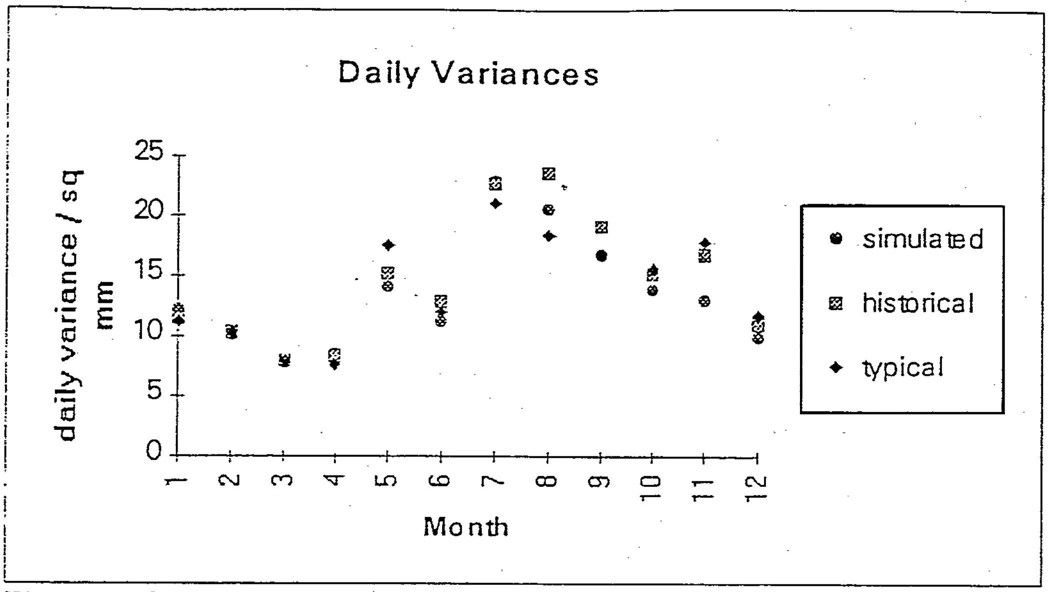

Figure 5: Comparison of monthly daily variances

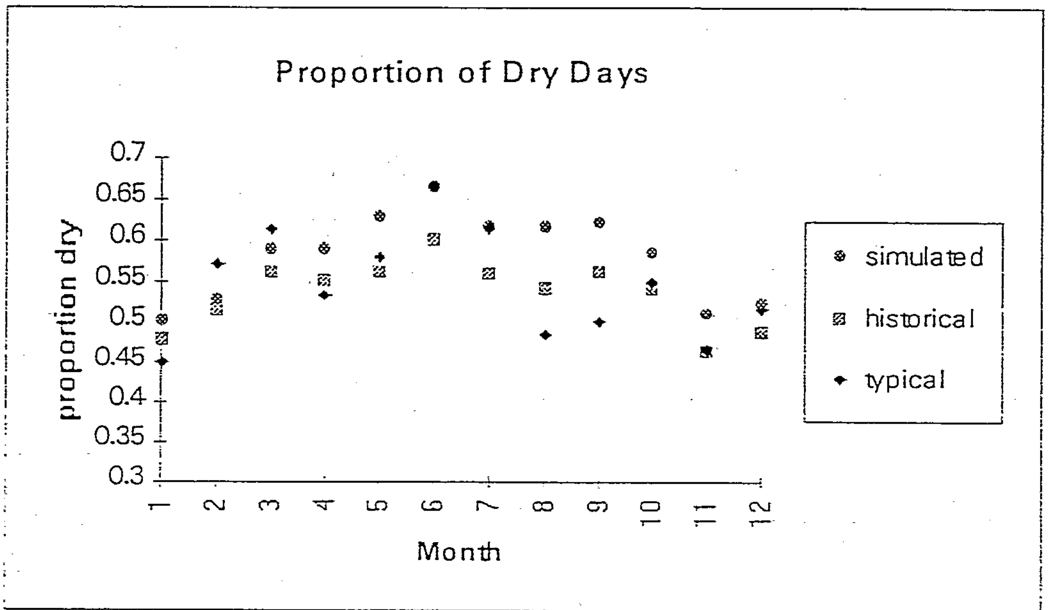

Figure 6: Comparison of monthly proportion of dry days

3.4 Assessing overflow performance using a typical year

A typical year was extracted from the 100-year record, and the 100 most severe rainfall events (in terms of volume of rain) extracted from the typical year. These events were disaggregated into 5 minutely rainfall profiles and stored in a form suitable for Wallrus simulation, with UCWI values calculated using the method outlined in Section 7.9 of the *Wallingford Procedure* (1983). The events were then used in conjunction with a verified Wallrus model of the Esh Winning sewerage system and the number, volume, and duration of spills found by flow simulation (Table 3).

Table 3: The number, volume and duration of spills per year for Esh Winning, Durham

Overflow	Number of spills	Volume of spills / m^3	Duration of spills / minutes
1	0	0	0
2	0	0	0
3	82	2852.8	9957
4	65	6099.2	7733
5	47	2535.5	3669
6	62	3924.4	6226
7	36	1890	3607
8	100	8555.8	16674
9	0	0	0
10	100	31606	15373

3.5 A comparison with the WRc Northumbrian annual series

The WRc annual Northumbrian series was based on 8 years of minutely rainfall data taken from Bishop Auckland, County Durham. Bishop Auckland is close in altitude and location to both Esh Winning and Durham, and so a comparison of overflow performance when using the WRc annual series is of interest (Table 4). The percentage differences were found between the volume of spills predicted by the WRc annual series and the volume of spills predicted when using the SRG (Table 5). From Table 5, it is evident that there is a tendency for the SRG to under-estimate the smaller spill volumes (overflows 4, 5, 6, and 7) when compared with the WRc annual series. For the more critical overflows (8 and 10) the SRG tended to over-estimate the spill volumes. Overall, the SRG showed no tendency to over- or under-estimate the spill volumes given by the WRc annual series. The differences between the WRc annual series and the SRG can probably be attributed to the sampling variability in the WRc annual Northumbrian series, which was only based on 8 years of historical rainfall data.

Table 4: Number and volume of spills when using the WRc annual series

Overflow	Number of Spills	Volume of Spill / m^3	Duration of Spill / minutes
1	3	39	291
2	0	0	0
3	45	2402	7025
4	41	6927	6475
5	44	2562	3248
6	46	4646	5881
7	24	3046	3743
8	46	6478	8797
9	0	0	0
10	100	26057	8385

Table 5: Percentage differences in spill volumes between the SRG and WRc annual series

Overflow	SRG nearest neighbour
3	19
4	-12
5	-1
6	-16
7	-38
8 *	32
10 *	21

(a minus sign indicates that there is a tendency for the SRG to under-estimate spill volume when compared with the WRc Northumbrian annual series; * = critical overflow)

3.6 Deciding which version of the SRG to use

The *regionalised* version of the SRG can be used with confidence to generate rainfall time series for any location in the UK. However, if site data are available or a nearest neighbour exists, then the nearest neighbour version of the SRG would provide more accuracy. If there is some doubt as to whether the nearest available data is close enough to the sewerage system under study then it would be appropriate to perform some statistical tests on published monthly rainfall totals (available in the Meteorological Office publication entitled 'Rainfall'). This could be achieved using standard t-tests to see whether differences in mean monthly totals between two rainfall gauges close to the sewerage system under study are statistically significant.

For the site in this study (Esh Winning) there was no statistical evidence to suggest that the nearest neighbour (Durham) had mean monthly totals significantly different from those of another neighbouring site (West Auckland, which was about 15km to the south of Durham and Esh Winning). Hence, the nearest neighbour method was used in preference to the regionalised method.

3.7 Using the SRG with the re-designed Wallrus model

The verified Wallrus model of the Esh Winning sewerage system was adapted in an attempt to improve the performance of the most critical overflow (number 10). The typical year of the nearest neighbour version of the SRG was used for the Wallrus simulation. The spill volume was found to be $1563m^3$ when using the revised Wallrus model (compared with $31606m^3$ previously), which was judged to be a satisfactory improvement by the engineer responsible.

5. Summary and recommendations

Two versions of the SRG are available for simulating *hourly* rainfall time series: 1) a *regionalised* version, and 2) a *nearest neighbour* version. The regionalised version provides an effective way of *interpolating* between the available site data, and should be adequate for most studies. If a greater accuracy is required then the nearest neighbour method could be used, either by calibrating the SRG using data bought from the Meteorological Office (at about £400 for 30 years), or using the nearest available station in the database (shown in figure 1).

A method of simulating a typical year using the SRG is available and can be used with confidence for assessing overflow performance on a yearly basis (i.e. number, volume, and duration of spills per year).

References

Cowpertwait P.S.P. (1991a), Further Developments of the Neyman-Scott Clustered Point Process for Modeling Rainfall, Water Resources Research, 27(7), pp1431-1438.

Cowpertwait P.S.P. (1991b), The Stochastic Generation of Rainfall Time Series, PhD Thesis, Department of Civil Engineering, University of Newcastle upon Tyne.

Cowpertwait P.S.P., Metcalfe A.V., O'Connell P.E., Mawdsley J.A., and Threlfall J.L. (1991), Stochastic Generation of Rainfall Time Series, Foundation for Water Research, Report Number FR0217.

Cowpertwait P.S.P., O'Connell P.E., Metcalfe A.V., and Mawdsley J.A. (1992), Stochastic Modelling of Rainfall Time Series: 1. Fitting a Neyman-Scott model to historical rainfall data, 2. A regionalised hourly model for the UK with disaggregation to shorter intervals, Journal of Hydrology (to appear).

Gill E.J., Threlfall J.L. and Dempsey P. (1992), Testing of the Stochastic Rainfall Generator (SRG) Model: Supplementary Report, Foundation for Water Research, Report FR0318.

Henderson R.J. (1986), Rainfall Time Series for Sewer System Modelling, Report number ER195E, WRc Engineering, Swindon, UK.

Rodriguez-Iturbe I, Cox D.R. and Isham V (1987), Some Models for Rainfall based on Stochastic Point Processes, Proceeds of the Royal Society London, 410, pp269-288.

The Wallingford Procedure (1983), Design and Analysis of Urban Storm Drainage, Department of Environment, National Water Council, Standing Technical Committee Reports Number 28, Volume(1).

TESTING THE STOCHASTIC RAINFALL GENERATOR MODEL

J L THRELFALL

WRc Swindon, Frankland Road, Blagrove,
Swindon, SN5 8YF

1. INTRODUCTION

The Stochastic Rainfall Generator (SRG) model synthesises rainfall time series for urban areas in the UK. It was developed, specifically to provide the input for flow simulation models for pollution control purposes.

A testing programme, funded by the Foundation for Water Research, has just been completed[1] in which the output from the model is compared with that of historical rainfall data in terms of practical parameters such as spill volumes and frequencies.

2. THE STOCHASTIC RAINFALL GENERATOR

The SRG produces hourly rainfall time series for any specified location in the UK. Two versions are available: a Regional version which requires site specific variables as input to the models, such as, altitude, grid reference and mean daily rainfall data for each month; and a Daily Regionalised version, which additionally requires a long local series of daily rainfall data as input. Throughout the testing programme reported in this paper the Regional version of the SRG was used.

A separate module has also been developed as part of the SRG which disaggregates hourly rainfall data to finer resolution values, for example five minutely, suitable for input to models such as WALLRUS.

Both versions of the SRG and the Disaggregator are described more fully, together with details of preliminary validation testing, in an earlier FWR report[2].

3. TESTING PROGRAMME OBJECTIVES

The main objectives of this testing programme were:

- To assess, for the purposes of sewer flow modelling, how well the SRG simulates storm events at different UK locations.

- To assess how well the Disaggregator creates realistic five minutely rainfall profiles from hourly rainfall profiles.

The magnitude of storm events tested ranged from between a two year and approximately a one month return period.

4. TESTING METHODOLOGY AND RESULTS OF THE SRG

The testing methodology can be summarised by the following main steps:

1. Preparation of rainfall series
2. Preparation of sewer models
3. Comparison of spill characteristics

4.1 Preparation of Rainfall Series

Four rainfall sites were selected (Squires Gate, Lancashire; Manston, Kent; Exeter, Devon; Shoeburyness, Essex). Historical rainfall records, for up to 15 years, were processed to identify rainfall events. The SRG was employed to produce hourly rainfall series which matched each of the historical series for both location and length of record. The SRG data were then processed, in a similar way to the historical data, to obtain rainfall events.

4.2 Preparation of Sewer Models

The SMARTS(3) model was used for comparing spills predicted by the different rainfall series. This model was selected because it works on an hourly time step and accepts a continuous hourly rainfall input. Two catchments were modelled: Catchment A represented a large and relatively flat system; and Catchment B represented a small and steep system. The purpose of this choice of catchments was to see how sensitive the results would be to contrasting catchment characteristics.

4.3 Comparison of Spill Characteristics

The historical and SRG events for a given site were run through the sewer flow models to give two series of spill volumes. Two main comparisons were made which considered design implications:

1. The storage volume needed to meet a particular spill frequency standard.
2. The spill frequencies that would have been experienced during the rainfall period if that storage had been installed.

4.4 Summary of Results

Table 1 summarises the extent to which the SRG model causes storage volume to be under or over predicted when meeting selected design standards. Results are displayed for Catchments A and B for 3 potential design standards.

The full range of variability found in the storage volumes in Catchment A was +4% to -16% and Catchment B +25% to -28%. It should be noted that percentage errors are likely to be greater when considering the small storage volumes associated with Catchment B.

Table 1 - Summary results of spill frequency analysis

	% Difference in SRG Storage Volume			
Design Standard	**Squires Gate**	**Manston**	**Exeter**	**Shoeburyness**
Catchment A				
6 spills/year	2.04	-5.65	-11.61	-
4 spills/year	2.02	-2.80	-15.90	-16.00
2 spills/bathing	3.76	-6.90	-14.51	-9.30
Catchment B				
6 spills/year	19.60	-10.00	-16.61	-
4 spills/year	24.66	2.63	-22.88	-5.63
2 spills/bathing	17.09	0	-27.84	-8.57

5. TESTING METHODOLOGY OF THE DISAGGREGATOR

5.1 Preparation of Rainfall and Sewer Model

The Disaggregator was tested using two historical minutely rainfall records. Ten winter and ten summer rainfall events were selected from both series, providing a total of 40 events. The events represented a range of return periods from 1 in 1 month to 1 in 2 years.

A reference series was created by aggregating the minutely historical data to 5 minutely data in WALLRUS format. A second series was created to test the disaggregation procedure. The historical minutely data were aggregated to hourly data and then disaggregated using the Disaggregator to 5 minutely data.

The accuracy of the Disaggregator model was tested by running both series of 40 events through a WALLRUS model of a flat/average catchment with one overflow.

5.2 Summary of Results

The results of the testing programme showed that the total volume spilled from each set of disaggregated storms was within -0.2% and +1.4% of the reference series. In addition, spills for individual events were in good agreement, with differences generally within +/-5%. No particular bias could be detected between regions or seasons.

6. CONCLUSIONS AND RECOMMENDATIONS

Based on the results from the testing programme the Regional version of the SRG in its current form is considered to be a useful and usable method of generating rainfall data for pollution assessment purposes. The performance of the SRG has been found to be good and has shown only the anticipated range of errors that are associated with any regionalisation procedure. However, it is also recognised that the Regional version of the

SRG has some limitations and further work will be necessary to broaden its scope for effective application.

The disaggregation model has been shown to be an accurate technique. This method can be used if 5 minutely data are required from either hourly historical or SRG rainfall data.

An interim procedure for the application of the SRG has been proposed, as follows:

(a) If a very long historical record of hourly data is readily available (15 years or more) this should be used in preference to the Regional SRG.

(b) If over 20 years of daily rainfall data are available for the site in question, the Daily Regionalised SRG should be used in preference to the Regional SRG. Whilst this version has not been fully tested, it should improve the regionalisation procedure.

(c) When using the Regional SRG to estimate storage volumes to meet a given spill frequency standard, 5% should be added to the volume estimate. This pragmatic approach should compensate for the bias detected in the testing programme on the larger catchment. The resultant storage volume can then be anticipated to be between approximately +/-10% of that which would have been specified with the use of historic rainfall data.

7. FURTHER DEVELOPMENTS

Development of the SRG is required in two areas:

1. Application of the model.

2. Improvements to the model capabilities and accuracy.

7.1 Application of the Model.

The SRG and Disaggregator will soon be available from WRc as a fully supported and documented software package. This will allow the user to:

- generate hourly rainfall time series using the Regional or Daily Regionalised versions of the SRG.

- identify rainfall events and rank events chronologically, by rainfall depth or intensity. The output will be in a form suitable for SMARTS input.

- disaggregate hourly rainfall events (either from SRG or historical rainfall) into 5 minutely values, in a format suitable for input to WALLRUS, WASSP or MOUSE.

The selection of storms to use in modelling work will depend on the particular application. Various options are available, such as ranking the events on the basis of rainfall characteristics, looking at the frequency distribution of the rainfall, and ranking on spill volumes produced by simple sewer flow models. Return periods can be allocated to the storms from the ranked set as an aid to selection. Further guidance on how to use historical rainfall or the SRG, and on the sampling strategy, is being developed and will have to be completed before the practising engineer can fully utilise the technology.

7.2 Improvements to the Model Capabilities.

The SRG model is amenable to further development work should practical application experience suggest it to be necessary. At present it is thought that such work should focus on the simulation capabilities for rainfall events of return periods in excess of 2 years. The accuracy of the simulation parameters could also be improved, to reduce the existing error band. In addition, an effort should be made to improve knowledge of the spatial variation in rainfall to the point where it can be incorporated into a practical design methodology at some point in the future.

8. REFERENCES

1. GILL E J, J L THRELFALL and P D DEMPSEY. **Testing of the Stochastic Rainfall Generator (SRG) Model: Supplementary Report.** Foundation for Water Research Report No. FR0318. 1992.

2. COWPERWAIT P S P, A V METCALFE, P E O'CONNELL, J A MAWDSLEY and J L THRELFALL. **Stochastic Generation of Rainfall Time Series.** Foundation for Water Research Report No. FR0217. 1991.

3. HUTCHINGS C J. **SMARTS User Guide - Version 1.** Report No UC1226, Water Research Centre, Swindon, UK. 1991 (Restricted).

2.2 Stochastic Rainfall Generator **P.S.P Cowpertwait (Model Solutions) and J.L.Threlfall (WRc)**

Question

Bernard Welsh Wilde Allison

Could the speaker comment on the accuracy of the testing method in steep catchments ?

Answer

The accuracy of the testing method in the steep catchments is limited by the SMARTS model which runs on a 1 hour time step, which may be greater than the time of concentration of the catchment. If the model is run in detail with WALLRUS then it gives better results.

Question Rob Henderson Wessex Water

Are there applications for this rainfall generator in water resources.

Answer

There may well be, at the moment the method is being tested on sewerage systems after the successful completion of this study they will be looking to other applications.

Question

Richard Kellagher Integrated Hydro Systems

In the assessments of the disagregation routine the volume of storage is a key factor in the disagregation process. If the volume is large then it is not a true test as the system will even out the disagregation effects negating the comparison and possibly why the results were so close.

Answer

The volume of storage was not excessively large, but it is a valid point. Other results from the ongoing work not reported here have given equally good results.

Question

Brian Sharman North West Water

The work looks promising but why were only sites round Airports used (these are surely not typical) and there is only one raingauge site in the Lake District ?

Answer

Airports tend to be the only good sources of detailed hourly rainfall data. Basically you are limited to what the Met office can provide, you work with the best available tools. Although there is only one gauge in the Lake District wet areas can be further calibrated by using local daily data. This will give better answers than by just using the built in regionalisation features.