Pergamon

Wat. Sci. Tech. Vol. 39, No. 9, pp. 285–292, 1999

0273–1223/99 $20.00 + 0.00

PII: S0273-1223(99)00243-7

MODELLING DRAINAGE PERFORMANCE IN SLUMS OF DEVELOPING COUNTRIES: HOW GOOD IS GOOD ENOUGH?

P. Kolsky*, D. Butler**, S. Cairncross*, U. J. Blumenthal* and M. Hosseini*

* *London School of Hygiene & Tropical Medicine, Keppel St, London WC1E 7HT, UK*

** *Imperial College of Science, Technology & Medicine, Imperial College Road, London SW7 2BU, UK*

ABSTRACT

A model for the dual drainage system in a flat 20 hectare slum in Indore, India was developed to evaluate the factors which influence drainage performance. Performance was defined by the depth, extent, duration, and frequency of flooding during the 1994 monsoon. This paper reports on the conceptual problem of model validation, particularly for "flooding" or "not flooding". The model's predictions for maximum depth during verification were reasonable, as 72% of maximum depth predictions were within +/- 0.10 m of observed values. The predictions for the binary variable of flooding at a point were, however, poor. The model gave good predictions of freeboard violation, where the freeboard level was set at 50 mm below the lowest kerb. The paper offers some intuitive reasons as to why prediction of freeboard violation may be expected to be better than prediction of flooding. The authors conclude that where models cannot be validated for flooding prediction, it may be better to redefine performance criteria and validate the model for freeboard violation.

KEYWORDS

Developing countries; dual drainage; flooding; performance; verification.

CONVENTIONAL DRAINAGE AND PERFORMANCE

Design of urban storm drainage is conventionally based on the frequency with which, on average, one would expect the system's capacity to be exceeded. Engineers thus design a system so that its capacity is equalled or exceeded only once in *n* years, where *n* is defined as the system's return period. In the North, flooding is fairly rare in practice, and most designers do not analyse what happens when flooding actually *does* take place. Conventional design and analysis also neglect the effects of solids deposits in the drain upon hydraulic capacity. Instead, designers assume that small solids (e.g. sand and silt) are carried along the conduits by the "self-cleansing velocities" for which the drains are designed. Larger solids are assumed not to enter the drain, but to be removed either by good solid waste management, or by well-designed and maintained inlets and gully traps.

There are three major characteristics of developing countries which make flooding more frequent than in the North: (i) higher rainfall intensities, (ii) scarcer resources for drainage construction and maintenance, and (iii) less effective solid waste management, resulting in large solids entering and blocking the drainage system. These three factors call into question both the conventional approach to drainage design, and the wisdom of defining performance solely in terms of flood frequency.

A definition of drainage performance

If frequent flooding is inevitable in the South, it would seem worthwhile to understand more clearly what happens when flooding *does* occur. How much does it matter that a street floods to a depth of five centimetres for five minutes? Does it make a difference if only one house is affected, or if 100 houses are flooded? How can we reflect such concerns in design criteria and the modelling of drainage in slums?

In 1992, research funded by the UK Overseas Development Administration began on the evaluation of alternative drainage systems in India, based on their performance during the monsoon. At the outset of the study, a working definition of drainage performance was adopted in terms of the **depth**, the **area (or extent)**, the **duration**, and the **frequency** of flooding. This definition was used in the evaluation of various drainage systems in Indore, a city in the state of Madhya Pradesh in India.

Sensitivity of performance to solids blocking conventional drains

Several of us have already reported our main findings on the effects of solids upon performance in a conventional open channel drainage network in one slum of Indore (Kolsky *et al.*, 1996). In this work, a SPIDA model from Wallingford Software, (Richens, 1985) was developed, calibrated, and verified for the catchment, using rainfall, runoff, and site data collected during the 1994 monsoon. Solids were also sampled from the drain, and subjected to sieve analysis to determine the approximate particle size distribution. This analysis led to the following conclusions for the open channel drainage network studied:

- *Significant levels of solids were present* in the drain, to an average depth of 30% at the start of the study. These solids were large, with over 10% by mass greater than 80 mm in diameter.
- *Solids have little effect upon maximum flooding depth*, once flooding takes place.
- The presence of *solids can have a substantial effect upon flood duration*, especially for those parts of the catchment where the drain is the only outlet for flood waters.
- The *extent and frequency of flooding can also be significantly affected by solids blockage.*
- *Surface routing of runoff deserves further attention* in the drainage problems of Southern cities.

DUAL DRAINAGE SYSTEMS

The principle of dual drainage recognises the existence of two distinct systems in any urban catchment (WEF/ASCE, 1992). The *minor drainage system* is the conventional network of pipes and inlets which can effectively manage small and frequent events without flooding, thus minimising nuisance. By contrast, the *major system* consists of surface routes such as roads, streets, and open space, which come into action whenever the capacity of the minor system is exceeded and runoff spills onto the surface. In conventional drainage practice, only the minor system is actually designed; the major system is usually ignored. Where flooding is frequent, however, it would seem imperative to consider both systems.

The infrastructure in a number of Indore slums was upgraded in 1992-1995, as part of the ODA-funded Slum Improvement Programme. The consulting engineer Himanshu Parikh applied the dual drainage philosophy to the improvement of many Indore slums (Parikh, 1990), including Motilal ki Chal, a flat 20 ha slum which was chosen as a study area for our research. In addition to a conventional minor drainage system, Parikh designed the streets of Motilal ki Chal to act as drains, through choice of appropriate cross sections and consistent slopes towards the outlet. An important part of the research work, therefore, was to study the viability of this approach.

Construction of model for Motilal ki Chal

The HydroWorks package from Wallingford Software was used to develop hydraulic and hydrologic models of both the major and minor drainage systems in the catchment. The minor system consists of parallel 450 mm and 300 mm drains along the area's main road, and tributary 300 mm pipes in every other cross street.

The major system consists of street sections which sloped consistently towards the main road and to the west.

Rainfall, velocity and depth data were collected for six weeks during the 1994 monsoon using Montec recording raingauges and DETEC 3510 Surveyloggers. Chalk "telltale" gauges were used to measure maximum water levels at a number of points in the system during major storms, while field staff checked the variation of water level over time at manholes and in the streets during several events. Detailed topographic surveys were performed to confirm the dimensions and levels of both major and minor drainage networks, and catchment surveys to determine the tributary area and the types of cover were also performed.

Calibration of model

Twenty-two events, in which distinct hydraulic responses to rainfall were observed at the outlet, were defined over the interval of recorded data. These events were then divided into Calibration and Verification events, with better quality data usually reserved for verification. Eight events were used to calibrate the model, through adjustment of various hydrologic and hydraulic parameters, while the rest were kept to verify the quality of the model's predictions. Runoff was predicted from rainfall using a simplified form of the "fixed-PR" model of the Wallingford Procedure (DoE/NWC, 1983), in which the percentage of rainfall converted to runoff is fixed throughout the event, and routed through two linear reservoirs. Adjustments were made to the calibrated model until no further improvement on discharge and level predictions appeared likely. Model predictions were then compared with a variety of measures taken in the calibration events, including maximum depths of flooding, and water levels over time at manholes and various points above ground. These comparisons confirmed the impression that there was little to gain from further adjustment, so the model was deemed ready for verification.

Verification of model

The calibrated model was then tested using the data from other events set aside for verification. The testing consisted of comparison of predicted values of peak rates of runoff, values of depth over time at outlets and points where depth was observed, and maximum recorded depths at a number of chalk "telltale" gauges. Results of the verification for maximum depth prediction are summarised in Figure 1. Note that 72% of the results are within +/- 0.10 m.

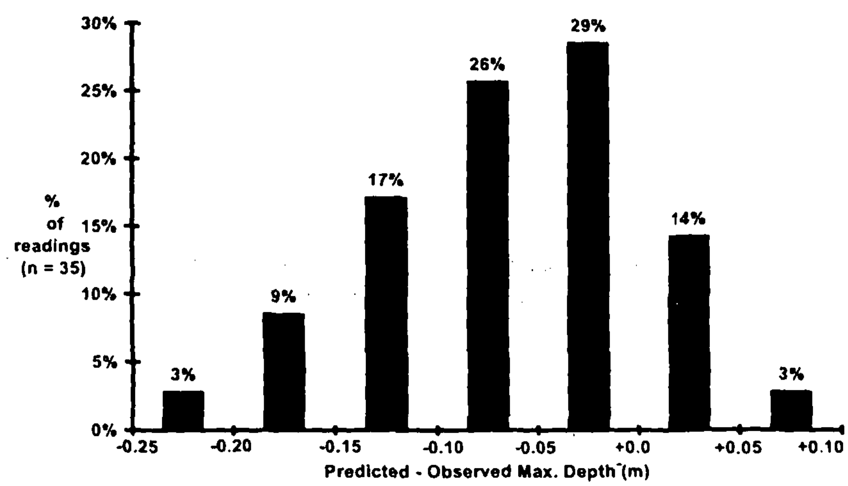

Figure 1. Error distribution for maximum depth in verification events.

Use of model

The model was developed to answer questions similar to those in the earlier open drain study. How would the blockage of inlets or pipes in the minor drainage system affect performance? How would the major drainage system perform alone if there were *no* minor system? Can the major system be redesigned to eliminate the need for a minor system?

As before, performance was to be designated in terms of the depth, duration, frequency and extent of flooding. Model simulations were performed for each recorded event under a variety of conditions of blockage (both of inlets and conduits), and results were compared. During comparison of these results, an important question of model validity arose that had not been directly tested in the calibration or previous "verification" of the model.

Validation of the model's ability to predict both depth and duration of flooding at various points in the network is conceptually straightforward. Verifying predictions of the "frequency" and "extent" of flooding, however, requires some care, as the verification will depend upon the chosen definition of flooding.

MODEL VALIDITY FOR FLOODING

Intuitively, a definition of flooding seems straightforward; flooding occurs when depth rises above a defined level. In most drainage studies, this would be taken as the road level; if water rises above the road surface, flooding has occurred. In a network where roads are designed as drains, however, this makes little sense; every event would be a flooding event whenever runoff flowed down the street as intended.

Kerb level is an alternative cut-off with which to define flooding; if the maximum water level rises above the lowest kerb elevation at a point, then flooding has occurred. Using kerb levels, it is now possible to define *the frequency of flooding at a point over the season*, as the number of events in the season in which the water rose above the lowest kerb level at that point. It is also reasonable to define *the extent of flooding in any given event* as the number of evenly-spaced nodes which flooded in a given event. Both frequency of flooding and extent of flooding can now be predicted from model estimates of maximum water levels, and observed frequency and extent of flooding can be computed for those nodes and events where water levels were observed.

The model had been considered "verified" on the basis of its predictions of maximum depth, and its overall ability to predict level and flow hydrographs. The depth predictions of this "verified" model were then used to predict flooding or not flooding relative to their corresponding kerb levels at 113 node-events for which suitable observation data were available. (A node-event consists of a prediction or observation at a node, in a given event.) Observed depths were analysed for the same node-events to determine whether flooding had in fact occurred relative to the kerb. Predictions are compared with observations in table 1 and indicate poor capacity to predict flooding at a node for a given event.

Table 1. Sensitivity and specificity of flood predictions

		Field Observations			**Predictive Values**	
		Flood	No flood	Total	(PV)	
Model Predictions	Flood	5	11	16	31%	+PV
	No flood	26	71	97	73%	-PV
	Total	31	82	113		
		16% **Sensitivity**	87% **Specificity**			

Table 1 includes the two statistical terms "sensitivity" and "specificity" which are perhaps more widely known in public health (e.g. Hennekens & Buring, 1987) than in drainage engineering. Sensitivity gauges how well a model predicts observed "positive" results (e.g. flooding), and is defined as the percentage of total observed "positives" which are successfully predicted. A model could, however, be very sensitive if it predicted a flood for every event; all observed floods would thus be successfully predicted! A good model must therefore also be "specific", and successfully predict observed "negatives" (e.g. "non-floods"). Specificity is therefore defined here as the percentage of total observed "negative" results which are successfully predicted. Positive and negative predictive value (+PV, -PV) are similar concepts, but gauge the validity of positive and negative predictions as a percentage of the *predictions* made, rather than of the *observations*.

Applying these concepts, and reviewing Table 1, suggests the following discouraging results:

- Less than one third of the predictions of flooding at a node are correct. (+PV)
- Nearly 1 in 4 model predictions of no flooding were *incorrect.* (-PV)
- Out of 31 observed floods at nodes, only 5 were successfully predicted. (Sensitivity)
- The one good result is that only 1 in 8 nodes which did not flood were erroneously predicted to flood. (Specificity)

Is it true that "a miss is as good as a mile"?

While depth is a continuous variable, the above definition of flooding is binary; in a given event, a node either floods or it doesn't. This means that predicting depth to "within 5 cm" may not be good enough, e.g. if the model predicts a water level 2 cm *below* the kerb when the actual level is 1 cm *above* the kerb. The problem is a common one of classification when switching between continuous and binary variables. While the model may seem "good" at answering questions about depth, it may not be "good" at answering questions about the binary state of flooding. The data were reviewed to see if poor predictions of "flooding at a given node in a given event" averaged out to better predictions of "frequency of flooding at a node over a season", or "extent of flooding in a given event." These averaged predictions were, however, still poor.

When such problems arise in converting between continuous and binary variables, a compromise of "banding" is often adopted. The idea is to recapture some of the lost information from the continuous variable by increasing the number of categories. The simplest approach seemed at first to be to move to three categories: "Maximum water level at least x mm *below* the kerb", "Maximum water level within +/- x mm of the kerb" and "Maximum water level at least x mm *above* the kerb." After some experimentation with trial values of x a further simplification emerged. It soon became apparent that the last category of "definitely flooded" contained very few data, and could be combined with the middle band. The question thus returned to a binary one: "Did the water level rise above a level x mm below the lowest kerb, or not?"

This form of the question is actually a traditional one for engineers who design open channel drains. Design criteria frequently specify a *minimum freeboard requirement*, where freeboard is the design height of an open section above the computed water level. Freeboard serves as a factor of safety to reflect all the uncertainties and simplifications implicit in hydraulic design and computation. A freeboard requirement of x mm is equivalent to requiring that "design computations must predict that the water level will not rise above a level x mm below the top of the channel." (see Figure 2.) The original question of "flooding at a node" is thus replaced by one of "freeboard violation"; validation then determines how well the model predicts "the frequency of freeboard violation at a node", and "the extent of freeboard violation for a given event." The previous definition of "flooding" can thus be seen as the special case of "violation of a freeboard of 0 mm."

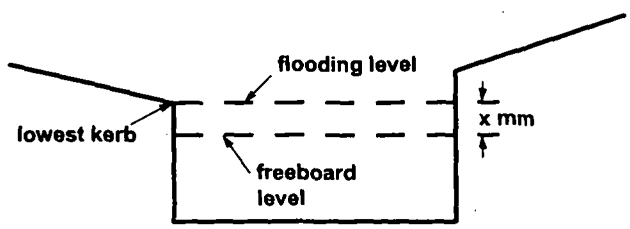

Figure 2. Definition sketch of flooding and x mm freeboard levels.

Model validity for freeboard violation

After some trials, a freeboard of 50 mm was adopted; adopting smaller freeboards made worse predictions, while greater freeboards only improved predictions marginally. The higher the value of freeboard adopted, the further the predictions move away from flooding, so 50 mm was selected as an appropriate compromise. Table 2 shows the model's ability to predict violation of a 50 mm freeboard.

Table 2. Sensitivity and specificity of predictions of 50 mm freeboard violation

		Field Observations			Predictive Value (PV)	
		Above	Below	Total		
Model Predictions	Above	29	3	32	91%	+ PV
	Below	14	57	71	80%	-PV
	Total	43	60	103		
		67%	95%			
		Sensitivity	**Specificity**			

These results are more encouraging; as they suggest the following:

- 9 out of 10 of the predictions of freeboard violation at a node were correct. (+PV)
- Only 1 in 5 model predictions of no freeboard violation were *incorrect*. (-PV)
- 2 out of 3 observed violations of freeboard were successfully predicted. (Sensitivity)
- The model correctly predicted no freeboard violation in 19 out of 20 of the observed cases where freeboard was not violated. (Specificity)

An alternative statistical question asks the following: given the above predictions and observations, how much better is the model than chance? One answer would be to compare the number of true positive and true negative predictions with the number which one would expect by chance for a random set of predictions with the same total numbers of positive and negatives. If x = the percentage of true predictions likely to be found by chance, then 100 - x equals the potential scope for improvement above chance of a good model. The Kappa statistic (Siegel and Castellan, 1988) is computed as the ratio of the model's *actual* agreement beyond chance to its *potential* agreement beyond chance. Table 3 shows how Kappa varies with the freeboard criterion adopted, and confirms the impression that the model's predictions are significantly better than chance for freeboard values of 50 and 75 mm.

Table 3. Kappa statistics for freeboard violation predictions as a function of freeboard

Freeboard Criterion	Observed model agreement	Agreement expected by chance	Agreement beyond chance	Potential Agreement Beyond Chance	Kappa	Hypothetical Significance Level (Ho: Kappa = 0)
0 mm	67%	66%	1%	34%	3%	Not significant.
25 mm	68%	61%	7%	39%	17%	Not significant
50 mm	83%	53%	30%	47%	65%	$p < 0.00001$
75 mm	85%	50%	34%	50%	69%	$p < 0.00000001$

The kappa statistics and significance levels computed above implicitly assume that each paired observation and prediction is drawn from a homogeneous population. The statistics are valid, in other words, only if all gauges and events produced similar types of discrepancies. More detailed analyses reveal that this is not the case, but kappa may still serve as a plausible heuristic indicator of whether or how the variation in cut-off level is affecting the predictive quality of the model.

Flooding: on the border between two regimes

There are intuitive reasons why predictive abilities for flooding may be poor. Flooding occurs when a freeboard of 0 mm cannot be maintained. The kerb (or "zero freeboard") level is on the border between two distinct hydraulic regimes, that of channel flow and floodplain flow, as shown in Figure 3. Each regime is characterised by a reasonably well-defined cross-section and relationship between depth and area of flow. Small errors in water level predictions around kerb level, however, (be they underestimation or overestimation) are unusually important in that they change the hydraulic regime being modelled from channel flow to floodplain flow, or vice versa. The adoption of a finite freeboard criterion is likely to permit more accurate predictions, as the "borderline" cases where small errors can lead to misclassification around the cut-off are all in the more stable and well-defined channel flow regime.

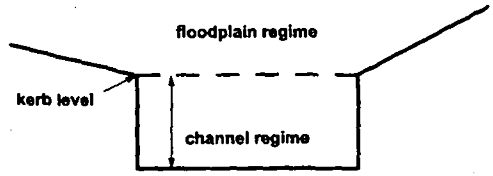

Figure 3. Definition sketch showing the hydraulic significance of the kerb level.

This improved predictive accuracy is not without cost. After all, we (and the residents of the slum!) are intuitively more interested in actual flooding, rather than in whether or not a given freeboard is violated. Nevertheless, with the data available, it would seem wiser to accept good predictions about violation of a small freeboard than to use bad predictions of flooding as a basis for gauging performance.

CONCLUSIONS

This study and previous work (Kolsky *et al.*, 1993; Heywood *et al.*, 1997) have been based on the premise that where frequent overloading of the minor system is inevitable, engineers need to understand what happens when it floods. The recent development of sophisticated computer software allows us to build and test models that help us move towards this goal. As we use models to address more questions, however, we also need to test their validity in answering them. As described in this article, such testing has forced us to distinguish between reasonable predictions of maximum depth, (which were achieved relatively easily), and reasonable predictions of flooding, which proved more difficult and less certain. Engineers have traditionally addressed uncertainty by the use of factors of safety, and we found the adoption of freeboard criteria to be helpful not only as a factor of safety in design, but also in defining the limits of what we can predict with confidence about performance.

We would naturally be happier with a model which did a better job of predicting flooding, so that we could use flooding as the basis for performance criteria, rather than freeboard violation. It is conceivable, although not obvious, that different calibration procedures might be more appropriate to maximise validity in the prediction of flooding, rather than in the prediction of level and flow, the traditional gauges of model validity (Wallingford Procedure Users Group, 1993). Where model predictions of flooding can be validated, we believe it remains the most appropriate basis for performance criteria. Where, as in this study, model predictions for flooding are poor, a different basis must be sought, as it would seem unwise to base analysis on predictions which cannot be validated. In this study, we have found that our model can make good predictions about violation of a 50 mm freeboard, and that such a basis for performance criteria is both intuitively clear and appealing to the engineer. Analysis of the data from Motilal ki Chal on this basis is nearing completion, and will be reported shortly.

ACKNOWLEDGEMENTS

The research described in this paper was funded by Research Project R5477 "Performance-Based Evaluation of Surface Water Drainage in Low Income Communities" of the Engineering Division of the British Overseas Development Administration (now the Department for International Development-DFID). However, DFID can accept no responsibility for any information provided, or views expressed. Wallingford Software Ltd. have kindly made the HydroWorks drainage analysis software available to us without charge for research purposes, and have given generous moral and practical technical support throughout the work. Mr Jon Parkinson played a key role in earlier stages of the research, including the fieldwork. The authors also wish to acknowledge the collaboration of Professor T. A. Sihorwala of the Shri Govindram Seksaria

Institute of Technology and Science in Indore, and his effective local management of the project. Finally, we wish to thank the staff of the Indore Drainage Evaluation Project for their hard and careful work in gathering high quality field data upon which these studies have depended.

REFERENCES

DoE/NWC (1983). *The Wallingford Procedure: Design and Analysis of Urban Storm Drainage. (Volume 1: Principles, Methods and Practice.)* Department of the Environment/National Water Council Standing Technical Committee Report No. 28. Hydraulics Research, Wallingford.

Hennekens, C. H. and Buring, J. E. (1987). *Epidemiology in Medicine.* Little-Brown, Boston, pp 331-335.

Heywood, G., Kolsky, P. J. and Butler, D. (1997). Modelling drainage performance in an Indian catchment. *J. CIWEM*, **11**, February, pp. 31-38.

Kolsky, P. J., Butler, D. and Sihorwala, T. A. (1993). Performance-based evaluation of urban drainage in developing countries. In: *Proceedings of the Sixth International Conference on Urban Storm Drainage*, J. Marsalek and H.C. Torno (ed.), Seapoint Publishing: Victoria B.C., pp. 1860-1865.

Kolsky, P. J., Parkinson, J. N. and Butler, D. (1996). Third-World Surface Water Drainage: The Effect of Solids on Performance. Chapter 14 in: *Low-Cost Sewerage*, D. D. Mara, (ed.) John Wiley & Sons, Chichester, pp. 189-214.

Parikh, H. (1990). *Aranya-An Approach to Settlement Design: Planning and Design of Low-Cost Housing Project at Indore, India.* Housing and Urban Development Corporation, New Delhi.

Richens, P. J. (1985). *Development of an Accurate Simulation Model for Urban Storm Sewers*, Report No IT297, DoE Contract PEDC 7/7/053, HR Wallingford Ltd, Wallingford.

Siegel, S. and Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioural Sciences*, 2nd Edn. McGraw-Hill, New York, pp. 284-291.

Wallingford Procedure Users Group (1993). *Code of Practice for the Modelling of Sewer Systems.* Water Research Centre, Swindon.

WEF/ASCE (1992). *Design and Construction of Urban Stormwater Management Systems.* American Society of Civil Engineers/Water and Environment Federation, New York and Alexandria, VA.