
Accuracy in Stormwater and Floodplain Management

Ben Urbonas

“To see our perfect ignorance is to exonerate others from imperfection”—Hugh Prather, *How to Live in the World and Still be Happy*, 2002

7.1 Introduction

Multi-million if not multi-billion dollar decisions are being made on the basis of computer modeling results each year. The question that is often not explored is How reliable and accurate are these results? Let's explore many of the issues related to the accuracy of modeling from a perspective of an engineer involved in municipal flood and stormwater management and design, research in stormwater technology, a part-time software developer, user of software, and reviewer of many hundreds of computer modeling runs by a variety of engineers in the fields of hydrology and hydraulics.

The accurate and reliable modeling of stormwater runoff (i.e. hydrology) and associated phenomena has been and continues to be a challenge, despite the advances in the models themselves, their interfaces and even their math engines. The market offers a wide variety of products with wonderful user interfaces and input–output management systems, facilitating their use by less and less skilled professionals.

All rainfall–runoff models rely on calibration to achieve accurate representation of hydrology. However <5% of all hydrologic simulation in the United States (off-the-cuff estimate/guess by the author based on observation over the last >30 y) has no simultaneous rainfall–runoff data to calibrate

against. Even when such data is available, it may not have sufficient time of record and may not be of sufficient temporal or spatial density. The result is often the pursuit by modelers of precision, with little understanding of accuracy. The big question that remains is: Is it possible to guide model users to pursue accuracy (reality) and to deemphasize the pursuit of precision?

Everyone seems to agree that the most important elements in accurate stormwater modeling are the following:

- the modeler's skill;
- selection of the appropriate model for the job;
- some are best for urban areas with small subcatchments, others are best for large non-urban catchments;
- the challenge is to select models that do well in both conditions;
- the availability of credible and appropriate rainfall and runoff data for calibration; and
- the math engine of the model.

"I think there is a world market for maybe five computers."—often attributed to Thomas Watson, chairman of IBM, 1943

7.2 Effects of Spatial Rain Gauge Density

When continuous simulations are performed over extended periods of time we often use rainfall data that is available through the National Weather Service or other sources. Unfortunately, the data is rarely collected at the site being investigated and is often only available in 60 min and 15 min clock time increments. We are very lucky if we get data at the study site itself and even luckier if we have more than one rain gauge at that site with a period of record $>3y$.

The United States Geological Survey (USGS) has been collecting simultaneous rainfall–runoff data for the Urban Drainage and Flood Control District (UDFCD) in Denver at the Harvard Gulch catchment since 1980 at 5 min time increments during the rainfall seasons of April through September. Unfortunately, because of computer system changes, USGS has lost access to this data prior to 1990 and only 15 y of 5 min rainfall and flow data was available at the time it was last examined by the author.

This Harvard Gulch catchment has a 3.1 mi^2 (7.9 km^2) area, mostly covered by single family residential neighborhoods. It also has significant areas of very dense commercial development and several small parks. USGS has been collecting continuous data at two flow gauges and five continuous rain

gauges. On occasions a rain gauge was out of operation during a specific event, and in order to maintain continuous record, the data from the closest rain gauge was used to fill in its record. This occurred in <2% of the storm events from 1990 to 2005.

A SWMM model was set up containing 59 subcatchments to describe areas of similar land uses and types of developments (see Figure 7.1). Impervious cover and drainage systems were field and aerial photo verified. All runoff elements were linked with routing elements, mostly open channel and several storm sewers, to model overflow paths when the initial design capacity of the routing element was exceeded. Figures 7.2 and 7.3 below illustrate the results of this calibration.

Rainfall data from one, two, three and four rain gauges were used to compare against the calibrated five gauge model's results in order to test the effects of rain gauge data spatial density. Figures 7.4 and 7.5 illustrate the peak flow variances for the two scenarios of different rain gauge placement within the catchment against the five gauge calibrated model. It was concluded that for the Denver region, rain gauge density of approximately 1/mi² is needed to achieve the most accurate results on a storm-by-storm basis.

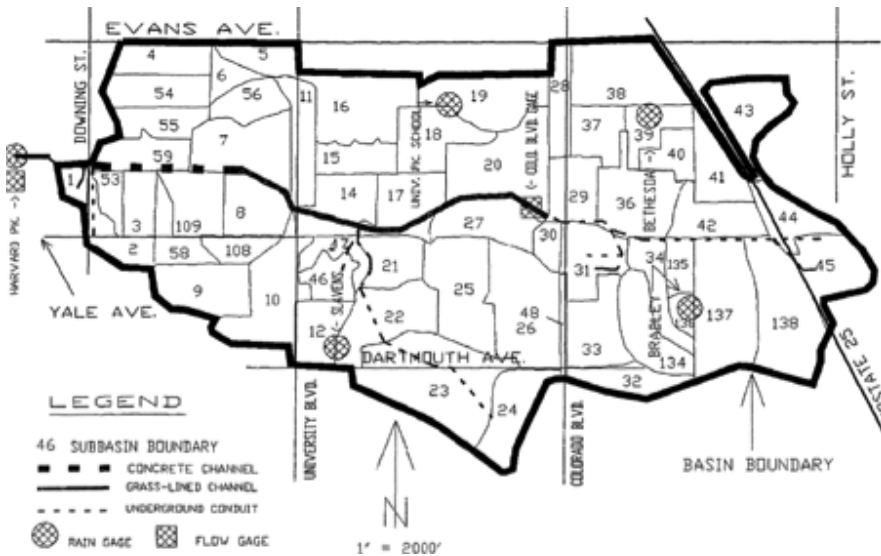


Figure 7.1 Harvard Gulch monitored catchment.

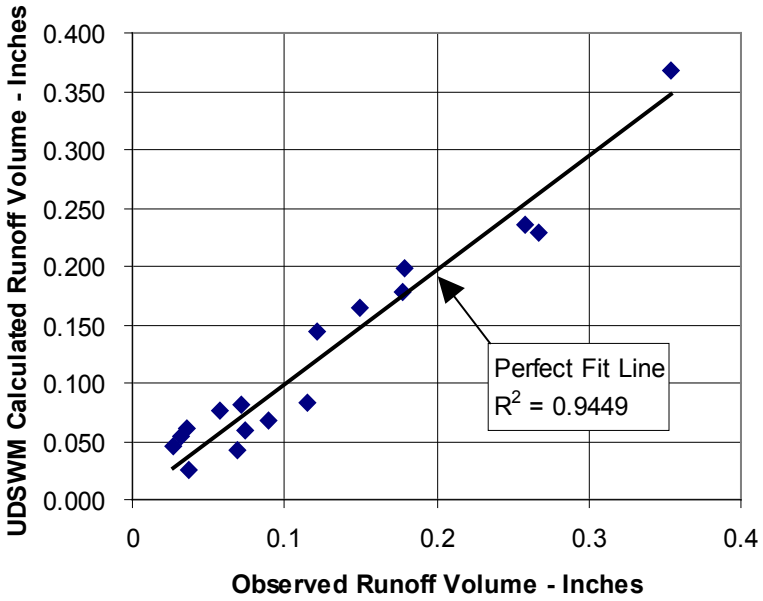


Figure 7.2 Calibration results for volumes.

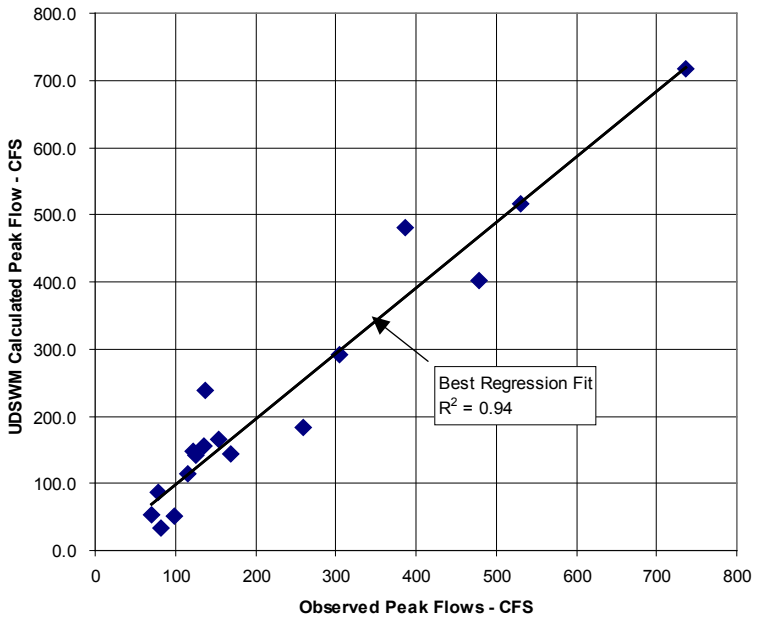


Figure 7.3 Calibration results for peaks.

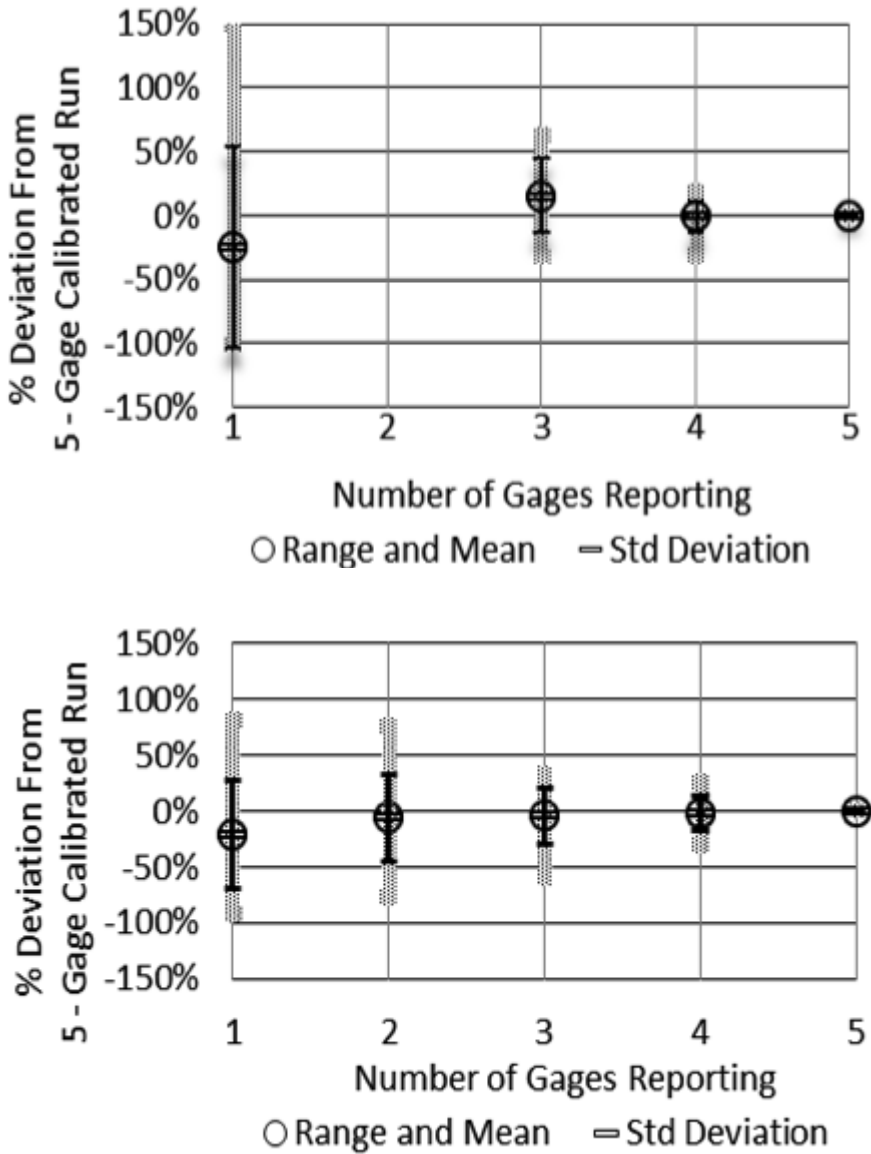


Figure 7.4 Peak flow variability, calibrated model (scenario 1 upper, scenario 2 lower).

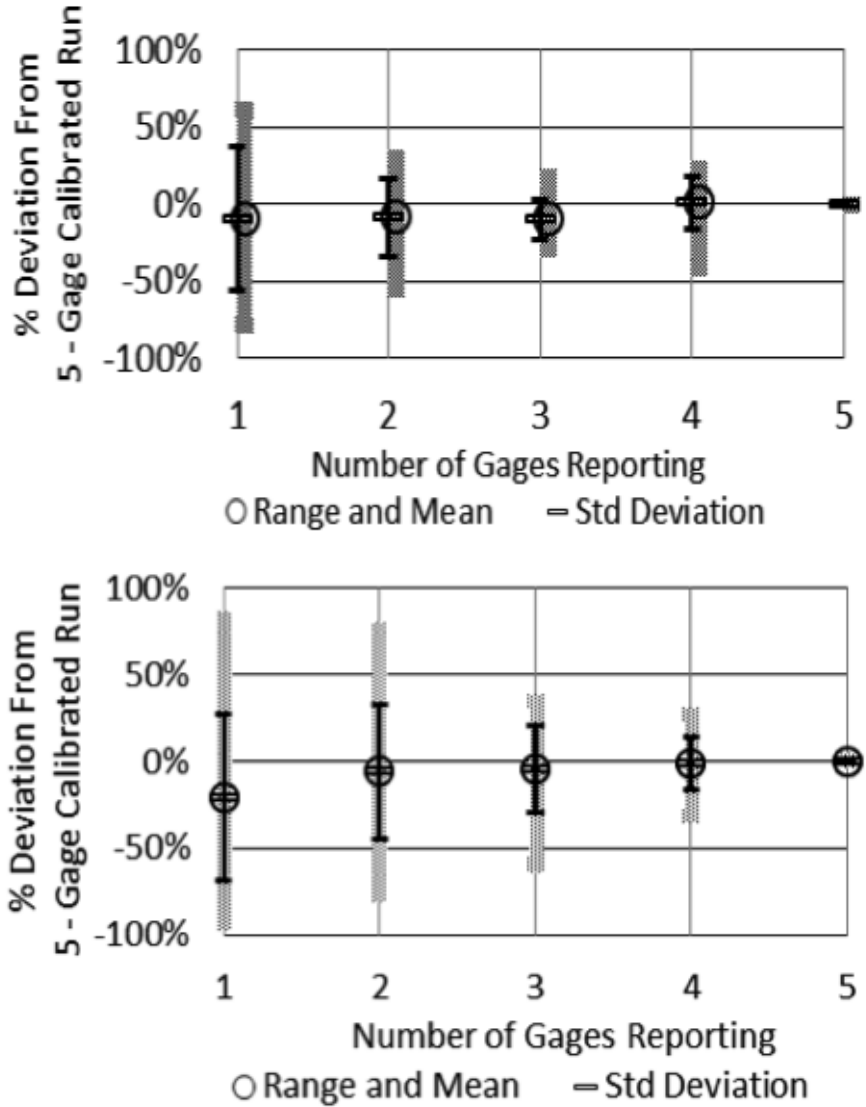


Figure 7.5 Peak volume variability, calibrated model (scenario 1 upper, scenario 2 lower).

If only one rain gauge was to be used, locating it near the centroid of this 3.1 mi² (7.9 km²) catchment gave the most accurate results and if two gauges were used, placing them at the opposite ends of the catchment gave the best results. The accuracy of continuous simulations can significantly be affected by lack of sufficient rain gauge density or even the lack of rain gauges

in the modeled catchments. As a result, the trends developed on the basis of continuous simulations have to be questioned to determine whether they are any better than using well thought out and region specific design storms when judging the sufficiency of the existing conveyance systems or the sizing for new ones. This is especially true when calibration data is unavailable.

7.3 Effects of Temporal Rainfall Data Density

The effects of temporal density were also tested using the same calibrated model. The 5 min data for all five gauges was agglomerated into 15 min and 60 min data. Effects on the peak flows are illustrated in Figure 7.6 below for the 3.1 mi² (7.9 km²) catchment.

The differences in peak flows between the 5 min and the 15 min data are quite noticeable for the larger events. When comparing the results of 5 min with 60 min data we see that the simulated peak flows differ by a factor of two. Similar, but not nearly as severe, differences were observed for the runoff volumes under urbanized conditions. Clearly, if we are to achieve a credible level of accuracy in urban runoff simulations, we need rainfall data at ≤ 15 min increments, preferably 5 min increments. In other words, temporal density of rainfall data can have a significant impact on the accuracy of simulations.

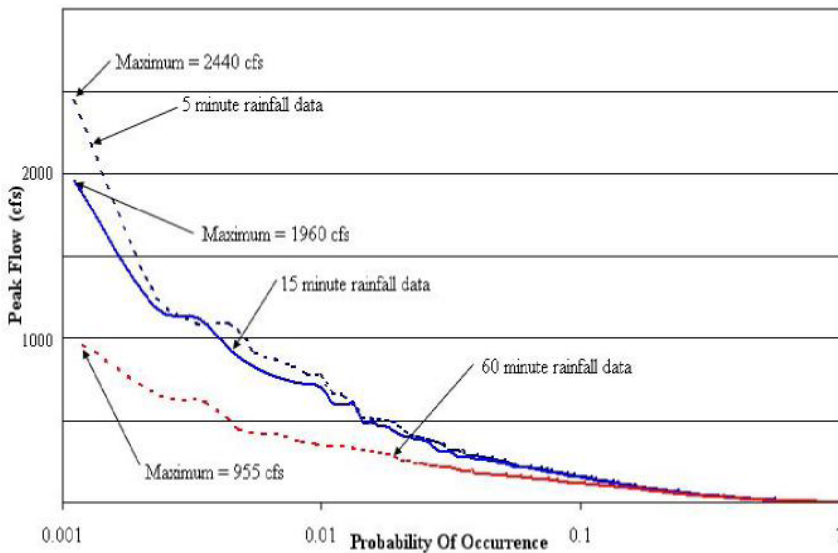


Figure 7.6 Effects of temporal rainfall data on peak flows, 3.1 mi² (7.9 km²) tributary area.

“There is no reason for any individual to have a computer in his home.”—Ken Olsen, president, chairman and co-founder of Digital Equipment Corporation, 1977

7.4 Use of Radar to Address Spatial and Temporal Rainfall Data Density Needs

High resolution radar has become popular during the last 20 y for estimating rainfall values reaching the ground. A number of companies now provide rainfall data at temporal resolutions of 5 min and spatial resolutions of 1 km². Radar data were used at the Harvard Gulch test catchment to see what accuracies simulating peak flows and volumes gave when compared to using rain gauge data. Each radar generated rainfall pixel of approximately 1 km² was input as a separate rain gauge in the calibrated SWMM model. Figure 7.7 compares the rainfall depths recorded at the five rain gauges against the radar generated rainfall totals provided by the experts in interpreting the radar reflections for the 2001 07 08 storm, the largest storm in a number of years at this site.

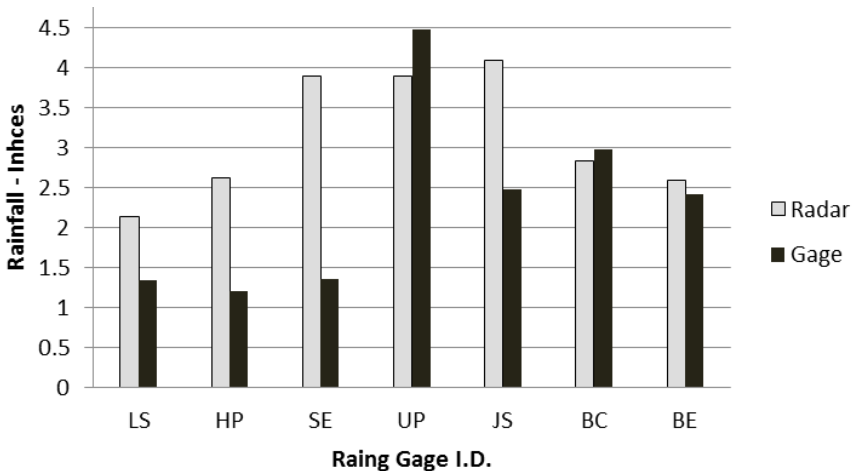


Figure 7.7 Rain gauge vs radar generated totals 2001 07 08 Harvard Gulch catchment.

The USGS monitors this catchment for the UDFCD and field estimated the peak flow at the downstream monitoring site to be 2 040 ft³/s (57.8 m³/s). An independent estimate by UDFCD using culvert backwater levels was within 5% of the USGS estimate. Both rain gauge and radar generated

rainfall data were used with the calibrated model and compared against the field estimates of peak flows. The resultant peak flows are listed in Table 7.1, where we see that the runs using rain gauge data came close to the observed flows while the radar data produced high estimates of peak flow. Clearly the use of radar generated rainfall data in this case overestimated the peak runoff rates during this storm event by a very large margin. This is not to say that radar information should not be used, but it does need to be calibrated against rain gauge data on the ground to give credible results, as has been done by many investigators in recent years.

Table 7.1 Simulation of 2001 07 08 storm using rainfall data.

Modeling Rainfall Data Scenario	Model Results (cfs)	Recorded Flow (cfs)
Using 5 min data from 5 rain gauges	2 140	2 040
Using 5 min radar generated rain data	3 250	2 040

7.5 Use of Models with No Calibration Data

7.5.1 Need to Verify Flow Measurement Data Before Accepting It

When field data is available for calibration, rarely do we ask ourselves how accurate the data is. For the Harvard Gulch catchment the data originally delivered some 18 y ago by USGS to UDFCD was tested for consistency. Initial comparisons of runoff volumes at Harvard Park and at Colorado Blvd flow gauges, when compared with rainfall volumes, were found to have runoff coefficients of 0.15 and 0.34 respectively. The coefficient at Colorado Blvd was reasonable for the land uses within the tributary area, but the coefficient for Harvard Park was clearly too low. This prompted careful analysis of the flow gauge rating curves at this location. Based on detailed field surveys of 16 channel cross sections within an 850 ft (259 m) reach near that gauge, and careful assignment of roughness coefficient for each portion of each section, and follow-up calibration against profiles of various flows, a new set of depth–runoff rating curves was developed. Based on this in-depth analysis, a runoff coefficient of 0.25 resulted, which was judged to be much more representative of the upstream land uses. Following this analysis, USGS was prompted to conduct its own depth–flow rating studies and the data received since then appear to be quite reasonable.

Lesson learned: Never use rainfall–runoff data for calibration without checking it for integrity and reasonableness. This may appear as a trivial recommendation, but there are many examples of the blind use of data. Just because data came from government or academic organizations does not

necessarily meant it has been tested for accuracy or reasonableness. We need to recognize that not all data is scrutinized for accuracy before publication.

7.5.2 Use of Uncalibrated Models

Shortly after the modeling of radar generated rainfall data at the Harvard Gulch catchment, a separate runoff model was set up independently by a consultant working for the City and County of Denver to help Denver assess their drainage system needs. They used local criteria without any calibration, a practice that is typical for most of the storm system analysis and design throughout the United States. Their model was used to simulate the 2001 07 08 runoff with rain gauge and radar generated rainfall data. Table 7.2 summarizes the comparisons between the field estimates, calibrated model peak flow estimates and the consultant's model results using the two sets of rainfall data.

Table 7.2 Simulation of 2001 07 08 storm using rainfall data.

Peak Flow source	Peak Flow	% Difference from Field Estimate
Field estimate of peak flow	2 040 cfs	
Calibrated SWMM with 5 gauges	2 140 cfs	+5%
Uncalibrated model with 5 rain gauges data	5 780 cfs	+180%
Uncalibrated model with radar generated rainfall	7 090 cfs	+250%

This illustrates the case that the blind use of off-the-shelf models and published criteria or standards can lead to costly mistakes in analysis and design. Unfortunately, most urban catchments do not have data to calibrate against and often the engineer or modeler falls back on local modeling/calculating criteria. This blind use of criteria, without much thought about whether the results make any sense or if they have any relation to observations in the field, is unfortunately a common practice. Putting it another way, it is not the model but the modeler that can make the greatest difference in whether the results are reliable or not. Lesson learned: if local calibration data is not available, at least check for the reasonableness of the results from regionally available data or even from observations by local residents.

7.5.3 Blind Test of Models by UWRRC

Nine uncalibrated rainfall–runoff models were tested under the sponsorship of the ASCE Urban Water Resources Research Council in the mid-1990s using data from two gauged sites. Phillip Zarriello of the USGS Ithaca, NY office provided rainfall data and catchment parameter data to volunteers experienced with each of the models, namely: CASC2D, CUHP, CUHP/SWM, DR3M,

HEC-1 (KW option), HSPF, PSRM, EPA SWMM, and TR20. The resultant findings for the peak flows and comparisons against the measured peaks for each of the events at the two sites are presented in Figure 7.8 (Zarriello, 1998).

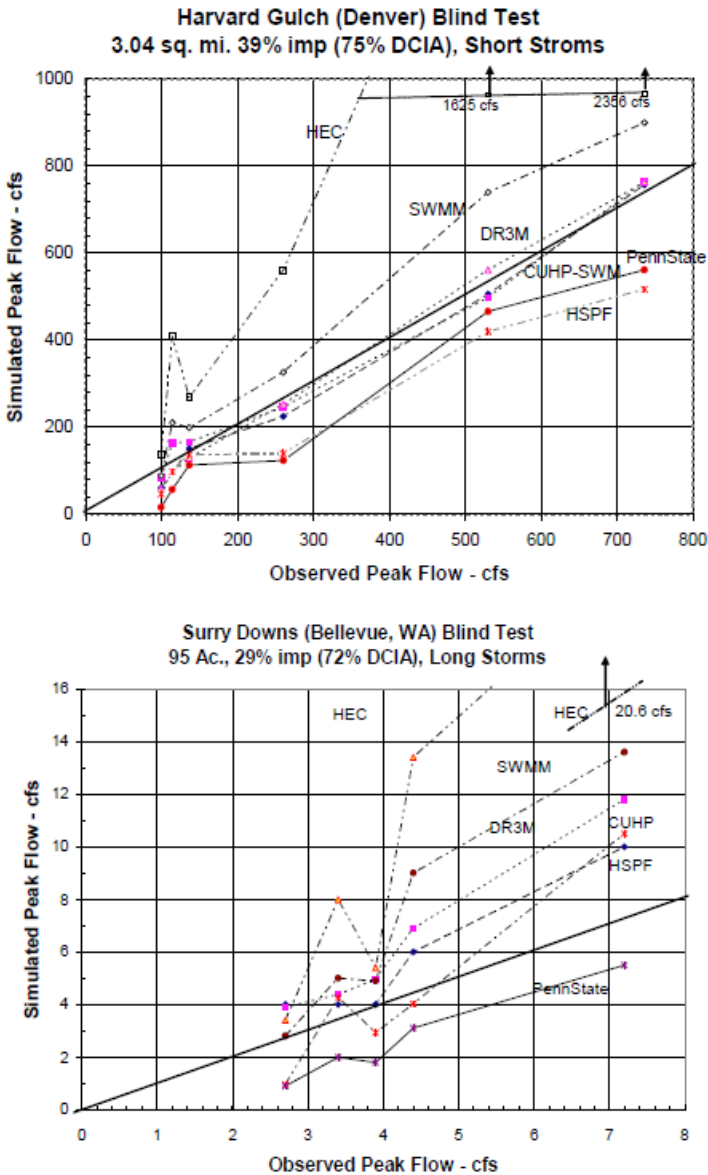


Figure 7.8 Results of blind testing: simulated vs recorded peak flows at two locations.

Unfortunately the TR20 results are not shown on this figure, but simulations by TR20 consistently underestimated the peak flows by a factor of about 50% or more. What the findings illustrated in Figure 7.8 tell us is that uncalibrated models can either be on target or can be off by a very large factor from the recorded peak flows.

“640K ought to be enough for anybody.”—rumoured to have been said by Bill Gates, 1981

7.6 Uncertainties Resulting from Rainfall Data Length of Record

Figure 7.9 illustrates the mean, 5% and 95% confidence bands for the mean when the continuous simulation peak flows are processed using Log Pearson Type 3 flow frequency analysis using annual duration series (i.e. annual peaks) for the 15 y recorded simulation results using the 5 rain gauge data set for Harvard Gulch.

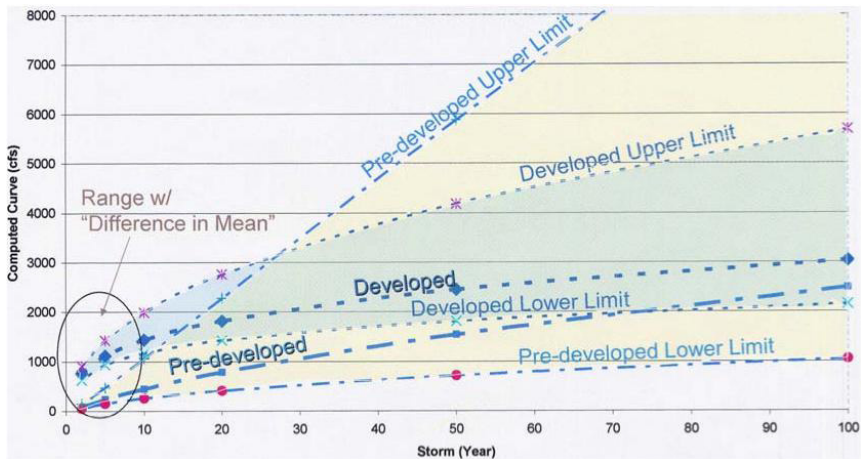


Figure 7.9 Peak flow and confidence limits vs return period at Harvard Gulch.

Note how the means for the pre- and post-developed catchment conditions overlap past the 10 y return periods. What is even more interesting is how the uncertainty of the mean increases with increasing return period to a point where the mean loses credibility at higher return periods. Keep in mind that this was done using a calibrated model with high spatial and temporal density of rainfall data, something that was argued for earlier in this chapter.

Can we really claim that continuous simulation results are more accurate for the very large events than ones obtained using a design storm? In this example, because the field record at this site has only 15 y data, we are hard pressed to make that claim. If, on the other hand, the long term rainfall record had 60 y of 15 min or 5 min data from the National Weather Service, the confidence bands would become much narrower and the high return period peaks would have more credibility.

7.7 Final Observations

The challenge is to distinguish between accuracy and precision in modeling. Very often we see discussions and claims of how much more accurate one model is over another. Often that claim really is that one model provides more precision in how and what input parameters are used and not how accurate the final results may be. The truth is that good hydrologic computer modeling is more dependent on the expertise of the modeler than the model used. However, some models are better suited for urban runoff simulation while others for simulating runoff from large undeveloped watersheds. The above examples were used to illustrate that there are many pitfalls that can lead to conclusions or recommendations that are not representative of reality or of what nature may produce.

Is modeling needed? Absolutely!

Does it serve a useful purpose? Without a question!

Does it produce accurate and defensible results? Sometimes, but is mostly a function of the expertise of the modeler and not of the specific model used.

References

- UDFCD (2005). Colorado Urban Hydrograph Procedure Excel-Based Computer Program – User Manual, Urban Drainage and Flood Control District.
- Zarriello, P. J. (1998). “Comparison of Nine Uncalibrated Runoff Models to Observed Flows in Two Small Urban Catchments,” Proceedings First Federal Interagency Hydrology Model Conference, Las Vegas, NV, April, 1998: Subcommittee on Hydrology of Interagency Advisory Committee on Water Data, p. 7–163 to 7–170.

