UNCERTAINTY ANALYSIS IN STORM SEWER COLLECTION SYSTEMS USING MONTE CARLO SIMULATION AND PARALLEL COMPUTING

Ahmad Malekpour, Fabian Papa, Djordje Radulj, HydraTek & Associates*

HydraTek & Associates, 216 Chrislea Road, Suite 204, Vaughan, Ontario L4L 8S5 www.hydratek.com a.malekpour@hydratek.com

ABSTRACT

This paper aims at showing how parallel computing can make Monte Carlo simulation practical in the context of sewer system design. To this end, SWMM 5 code was adjusted and used as the computational engine in a mother program distributing numerous individual SWMM 5 simulations onto a cluster consisting of 200 computational cores. The Monte Carlo simulation is then performed for an anonymous sewer system to determine the probability distributions for the maximum water levels in some key manholes as well as for the maximum flow discharging through the sole outfall of the system. More than 10,000 simulations were performed to reach the convergence for the case under study. The results were then used in preparing probability distributions for maximum water level at the shafts and manholes, as well as the maximum flow at the outfall. The results show that the parallel processing in this case is highly scalable and can be employed to perform such analysis over a practical time period if the proper number of processors are employed. Fortunately, the associated costs with rental machines on the cloud are reducing considerably over time and, in the near future, such analysis is expected to be conducted at a very low cost.

INTRODUCTION

Sewer systems are key infrastructure elements supporting modern society and, if not comprehensively designed, they have the potential to compromise the health and safety of the public, as well as cause considerable damage to property. Inadequate sizing of conduit and storage systems, for example, may result in large volumes of polluted sewer flow to be spilled over the streets or otherwise to the environment during a storm event and, in the process, producing health risks, damage and inconvenience. Particularly in the case of deep sewers, the compression of large entrapped air pockets may concentrate a significant amount of energy at a local point of the system and, when released suddenly through shafts or manholes, significant geysering may occur. The resulting transient pressures may be strong enough to result in the explosive release of a high velocity air-water mixture at the surface with the associated risks to health and safety, as well as potential damages at the surface and to the sewer system structures themselves.

To prevent the aforementioned situations, special attention should thus be paid to the design of sewer system components. However, due to some technical issues the design of sewer systems could be quite challenging. During an extreme storm event, the system must response to the time-varying flow generated in the sub-catchments, and the nature of the resulting flow in the sewer system is of a complex transient type which is dependent upon many parameters of the system. Considering that these parameters mostly contain uncertainties, the challenge to the decision maker is related to understanding the extent to which uncertainty in system parameters affects the transient responses of the system. Due to the complex nature of transient flow in sewer systems, the prediction of how changes in different input parameters for the system affect the transient response of the system is almost impossible even in small systems with a few pipes and storage components. This complexity is rooted in the fact that transient flow increases the degrees of freedom in the analysis of the system far beyond that in a steady state flow condition.

To overcome this problem in a real engineering project the designer either choses to be conservative in estimating the design parameters, or to do some sort of sensitivity analysis to earn some physical insight as to how the system responses to uncertain design parameters. The first option may converge on an overdesigned solution. In addition, due to the complex nature of transient flow, it is not always easy to predict what combination of input parameters produces the worst conditions in the system (and its various components).

In the second approach, the sensitivity of responses of the system with regard to some design parameters can be tested by performing some transient analyses in which the target parameter is changed while other parameters are kept at their original values. However, this approach also seems to be inefficient because, due to time and budget restrictions in a practical setting, only a limited number of test cases, out of a very large number of potential combinations of design parameters, can be practically considered, so it is very unlikely that the worst case condition for each design element would be captured using such a simplistic approach.

During the last couple of years the authors of this paper have been involved in the design and analysis of large scale sewer flow pipe systems in Ontario and British Columbia, and their experience has revealed that finding a safe and simultaneously economical design is very cumbersome, if not impossible, using currently available methods. This fact has encouraged the authors to search for a practical way to handle the aforementioned challenges.

A Monte Carlo simulation approach appears to be a theoretically possible solution to resolve the challenge. In this approach, many thousands of simulations are performed by the design parameters which are randomly changed within their possible range. The outcome of such analysis is a continuous range of transient responses in the probabilistic domain assisting designers and decision makers to make trade-offs between safety and reliability. A significant practical hurdle, however, is the time-consuming nature of the Monte Carlo analysis. Depending the size of the system and simulation time required, performing many thousands of transient analyses may take in the order of a few days for smaller systems and perhaps much longer in larger systems if the simulations are conducted with the aid of conventional PCs typically used in engineering offices.

Considering that the rapid advancement in computer technology is expected to make high volume calculation and cloud computing on a rental basis viable in the near future (AWS, 2017), the aforementioned challenges could be resolved with the aid of parallel computing. Thus this paper examines how to apply Monte Carlo simulation in context of sewer system performance analysis and to explore to what extent parallel computing can reduce the associated time requirements.

COMPUTATIONAL TOOL

The key tool for performing Monte Carlo simulation is a reliable code that can accurately capture transient flow in sewer system. To this end, the SWMM 5 code developed by US EPA was adjusted to the needs and purpose of this paper. This code is widely used to analyze the responses of sewer systems to hydrological events. Different features of the software include the handling of sub-catchment flow hydrographs, snow melt flows, LID controls, unsteady flow analysis of closed conduits and open channels for both prismatic and non-prismatic flow cross-sectional areas, real-time control logic (RTC), varying hydraulic control structures, amongst others. It allows the user to simulate complex sewer systems as they wish. Moreover, it is shown that the Dynamic Wave Routine (DWR) of SWMM 5 can accurately handle transient flow in both open channel and closed conduit systems, provided that the channel or conduit is broken down into an adequate number of short segments. The DWR, however, cannot account for the elastic feature of the flow or waterhammer effects in pressurized systems, which is the only shortcoming of the model. Nevertheless, mass oscillations dominate the flow regime in sewer systems and the elastic feature of the flow is of lesser importance. This is because the large amount of storage and space in the system prevents the liquid from becoming either compressed or stretched in the conduit to such an extent that the production of waterhammer pressures is typically not significant.

It is worth mentioning that SWMM 5 can capture mass oscillation in sewer systems much better than the software using the Preissmann Slot Method approach in calculating mixed transient flow in close conduits (e.g. InfoWorks). The main reason for this is due to issues with control of the numerical instability associated with the Slot Method for which a very low acoustic wave velocity (e.g., 10 to 30 m/s) has to be employed which, in turn, often necessitates a relatively wide slot. Since the slot extends all the way up to the ground level, the wider slot produces non-physical storage which may significantly attenuate the surge amplitude. The unrealistic attenuation becomes more intensified in deep conduits such as tunnels in which the tall slot results in large available storage volumes not representative of reality. Fortunately, SWMM 5 does not suffer from this shortcoming because its engine employs a link-based approach for solving the continuity and momentum equations. Ridgway and Kumpula (2008) compared the performance of SWMM 5, InfoWorks and some other commercial software packages against experimental data and their results demonstrated the superiority of SWMM 5 in capturing mass oscillation in a simple pipe-reservoir system.

In order to perform parallel computing using SWMM 5, the code was adjusted to distribute the computational load over several computer clusters. To this end, MPI API functions are used to send different individual SWMM 5 simulations amongst the clusters. Fortunately, the nature of the problem is such that there is no need for these individual simulations to communicate through the simulation process and thus the problem is of an embraced parallel type (i.e., thread-safe). This, as will be shown later in the paper, makes the scalability of the parallel processing to be almost linear. The practical implication is that the required run time will decrease (roughly) linearly with an increasing number of CPUs.

HYPOTHETICAL SEWER SYSTEM

In order to perform the Monte Carlo simulation, the hypothetical sewer system shown in Figure 1 is considered. This system is inspired by a real engineering problem the authors were engaged upon, however, the sizing and incoming flow was adjusted in such a way that the resulting transient flow in the system becomes harsher. This system receives sewer flow from a large sewer collection system via storage shafts S1 to S6, and its function is to attenuate (route) the flow before releasing it to the river. Considering the flow hydrographs shown in Figure 2 reveals that that the total instantaneous flow could be as high as 65 m^3/s, being far greater than the flow permitted to be released to the river, being 16 m^3/s.

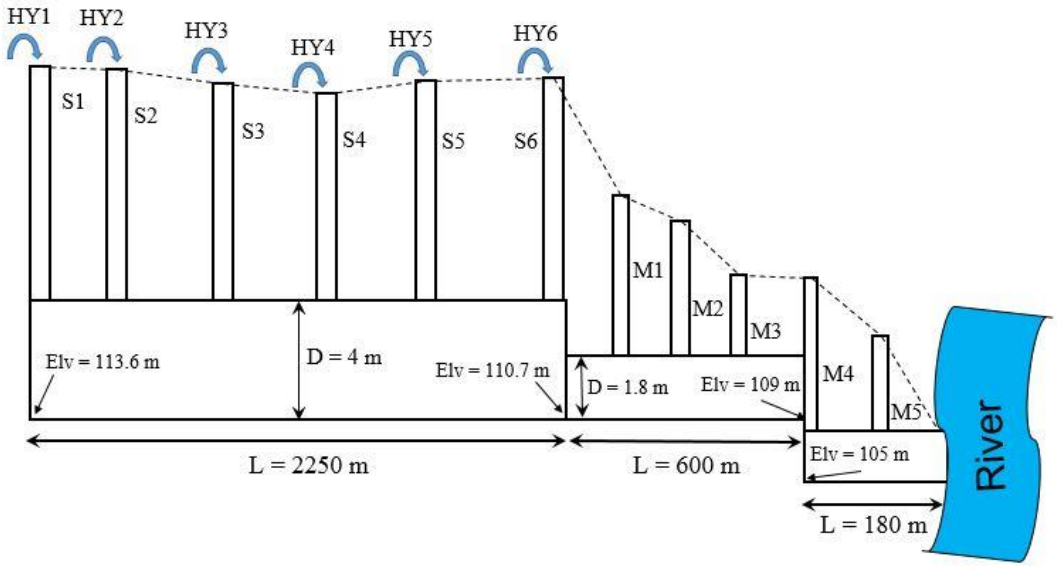

FIGURE 1 SCHAMTIC OF SEWER SYSTEM

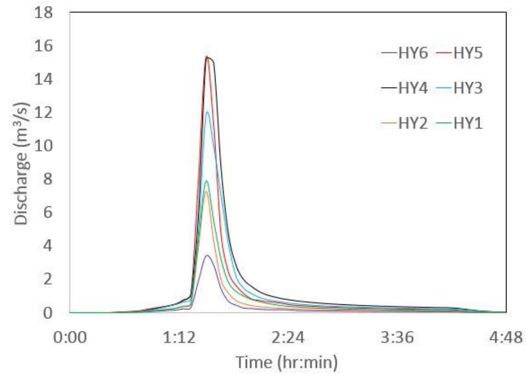

FIGURE 2 INFLOW HYDROGRAPHS

To attenuate the instantaneous flow, some storage is provided through a large size tunnel with a diameter and length of 4 m and 2,250 m, respectively. Further storage is provided in the storage shafts S1 to S5. This large tunnel is connected at its downstream end to a smaller tunnel with an internal diameter of 1.8 m. This sudden decrease in conduit diameter is made to induce a local flow restriction in the system. This local restriction causes the flow to back up and for the tunnel and shaft storage come into action during the event.

In order to accurately capture the transient response of the system, the conduits are broken down to small segments of 5 m each. To examine what combination of storage shaft size is capable of controlling the flow released to the river, several transient analyses were conducted. In all analyses, it was assumed that Manning's roughness coefficient was 0.014 for all conduits, and that the allowable freeboard in the storage shafts and manholes should never be less than 0.60 m. Table 1 summarizes the outcome of this exhaustive analysis. As can be seen, no parts of the system experience overflow, and the freeboard guideline is generally satisfied. Figure 3 and Figure 4 also depict the time history of the water level in the storage shafts and of the flow release to the river. As the results show the flow release to the river is adequately controlled.

TABLE 1 STORGE SHAFTS AND MANHOLES INFORMATION

Shaft / Manhole	Diameter (m)	Top Elevation (m)	Max. Water Level (m)	Freeboard (m)
S1	12	140.14	139.21	0.93
S2	6	140.13	138.95	1.18
S3	6	138.72	138.02	0.7
S4	6	138.36	137.71	0.65
S5	11	138.97	137.98	0.99
S6	11	139.13	138.44	0.69
M1	2.5	130.45	128.36	2.09
M2	2.5	127.04	124.51	2.53
M3	2.5	122.7	120.67	2.03
M4	2.5	122.1	115.07	7.03
M5	2.5	112.71	112.33	0.38

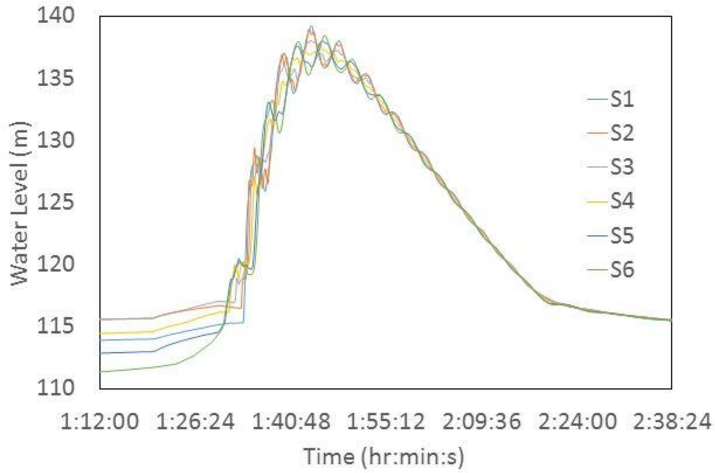

FIGURE 3 TIME HISTORY OF WATER LEVEL IN SHAFTS

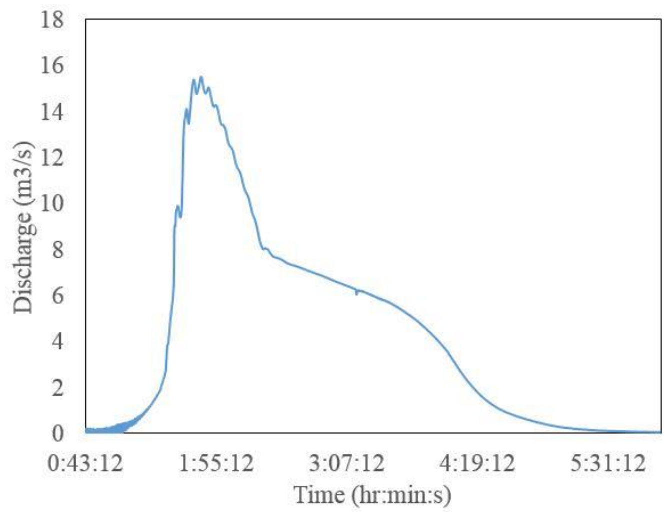

FIGURE 4 TIME HISTORY OF RELEASE RATE TO RIVER

In summary, the results of this typical analysis and design process illustrates that, if the input data is exact, the system is expected to perform satisfactorily during the design storm. However, the concern is whether the input data contains some sort of uncertainty, and how the transient responses of the system would be affected by this uncertainty. This matter is addressed through the results obtained from a Monte Carlo analysis in the subsequent section.

MONTE CARLO SIMULATION

Monte Carlo methods are a broad class of computational algorithm which depends on repeated random sampling to generate numerical results. The Monte Carlo methods in general consist of a few steps including: 1) defining a domain of possible inputs; 2) generating inputs randomly from a probability distribution over the domain; 3) performing deterministic computation on the inputs; and 4) aggregating the results into a probability distribution.

To perform Monte Carlo simulation in the context of the current problem, it is assumed that the hydrographs influent to the storage shafts, and the conduits' roughness, are the input parameters of the system containing uncertainty. If the inflow hydrograph ordinates are assumed to contain an error tolerance of ±20% and the roughness coefficient varies with a tolerance of ±25%, the objective of the Monte Carlo simulation is to calculate how the combination of the uncertain parameters affect the probability distribution of transient responses of the system. To this end, a large number of numerical inputs are generated by randomly changing the input parameters of the system within its predefined tolerance. In this case, uncertain input parameters are assumed, based on uniform probability distributions, for the six hydrographs suppling the system and the roughness coefficient of the conduits. For each sample, a transient analysis is performed and the transient responses of the system at given location are stored. Aggregating the transient responses of the system at given locations produces a probability distribution of the transient responses at those points.

It is worth mentioning that a minimum number of samples are required in order for the Monte Carlo simulation to converge. The convergence depends on some factors, the most important of which are the number of uncertain input parameters and the sampling method. One approach to test if convergence has been achieved is to incrementally increase the number of samples until no significant changes are further observed in the results.

Numerical results

Monte Carlo simulation for the current problem is conducted through performing transient analyses for 10,000 randomly generated samples on a computer cluster with 25 nodes each consisting of 8 CPUs. The minimum freeboard in the shafts and manholes as well as the maximum released discharge to the river for each sample are stored in each simulation. Considering the number of samples, the Monte Carlo simulation results in 10,000 minimum freeboards for each shaft and manhole, and the same number of results for the maximum flow release into the river.

To evaluate the probabilistic characteristics of the results, each series of data is first sorted in descending order and each is indexed form 1 to 10,000. The cumulative probability associated with each data point is then calculated using the following equation:

$$\mathrm{p}[X > x] \ = 1 - \frac{\text{Index}}{\text{Number of Samples} + 1}$$

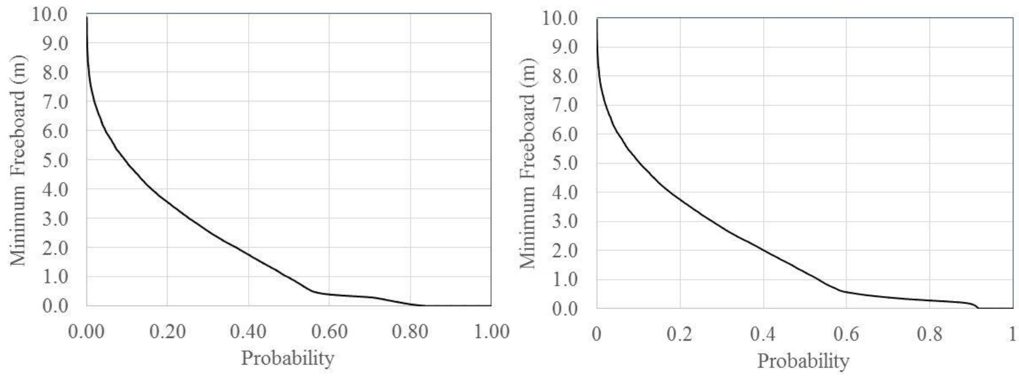

FIGURE 5 MINIMUM FREEBOARD: LEFT S1; RIGHT S2

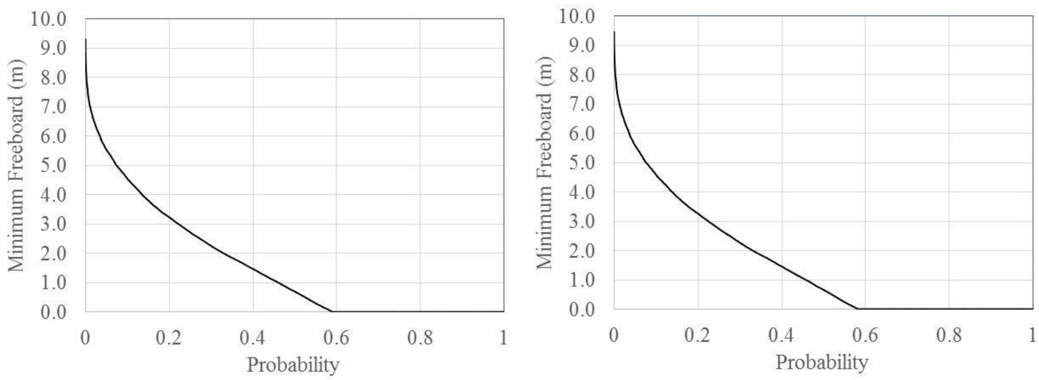

FIGURE 6 MINIMUM FREEBOARD: LEFT S3; RIGHT S4

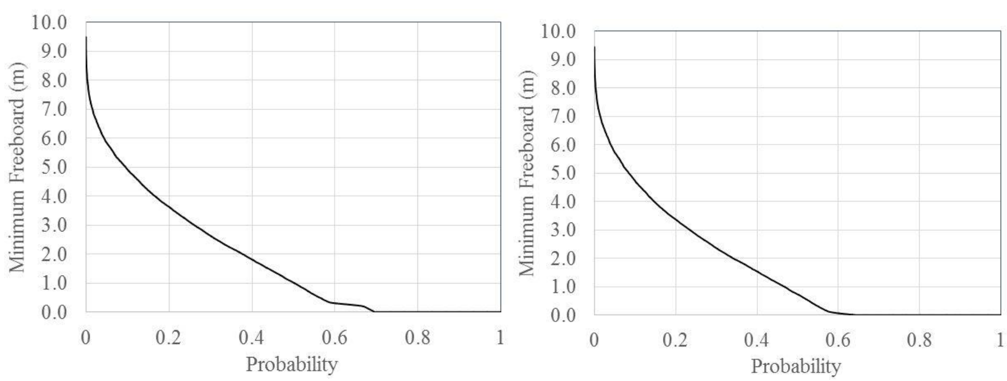

FIGURE 7 MINIMUM FREEBOARD: LEFT S5; RIGHT S6

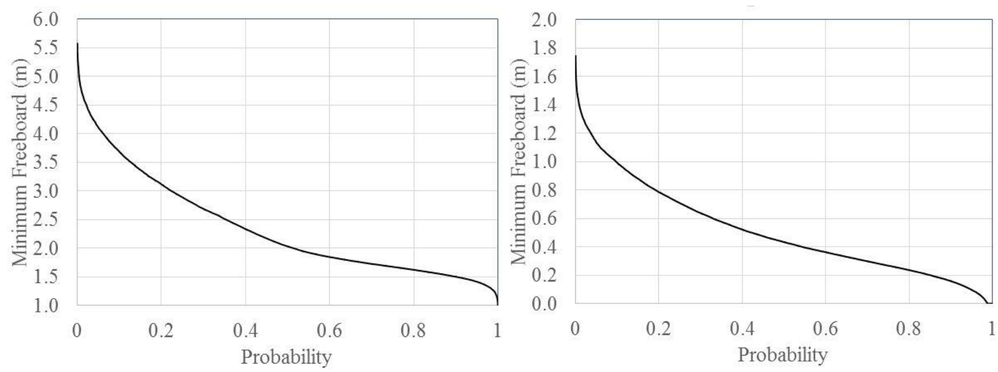

FIGURE 8 MINIMUM FREEBOARD: LEFT M3; RIGHT M5

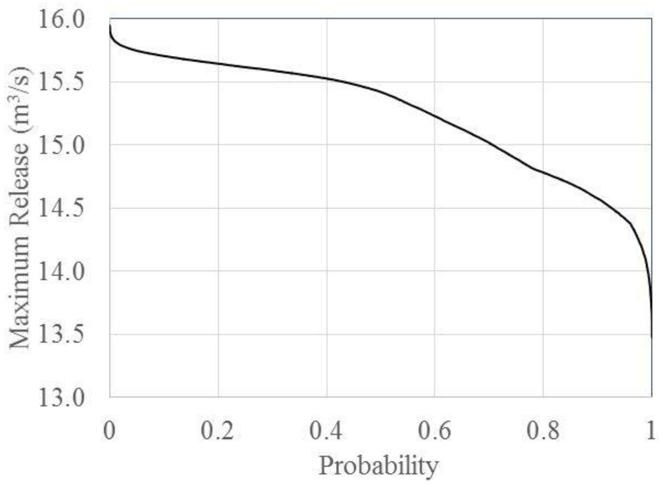

FIGURE 9 MAXIMUM DISCHARGE RELEASED TO THE RIVER

The minimum freeboard versus probability for shafts S1 to S6 are shown in Figures 5 to 7. Interesting to note is that these figures provide valuable information which can be used in assessing the performance of the system with uncertain parameters. As can be seen in Figure 5, shafts S1 and S2 do not experience overflow with a reliability of approximately 80% and 90%, respectively, while this reliability drops to slightly less than 60% for both shafts S3 and S4 (see Figure 6). Shafts S5 and S6 as shown in Figure 7 experience no overflow with reliability of approximately 70% and 60%, respectively.

The results also show that manholes never experience overflow, as indicated in Figure 8. It is worth mentioning that the flow restriction at the connection point of large diameter tunnel with the smaller one results in the transient flow at the downstream side of the restriction where the manholes are located to be significantly smoother than that in the storage shaft area, explaining the high reliability at the manholes (i.e., no overflow/spillage).

Overall, the reliability of the storage shafts against overflow are comparatively low and the design seems to contain inherent risks associated with input parameter uncertainties and which are not otherwise visible using current design approaches employing few scenarios, though without considering uncertainty the minimum freeboards are found to be in acceptable level. By further increasing shaft sizes the reliability of the system against overflow can be increased to the desired level. This goal can be accomplished through a trial and error procedure. It is noted, however, that a rational approach to design should duly consider an analysis of costs and benefits, as well as an assessment of what an acceptable level of risk might be so as to appropriately balance performance (level of service) with economy of its delivery.

Unlike the freeboard in the shafts, the maximum flow released to the river presented in Figure 9 shows no significant sensitivity to the uncertain input parameters and the maximum flow remains bounded between 13.5 m^3/s and 16 m^3/s. This implies that the proposed design is quite robust in controlling the flow release to the river.

The smoothness of the response curves confirms the convergence of the Monte Carlo simulation. Nevertheless, to independently check the convergence, another simulation was performed with 20,000 samples and it was found that the results are in excellent agreement with those obtained from 10,000 samples.

It is worth mentioning that, with conventional PCs typically utilized in engineering offices, performing 20,000 transient analyses for such a simple case would take approximately 44 days, but this time is significantly reduced to approximately 5 hours for the computer cluster arrangement employed. Given that convergence was obtained for 10,000 samples, the actual time needed to arrive at an acceptable solution to this problem is halved. The numerical experiments shows that scalability of parallel processing is almost linear for this kind of problem, meaning that further increasing the number CPUs, the run time can be further decreased in a near-linear manner. This is a tempting finding, signifying that this approach can be used in practice irrespective of how big the system is and regardless of the time-consuming nature of the problem. Performing conventional engineering simulations on a computer cluster is no longer a fantasy; rather, it could be a practical approach considering that the continual advancement of computer technology makes the usage of rental machines on the cloud to become more practical and affordable.

Finally, the results presented show that Monte Carlo simulation coupled with parallel computing is a strong tool which can provide decision makers and designers with invaluable information. In addition, this tool can be used in the value engineering stage to evaluate the performance of an already designed sewer system.

APPLICATION IN ADDRESSING CLIMATE CHANGE UNCERTAINTY

In addition to evaluating system vulnerabilities to uncertainties in input parameters under current climatic conditions, this approach can also be used to understand potential vulnerabilities given climate change as an additional uncertainty to consider in this regard. The case study used in this paper, as it relates to input hydrograph uncertainty, is a direct example of such an application.

CONCLUSIONS

In this research, Monte Carlo simulation was performed on a hypothetical sewer system by means of parallel processing on a computer cluster with 200 CPUs. SWMM 5 code was adjusted to generate the required samples and to distribute the computational load amongst the clusters. The results show that Monte Carlo simulation provides invaluable information helping the designers and decision makers identify how uncertainty in input parameters affects the performance of a sewer system. The results also reveal that the problem is of a parallel embraced type with near-linear scalability. This implies that, regardless of the size of the sewer system, such analysis can be performed in a practical time period by increasing the number of CPUs. Fortunately, rapid advancement in computer technology is expected to allow engineers to perform such analyses in the near future with rental computational resources on the cloud within a practical time period and at a reasonable price.

BIBLIOGRAPHY

Amazon. (2017, February 10). *Amazon Web Services (AWS)*. Retrieved from Cloud Compute with AWS: https://aws.amazon.com/products/compute/?nc2=h_l3_db

Ridgway, R. E., & Kumpala, G. (2008). Surge Modeling in Sewers Using Alternative Hydraulic Software Programs. *Journal of Water Management Modeling* . doi:10.14796/JWMM R228-10.