

CE-791

SeyedMojtaba Noghabaei

North Carolina State University

CE-791

Pipe Overflow Prediction Model in Stormwater Networks and Visualization Using Mixed Reality

# Abstract

**Objective**: The main objective of this paper is to predict overflows in stormwater networks using data-driven models and improve collaboration between utility managers and technicians using smart visualization techniques.  
**Methods**: In order to achieve better communication and visualization for utility managers and develop overflow prediction models, the researchers in this paper used digital twin models in a Mixed Reality (MR) environment which is connected to a real-time overflow prediction model based on forecasted precipitation or rainfall hydrograph.  
**Results**: The results of this paper indicate that it is possible to predict pipe overflows in stormwater networks using neural networks with a very high accuracy and connect that model to an MR environment in a 3D engine.  
**Conclusion**: The prediction model which is proposed by the researchers in this paper can improve the efficiency of stormwater management systems and help utility managers in hard decision-making processes and improve their performance during disasters.

Keywords: Mixed Reality; Smart Cities; Digital Twin Model; Artificial Neural Network (ANN); Overflows prediction model

# 1. Introduction

**One of the new and growing technologies in the U.S. and all over the world is smart city technologies. The smart cities technologies can save over $5 trillion for enterprises, governments, and citizens annually over the world (Ismail 2017). For example, there is a $14 billion saving opportunity for enterprises in areas like more efficient transportation options such as drones, robots or driverless cars and trucks and there is a $5 billion annual saving opportunity for governments using street lighting and smart building technologies. Using smart street lights it is expected to reduce repair and maintenance costs by 30 percent. It is possible to utilize smart city technologies in new areas in order to achieve more savings.**

Many researchers tried to utilize smart cities technologies to achieve more savings in different areas like smart transportation, smart infrastructure, smart Disaster Management System (DMS), and smart visualization. Two areas that still need more developments are smart DMS and smart visualization. Many researchers tried to develop emergency response system in order to reduce costs during a disaster (Alazawi et al. 2014), but still, there is a big gap between other sections of smart cities and smart DMS. Furthermore, smart visualization technologies can help utility managers to reduce costs as well. (Data, Life, and Better 2017)

Overflow in wastewater and stormwater networks is one of the old problems of big cities with old networks. Especially in cities like Chicago because of old combined sewer networks. (Jim Gudas 2018) Detection of these overflows and the reasons behind might be very helpful in the post and pre-disaster responses, improve the quality of citizens’ lives and help to achieve smart cities ideas.

Besides, it is very important for future smart cities to have a smart visualization. “*Smart visualization is the first step toward becoming a truly smart city*” (Data, Life, and Better 2017). Without proper visualization, it is almost impossible to achieve smart cities so smart visualization is the key to achieve smart cities and give a utility manager better understanding about their cities. This smart visualization can help the managers to make better decisions and more efficient management during disasters.

The researchers in this paper tried to develop a real-time overflow prediction model based on the forecasted rainfall hydrograph for the city of Raleigh and utilize smart visualization techniques by developing an MR-based twin model. They used different Artificial Neural Networks (ANN) models to achieve best real-time prediction model for overflow prediction with a reasonable error.

# 2. Literature Review

## 2.1 Pipe overflow problems

In the past, the communities built sewer systems to collect both stormwater and sewage using the same pipe network. During dry weather conditions, these systems transport the wastewater directly to the sewage treatment plant. In the periods of rainfalls, the wastewater volume increase and can increase the capacity of the sewer system and as a result, several overflows might occur across the network. (U.S. EPA 2017)

EPA estimations are that between 23,000 and 75,000 sanitary sewer system overflows happens annually in the United States. The results of these overflows are 3 to 10 billion gallons of untreated wastewater and these events take place throughout the United States. The figure below shows demographic maps of the U.S. and places with overflow problems and basement flooding complaints documented in Chicago during 2004 caused by this problem. (U.S. Environmental Protection Agency (EPA) 2004)

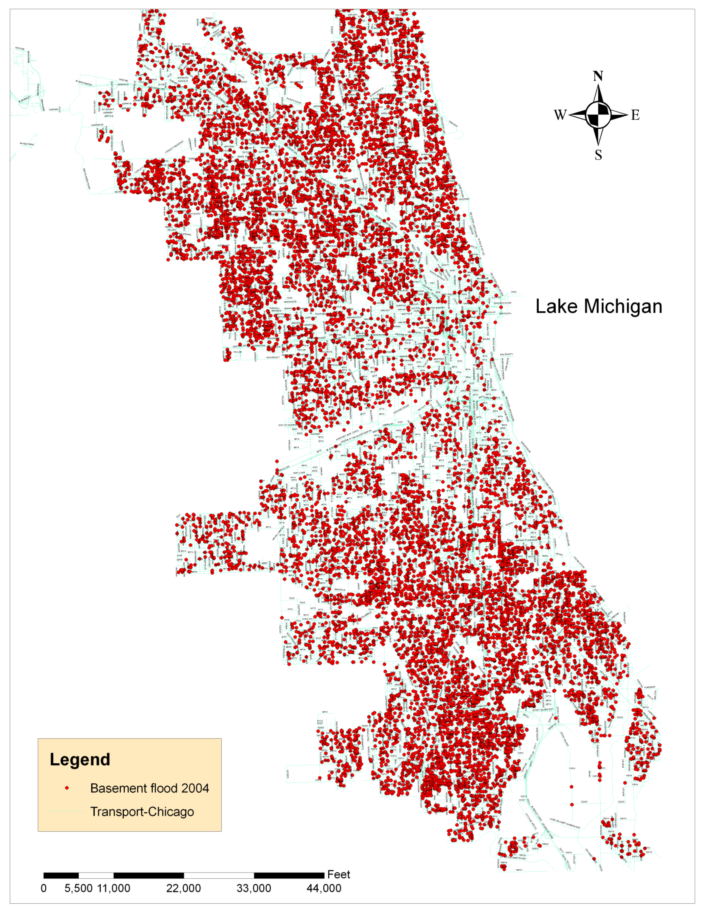
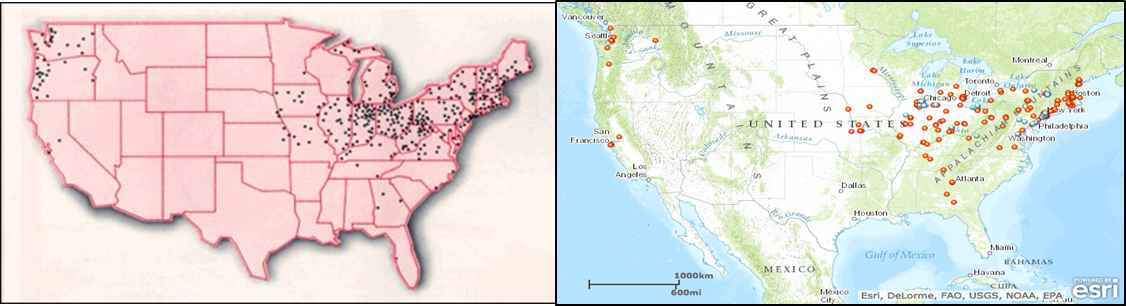


Figure 1: basement flooding complaints documented in Chicago during 2004 in right picture and overflow problems across the United States in the left picture (Demographics by U.S. EPA)

It is very important to predict, monitor, and control these overflow problems in a smart city so the smart cities can reduce disaster costs and achieve better management.

Many researchers tried to use different algorithms such as LASSO algorithm to make a prediction model for sewer overflows. They tried to achieve accuracy more than 90 percent and predict overflows 15 minutes before overflow occurs. Although they achieved a relatively good accuracy, the prediction cannot be done in a day before so there is a need for better prediction models to be able to predict overflows earlier. (Zhao, Beach, and Rezgui 2017)

## 2.2 Data visualization challenges

In addition to overflow problems as one inseparable problems of smart cities with old infrastructure, data visualization challenges and smart visualization plays a vital role in having a smart cities. In order to attain this intention, it is necessary to understand data visualization challenges. A digest of main challenges represented as follows.

1. **Limited field of view of object:** Most of the data which is generated in a smart city is multidimensional data, therefore it is hard to probe these data effectively using old visualization techniques. For instance it is hard to visualize “when was traffic data highest” or “How traffic data is related to weather” in a “vehicular traffic database” from a smart city on 2D flat screen. To understand the relationships of some data to other data, it is necessary to run analysis several times. Often times, it is hard to remember previous relations and it is difficult to get a complete picture of a situation simultaneously. (Jake Hissitt 2017)(Data, Life, and Better 2017)

2. **Lack of features enabling collaboration:** Current visualization and analysis platforms encourage collaboration between different departments such as water and waste management departments. It is very important for them to have access to the data simultaneously from remote locations. Government agencies prevent unwanted events and warn citizens beforehand any disaster to protect those using real-time data insights. Any inefficiency to handle data not only make the smart cities ineffective but also pose severe risks to citizens’ lives. (Jake Hissitt 2017)(Data, Life, and Better 2017)

3. **Lack of real-time data dissemination:** Assume that there is a leakage in a pipe at some point of a network. The technicians must rely on a 2D static map without any detailed information of the equipment installed beneath the ground. The technician must rely on voice-based information from his senior which is not flawless. Although, the smart city sensors and digital twin models are able to capture and visualize all the information about the pipeline. But, the current platform fails to disseminate critical information to remote workers. (Jake Hissitt 2017)(Data, Life, and Better 2017)

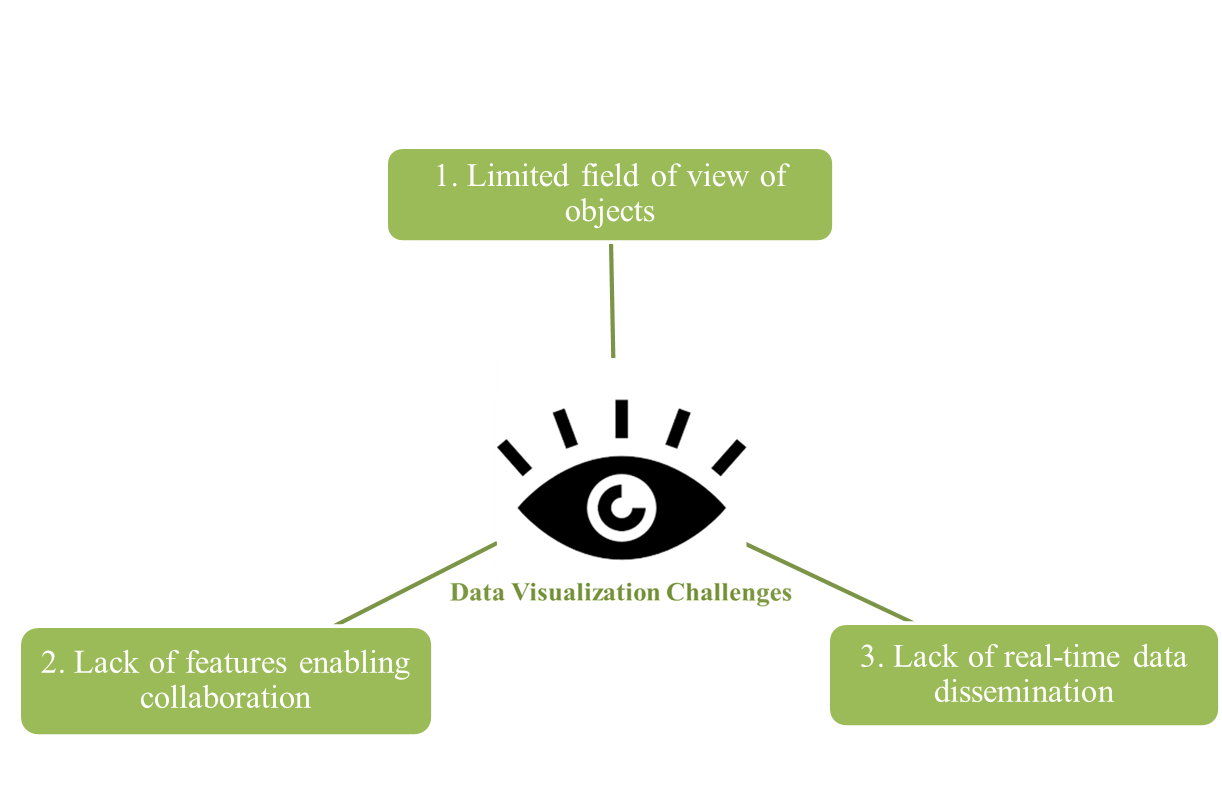


Figure 2: Main data visualization challenges

## 2.3 MR in civil and smart cities industry and other industries

One of the technologies that have the potentials of answering data visualization challenges is Mixed Reality (MR) technology. The main question here is what is MR?

“*MR is the merging of real and virtual worlds to produce new environments and visualizations where physical and digital objects co-exist and interact in real time. MR takes place not only in the physical world or the virtual world but is a mix of reality and virtual reality, encompassing both augmented reality and augmented virtuality via immersive technology*.”(de Souza e Silva, Adriana; Sutko 2009)

MR with a combination of other technologies like digital twin can help to address data visualization challenges and improve the visualization and collaboration.

The concept of digital twin introduced around 2002 and by improving technologies like the Internet of Things (IoT). This concept tries to develop a digital twin model on a system in order to understand how it works and how to have more cost and schedule savings. This concept has become more cost-effective to implement and can create more opportunities for cost and schedule savings and proved itself in many oil industries. (Bernard Marr 2017) Furthermore, this technology named as one of the top 10 strategic trends for 2017 by Gartner. (Kasey Panetta 2016) (DNV GL Inc. 2016)

The combination of digital twin technology and MR can open whole new opportunities for cost-savings in smart cities. One of the samples of these technologies is that construction workers use MR headsets and Building Information Modes (BIM) to improve inspection and reduce re-work in construction. The following picture shows the workers mounted MR headsets are inspecting construction site.



Figure 3: Workers using MR headsets for inspection in a construction site

MR in combination with digital twin can help visualization and decision making in several ways. These technologies can improve cost savings, help you in the decision-making process, allow you to have live experience in augmented and virtual environments, and improve the way that utility managers think about their cities. (Data, Life, and Better 2017)

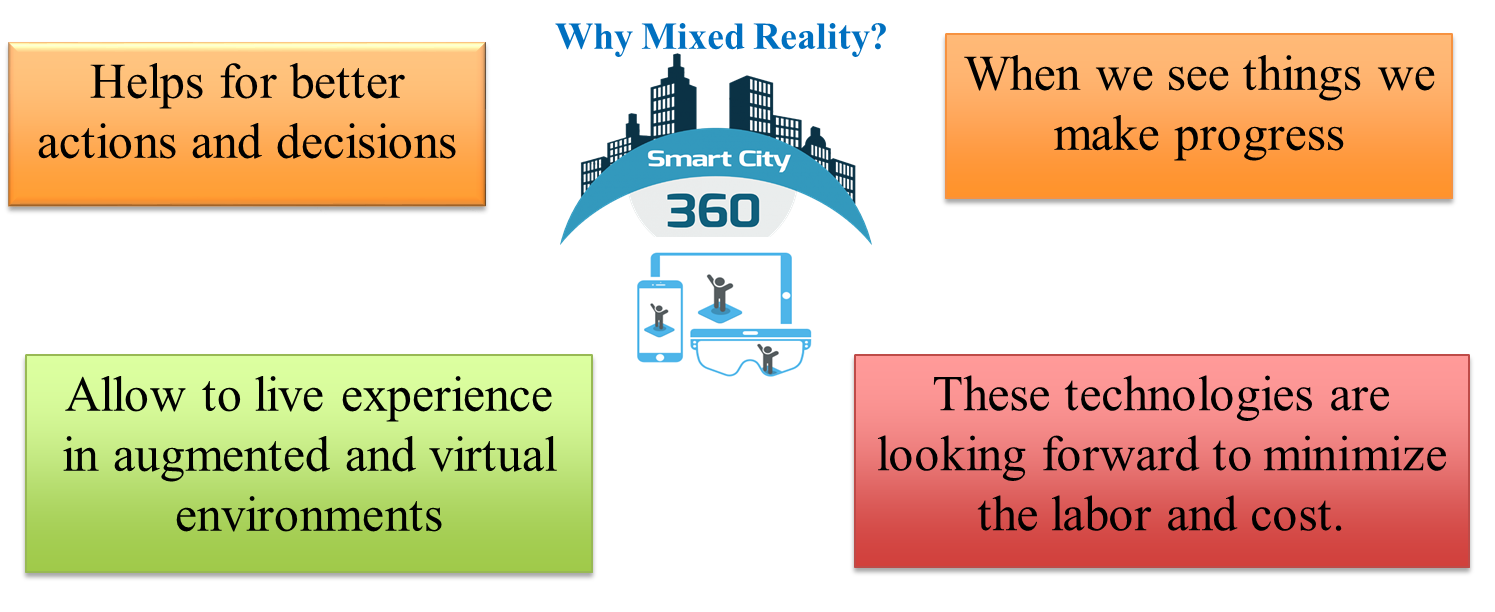


Figure 4: main reasons for using MR as a better approach for visualization

# 3. Research Objectives

The main objectives of this study are to predict pipe overflows in Stormwater networks and create an MR digital twin in order to improve visualization and communication between utility managers. The key steps were as follows:

1. Create a framework to predict overflows in order to help utility managers before disasters
2. Create a digital twin for the Stormwater network and enable utility managers for better decision making using Mixed Reality technology

# 4. Method

In order to achieve the objectives, the researchers in this paper used the following approach. At the beginning the researchers 1) gathered the overflow data from Storm Water Management Models (SWMM) for the city of Raleigh. After that, the researchers 2) created an artificial neural network model in order to predict overflow positions based on the SWMM data. Finally, the researchers 3) visualized the overflow problems in an MR environment in order to achieve better visualization and collaboration for utility managers.

1. Gather data from SWMM
2. Create models for predicting overflows
3. Visualize in a twin MR model

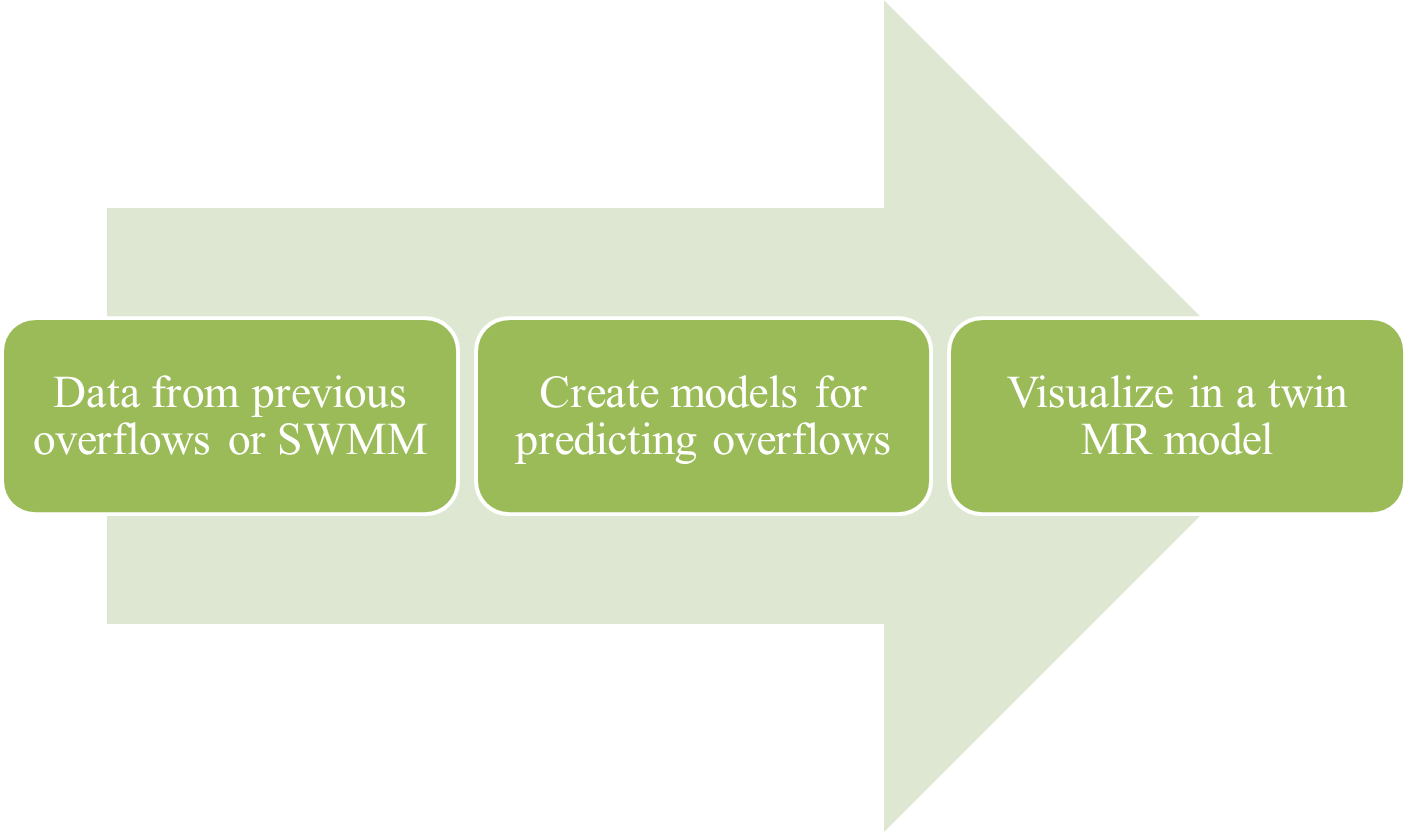


Figure 5: research method main steps

The researchers in this paper proposed the following system architecture for the real-time prediction model for overflow detection in MR environment.

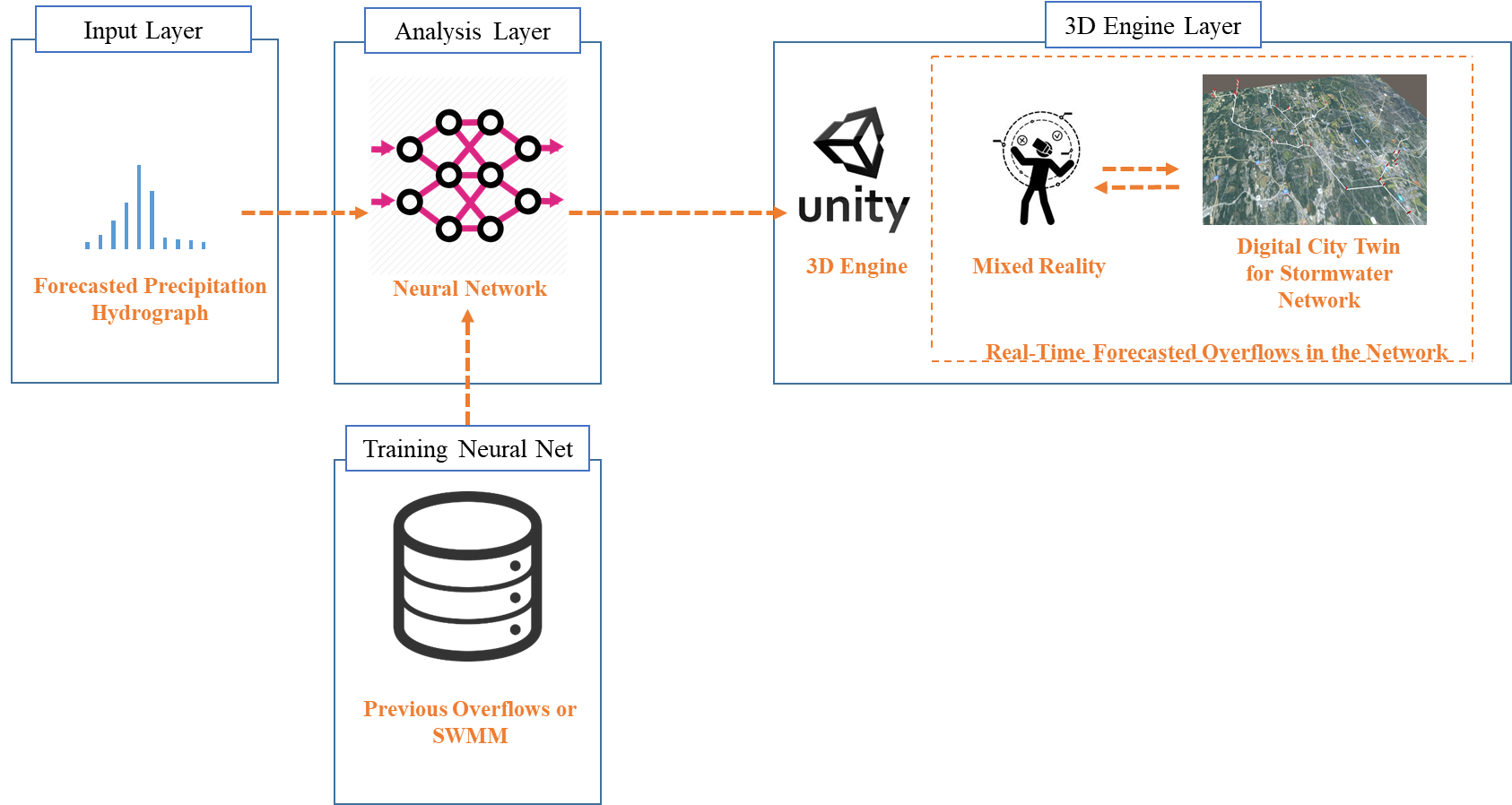


Figure 6: system architecture which represents the prediction model and MR

The system architecture consists of three different layers. These layers are as follows:

**Input layer:**

As the input of the prediction model, researchers in this paper used forecasted rainfall hydrograph. The main factor and reason for overflow in stormwater networks is rainfall with a big peak. The reason behind is when there is a high peak in rainfall hydrograph, the load on the stormwater network is high and the level of water increase significantly inside the pipes which lead to overflows and flooding of the network.

**Analysis layer:**

In this layer a database was used to train a neural network to predict overflows inside stormwater networks. The main benefits of using such models is that it will enable you with real-time features.

**3D engine layer:**

In the last step, the output data from the neural network will be sent to a 3D engine. The 3D engine is able to analyze that data and show them in the correct format. For instance the 3D engine is able to change the color of the pipes that have overflow problems to red. The 3D engine can also give better interaction to the users and improve the visualization for the users.

In the following section, the researchers will introduce the data which is used to train neural network and the specifications of the neural network which was used.

## 4.1 Data from previous overflows or SWMM

Input data for the prediction model extracted from city of Raleigh SWMM. The researcher in this paper gave 150 different rainfall hydrograph with 5 minute sampling rate to SWMM and find the pipes with overflow problems from SWMM analysis. The output data was the volume of the surcharge so the researchers annotated that data in the following way. The pipes which have overflow problems are 1 and the pipes with no overflow problems are 0. A map of Raleigh stormwater network and a 5 minute sampling hydrograph are demonstrated in the figure 8.

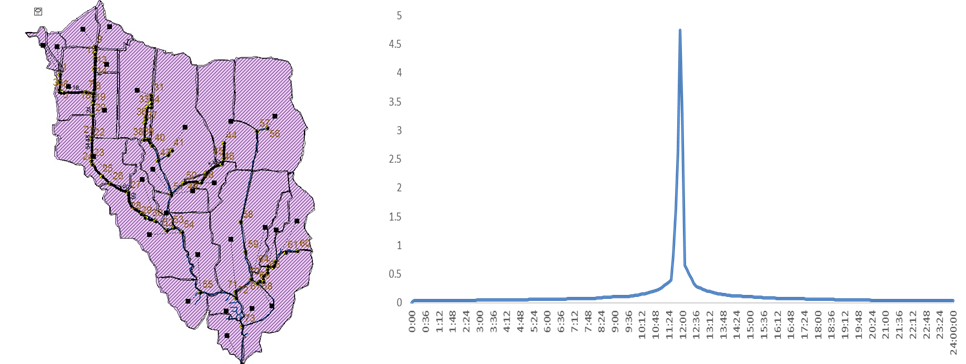


Figure 7: Raleigh stormwater network map on left and sample precipitation hydrograph on right

The 5 minute sampling rainfall hydrograph has a very high dimensionality. It is very important to decrease this dimensionality, otherwise it is very important to increase database significantly. It is very hard to increase database size so the researchers tried to decrease the dimensionality by feature extraction technique. The researchers picked 4 different variables to identify each hydrograph. This approach reduce dimensionality significantly and reduce required data point in the database. Variables are listed as follows:

1. Variable 1: first report in rainfall hydrograph which is more than one inch per minute. This variable shows the gradient of the beginning of rainfall indirectly. The following formula demonstrates the direct relationship between variable 1 and the gradient.

1. Variable 2: highest point in the rainfall hydrograph is the second important variable. As aforementioned earlier this variable has a direct relationship with the overflow problems in stormwater networks.
2. Variable 3: last report in rainfall hydrograph which is more than one inch per minute. This variable shows the gradient of the rainfall end indirectly. The following formula demonstrates the direct relationship between variable 3 and the gradient.

1. Variable 4: this variable shows the duration of rainfall between variable 1 and variable 3. This duration shows the rainfall and it can have an important impact on the overflow problems.

Four different variables are demonstrated in figure 9.



Figure 8: four variables which represent a hydrograph

In addition to decreasing dimensionality, it is very important to have a good variation in the database so neural network algorithm can predict different circumstances correctly. The database has four different inputs, it is not possible to draw these four inputs in a chart, but in order to show that the database has enough different data, the researchers picked 3 first variables and draw in on a chart. Figure 10 demonstrates the input data distribution over the space. This chart shows that there is a good distribution of input data in the database so this database can be used for a good prediction.

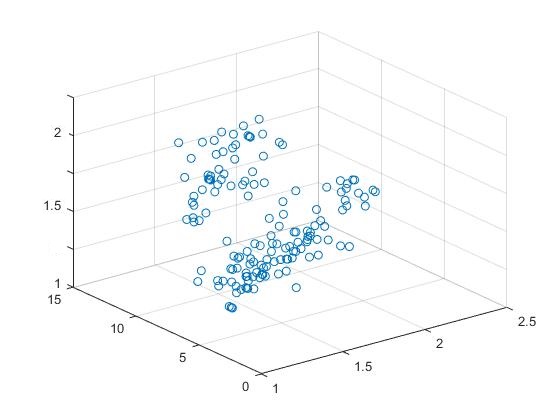


Figure 9: input data distribution

## 4.2 Create models for predicting overflows

The researcher in this paper used different neural network models and get the best result with the neural net fitting approach. The researchers used a neural network with the following specifications.

The neural network consists of 4 input variables as main parameters of a precipitation hydrograph, 3 hidden layers, and 73 outputs which represent 73 junctions in the network. The output values are either zero or one which means flooded or not flooded.



Figure 10: neural network overview

The researcher used 70 percent of total data for training, 15 percent of the data for validation, and 15 percent of the data for the test. Finally, the results of this model will be discussed in the results section.

## 4.3 Visualize in a twin MR model

In order to have a dynamic digital twin model for stormwater networks, the researchers in this paper developed a database approach to create the digital twin model for the stormwater network. The only requirement for using this model is to have all the network data in one database. The database should consist of all aspects of a stormwater network for example pipes length, pipes diameter, pipes section, the location of the manholes and junctions, pipes materials, and all of the specification about a stormwater network. The researcher in this paper developed a plugin for the 3D engine to import the database and generate twin model.

In the next step, the researchers are able to move the model into an MR environment to give and demonstrate the model to the utility managers.

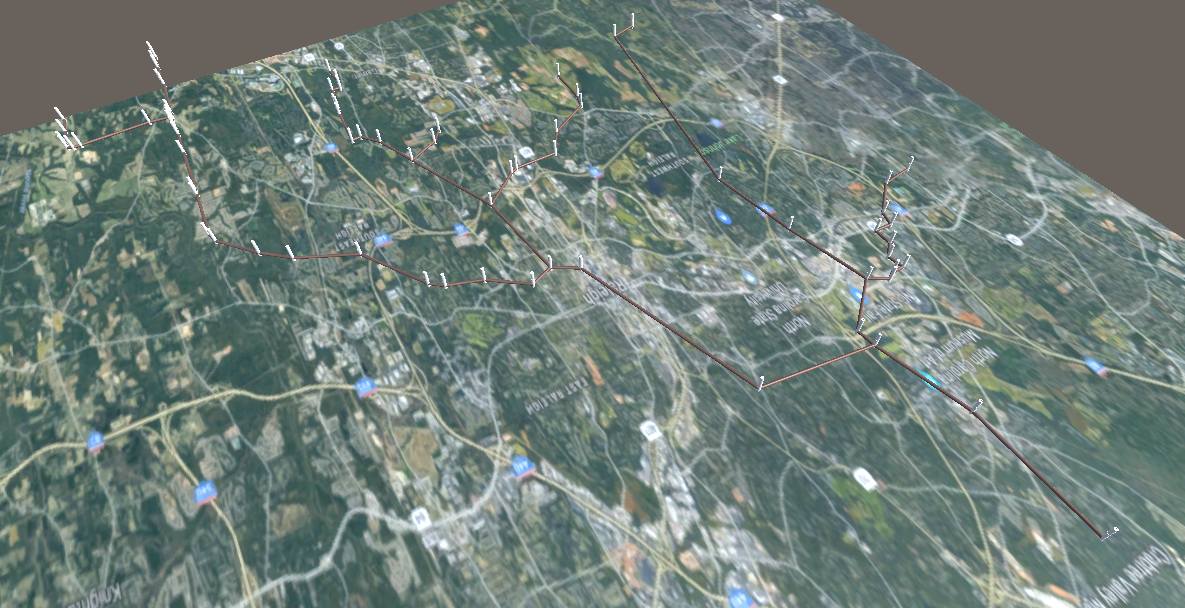


Figure 11: digital twin model of Raleigh stormwater network in MR environment

# 5. Results and Validations

The results from the training of the neural network were very good. The error histogram shows that there is a low error and high accuracy for the model in the test data.

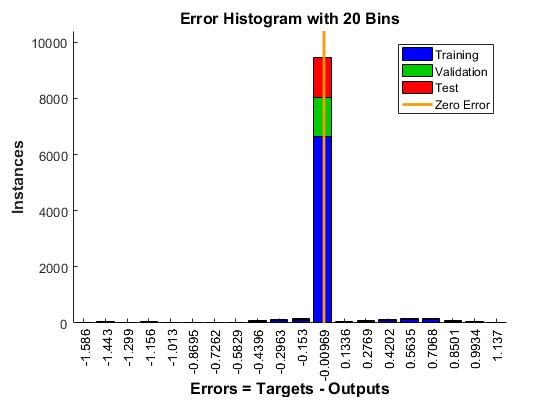


Figure 12: error histogram for the trained neural network

The results show that the neural network works very well in predicting overflows in stormwater networks. This platform can predict pipe overflows in real-time.

In order to check any problem with the neural network, the researchers give a separate hydrograph from the database to the model and test the model based on that. It had only one misclassification in one of the junctions which shows that the model is a well-trained model.

The model of which was used represented in the figure below. The red junctions are the junctions which are identified by prediction model as overflow junctions and white junctions are junctions without any problems.

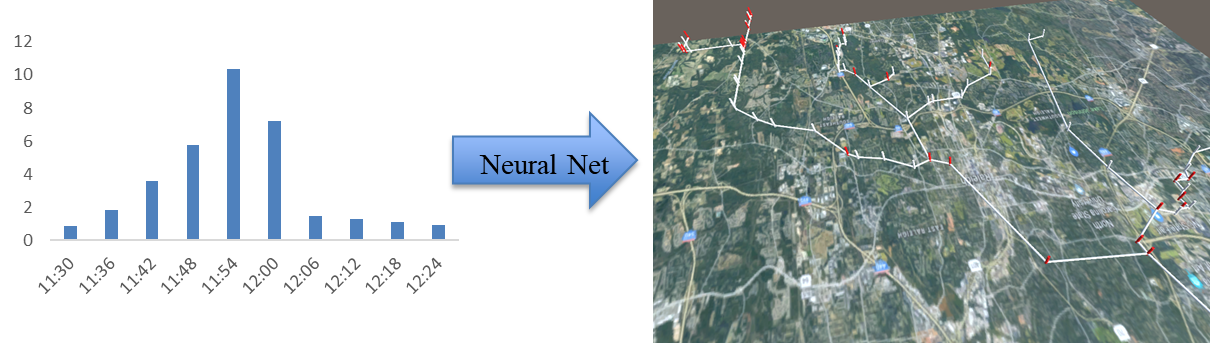


Figure 13: real-time MR environment which can predict overflows and show them by red color

As one of the results of this paper, the proposed visualization system is a totally new system. In the following figure, you can compare the old smart city dashboard vs. new digital twin model in MR environment. This technology can change the way utility managers and technicians about the cities and help them during decision-making process.

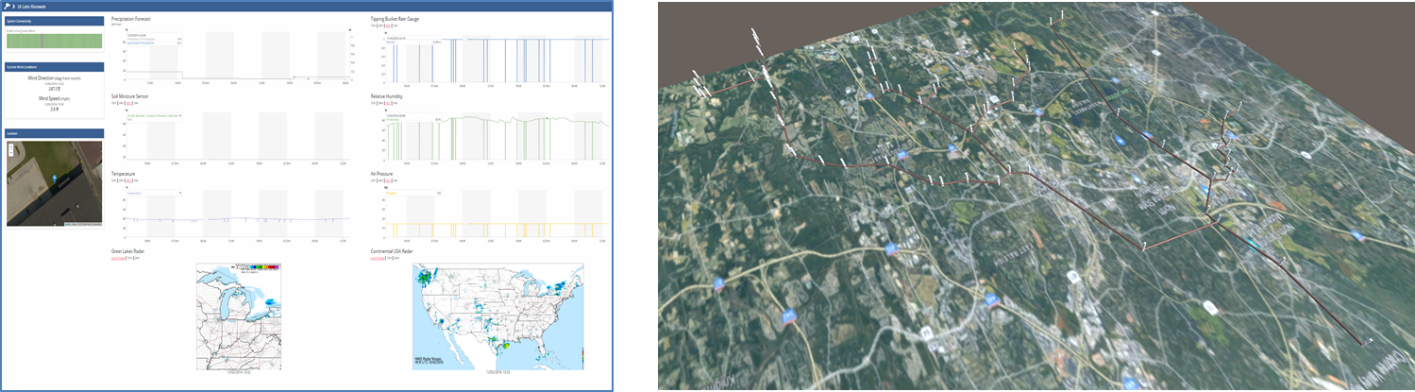


Figure 14: the transition from old smart city dashboard into the new MR digital twin model

The results of this study show that it is possible to use data which collected by different sensors to improve management and predict overflows.

# 6. Discussion

The main criticism of smart city which is discussed in this research is how using sensor data and models help the cities to predict outcomes of a disaster and help the citizens to reduce the damages. In addition to these predictions that can save citizens’ life, this research showed it is possible to have technologies like MR and digital twin in order to improve collaboration and many other aspects of smart cities.

# 7. Limitations and Future works

## 7.1 Limitations

One of the main limitations of the overflow prediction models is that it cannot detect issues caused by extraordinary events. For instance, if a tree falls into one the conduits in the network, the prediction model cannot predict the upstream overflow for that specific conduit and it may give you false results for the downstream of that pipe.

In addition to the first limitation, this platform cannot give you any recommendation during a disaster. For example, if you have 5 teams of technicians that can solve overflow problems and there are 10 different overflow locations, how you can prioritize the overflows. So on these occasions, the utility managers should make a decision based on their experience and the real-time platform cannot help them in the decision-making process in the post and pre-disaster responses.

## 7.2 Future works

It is possible to connect several MR devices in order to have a multiplayer experience for utility manager. For instance, if one of the managers changes something in the stormwater network, all of the utility managers and technicians can see that change simultaneously. This improvement can cause better communication and collaboration between utility managers and technicians.

In this paper,the researchers used data which is generated by SWMM, but it is possible to use previous overflow data in order to achieve more realistic results. The reason behind is using real data can give you better information because the real network might be slightly different from SWMM or even some part of the network might be changed over time.

In order to improve digital twin technology, it is possible to combine current stormwater twin models in smart city models. And even generating twin models for other infrastructures like water network, electric network, and gas networks. Having a digital twin model which contains all of the data related to the networks can give utility managers better visualization and help them in the decision-making process.

# 8. Conclusion

The problem which was represented in this paper was to predict overflows in stormwater networks and improve the visualization and collaboration using smart visualization techniques. The proposed platform in this paper is able to predict the overflows in stormwater networks based on predicted rainfall hydrograph and visualize the results in a real-time MR environment for better visualization and collaboration. This platform can improve efficiency in stormwater management systems and give the utility manager ability to have a rapid and ahead of time response in emergencies and disasters which can lead to effective disaster management.

The researchers in this paper used different neural network approaches to predict overflows and represented the best results with the highest accuracy. This model is able to predict the overflows with high accuracy in real-time.

# 9. References

Alazawi, Zubaida, Omar Alani, Mohmmad B. Abdljabar, Saleh Altowaijri, and Rashid Mehmood. 2014. “A Smart Disaster Management System for Future Cities.” *Acm 2014*, 1–10. https://doi.org/10.1145/2633661.2633670.

Bernard Marr. 2017. “What Is Digital Twin Technology - And Why Is It So Important?” Forbes. 2017. https://www.forbes.com/sites/bernardmarr/2017/03/06/what-is-digital-twin-technology-and-why-is-it-so-important/#275648922e2a.

Data, Bringing, T O Life, and F O R Better. 2017. “Immersive Data Visualization for Smart Cities,” no. November.

DNV GL Inc. 2016. “Making Your Asset Smarter with the Digital Twin.” 2016. https://www.dnvgl.com/article/making-your-asset-smarter-with-the-digital-twin-63328.

Ismail, Nick. 2017. “Smart Cities Could Lead to Cost Savings of $5 Trillion - Report Suggests.” 2017. http://www.information-age.com/smart-cities-lead-cost-savings-5-trillion-123469863/.

Jake Hissitt. 2017. “The 5 Biggest Challenges Facing Data Visualization.” 2017. https://channels.theinnovationenterprise.com/articles/the-5-biggest-challenges-facing-data-visualization.

Jim Gudas. 2018. “Heavy Rain Causing Some Flooded Basements And Streets In Chicago.” 2018. http://chicago.cbslocal.com/2018/02/20/heavy-rain-flooded-basements-streets/.

Kasey Panetta. 2016. “Gartner’s Top 10 Strategic Technology Trends for 2017.” 2016. https://www.gartner.com/smarterwithgartner/gartners-top-10-technology-trends-2017/.

Souza e Silva, Adriana; Sutko, Daniel M. de. 2009. *Digital Cityscapes: Merging Digital and Urban Playspaces*.

U.S. Environmental Protection Agency (EPA). 2004. “Report to Congress on Impacts and Control of Combined Sewer Overflows and Sanitary Sewer Overflows Fact Sheet.” *Water*, 1–2. https://doi.org/10.1227/01.NEU.0000028680.51299.00.

U.S. EPA. 2017. “Combined Sewer Overflow Frequent Questions.” 2017. https://www.epa.gov/npdes/combined-sewer-overflow-frequent-questions.

Zhao, W, T H Beach, and Y Rezgui. 2017. “Automated Model Construction for Combined Sewer Overflow Prediction Based on Efficient LASSO Algorithm.” *IEEE Transactions on Systems, Man, and Cybernetics: Systems* PP (99):1–16. https://doi.org/10.1109/TSMC.2017.2724440.