MID-AMERICA TRANSPORTATION CENTER

Report # MATC-MS&T: 130

Final Report
WBS: 25-1121-0005-130

Predictive Deep Learning for Flood Evacuation Planning and Routing

Steven Corns, PhD
Associate Professor
Engineering Management and Systems Engineering Department
Missouri Univeristy of Science and Technology

Suzanna Long, PhD
Professor and Department Chair

Bhanu Kanwar
Graduate Research Assistant

Vinayaka Gude
Graduate Research Assistant

Jacob Hale
Graduate Research Assistant

2020

A Cooperative Research Project sponsored by
U.S. Department of Transportation- Office of the Assistant Secretary for Research and Technology

MATC

Predictive Deep Learning for Flood Evacuation Planning and Routing

Dr. Steven Corns, PhD
Associate Professor
Engineering Management and Systems
Engineering Department
Missouri University of Science and
Technology

Dr. Suzanna Long, PhD
Professor and Department Chair
Engineering Management and Systems
Engineering Department
Missouri University of Science and
Technology

Vinayaka Gude
Graduate Research Assistant
Engineering Management and Systems
Engineering Department
Missouri University of Science and
Technology

Jacob Hale
Graduate Research Assistant
Engineering Management and Systems
Engineering Department
Missouri University of Science and
Technology

Bhanu Kanwar
Graduate Research Assistant
Engineering Management and Systems
Engineering Department
Missouri University of Science and
Technology

A Report on Research Sponsored by

Mid-America Transportation Center

University of Nebraska-Lincoln

January 2020

Technical Report Documentation Page

1. Report No. 25-1121-0005-130	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Predictive Deep Learning for Flood Evacuation Planning and Routing		**5. Report Date** January 2020
		6. Performing Organization Code
7. Author(s) Steven M. Corns, PhD ORCID: 0000-0002-3685-2892; Suzanna K. Long, PhD, PEM, F.ASEM, ORCID: 0000-0001-6589-5528; Jacob Hale; Bhanu Kanwar; Vinayaka Gude		**8. Performing Organization Report No.** 25-1121-0005-130
9. Performing Organization Name and Address Department of Engineering Management and Systems Engineering Missouri University of Science and Technology 600 W. 14th Street, Rolla, MO 65409		**10. Work Unit No.**
		11. Contract or Grant No. 69A3551747107
12. Sponsoring Agency Name and Address Missouri Department of Transportation (SPR) Construction and Materials Division P.O. Box 270 Jefferson City, MO 65102 Mid-America Transportation Center 2200 Vine St. PO Box 830851 Lincoln, NE 68583-0851		**13. Type of Report and Period Covered** Final Report (January 1, 2019-December 31, 2019)
		14. Sponsoring Agency Code MATC TRB RiP No. 34760
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration. MoDOT research reports are available in the Innovation Library at https://www.modot.org/research-publications.		
16. Abstract This research was completed in tandem as a project funded through MoDOT and the MidAmerica Transportation Center. It used deep learning methods, along with geospatial data from the USGS National Map and other public geospatial data sources, to develop forecasting tools capable of assessing water level rate of change in high risk flood areas. These tools build on existing models developed by the USGS, FEMA, and others and were used to determine evacuation routing and detours to mitigate the potential for loss of life during flash floods. The project scope included analysis of publicly available flood data along the Meramec River basin in Fenton as part of a pilot project in Missouri. These data were then used to determine the rate of rise in order to model evacuation or detour planning modules that can be implemented to assure the safety of the community and highway personnel, as well as the safe and secure transport of goods along public roadways. These modules were linked to existing real-time rainfall gauges and weather forecasts for improved accuracy and usability. The transportation safety or disaster planner can use these results to produce planning documents based on geospatial data and information to develop region-specific tools and methods.		
17. Key Words Deep learning; flood management; evacuation routing and planning	**18. Distribution Statement** No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.	

19. Security Classif. (of this report) Unclassified.	**20. Security Classif. (of this page)** Unclassified.	**21. No. of Pages** 52	**22. Price**

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Advanced Hydrologic Prediction Service (AHPS)
Area of Interest (AOI)
Autoregressive (AR)
Autoregressive Integrated Moving Average (ARIMA)
Autoregressive Moving Average (ARMA)
Corps of Engineers' Hydrologic Engineering Center (CEIWR-HEC)
Flood Inundation Mapper (FIM)
Geographic Information System (GIS)
Hydrologic Engineering Center Flood Damage Reduction Analysis (HEC-FDA)
Hydrologic Engineering Center Flood Impact Analysis (HEC-FIA)
Long Short-Term Memory (LSTM)
Mean Absolute Error (MAE)
Mean Squared Error (MSE)
Mid-America Transportation Center (MATC)
Missouri Department of Transportation (MoDOT)
Moving Average (MA)
Multi-Agent Transport Simulation Toolkit (MATSim)
National Oceanic and Atmospheric Administration (NOAA)
National Transportation Dataset (NTD)
National Weather Service (NWS)
Nebraska Transportation Center (NTC)
River Analysis System (HEC-RAS)
Seasonal Autoregressive Integrated Moving Average (SARIMA)
Simulation of Urban Mobility (SUMO)
Spatial Computable General Equilibrium (SCGE)
Supply Chain Infrastructure Restoration Calculator (SCIRC)
United States Army Corps of Engineers (USACE)
United States Geological Survey (USGS)

Acknowledgments

This project builds upon previous work done by the United States Geological Survey. The gauge height data and the Flood Inundation Mapping tools that were integrated into our framework was created by the United States Geological Survey and is available for public use. The traffic simulator used for this work (Simulation of Urban Mobility) is an open source package made available by the German Aerospace Center, Institute of Transportation Systems.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Abstract

This research was completed in tandem as a project funded through MoDOT and the Mid-America Transportation Center. It used deep learning methods, along with geospatial data from the USGS National Map and other public geospatial data sources, to develop forecasting tools capable of assessing water level rate of change in high risk flood areas. These tools build on existing models developed by the USGS, FEMA, and others and were used to determine evacuation routing and detours to mitigate the potential for loss of life during flash floods. The project scope included analysis of publicly available flood data along the Meramec River basin in Fenton as part of a pilot project in Missouri. These data were then used to determine the rate of rise in order to model evacuation or detour planning modules that can be implemented to assure the safety of the community and highway personnel, as well as the safe and secure transport of goods along public roadways. These modules were linked to existing real-time rainfall gauges and weather forecasts for improved accuracy and usability. The transportation safety or disaster planner can use these results to produce planning documents based on geospatial data and information to develop region-specific tools and methods.

Executive Summary

This research uses deep learning methods, geospatial data from the USGS National Map, and other public geospatial data sources to create a methodology capable of assessing water level rate of change in high risk flood areas. This methodology is used as part of a framework that can determine routing decisions based on safety constraints. These tools build on existing models developed by the USGS, FEMA, and others. The project scope includes analysis of publicly available flood data along the Meramec River basin in Fenton (intersection of I-44 and Route 141) as part of a pilot project in Missouri. The data was integrated to provide an indication of roads affected by flooding and suggest rerouting schemes (fig. E.1) along with the indirect costs that would be incurred. This framework introduces a methodology to integrate flood data with transportation data that has not existed prior to this research. This provides combined information on water rise, including when a road will be overtopped. Prior tools only updated flood inundation information at six hour intervals using data provided by NOAA.

While useful for response planning for a widespread flooding events, this is of limited use in traffic rerouting. As this inundation data was not designed with transportation systems in mind, there is no existing methodology available to match the flood data to transportation system data. Neither is there a methodology available specifically to determine how to best redirect traffic based on flood data while providing the added costs for this rerouting.

The algorithms developed in this research identify the patterns in river gauge behavior to predict gauge height at 15 minute intervals, allowing plans to be made based on events occurring every quarter of an hour rather than quarter of a day. In addition to proving updates at a time step that better maps with traffic planning, the root mean square prediction error from these deep learning algorithms is 0.453; less than half the RMSE of current techniques based on physics based models employed by the USGS, which is 1.065. This corresponds to an improved accuracy of more than seven inches of water level, provided on a time-scale increase of twenty-four.

The framework also provides precise information on how traffic can be rerouted the most effectively to avoid high risk areas and takes into account traffic that is already in route, making it possible to dynamically change traffic based on changes in flood predictions. This makes it possible for transportation safety or disaster planners to create tools specific to their region in the event of a flooding event and when an event occurs, it provides a method to predict when roads are no longer safe for motorists. In addition to the actual rerouting of traffic, this framework determines the time added by modifying the route and calculates the total indirect costs incurred, a capability that did not exist prior to this project.

Figure E.1 Data Integration in GIS Software

Chapter 1 Literature Review

There have been numerous studies investigating flood prediction, flood inundation, traffic simulation, and indirect cost calculations related to flood disasters. However, there is no methodology in the literature that unifies these approaches. To address this gap, a review of the research in each of these fields is presented and used to validate a unified framework.

1.1 Gauge Height Prediction

Flood prediction is one of the critical research areas due to its severe impact on the economy and loss of lives. Several models have been developed in the last decade that take advantage of the historical gauge height and discharge data published by the United States Geological Survey (USGS).

Several research groups have examined the prediction of the water discharge to predict floods in different locations (Damle & Yalcin, 2007; Elsafi, 2014; Chiari, Delhom, Santucci & Filippi, 2000 and Xiong, Shamseldin & O'connor, 2001). Some of these studies have used artificial neural networks (Elsafi, 2014; Chiari et al., 2000 and Xiong et al, 2001) and other traditional statistical approaches (Damle et al., 2007 and Campolo, Andreussi & Soldati, 1999). Three significant approaches for time series prediction are physical models, statistical models, and artificial intelligence models. Physical models include mathematical equations developed from physical information such as river profile, weather condition, basin area, etc. to forecast gauge height. This approach is data intensive and difficult to generalize as the model would be specific to the region selected. Statistical models such as autoregressive models (AR), moving average models (MA), autoregressive moving average models (ARMA), autoregressive integrated moving average models (ARIMA), and seasonal autoregressive integrated moving average models (SARIMA) can be used to forecast data. As the data size and complexity increases, artificial intelligence techniques such as Multi-Layer Perceptron, Recurrent Neural Networks, and Long Short-Term Memory (LSTM) Networks have performed better. Because of the complexity and size of the data, an artificial intelligence approach was chosen.

A few papers used Groundwater, precipitation, temperature, wind speed, and tides for calculating water discharge to predict floods (Tsakiri, Marsellos, & Kapetanakis, 2018). In a given region, precipitation, flow, and geographic characteristics are maintained by different organizations resulting in a non-uniform time frame for these parameters and also lack of availability in some places. One approach to avoid dependence on this variation of features is to choose an algorithm which can extract relevant features on its own based on the data, so deep learning time series prediction technique is the ideal choice.

1.2 Flood Inundation

Flooding is a global phenomenon that is responsible for numerous deaths, injuries, and extensive loss of property. The magnitude of the problem has warranted scholastic attention. Concerted research efforts have pervaded the literature since 1970 and the results have improved flood inundation capabilities (Teng et al., 2017). A majority of flood inundation studies fall into three categories: empirical models, hydrodynamic models, or simplified models.

Empirical models consist of synthesizing flood related data observations to create a limited representation of reality. Traditional approaches such as streamflow gauging are waning due to the resources required. Conversely, the use of remote sensing technologies are growing. Several satellites have been launched in recent years that possess improved capabilities related to sensing and processing (Teng et al., 2017). Complementing this improvement in data gathering are the advances of computational approaches such as algorithms and data mining. One such example is the use of a maximum entropy model for spatial extent of flood inundation over a study area in Sheikh et al. (2019). Empirical modelling was used to develop an equation for peak discharge using total precipitation, standard deviation of precipitation, and duration of storm in Aristeidis and Tsanis (2010).

Hydrodynamic models use computation to replicate fluid motion. The three dominant approaches are 1D, 2D, and 3D models pertaining to their dimensionality. One-dimension modelling provides the simplest representation by treating the flow along the centerline as one-dimensional (Brunner, 2016). Two-dimensional modelling represent mass and momentum conservation in a plane and operates under the assumption that the depth is shallow in comparison (Roberts et al., 2015). A flood inundation simulation is conducted and compared using both one-dimension and two-dimensional approaches in Bates and Roo (2000). Three-dimensional modelling addresses the vertical dimension of a flooding scenario and is particularly useful in capturing flood dynamics during catastrophic flood events related to dam breaks, tsunamis, or flash floods to name a few (Monaghan, 1994; Ye and McCorquodale, 1998). Integration of a 3D mesh with surrounding topography for flood inundation is presented in Merwade et al. (2008). Generally, 2-D models are the most commonly used due to data quality and availability for model building and validation (Alcrudo, 2004).

Simplified methods are based on simplified hydraulic concepts and do not involve the simulation of physical processes of inundation. One example of this approach is the planar method, often referred to as the “bathtub method”. This approach involves intersecting a series of planes with a digital elevation model to link the water stage/volume of the flooded area. Readers are directed to the review completed by Teng et al. for further elaboration on the approaches presented.

United States Army Corps of Engineers' Hydrologic Engineering Center's (CEIWR-HEC) River Analysis System (HEC-RAS) software tool has also been explored by researchers to conduct hydraulic analysis and flow simulations for a selected test area (HEC-RAS). This software tool can also be used to execute various hydraulic study tasks like one-dimensional steady flow, one and two-dimensional unsteady flow calculations, sediment transport/mobile bed computations, and water temperature/water quality modeling for different water channels.

1.3 Traffic Simulation

Traffic analysis is required for evaluating the effect of road closures and corresponding indirect cost estimation. Linear Optimization is one of the simplest techniques used for traffic analysis (Gartner, Little, & Gabbay, 1975). The problem space for this scenario is a graph. The objective for the problem is minimizing the travel time for all the passengers and the constraints are the traffic limits on the edges (roads).

Simulation is another approach that is increasingly being used for traffic analysis. Simulation involves generating traffic that includes types of vehicles, start times, corresponding origins and destinations. Based on which, the shortest path is assigned for the vehicles using Dijkstra algorithm (Dijkstra, 1959). Dynamic Traffic Assignment is the most commonly used methodology for traffic simulations (Astarita, Er-Rafia, Florian, Mahut, & Velan, 2001). It is an agent based model where all the cars in the networks are agents. The objective is to evaluate an equilibrium solution where no traveler can reduce the travel time by switching paths. Most transportation related articles focused on traffic optimization, urban planning, and individual vehicle behavior is not given much importance.

The granularity of the simulation can be selected to suit the corresponding problem. A macroscopic simulation focuses on traffic flow that can be useful for high level traffic analysis.

A microscopic simulation is more detailed and can capture traffic signal changes and vehicle positions every second. This simulation is more computationally intensive than the macroscopic model due to the calculations required for evaluating individual vehicle dynamics every second. For evaluating congestions and indirect costs associated with rerouting, assessing the delays for individual cars can be crucial.

Multi-Agent Transport Simulation Toolkit (MATSim) and Simulation of Urban Mobility (SUMO) are some of the most popular open source traffic simulation applications. Looking at the overall integration capability of the project, SUMO can be ideal for the project as it is coded in Python and can interface with other Python libraries and files.

1.4 Economic Impact

Research efforts have been made to find the economic losses associated with natural hazards such as floods. Various researchers have utilized their resources to study different costs involved when a minor or major flood impacts a certain region. High-resolution remote sensing data was used by Gerl, Bochow, and Kreibich (2014) to develop detailed spatial information on different types of residential buildings. The information obtained was then used to model flood losses and improve existing risk analysis techniques. The use of remote sensing datasets by the team of researchers also compensated for the lack of availability of detailed map datasets for certain sections of the test area, thereby aiding in the development of comprehensive models. Jonkman et al. (2008) also relied on a hydrologic dataset to develop an integrated hydronomic-economic model to thoroughly study the damages and economic losses caused by floods in the Netherlands. The insights from the proposed damage functions were also used to improve the existing risk assessment and safety protocols in the respective study areas.

Floods adversely affect the road network in the flood-stricken areas which paralyzes the flow of goods and services along the network resulting in huge monetary losses for the entities in charge of freight transportation. Hundreds of thousands of vehicles ply on our roads every day and delay of even a few hours might result in extra costs borne by the drivers. Therefore, it is important to dedicate resources to understanding the economic costs of various disruptions to the traffic. Janic (2007) proposed a model to find the direct and indirect economic costs generated due to disruptions in the flow of freight on an intermodal road transport network. European Union data was used to find the total cost of the selected intermodal road freight transport network. Tatano and Tsuchiya (2007) have developed a spatial computable general equilibrium (SCGE) model to gauge the economic losses related to disruption of a transportation network. Forecasting models were developed by Chu (2016) to analyze the relationship between extreme weather conditions and freight movement along road, rail, air, and water corridors. The impact of several weather factors on road traffic flow was also explored by Keay and Simmonds (2005). As per their investigation, traffic volume decreases by 1.35% and 2.11% during rainy days in winter and spring respectively which results in expensive delays for commuters. Suarez et al. (2005) reproduced the flooding scenarios in the urban infrastructure of Boston Metropolitan area to probe the impact of resulting delays and lost trips. A computable general equilibrium (CGE) model was also proposed by Tirasirichai and Enke (2007) to inspect the indirect economic losses suffered due to the closure of damaged bridges post-natural disaster.

Floods are known to trigger huge monetary losses that have a negative impact on the economy of the affected regions. Lada (2018) reports that during the Great Flood of 1993, high water levels in both the Mississippi and Missouri rivers forced the closure of respective waterways that resulted in a loss of $2 million per day for the shipping industry. Since many regions remained flooded for several weeks, the indirect economic losses also skyrocketed, adversely affecting the lives of the people suffering from this catastrophic event. Also according to Lada (2018), dangerously high-water levels damaged several bridges and as a result, people had to travel nearly 100 extra miles to reach the other side of the swollen Mississippi River. Since natural disasters like floods generally damage and destroy different infrastructure elements and subsequently lead to huge economic losses, it is important to develop a methodology which can help us in determining associated direct and indirect economic losses. An ability to generate accurate cost estimates will be beneficial to the authorities responsible for performing respective disaster management tasks.

Flood-damaged road sections not only affect the daily lives of commuters, it also plays havoc with bottom line of various freight operators who use those roads to transport precious cargo across the country. A damaged road section can also slow down various rescue and relief operations undertaken by the concerned personnel in the wake of a flooding event that might prove detrimental to the safety of the people affected by the flood. Thus, it becomes very important to have a cost calculating methodology at one’s disposal that can also be used by the authorities to develop efficient routing protocols. This report presents a methodology that can be used to calculate indirect costs incurred by various drivers when they are forced to take detours to reach their destination points.

1.5 Summary

The gauge height predictions developed by the Advanced Hydrologic Prediction Service (AHPS), managed by the National Weather Service (NWS), are provided on the USGS website (https://waterdata.usgs.gov). The results are based on a physical model which is developed from digital elevation maps, weather, and other geophysical properties of the given region. The problem with these predictions are that they are 6 hours apart and do not provide sufficiently accurate results to conduct flood inundation mapping and subsequent traffic simulation. Further, physical models cannot be generalized and have to be developed from scratch for each new region. Therefore, there is an opportunity to develop a model with improved prediction time frame, accuracy, and generalizability. A Long Short-Term Memory Network is an ideal solution for the problem.

Calculating indirect cost is not a straightforward process. Since most of the privately-held data is not easily available to the general public, the calculation process can become a cumbersome activity. However, the implications of ignoring the post-disaster indirect costs can be a lot more damaging than expected. Research (National Research Council, 1999) has shown that the percentage of indirect losses increases in large natural hazards and may constitute a large portion of the economic losses suffered by a region. Even though there are several models and tools available to calculate direct costs, there is a dearth of options when it comes to determining the indirect costs. Therefore, it is crucial to develop different tools and methodologies which can prove useful in determining the indispensable indirect costs.

A breadth of studies have been conducted in each of the reviewed fields, but no unified framework has been presented. Therefore, the following methodology is presented to address this gap. The next section presents the model framework and gives a detailed description for each of the sub model processes.

Chapter 2 Methodology

The primary purpose of this study is to provide emergency planners with the ability to preemptively reroute traffic in the event that a flood event will compromise certain road segments. The following methodology mirrors the sequential relationship between the sub models as presented in the literature review section and is illustrated by figure 2.1. Step one uses time series river stage data and predicts future stage values using a computational intelligence model (an artificial neural network). Step two uses the stage prediction value from step one as a model input for flood inundation profiles generated by USGS' Flood Inundation Mapper and processes them to acquire a list of affected road segments. Step 3 uses the set of affected road segments and conducts a traffic simulation using SUMO. Step 4 conducts an indirect cost analysis of the traffic simulation between normal conditions and the altered conditions after the road network has been affected by the flooding event. The remainder of this section discusses each of the steps in greater detail.

Figure 2.1 Model Framework

2.1 Gauge Height Prediction

Gauge Height prediction is a Time Series prediction problem as the future values can be estimated from the historical data consisting of temporal information. The objective of time series prediction is to forecast the value of a given parameter at a future time *(t+h)* using available observation at time *t.* Some of the characteristics of time series models are time interval, recursion, and number of variables. The time interval is the difference in time between any two given consecutive observations. If the difference is uniform across all the data, then it is a uniform time series and when the difference is not the same throughout, then it is a non-uniform time series. A recursive prediction strategy is used when more than one value is being forecasted into the future. Recursion means the process of prediction is repeated until the desired number of forecasts is reached. It is important to note that the further into the future predictions are made, the less accurate the results are. Another characteristic for a time series prediction model is the number of variables. If only one variable is being forecasted by the algorithm, then it's a univariate model and if more than one variables are being predicted, it is a multivariate model.

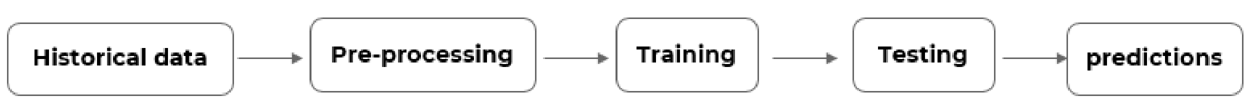

Figure 2.2 Time Series Prediction Methodology

There are 5 major steps for time series prediction (fig. 2.2): data collection, pre-processing, training, testing, and making predictions. Data Collection involves collecting data from the sources (data repositories, websites or API's). Once all the relevant data are acquired, the data are transformed and modified to work with the algorithm being used. This process is called pre-processing and includes changing the format of the file (xlsx, csv, xml, etc.), removing unwanted values of the data, removing errors and handling missing data. In most cases, the missing data are obtained by interpolation. As mentioned above, the idea behind time series prediction is to use past temporal data to forecast future values, and this is achieved in the training process. Often 80% of the entire data is used for training and 20% for testing. The algorithm used for forecasting takes in the historical data and identifies features that can aid in predicting future values. Validation is a key concept in training computation intelligence algorithms that indicates how well the algorithm is being trained. A simple way to perform validation, is by using 80% of the training data for training (64% of the entire data) and rest of 20% (16% of the entire data) of the training data to validate. Once the algorithm is trained effectively, it needs to be tested with 20% of the data which was initially separated. Mean Absolute Error (MAE) and Mean Squared Error (MSE) are the metrics most commonly used to evaluate the performance of the algorithm. MAE is the average error of all the predictions and MSE is the square root of the average of squared differences between prediction and actual observation.

Regression based models are popular for time series prediction as they are well understood. These models tend to have lower performance as non-linearity and data size increases. Deep learning models are capable of capturing complexity automatically and have become increasingly popular in complex time series forecasting problems. This reduces the need to perform feature engineering on the data as the algorithms can learn to extract features by themselves during training.

2.2 Flood Inundation

The gauge height predictions are used as an input to consider a given flood inundation profile to determine the extent of the flood event. These flood inundation profiles are created by USGS' Flood Inundation Mapper (FIM) and correspond to several monitoring sites around Missouri (denoted by triangles in figure 2.3). The FIM provides a web-based tool where the user can select a monitoring site to explore the flood inundation profiles (fig. 2.4). The slider bar allows the user to view flood inundation profiles for river stage values in half-foot or one-foot increments depending on the underlying data. USGS publishes a complementary report for each of the study areas that provides more information on model building, calibration, and validation (USGS, 2019). This also includes a description of how to create a raster data layer for each of the flood inundation profiles. These layers can be uploaded to geographic information system (GIS) software for integration with other data layers.

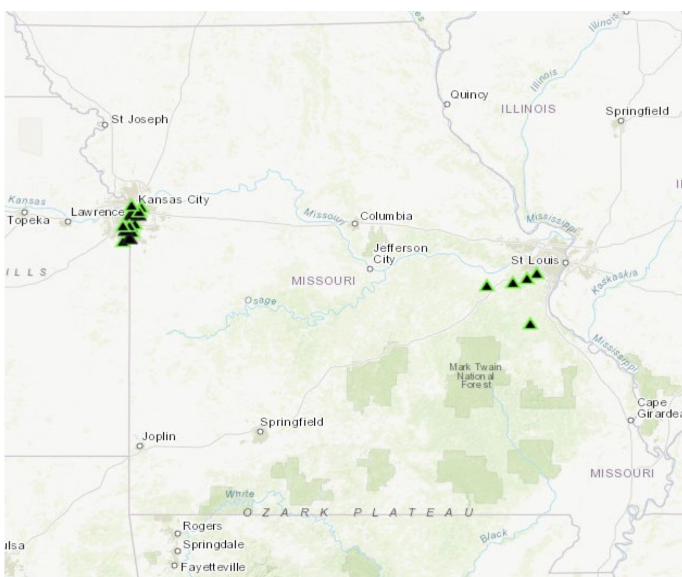

Figure 2.3 Missouri Flood Inundation Mapper Sites

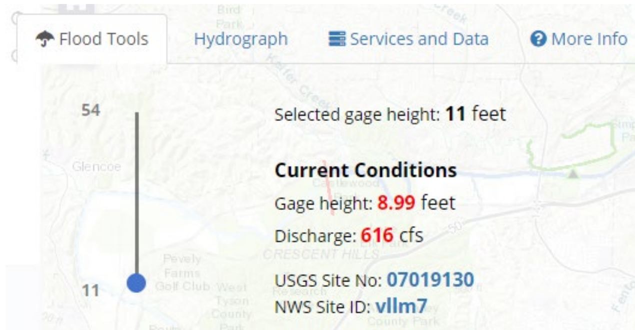

Figure 2.4 Interaction Flood Inundation Tool

To determine the road network affected by the flood event the raster layer needs to be combined with road network data. This information is acquired from USGS' National Transportation Dataset (NTD) and integrated with the flood inundation profiles created by the FIM. NTD consists of all road networks within the state. Because of the amount of data, it is computationally advantageous to select an area of interest (AOI) within the transportation network to minimize the time needed to conduct geoprocessing. This can be accomplished by using the software's selection capability and creating a new layer out of the selected features. Integration of the two layers is then done using a GIS software package (ArcMap, fig. 2.5) for the case study location addressed in the Results section.

Figure 2.5 Example of Data Integration in GIS Software

Visual evaluation and manual manipulation allows a user to determine which road segments are affected by a given flood stage. After verifying this provides accurate results, the process was then streamlined within the GIS software. Data from the FIM is stored in a raster format whereas transportation data is stored in vector format. Conversion of one of the data layers is required before further study can be conducted. In this study, the raster layer is converted into a point layer using the conversion toolbox within ArcMap (fig. 2.6). Geoprocessing is then done between the discrete point layer and the transportation network using the intersection function. Resolution is user-specified and in the example (fig. 2.7) it is set to one meter. The output of these steps can then be used as an input for the traffic simulation sub model.

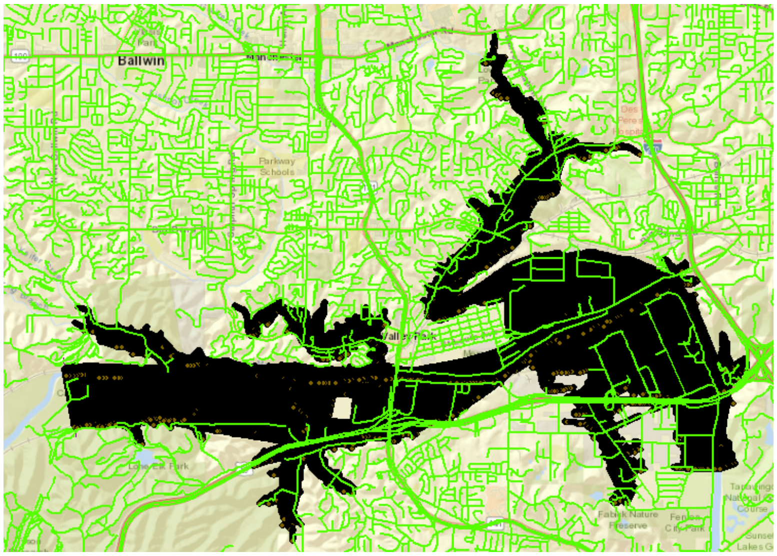

Figure 2.6 Example of Point Layer and Road Network Integration

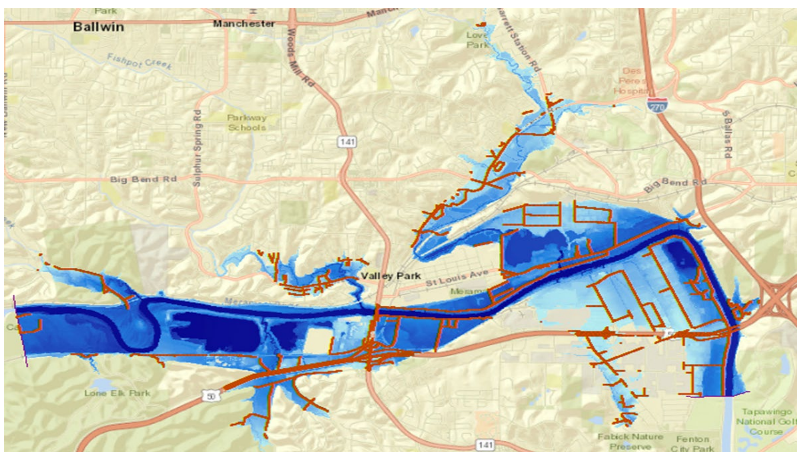

Figure 2.7 Example of Affected Road Network Given Flood Inundation Profile

<u>2.3 Traffic Simulation</u>

The GIS output showing which portion of the road network is closed is then used as input for SUMO, a traffic simulation package. SUMO is an open source python simulation package for performing different steps of the traffic simulation process, including generating traffic, identifying shortest paths, and assigning traffic. Sumo has both a graphical user interface (fig.

2.8) and the ability to work with command line options. The road network can be generated in SUMO from shapefiles, open street map, and by creating a manual network within the software using a Graphical User Interface called 'NETEDIT' (fig. 2.9). After the analysis is complete, the generated route can be saved in xml format.

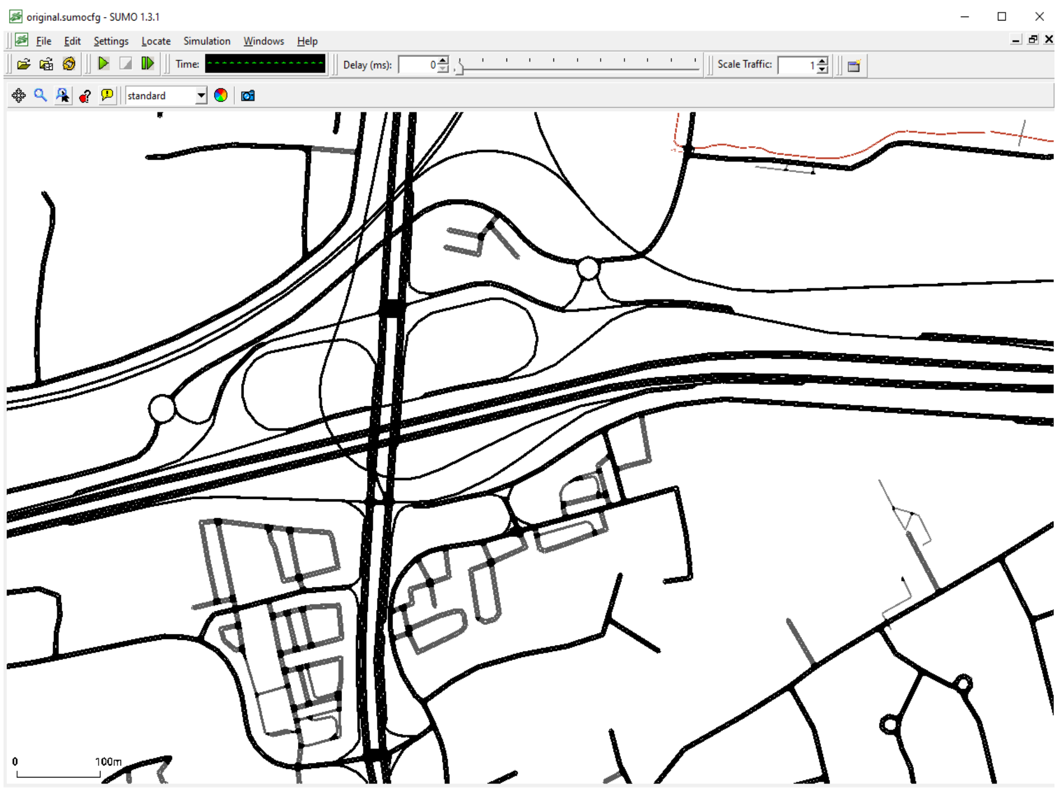

Figure 2.8 SUMO Interface

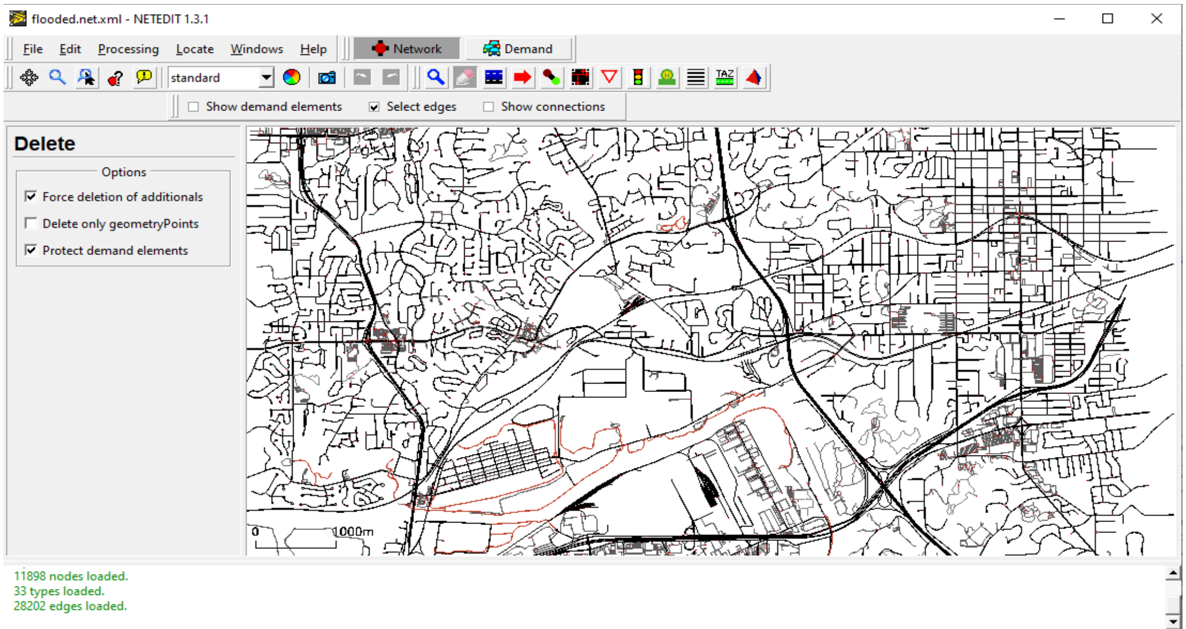

Figure 2.9 NETEDIT Interface

The simulation process begins with importing the road network into the NETEDIT application so it can be edited. With the roadway model the established simulation generates models of traffic. The traffic generated is matched to the average annual daily traffic data published by the Missouri Department of Transportation when possible to improve simulation accuracy. Traffic is generated randomly over a period of 3600 seconds. A single vehicle enters the simulation every second giving a total of 3600 vehicles per hour to represent the traffic. The origins and destinations for these vehicles are assigned at random with the generated trips recorded in xml format (fig. 2.10). These trips represent a single vehicle with the origination, destination, and departure time. A unique id is assigned to each vehicle for ease of tracking.

```
<trip to="480504288#0" from="646943268#7" depart="10.00" id="10"/>
<trip to="500186854#3" from="18716619#1" depart="11.00" id="11"/>
<trip to="-18711539#1" from="18714574#0" depart="12.00" id="12"/>
<trip to="43007267" from="465543192#0" depart="13.00" id="13"/>
<trip to="18708013#2" from="-18710603#2" depart="14.00" id="14"/>
<trip to="18696409#4" from="18696031#7" depart="15.00" id="15"/>
<trip to="-18708502#0" from="-351516568#0" depart="16.00" id="16"/>
<trip to="96165202#5" from="-18699574#3" depart="17.00" id="17"/>
<trip to="24266860#3" from="-25543414" depart="18.00" id="18"/>
<trip to="-18715951#1" from="18708481#3" depart="19.00" id="19"/>
<trip to="43388319#2" from="48556103#9" depart="20.00" id="20"/>
<trip to="-524713230#8" from="235485898#0" depart="21.00" id="21"/>
```


Figure 2.10 Traffic xml file

The road network and traffic data are imported into the SUMO simulation as net-file and route –files respectively. The simulation can generate different outputs depending on the user's preference. Some of the preprogrammed options are *'summary'* (provides information regarding the number of cars on the networks at a given time step, mean speed, mean waiting time, travel time, etc.), *'sumotrace'* (provides information regarding the location and speed of different vehicles every second) and *'tripinfo'* (provides information regarding the overall travel time, waiting time, origin and destinations for each vehicle). In this research, Sumotrace provided the information on the new route and Tripinfo provided the overall delay for each vehicle required for calculating indirect costs.

2.4 Economic Impact

The economic losses incurred from natural disasters like floods can be divided into two types of costs: Direct Costs and Indirect Costs. The economic losses from damages attributed directly to floods can be categorized under direct costs e.g. costs incurred from damage to infrastructure elements and agricultural land, loss of lives, etc. However, the indirect costs originate from disruptions to local businesses, transportation networks, and daily lives of people in the affected regions (Tirasirichai et al., 2007 and Enke et al., 2008). Unlike direct costs, the calculation of indirect costs can often be a time-consuming process. Nonetheless, it is important to get proper estimates for these indirect costs so that efficient relief and restoration plans can be developed.

Indirect costs can have a variety of sources indirectly associated with a disaster. For example, if a section of road is damaged due to floodwater, commuters will have to use an alternate route to reach their destination, resulting in extra fuel costs and/or lost wages. For freight operators, any required detour will increase both travel expenses and the shipping times of the transported goods. Thus, indirect costs have the potential to put a strain on both the supply chain transportation network and traveler's economic budget.

Indirect costs can be affected by various factors. In this study, average hourly wage and delay factors are used to calculate these costs incurred from taking detours when certain roads are damaged from floods. The relationship between these variables is captured empirically through the development of a mathematical model.

While calculating the detour costs, a delay factor is used to represent the costs incurred on a particular type of trip undertaken by the driver. Along with variable costs, there are three delay factors impacting indirect costs: work trips, social trips, and other trips. 'Work Trips' involve trips which are undertaken for work-related tasks such as commuting to the workplace, transportation of goods, etc. Trips undertaken for recreational purposes are categorized under the 'Social Trips' e.g. trip to the shopping mall, vacation trips, etc. All the remaining trips not falling under these two categories are designated 'Other Trips'. Table 2.1 shows the traffic delay factor values that affect the indirect costs based on the type or purpose of the trip taken by a given driver.

Table 2.1 Delay Factors

Delay Factors	
Low Time Savings (0-5 minutes)	
Work Trips	0.064
Social Trips	0.013
Other Trips	0.001
Medium Time Savings (6-15 minutes)	
Work Trips	0.064
Social Trips	0.013
Other Trips	0.001
High Time Savings (>15 minutes)	
Work Trips	0.064
Social Trips	0.013
Other Trips	0.001

The values of the different delay factors for different time savings were calculated by researchers with the United States Army Corps of Engineers (USACE) (Lampo et al., 1993). These delay factors represent the value of time saved adjusted to an hourly basis and are based on the percentage of hourly income of the driver. Since this value of time is an opportunity cost, it can also be used to calculate the money lost while undertaking a certain type of trip along an alternate route due to a traffic rerouting. During traffic rerouting, the driver must take an alternate route in order to resume the trip. This new alternate route typically increases both the expenses and time of the trip. As the driver does not gain the benefits provided by the preferred normal route, this opportunity cost represents a monetary loss incurred due to using an alternate route to reach the destination. These delay factor values are used to provide information regarding the losses in travel time and to calculate indirect costs while undertaking a specific trip on an alternate route. Based on the respective delay factors and average hourly wage, the indirect costs incurred as part of the trip can be calculated using equation (2.1).

$$Indirect\ Cost = \sum_{i=0}^{n} x_i * Average\ Hourly\ Wage. \qquad (2.1)$$

Where 'n' is the number of delayed cars obtained after running SUMO software package and x_i are the respective delay factor values from table 2.1. An average hourly wage of $20 is considered for these calculation purposes.

These indirect costs represent the total losses for each type of trip. They can also be used by the drivers to perform the cost-benefit analysis and plan accordingly in case of road congestion due to floods. If the information related to closed roads is available before the trip is undertaken, the driver can also use this information to choose the most economical and safe alternate route.

Chapter 3 Results and Discussion

To demonstrate the effectiveness of the methodology presented, a case study is presented for a major transportation sector near St. Louis, Missouri that has recently experienced multiple flooding events. The specific area of study is Valley Park, Missouri, situated at the intersection of Interstate 44 and State Route 141.

3.1 Study Area

The area around the intersection between Route 141 and Interstate I-44 at Valley Park, St. Louis County, MO (fig. 3.1) was selected for the study area as the roads in and around this intersection experience heavy morning and evening traffic flow and have been impacted by flood events in the last few years. Interstate I-44, which runs between the city of St. Louis and Oklahoma City, has an average annual daily traffic volume of over 20,000 vehicles per day (Traffic Volume Maps: Missouri Department of Transportation) and is also widely used by freight truck operators to transport goods along this major highway.

Figure 3.1 Intersection of I-44 and 141 during normal conditions in 2017 (Photos: Before & After Meramec River Flooding)

However, the nearby Meramec River has posed serious risks to the surrounding area during both minor and major flood events. During the floods of 2017, acres of private and public property along the I-44 and 141 intersection were damaged when the floodwaters overflowed the banks of the Meramec River (fig. 3.2).

Figure 3.2 Intersection of I-44 and 141 during Meramec River floods in 2017 (Photos: Before & After Meramec River Flooding)

As a result of sudden and frequent flooding in this area, St. Louis County has spent more than $200 million in repairing damaged roads over the last few years (Meramec Flooding Proposal). This busy road intersection has also suffered from various flooding events in 2019, making it a suitable candidate to test the presented methodology.

3.2 Gauge Height Prediction

The historical gauge height data was obtained from Water data published by the United States Geological Survey (https://waterdata.usgs.gov/). The location of the corresponding gauge (station number – 07019130) is shown by the green square in figure 3.3.

Figure 3.3 Gauge Location

The available data structures vary for different sites as they are operated in cooperation with different organizations. The Valley Park site is operated by the U.S Army Corps of Engineers – St. Louis District. The 15 minute time interval data for stage flow at the site is available from May 15th, 2016 5 PM onward, with September 1st, 2019 4 PM being the last data point used here. This gives a total of 113,994 samples that when plotted give insight to the number and degree of flood events at that location (fig. 3.4).

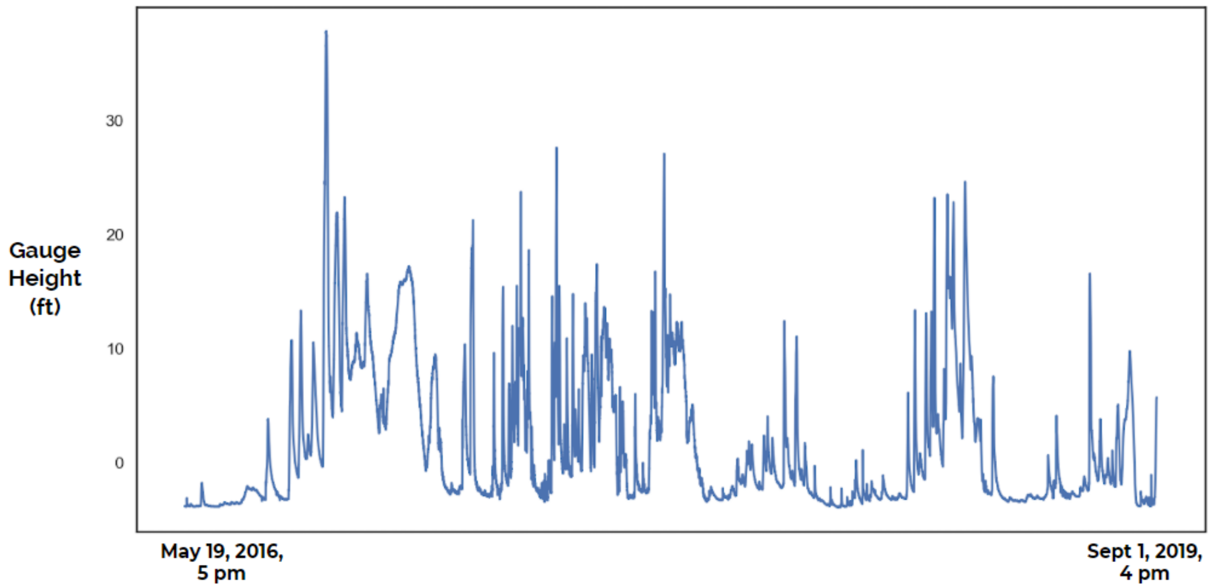

Figure 3.4 Historical Gauge Height data

Weather forecasts are provided by Advanced Hydrologic Prediction Service (AHPS) of National Weather Service (NWS), which is a part of National Oceanic and Atmospheric Administration (NOAA). The forecasts are 6 hour apart for the chosen testing location, which is not sufficient for developing real time traffic management solutions. Therefore, a model capable of predicting gauge height for every 15 minutes based on past data is was developed.

The algorithm performed best with 100 LSTM input layers, 1 dense layer, *'adam'* - optimizer, lookback - 70, batch size - 110 as the parameters (see appendix A for details). The algorithm performs well fitting the data during both training and testing (fig. 3.5).

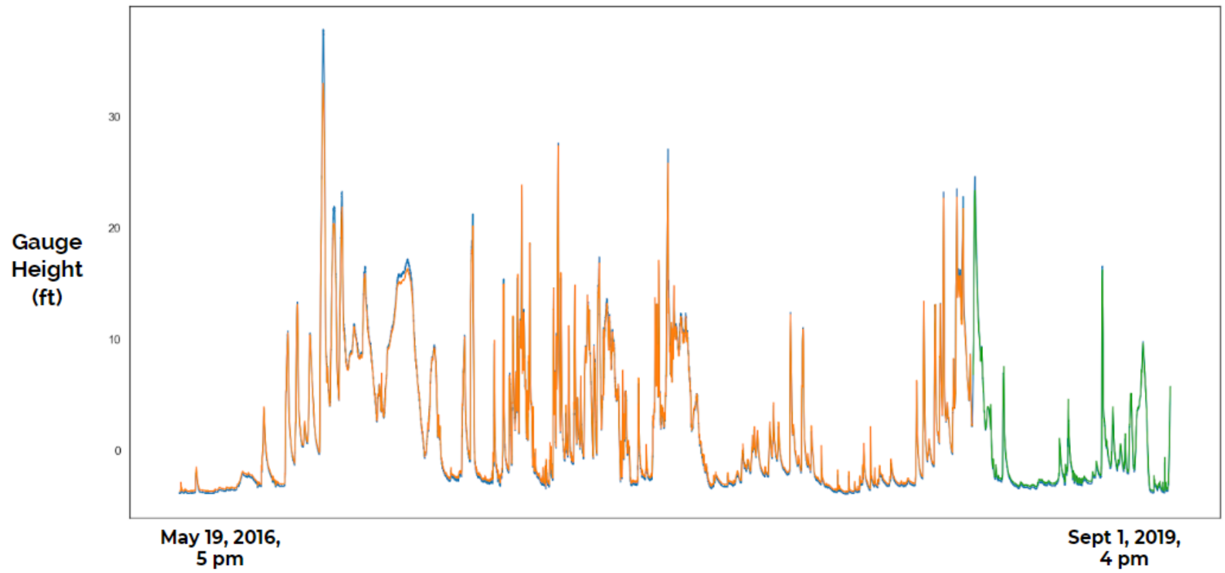

Figure 3.5 LSTM training and testing results. The input data is in blue, training results in orange, and testing results in green.

After the algorithm was trained and tested, it was used to make predictions from September 1st, 2019 6 PM until September 3rd, 2019, 6 AM at 15 minute intervals. These results were compared to the predictions from USGS website by the AHPS of the NWS and the actual data during this period (fig. 3.6). The root mean square error for USGS data was 1.065 and LSTM was 0.453. The error in the LSTM results was lower than the USGS predictions, even though the LSTM made 24 predictions in six hours compared to the USGS prediction once every 6 hours. The results of this algorithm are used as model inputs for the flood inundation sub model.

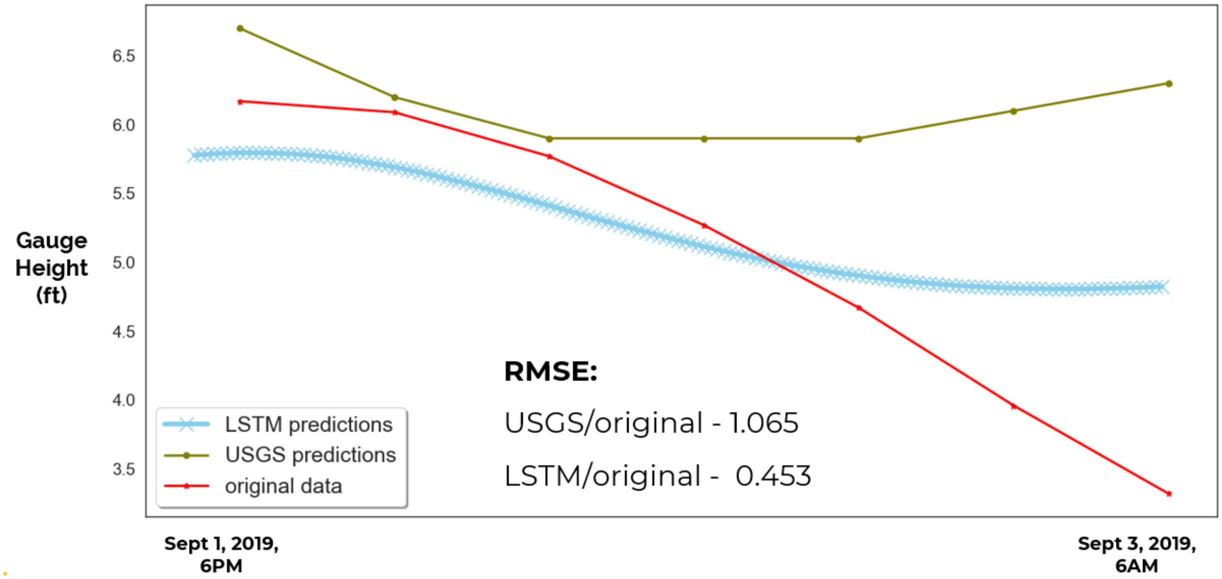

Figure 3.6 LSTM prediction results and comparison

3.3 Flood Inundation

Valley Park possesses 43 flood inundation profiles between 11ft and 54ft using the FIM tool. The historic crest for this location is 44.11ft recorded on 12/31/2015 and that profile will be used to prove the efficacy of the model framework presented (NWS, 2019). Figure 3.7 represents the flood inundation profile for a stage value of 45ft. Note that 45ft was chosen instead of 44ft to account for the historic crest value and to adhere to the one-foot increment limitation of the FIM at this monitoring site. Figure 3.8 represents the closed road segments as determined from geoprocessing and data layer integration. This set of closed road segments is used as a model input for the traffic simulation in SUMO.

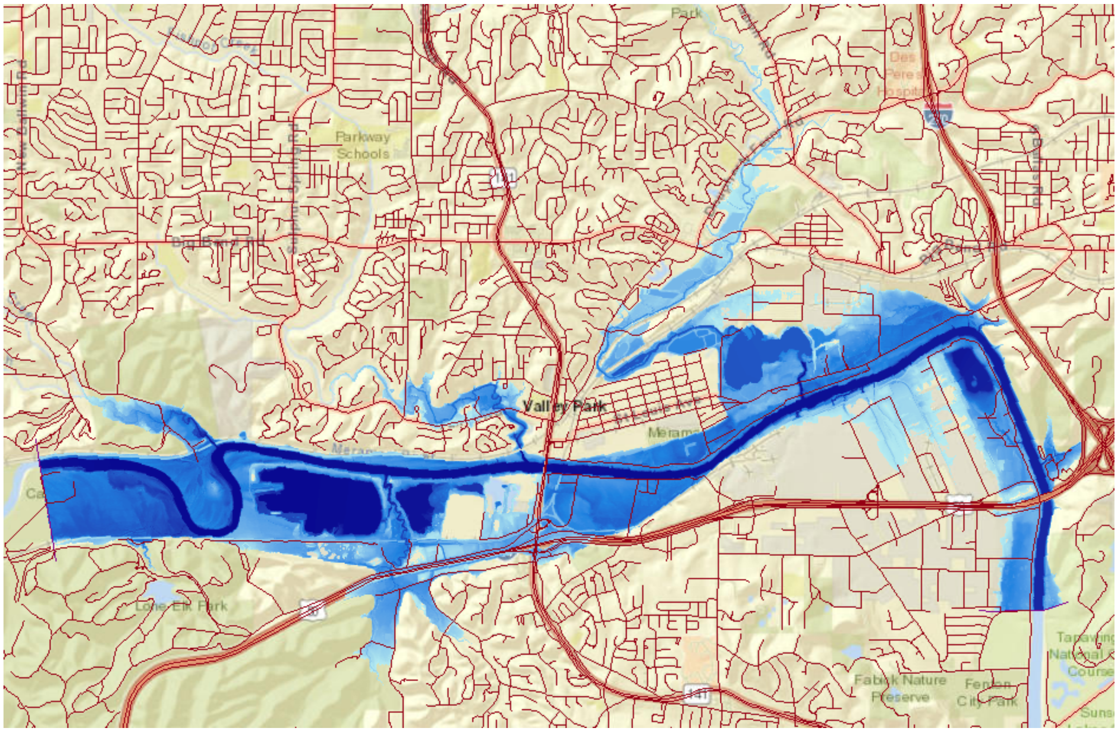

Figure 3.7 Flood Inundation Profile for 45ft Stage Value for Valley Park, Missouri

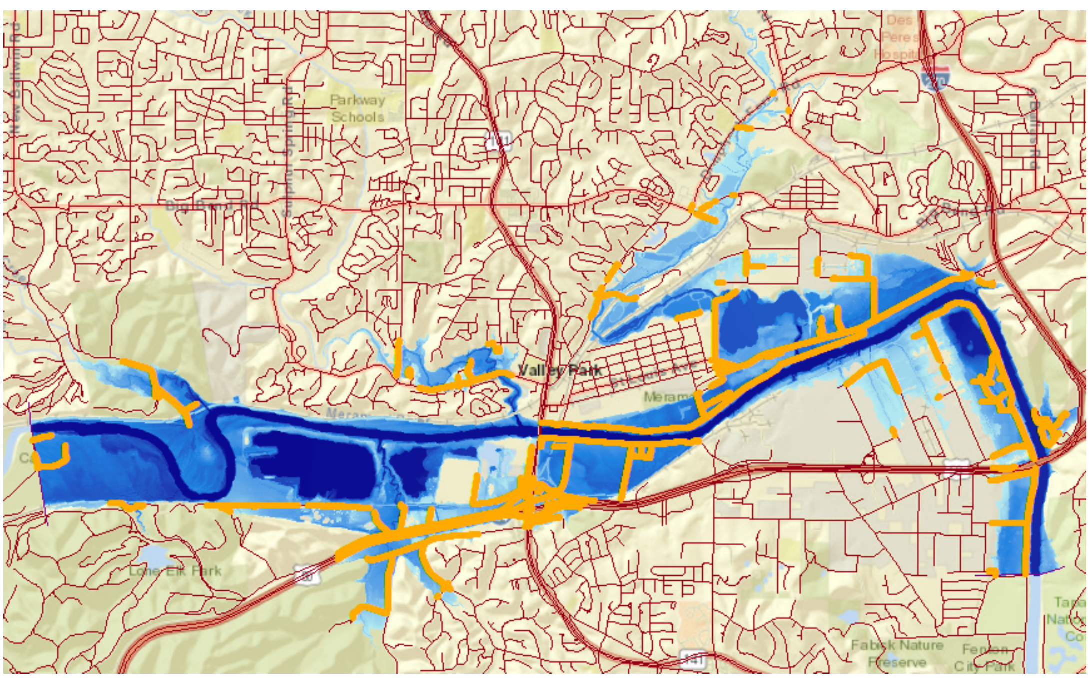

Figure 3.8 Closed Road Segments for Flood Inundation Profile of 45ft for Valley Park, Missouri

3.4 Traffic Simulation and Economic Impact

Based on the flood inundation profile and affected road segments, the relevant roads were removed from the network manually. Figure 3.9 compares the original road network (left) and the flooded road network (right).

Figure 3.9 Original road network (left), flooded road network (right)

For the traffic simulation, vehicles were generated with random assignment of origins and destinations for a period of 3600 seconds. Two simulations were performed, once on the original network and then on the flooded network. The number of cars running on networks at every second are plotted in figure 3.10.

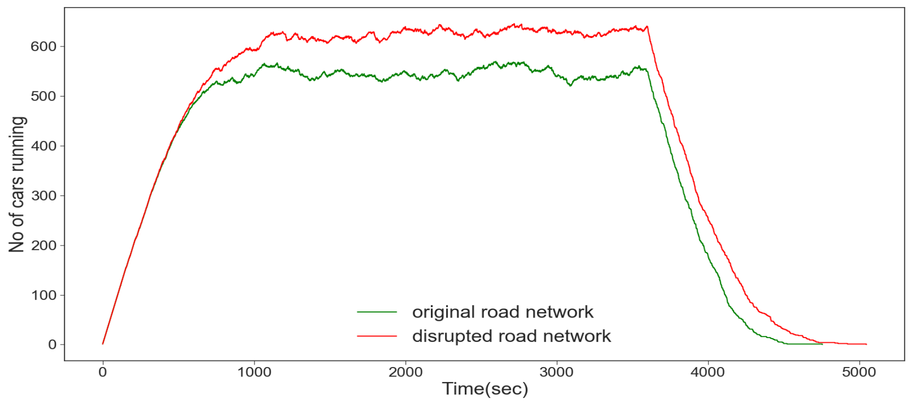

Figure 3.10 Comparison of number of cars on the networks

The overall travel time for each vehicle was evaluated both on the normal and flooded road networks to determine the corresponding delay time. To evaluate indirect costs the purpose of travel for the vehicles was assumed to be 40% for work, 40% for social activities, and 20% for others. Using an average hourly income of $20 per hour, the final indirect costs associated with rerouting was calculated.

For this scenario, the total indirect costs associated with delays were $5519/hour. The usual traffic disruption would last for many hours and in some cases several days. Therefore, this cost can quickly add up to a significant value during flooding events.

Chapter 4 Conclusions

The results reported here demonstrate a proof of concept for the integration of geospatial data, river gauge data, and traffic data to improve flood prediction and traffic routing information. The framework presented here provides updated information on the status of floodwaters and how they are affecting roads at 15 minute intervals. More importantly, it predicts future river gauge heights more accurately than current models, making it possible for decision makers to evaluate the future state of the road network and plan timely road closures accordingly.

The framework also provides precise information on how traffic can be rerouted most effectively to avoid high risk areas. Integrating the inundation mapping with the road networks provides information on roads that are likely overtopped at 15 minute intervals rather than six hour intervals if the NOAA data were used, giving motorists more advanced warning. The framework also takes into account traffic that is already in route, making it possible to dynamically change traffic based on changes in flood predictions. The traffic analysis also allows the planner to consider indirect losses associated with the road closures to better evaluate the overall economic impact of the flood event. The water level increments of one foot give a better indication of the road conditions than if only Digital Elevation Models were used, which is current practice. This makes it possible for transportation safety or disaster planners to create tools specific to their region in the event of a flooding event and when an event occurs, it provides a method to predict when roads are no longer safe for motorists to travel on.

Chapter 5 Limitations and Future Work

The results reported here demonstrate a proof of concept for the proposed methodology and there are some limitations. The FIM provides useful flood inundation profiles but only for discrete locations near St. Louis and Kansas City, Missouri. Further, profiles are only given in six inch to one foot increments and require that the next integer value be chosen to avoid underrepresentation. This shortcoming results in roads being closed that are unaffected by an actual flood event. Therefore, the methodology presented here is limited to those locations and incremental stage values. In addition, some areas within the FIM are labeled as areas of uncertainty. This designation can be due to several modelling limitations, but in this case they are attributed to a levee surrounding the Valley Park Industrial area. In the event that the river stage value reaches 44ft, the levee at Valley Park would be overtopped. The profile within the area is uncertain and therefore not represented in the profile generated.

The gauges are operated by different organizations and this results in varying data availability. This is a limitation on the generalizability of the model. The gauge in Valley Park, MO for the Meramec River is operated by the U.S. Army Corps of Engineers, St. Louis - District. They started collecting the gauge height for every 15 minutes starting May 19, 2016, this provided a significant amount of data to train the algorithm. However, many sites are still collecting data every hour. The time frame for gauge height prediction algorithm would be different in certain locations depending on the availability of data.

There is an uncertainty associated with every model and evaluating that uncertainty helps to define the limitations of the model. A methodology for evaluating the certainty measure associated with the prediction of a road being flooded can be developed in the future. The traffic data (origin, destination, etc.) used for simulations are randomly generated and may not be an exact match to current traffic patterns. Because of this, the actual delay times and indirect cost can vary significantly from the estimated values. The simulation was run for a period of 60 minutes or 3600 seconds using the SUMO library. Using longer run times and/or a larger road network in the simulation may provide a more robust model. Additionally, the traffic generated over a period of 3600 seconds does not match exactly with the traffic data collected by the Missouri Department of Transportation (MoDOT). Future work will involve making these changes in order to represent the interactions between the variables in a more efficient way.

A Supply Chain Infrastructure Restoration Calculator (SCIRC) software tool developed by researchers at Missouri S&T, Rolla and United States Geological Survey (USGS), Rolla can also be used to calculate direct costs and other resources needed to restore damaged road network post-flood events. Better protocols can then be developed once the estimates of material, costs, and number of restoration crews are made available to the responsible personnel using the software tool. Other tools developed by the US Army Corps of Engineers (USACE) such as Hydrologic Engineering Center Flood Impact Analysis (HEC-FIA) (FIA 2.2 Features) and Hydrologic Engineering Center Flood Damage Reduction Analysis (HEC-FDA) (FDA) can also provide information related to losses to agriculture, structures, and lives etc. from floods along with the additional economic parameters like expected annual damage and equivalent annual damages.

Currently, all of the models exist as independent solutions and manual input required to connect them. A single integrated flood and indirect cost estimation tool can be developed in the future by automating and integrating these models into a single cohesive application.

References

Alcrudo, F. (2002). A state of the art review on mathematical modelling of flood propagation. *EC Contract EVG1-CT-2001-00037 IMPACT Investigation of Extreme Flood Processes and Uncertainty: Proceedings of the First IMPACT Project Workshop, Wallingford, UK, May 2002 (CD-ROM)*, 1–22. Retrieved from http://www.impact-project.net/cd/papers/print/008_pr_02-05-16_IMPACT_Alcrudo.pdf

Astarita, V., Er-Rafia, K., Florian, M., Mahut, M., & Velan, S. (2001). Comparison of three methods for dynamic network loading. *Transportation Research Record*, *1771*(1), 179-190.

Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, *236*(1–2), 54–77. https://doi.org/10.1016/S0022-1694(00)00278-X

Brunner. (2016). Hydraulic Reference Manual Version 5.0, (February), 547.

Campolo, M., Andreussi, P., & Soldati, A. (1999). River flood forecasting with a neural network model. *Water resources research*, *35*(4), 1191-1197.

Chiari, F., Delhom, M., Santucci, J. F., & Filippi, J. B. (2000, October). Prediction of the hydrologic behavior of a watershed using artificial neural networks and geographic information systems. In *Smc 2000 conference proceedings. 2000 IEEE International Conference on Systems, Man and Cybernetics.'cybernetics evolving to systems, humans, organizations, and their complex interactions* (Vol. 1, pp. 382-386). IEEE.

Chistropher, O. (2015, August 27). Understanding LSTM Networks. Retrieved from https://colah.github.io

Chu, H.-C. (2016). Effects of Extreme Weather and Economic Factors on Freight Transportation. *Advances in Management & Applied Economics, Vol. 6, No.1*, 113–130.

Damle, C., & Yalcin, A. (2007). Flood prediction using time series data mining. *Journal of Hydrology*, *333*(2-4), 305-316.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, *1*(1), 269-271.

Elsafi, S. H. (2014). Artificial neural networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal*, *53*(3), 655-662.

Enke, D. L., Tirasirichai, C., & Luna, R. (2008). Estimation of Earthquake Loss due to Bridge Damage in the St. Louis Metropolitan Area. II: Indirect Losses. *Natural Hazards Review*, *9*(1), 12–19. doi: 10.1061/(asce)1527-6988(2008)9:1(12)

FDA https://www.hec.usace.army.mil/software/hec-fda/

FIA 2.2 Features https://www.hec.usace.army.mil/software/hec-fia/features.aspx

Gartner, N. H., Little, J. D., & Gabbay, H. (1975). Optimization of traffic signal settings by mixed-integer linear programming: Part I: The network coordination problem. *Transportation Science*, *9*(4), 321-343.

Gerl, T., Bochow, M., & Kreibich, H. (2014). Flood Damage Modeling on the Basis of Urban Structure Mapping Using High-Resolution Remote Sensing Data. *Water*, *6*(8), 2367–2393. doi: 10.3390/w6082367

Horitt, M. S., & Bates, P. D. (2002). Evaluation of 1D and 2D Numerical Models for Predicting River Flood Inundation. *Journal of Hydrology*, *268*, 87–89. https://doi.org/10.5059/yukigoseikyokaishi.64.628

Janic, M. (2007). Modelling the full costs of an intermodal and road freight transport network. *Transportation Research Part D: Transport and Environment*, *12*(1), 33–44. doi: 10.1016/j.trd.2006.10.004

Jonkman, S., Bočkarjova, M., Kok, M., & Bernardini, P. (2008). Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. *Ecological Economics*, *66*(1), 77–90. doi: 10.1016/j.ecolecon.2007.12.022

Keay, K., & Simmonds, I. (2005). The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia. *Accident Analysis & Prevention*, *37*(1), 109–124. doi: 10.1016/j.aap.2004.07.005

Koutroulis, A. G., & Tsanis, I. K. (2010). A method for estimating flash flood peak discharge in a poorly gauged basin: Case study for the 13-14 January 1994 flood, Giofiros basin, Crete, Greece. *Journal of Hydrology*, *385*(1–4), 150–164. https://doi.org/10.1016/j.jhydrol.2010.02.012

Lada, B. (2018). 25 years later: The Great Flood of 1993 remains worst river flooding US has ever seen. Retrieved from https://www.accuweather.com/en/weather-news/25-years-later-the-great-flood-of-1993-remains-worst-river-flooding-us-has-ever-seen/70005654.

Lampo, Richard G., et al. Corps of Engineers Technology Adoption Processes (CETAP) Study. Volume 1. Executive Summary. No. CERL-TR-FM-93/09-VOL-1. CONSTRUCTION ENGINEERING RESEARCH LAB (ARMY) CHAMPAIGN IL ENGINEERING AND MATERIALS DIV, 1993.

Meramec Flooding Proposal - Update https://dnr.mo.gov/env/meramecfloodingproposal.htm

Merwade, V., Cook, A., & Coonrod, J. (2008). GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. *Environmental Modelling and Software*, *23*(10–11), 1300–1311. https://doi.org/10.1016/j.envsoft.2008.03.005

Monaghan, J.J. (1994) Simulating Free Surface Flows with SPH. Journal of Computational Physics, 110, 399-406. http://dx.doi.org/10.1006/jcph.1994.1034

National Research Council. (1999). The Impacts of Natural Disasters: A Framework for Loss Estimation. *Washington, DC: The National Academies Press*. https://doi.org/10.17226/6425

Photos: Before & After Meramec River Flooding https://www.kmov.com/news/photos-before-after-meramec-river-flooding/article_fc16115e-12e2-54e6-bf84-ba1732a4dcbd.html

Roberts, S., Nielsen, O., Gray, D., Sexton, J., Davies, G., & Australia, G. (2015). ANUGA User Manual Release 2.0, (May).

Sheikh, V., Kornejady, A., & Ownegh, M. (2019). *Application of the coupled TOPSIS–Mahalanobis distance for multi-hazard-based management of the target districts of the Golestan Province, Iran*. *Natural Hazards* (Vol. 96). Springer Netherlands. https://doi.org/10.1007/s11069-019-03617-0

Suarez, P., Anderson, W., Mahal, V., & Lakshmanan, T. (2005). Impacts of flooding and climate change on urban transportation: A systemwide performance assessment of the Boston Metro Area. *Transportation Research Part D: Transport and Environment*, *10*(3), 231–244. doi: 10.1016/j.trd.2005.04.007

Supply Chain Infrastructure Restoration https://communities.geoplatform.gov/disasters/supply-chain-infrastructure-restoration/

Suvro, B. (2018, May 23). An Introduction to Recurrent Neural Networks. Retrieved from https://medium.com

Tatano, H., & Tsuchiya, S. (2007). A framework for economic loss estimation due to seismic transportation network disruption: a spatial computable general equilibrium approach. *Natural Hazards*, *44*(2), 253–265. doi: 10.1007/s11069-007-9151-0

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F. W., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling and Software*, *90*, 201–216. https://doi.org/10.1016/j.envsoft.2017.01.006

Tirasirichai, C., & Enke, D. (2007). Case Study: Applying a Regional Cge Model for Estimation of Indirect Economic Losses Due to Damaged Highway Bridges. *The Engineering Economist*, *52*(4), 367–401. doi: 10.1080/00137910701686996

Traffic Volume Maps: Missouri Department of Transportation. (n.d.). Retrieved from https://www.modot.org/traffic-volume-maps.

Tsakiri, K., Marsellos, A., & Kapetanakis, S. (2018). Artificial Neural Network and Multiple Linear Regression for Flood Prediction in Mohawk River, New York. *Water*, *10*(9), 1158.

United States Army Corps of Engineers, HEC-RAS. Retrieved from: https://www.hec.usace.army.mil/software/hec-ras/

United States Department of Commerce, & Noaa. National Weather Service Advanced Hydrologic Prediction Service. Retrieved from https://water.weather.gov/ahps2/hydrograph.php?wfo=lsx&gage=vllm7.

Vikas, G. (2017, October 9). Understanding Feedforward Neural Networks. Retrieved from https://www.learnopencv.com

Xiong, L., Shamseldin, A. Y., & O'connor, K. M. (2001). A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi–Sugeno fuzzy system. *Journal of hydrology*, *245*(1-4), 196-217.

Ye, B. J., & Mccorquodale, J. A. (1998). Simulation of Curved Open Channel Flows by 3D Hydrodynamic Model. *Hydraulic Engineering*, *124*(July), 687–698.

Appendix A

LSTM Networks:

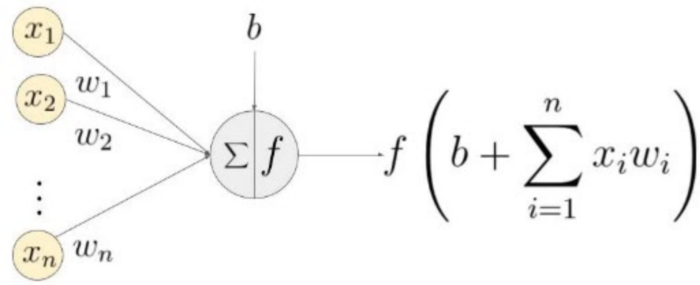

Figure A.1 An Artificial Neuron (Gupta, 2017)

A neural network is an artificial intelligence technique based on the functioning of the human brain. The Basic unit of a neural network is an 'Artificial Neuron' (fig. A.1). The Neuron takes the input variables (x_1, x_2….x_n), multiplies them with corresponding weights (w_1, w_2….w_n), and applies an activation function (f) to generate the output. A bias (b) can also be added as part of the function. The bias and the weights are the parameters that the neuron changes during the training step.

A neural network consists of a large number of such neurons (fig. A.2). The algorithm uses the current weights to predict an outcome and then adjusts the weights so that it can generate the corresponding output. Thus, a neural network represents a function that maps the input variables to the outputs.

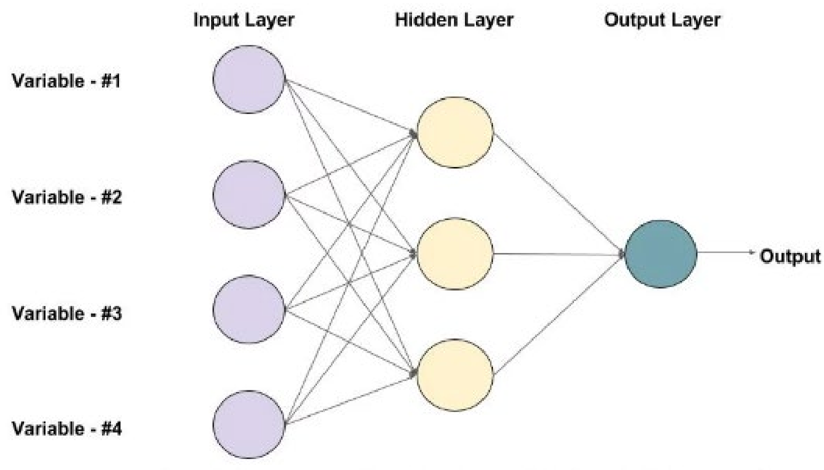

Figure A.2 Neuron Network (Vikas, 2017)

One shortcoming of traditional neural networks is that they cannot retain temporal information. To account for this shortcoming recurrent neural networks (RNN) were introduced. This network consists of loops which help in retaining information from previous time steps as shown in figure A.3. The information from the first time step (0) is passed to the next time step (1) and so on. This structure can make this algorithm effective for time series forecasting.

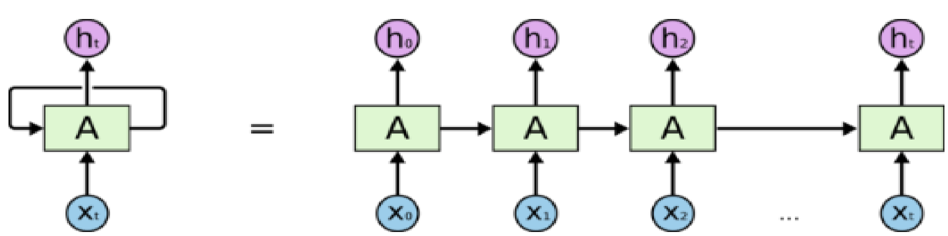

Figure A.3 Recurrent Neural Network (Suvro, 2018)

One issue with recurrent neural networks is that as the number of time steps increases the RNN cannot connect the information. For example, the information passed at time step (0) has little to no effect at time step (20). To address this inability to consider information many time steps prior the Long Short-Term Memory (LSTM) architecture was introduced. This is a type of recurrent neural networks as it still has a recurring chain-like structure, but uses a repeating module known as a 'LSTM cell' (fig. A.4).

LSTM cells can remove or add information regulated by the gates. It uses vector addition and multiplication to change the data. A sigmoid (σ) layer outputs either 1 or 0, which means it would 'let nothing go through' or 'let everything go through'. These gates ensure that the relevant information is being retained in the network over time.

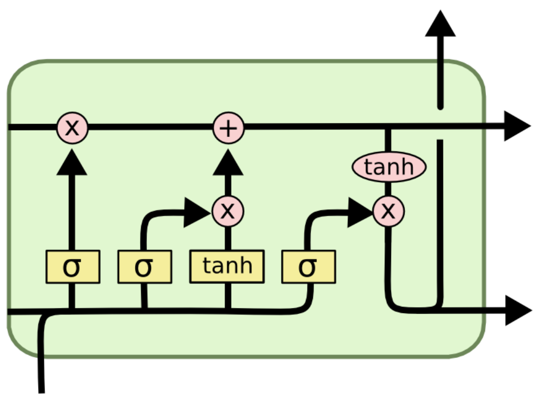

Figure A.4 LSTM cell (Christopher, 2015)

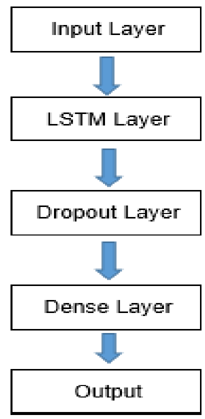

Figure A.5 LSTM architecture

A sequential model was used to develop the algorithm as it helps to stack up different layers required to build the LSTM linearly (fig. A.5). The 'Dropout' layer is placed after input layer, limiting the number of input neurons evaluated to reduce overfitting the training data. It can also be used to evaluate uncertainty in the predictions of LSTM. Dropout results in the algorithm learning more features and might take more iterations to converge. The final dense layer performs a linear operation to output a single forecast.