Creation of an All-weather Road Impact Prediction System (ARIPS)

January 2026	Project number TR202513
Final Report	MoDOT Research Report number cmr 26-002

PREPARED BY:

Heather D. Reeves

Andrew A. Rosenow

Jorge Duarte

Cooperative Institute for Severe and High-Impact Weather Research and Operations (CIWRO)

PREPARED FOR:

Missouri Department of Transportation

Construction and Materials Division, Research Section

Technical Report Documentation Page

1. Report No. cmr 26-002	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Creation of an All-weather Road Impact Prediction System (ARIPS)		5. Report Date January 2026 Published: January 2026
		6. Performing Organization Code
7. Author(s) Heather D. Reeves, https://orcid.org/0000-0003-1843-2218 Andrew A. Rosenow, https://orcid.org/0009-0003-8833-2690 Jorge Duarte, https://orcid.org/0000-0002-0463-4778		8. Performing Organization Report No.
9. Performing Organization Name and Address Cooperative Institute for Severe and High-Impact Weather Research and Operations (CIWRO) 120 David L. Boren Blvd. Suite 2100 Norman, OK 73072		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. MoDOT project #TR202513
12. Sponsoring Agency Name and Address Missouri Department of Transportation (SPR-B) Construction and Materials Division P.O. Box 270 Jefferson City, Missouri 65102		13. Type of Report and Period Covered Final Report (January 2025-January 2026)
		14. Sponsoring Agency Code

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 project is designed to examine the utility of using the database of flood-induced road closures collected by MoDOT to develop an Artificial Intelligence/Machine Learning (AI/ML) model. This was accomplished by assembling a database of predictors. These predictors include quantitative precipitation estimates, hydrologic model output, and a series of static fields relevant to flooding, such as ground permeability, elevation, and slope. These disparate sources of data were assembled into an AI/ML ready dataset.
The second component of the project was the evaluation of the utility of this dataset for the task of training an AI/ML model. In general, the bulk evaluation of the dataset showed promise for this task. Two important shortcomings were identified as part of this process. First, the closure database records only the day of the closure, which is a challenge for flash floods, which occur on the time frame of hours. Secondly, the MoDOT dataset represents only events where flooding occurred - for a model to predict a probability of an event, it must have a representation of times where the event did not happen to learn the difference between events and non-events. Thus, a database of non-flooding rainfall is needed alongside the MoDOT closure database. Both of these identified issues can be rectified in future research efforts using the database assembled in this initial project.

17. Key Words Flooding, Flash flooding, Rainfall, Machine learning, Road closures		18. Distribution Statement No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 28	19.

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Creation of an All-weather Road Impact Prediction System (ARIPS)

By

Heather D. Reeves, Andrew A. Rosenow, Jorge Duarte
Cooperative Institute for Severe and High-Impact Weather (CIWRO),
The University of Oklahoma, Norman, Oklahoma

Prepared for
Missouri Department of Transportation

January 2026

Final Report

TR202513

Copyright

Disclaimer

The contents of this report reflect the views of the author(s) who is (are) responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Missouri Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Declaration of Generative Artificial Intelligence (AI), and AI Assisted Technologies

The use of unapproved, **open AI systems** (such as ChatGPT, Bing Chat, and other publicly available AI systems) for research report writing **is prohibited.** Open AI systems such as these, capture data that is entered into it and then uses that data for training their systems to respond to all users of the system. **This is the equivalent of publishing information onto a public website – a violation of MoDOT policy.** Researchers inputting data and information into an AI tool are prohibited from disclosing confidential data or information belonging to MoDOT or MoDOT's partners. Researchers must comply with MoDOT's policies concerning data and record retention, and the proper storage, handling and sharing of sensitive information.

Acknowledgments

The authors acknowledge MoDOT for providing funding for this research.

Abstract

This project is designed to examine the utility of using the database of flood-induced road closures collected by MoDOT to develop an Artificial Intelligence/Machine Learning (AI/ML) model. This was accomplished by assembling a database of predictors. These predictors include quantitative precipitation estimates, hydrologic model output, and a series of static fields relevant to flooding, such as ground permeability, elevation, and slope. These disparate sources of data were assembled into an AI/ML ready dataset.

The second component of the project was the evaluation of the utility of this dataset for the task of training an AI/ML model. In general, the bulk evaluation of the dataset showed promise for this task. Two important shortcomings were identified as part of this process. First, the closure database records only the day of the closure, which is a challenge for flash floods, which occur on the time frame of hours. Secondly, the MoDOT dataset represents only events where flooding occurred - for a model to predict a probability of an event, it must have a representation of times where the event did not happen to learn the difference between events and non-events. Thus, a database of non-flooding rainfall is needed alongside the MoDOT closure database. Both of these identified issues can be rectified in future research efforts using the database assembled in this initial project.

Executive Summary

This project consisted of the exploratory step in identifying how a database of flooding closures kept by MoDOT can be combined with precipitation datasets being created within the Multi-Radar Multi Sensor (MRMS) operational meteorological system to develop and train an Artificial Intelligence/Machine Learning (AI/ML) model. This model would predict the probability that a given point is experiencing conditions that cause significant enough flooding to require the closure of at least some roads.

An evaluation of several MRMS products was conducted. These include quantitative precipitation estimates (QPEs) - instantaneous precipitation rates, and accumulated precipitation over hourly, three-hourly, six-hourly, and twenty-four-hourly time periods. In addition, products from a hydrologic model within MRMS were included to quantify the hydrologic response to those precipitation fields. Finally non-meteorological factors, such as the permeability of the ground, the land use, the slope of the ground, the elevation, and more were considered. By assembling these myriad fields, the AI/ML model will have access to the various factors that can exacerbate or mitigate potential flooding.

The first task was to assemble the data needed for eventual model training. Due to the massive size of the MRMS grids (24.5 million grid points), a Missouri-encompassing domain was created and applied to all the meteorological datafields to make the size of the entire database more manageable and reduce the train time of the AI/ML model. Once assembled, the combined dataflow for model training was queried.

The second task was to evaluate the gridded datasets to ensure they are suitable for eventual AI/ML applications. A few results of this combination were noted that require further work before training an AI/ML model. The first is the flooding closures only recorded with the day of the closure, not the hour. As the precipitation data has an hourly cadence, with precipitation rates generated every two minutes, a more precise estimate of the flooding time will improve the model considerably. The second is the lack of non-flooding events in the dataset. For a model to learn probabilities, it needs examples of near-events that did not cause the phenomenon in question in order to learn the difference between the two populations. In order to train a model, a dataset of rainy, but non-flooding events needs to be created.

These identified issues can both be solved as part of a future research effort. Moreover, the dataset developed here can be used to solve both issues, as a method could be developed to estimate the time of flooding from the rainfall and hydrologic model outputs, and a negative events dataset could be developed using those in conjunction with the observed closure locations.

With the dataset created in this research effort, and those identified issues overcome in a future research effort, the data assembled here has utility in training an AI/ML model to predict flash flood closures.

Table of Contents

List of Figures

List of Abbreviations and Acronyms

AI/ML	Artificial Intelligence/Machine Learning
CREST	Coupled Routing and Excess Storage
FFG	Flash Flood Guidance
FLASH	Flooded Locations and Simulated Hydrographs
MRMS	Multi-Radar Multi Sensor
QPE	Quantitative Precipitation Estimate

Chapter 1: Introduction

Flooding represents a significant risk to the road transportation network. What seems like a relatively straightforward problem - identifying locations with lots of water - ends up being far more complicated than simply determining how much rain falls. The characteristics of that rain - including the location relative to the road, as well as the speed at which it falls - matters. The challenge is likewise increased by the dependence of these events on non-meteorological factors, such as the topography surrounding the road. In short, these impacts to the road network depend on the interactions of a number of factors that are not necessarily related in a linear manner.

Flash flooding - a rapid rise of water into a normally dry area within minutes to hours of the causative event - is among the most dangerous types of flooding to the road transportation network (NOAA 2019). Flash floods also represent the majority of fatalities in the United States. (Ashley and Ashley 2008). By their nature, flash floods are substantially less predictable than long-period river flooding. Where river floods can be monitored over time with river gauges and visual observations of flooding, the quick onset of flash flooding conditions make the prediction of flash flooding more difficult. Moreover, where river flooding is generally confined to existing river valleys, flash flooding can happen anywhere.

Some tools already exist to diagnose and predict flash floods. Much of the development for these tools has happened as part of the Multi-Radar Multi Sensor (MRMS, Zhang et al. 2016) system. The MRMS system was designed from the ground up to measure precipitation accurately. It combines radar, rain gauges, numerical weather prediction, and more to produce products to quantify not only the amount of rainfall, but also to provide advanced decision support for flash flooding. This advanced decision support is provided through hydrologic modeling capabilities with the Flooded Locations and Simulated Hydrographs (FLASH) (Gourley et al. 2017).

As useful as MRMS and FLASH are for flooding, however, the fields are more applicable for generalized decision support, such as a forecaster's decision on whether to issue a flash flood warning. When looking at specific applications, such as whether a particular roadway will be inundated in a specific location, more information is needed. The effects of the rainfall and hydrologic fields from MRMS and FLASH happens in context of the roadway environment - some roads are more susceptible to flooding than others due to design, geography, soil type, or a number of other factors. These factors also have interactions with each other - for instance, soil's ability to retain water can be reduced by previous rainfall events.

In situations like this, where complex relationships between multiple variables exist, a machine learning technique is a useful way to develop a decision support system. While ML techniques are powerful, they also require careful consideration of the problem at hand, and the availability of data to answer the underlying research question. The interactions between these datasets need to be explored, including any challenges in relating them in a way that an AI/ML model will understand, to result in the best possible training and verification dataset. This project is designed to accomplish this preparatory work for the training of an AI/ML model.

There are two main objectives of this project. The first was to construct an AI/ML-enabling dataset based on diverse data sources including the Missouri Department of Transportation (MoDOT) data archive for road closures. The second was to perform a pilot study demonstrating the utility of AI/ML approaches to provide a state-wide probability of flooded roads. This report presents the results of these two research efforts. The second chapter describes the datasets and methodology for this project. The third chapter presents analyses of the AI/ML inputs. The fourth chapter describes how data could be used to develop an AI/ML model in future research. Finally, chapter five presents the conclusions of this research.

Chapter 2: Data and Methods

There are two categories of datasets that have been evaluated as necessary to the development of this model, predictor datasets and the verification dataset.

The predictor datasets are a series of data sources representing factors expected to cause or exacerbate roadway flooding. Rainfall is the most important contributor in the occurrence of flooding. Quantifying rainfall is accomplished by using output from the MRMS system.

The particular MRMS fields used as input for model training included four quantitative precipitation estimates: an instantaneous rate, a one-hour accumulation, a six-hour accumulation, and a twenty-four-hour accumulation. The accumulation fields are time-integrations of the precipitation rate. By including these different time windows, the model will incorporate both immediate rainfall for its flash flood potential, as well as rainfall at longer time lags that cause soil saturation, thus increasing runoff potential. In addition to the rainfall quantity, the output of the FLASH model within the MRMS suite is included. These model fields include maximum Annual Recurrence Interval (ARI), maximum unit streamflow, and maximum QPE to Flash Flood Guidance (FFG) ratio. These model outputs combine precipitation fields with other information to indicate the flash flooding risk.

Figure 1. A representation of the domain developed for this study. The green line represents the edge of the data mask applied to all the datasets processed for this project.

In addition to these precipitation fields, static fields related to flash flooding were acquired. These fields include: imperviousness, a measure of how readily the ground allows for water infiltration; a digital elevation model, containing the altitude of every grid point; slope, the change in elevation over a grid point; population density, the number of people per square kilometer within a grid point; flashiness, the tendency for flooding in a grid point to be a flash flood; road density, the density of primary or secondary roads within a grid point; and a historical flood frequency grid. These products are unchanging between individual flooding events and provide context for the MRMS fields.

All of these fields were processed, creating a subset of these fields on a common grid, and limiting the datasets to only those grid points in or near Missouri. This slice, shown by the green box in Figure 1, contains the entire state of Missouri plus a 10 km buffer zone. The use of kilometers in the buffer zone and other portions of this project is driven by the 1 km grid spacing in MRMS; this keeps data selections to whole numbers of grid points. By reducing the full MRMS datasets to this subset, it will be substantially easier and faster to train an AI/ML model. The addition of a small buffer zone is to enable the model to consider data in a window around a grid point, even at the edge of the state.

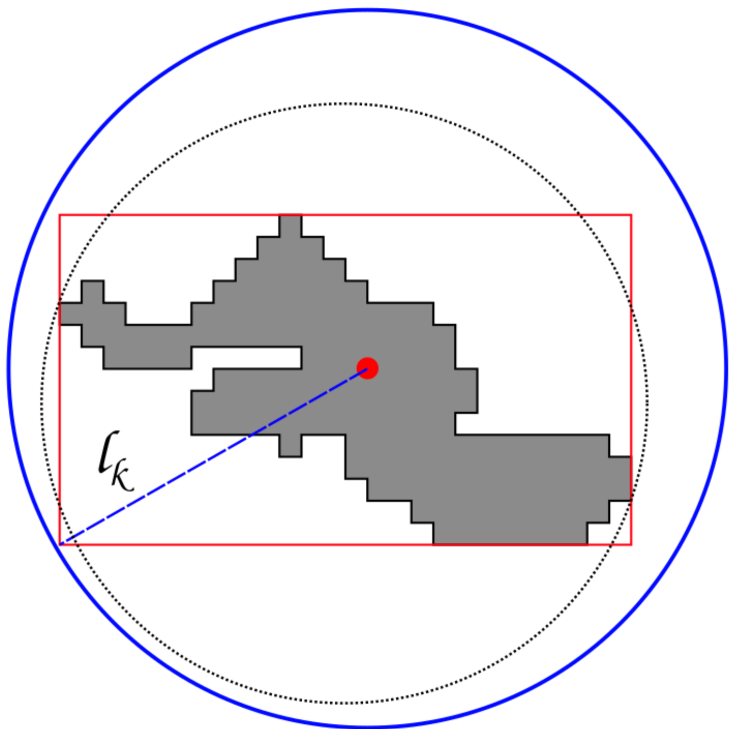

Figure 2. Schematic illustrating the bounding box (red rectangle), centroid (red circle), event circle (dotted line circle), and average radius (l_k).

With the meteorological and static datasets assembled, the closure dataset needed to be processed for linkage to these predictor datasets. These meteorological and static datasets are grid-based and were moved onto a common grid (the MRMS 1 km x 1 km grid). The road closures, on the other hand, are of arbitrary size and shape. Some road segments are short and could be relatively safely assumed to be a point. Other road segments are long, winding, and far larger than the grid points of the input dataset. Thus, a methodology was devised to uniformly handle the varying sizes of road closures. A schematic of this methodology is shown in Figure 2. The grey represents a shape of unknown size; here, it would be a road closure segment. First, a bounding box is created surrounding the road polygon that completely encompasses the shape. The centroid of this bounding box is then determined and recorded; this centroid is considered to be the location of the closure for the analyses presented in the rest of this report. In addition to determining a centroid, a consistent measure of the size of the road closure was desired for future use. To calculate this radius, the smallest possible circle that completely encloses that closure polygon was created, and the radius of this circle is considered to be the radius of the closure. This radius was recorded for all of the polygons from the closure dataset; then, an average radius across the dataset was calculated to represent an average "size" for all closures; here, that was 6 km when calculated across all closure events.

Chapter 3: Analyzing the AI/ML Inputs

When the datasets were successfully assembled, the properties of the combined data were examined to better understand the properties of the combined dataset, and what these properties mean for the development of an AI/ML model that predicts road inundations based on meteorological inputs. First, some of the bulk properties of the dataset will be presented.

The closures in this dataset happened over a range of time periods. Because only the date of closure was recorded, closures are measured in days. A road that is only closed for a few hours on the same calendar day will show up in the database as a zero-day closure. The vast majority of the closures are zero (23%) and one day (33%) in length, indicating that these are probably flash-flood events. Many of the longer-length events are probably longer-term river flooding but could also represent flash floods that caused damage to the road. In this dataset, 23% of closures are four days or longer.

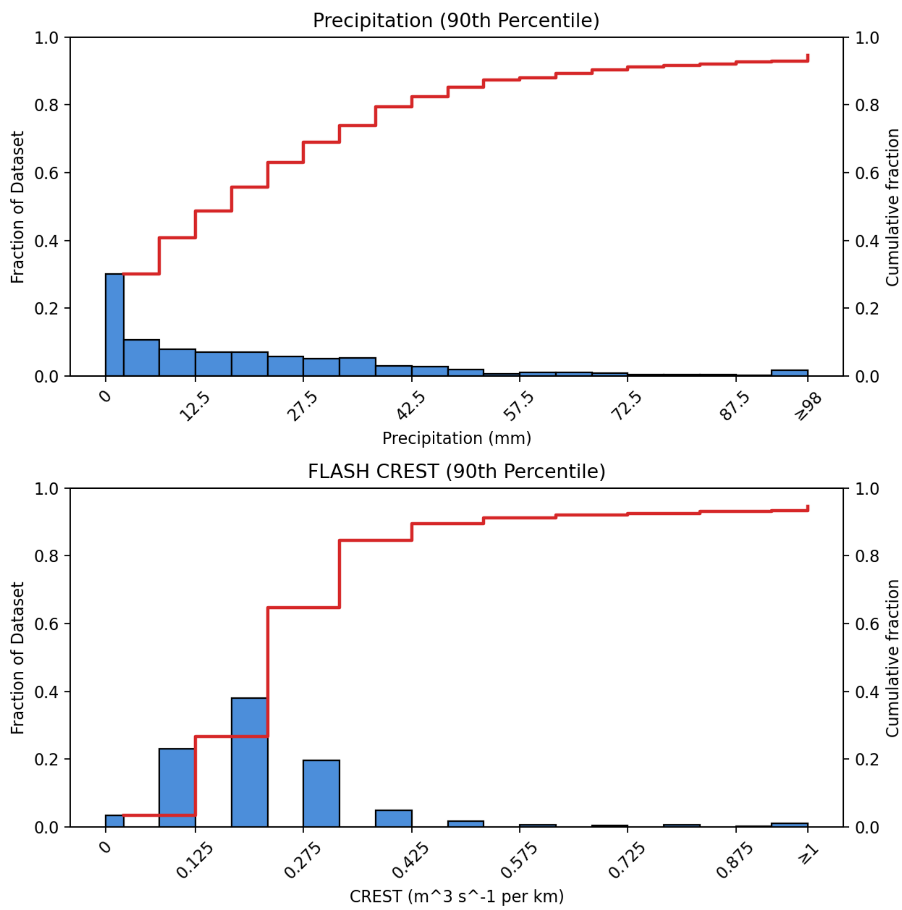

Figure 3. Histograms and cumulative distributions of the largest value within 20 km of flood event centroids for a) Twenty-four-hour QPE, and b) FLASH CREST unit streamflow.

Figure 3 contains histograms of precipitation fields associated with closure events. For these plots, the maximum values within 20 km were used to represent values in the region of the closure. The distribution of twenty-four-hour QPE maxima within 20 km of the closure in Figure 3a shows that many of these events have relatively low precipitation maxima, with 30% having a near-zero maximum. As mentioned previously, the underlying dataset only has a date of closure, necessitating an assumption of 0000 Coordinated Universal Time (UTC) as an hour of analysis for every closure. Some of these closures may be due to heavy rainfall after 0000 UTC; however, if the assumption were moved back to 2300 UTC as the hour of closure, other events would have their precipitation fall to before the lookback period. 0000 was chosen out of necessity due to the lack of any information within the closure dataset to specify the hour of inundation.

Conversely, many of these events have substantial precipitation nearby, with around 40% having a pixel of 25 mm in 24 hours or more nearby. Moreover, while the precipitation field has a substantial near-zero component, very few closures have FLASH Coupled Routing and Excess Storage (CREST) streamflow maxima near-zero, as shown in Figure 3b. The exact magnitudes of CREST are not important here; the AI/ML model should learn a relationship between these values and road inundation. What is important here is that the FLASH system is reacting to nearby precipitation, and this is with some events' precipitation occurring after the 0000 UTC assumed event time.

Another perspective on this precipitation-closure link is shown in Figure 4. Each dot represents the calculated centroid of a flooding closure in the MoDOT dataset, colored by the maximum precipitation within that 20-km window. There is no apparent geographic distribution in events with either heavy or near-zero precipitation - the variance is almost certainly due to the distribution of rain events around 0000 UTC. There are some geographic patterns in the closures - for instance, a clear signal of the Missouri River and its tributaries across central Missouri. Some of these closures may be due to river flooding, which in turn may not be due to rain anywhere near the closed road. In sum, Figure 4 shows that there are a significant number of observations that are ready for AI/ML model training, but also a substantial number of events that need more research to process to either make them ready as training data for a model or removed from the dataset as unlike the flash flood-caused closures.

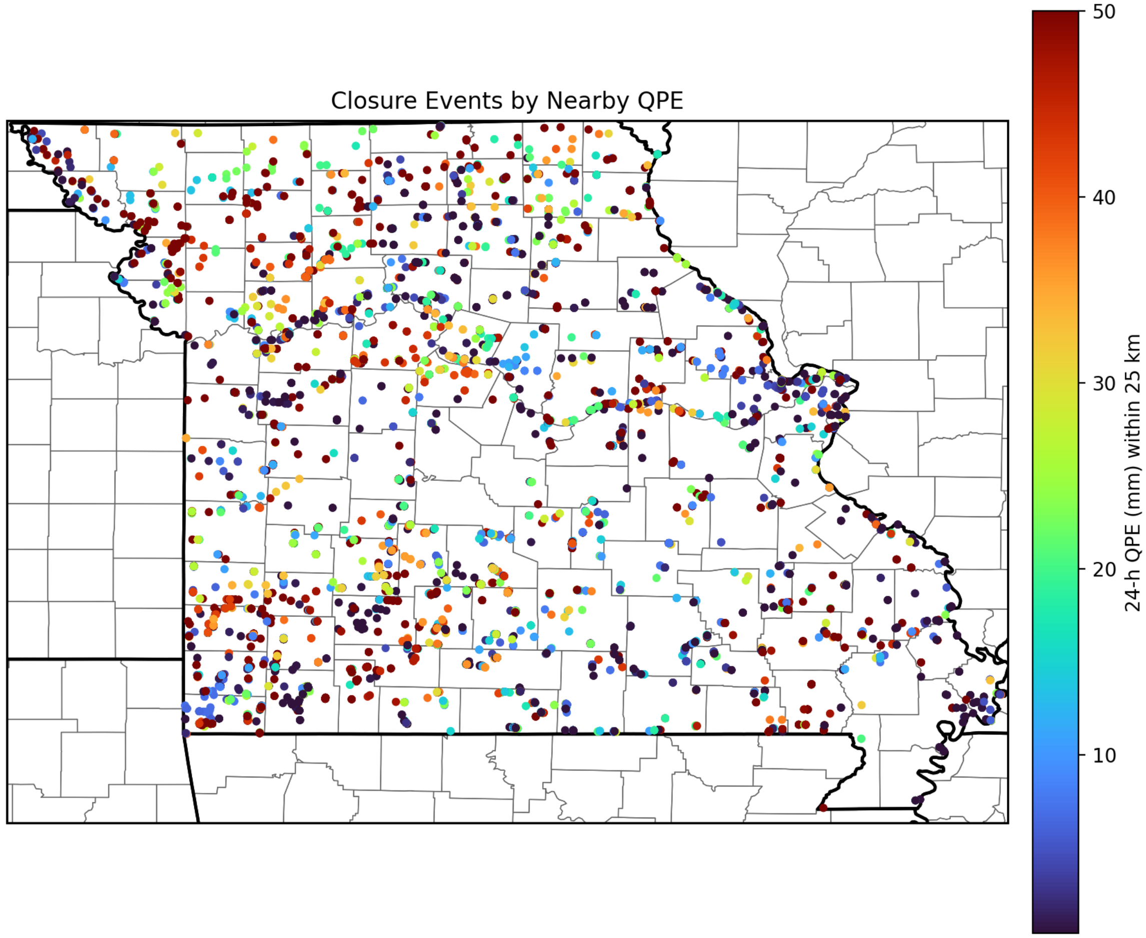

Figure 4. All centroids of closures in the MoDOT closure dataset, colored by the largest twenty-four-hour precipitation within 20 km of the location.

In this section, the data from the database will be examined for a couple of closures from the record generated here to illustrate how the meteorological fields interact and would inform an AI/ML model trained on the data.

Setting the stage for this event, a cluster of thunderstorms developed in eastern Nebraska on the afternoon of June 23, 2021. This storm cluster moved southeastward, dropping substantial rainfall across the northern portion of the state overnight into June 24th. This wetted the ground, limiting the possibility of any future rainfall to be absorbed into the soil, and exacerbating the potential for flooding in the region.

On June 25, 2021, a cluster of thunderstorms developed over northwestern Missouri by 0000 UTC. This cluster developed into a squall line that propagated over northern Missouri. This thunderstorm complex moved eastward across the state. However, the southern edge of the cluster, located roughly along Interstate 70, saw repeated development of thunderstorms over the same areas. The combination of these two rainfall events created a wide swath of 48 hour QPEs of four inches, with localized maxima above a foot of rain. Many of the closures in the MoDOT database are recorded as happening on June 25, which with the necessary assumption of 0000 UTC, places them as being effective at the beginning of the second round of precipitation.

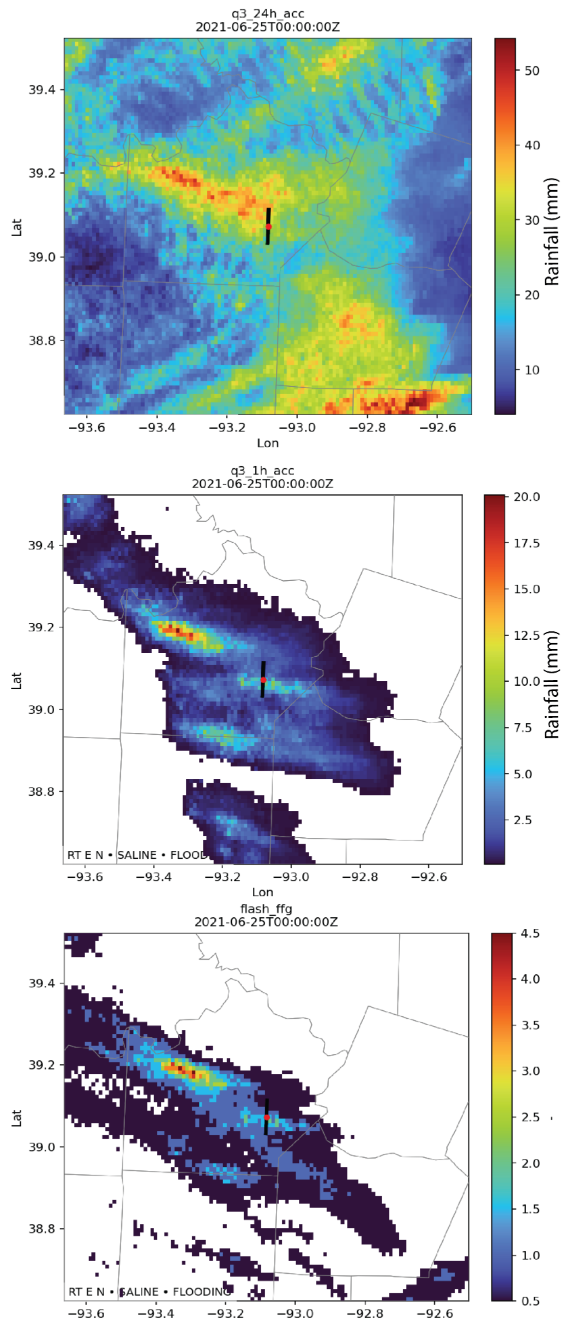

Figure 5. Meteorological datasets for a closure of Route E in Saline County on 06/25/2021. Fields are a) twenty-four-hour QPE (mm), b) one-hour QPE (mm), and c) QPE-to-FFG ratio. The black line represents the portion of the road that is closed, and the red dot represents the centroid.

The first closure that will be examined in more detail here is Route E in Saline County, Missouri, shown in Figure 5. Saline County is located between Columbia and Kansas City, along Interstate 70. The 24-hour precipitation panel in Figure 5a shows that areas near the road received between 30 and 40 mm of precipitation in the previous twenty-four hours. While the most significant precipitation was organizing to the northwest, a few isolated storms were moving across Saline County at 0000 UTC. This resulted in estimates of five-to-ten mm of rainfall in the previous hour, shown in Figure 5b. Because of the extensive rainfall that had happened over the previous day, the Flash Flood Guidance (FFG) from the FLASH hydrologic model indicated that the ratio of QPE to FFG was near or above one, indicating flash flooding was imminent, if not occurring at this time. This time represents one where the 0000 UTC assumption, while not ideal, ends up working due to precipitation ongoing at that time.

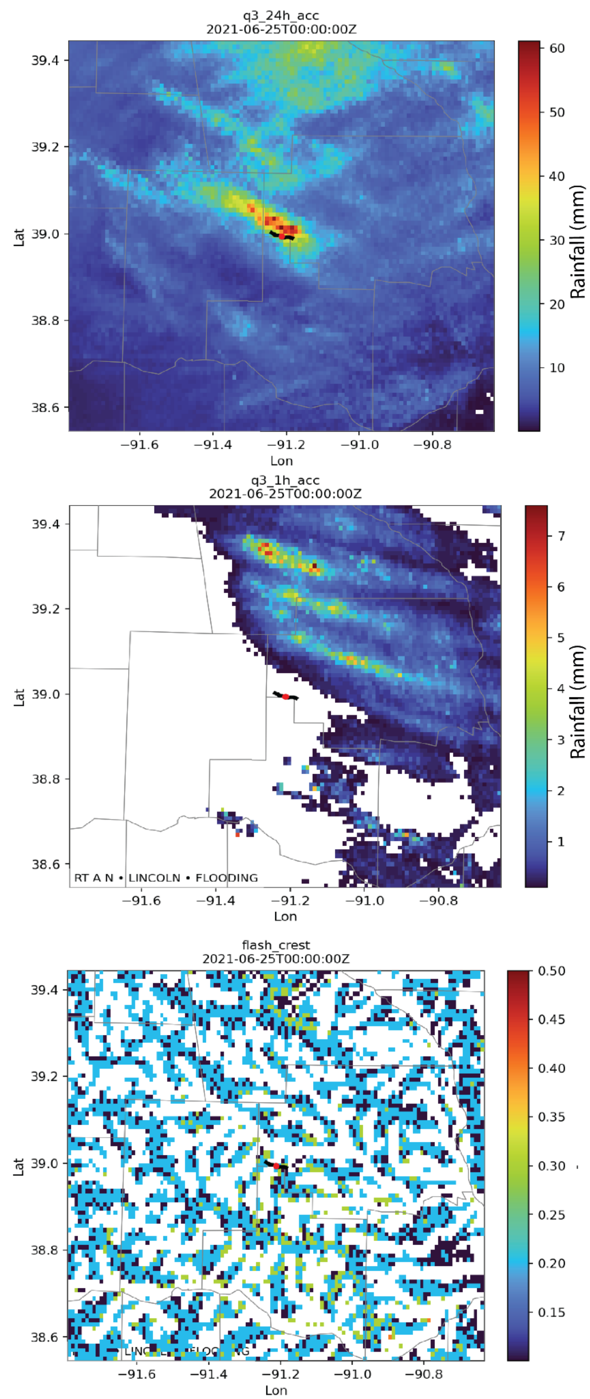

Figure 6. As in Figure 5, but for the closure of Route A in Lincoln County, and panel c) is FLASH CREST maximum unit streamflow (m^3 s^{-1} km^{-1})

The second example from that day is from Lincoln County, Missouri, where Route A was closed due to flooding on June 25. Meteorological data from this day is shown in Figure 6. This flooding closure happened in proximity to heavy rainfall from the day before, with 50+ mm QPEs located near to Rt. A in Figure 6a. However, by 0000 UTC, rain had ceased with the first round of precipitation, with only light showers (~5 mm hr^{-1}) to the north of the road site. If flooding had occurred by the start of June 25, it was not due to precipitation falling at 0000 UTC, and likely due to precipitation that fell substantially earlier.

The results of round one of precipitation is shown in the FLASH model's CREST streamflow in Figure 6c. CREST streamflows from 0.20 to 0.35 $m^3 s^{-1} km^{-1}$ are elevated from typical background values, indicating that the previous rainfall's runoff had swollen streams and rivers in the region. These streamflows would eventually double or more during the subsequent round of precipitation later on June 25. This illustrates how the longer QPE fields and the FLASH output present a signal of flood potential that an AI/ML model can be trained to recognize and include in its probabilities.

Chapter 4: Developing an AI/ML Model

In this chapter, the necessary next steps to developing an AI/ML model for flooding in roadways will be discussed, based on the results of the analysis in the previous chapter. The dataset assembled this year appears to be a sufficient basis for a model to learn the characteristics of a flooding event. While there are some events where the precipitation that drove road inundation is not clearly identifiable from bulk statistics (e.g. the events with near-zero precipitation within 20 km), most have notable precipitation within that window. Additionally, the hydrologic modeling (FLASH) shows nearby responses to this precipitation.

However, while the dataset is promising for AI/ML applications, there are two fundamental challenges that remain to be addressed before it is possible to train an AI/ML model. The first is to improve the handling of time, and ideally create an estimate of the time of flooding of the road. By its nature, the MoDOT closure database is a record of when the road was closed, not when it was flooded. Moreover, it does not contain the time of the closure, but simply the date it was closed. Because flash floods occur on the time scales of hours, this inherent lack of precision creates two distinct, but significant issues with training an AI/ML model.

The first is the present necessity of assuming a valid time. Whether it is the twenty-four-hour QPE, FLASH output, or any other precipitation field, the model needs data valid at a certain time to be associated with each closure. As discussed above, this was assumed to be 0000 UTC for this data analysis example out of the necessity of having to choose something. The bin of near-zero twenty-four-hour precipitation in Figure 3a is at least partially due to this assumption.

It is possible to use the data assembled for this project to develop a better methodology than a blanket assumption. This would involve using the granularity of the precipitation datasets (two-minute rate, hourly, six-hourly, and daily) in combination with the hydrologic model output from FLASH to determine a time when these fields indicate a combined response to a precipitation event on the date of the closure event. While this would still be an estimate, it would be better than a blanket assumption.

These estimates are critical to model development for two main reasons. The first is the ability to use cases where precipitation is not in the analysis window. Any assumption of a uniform flooding valid time will necessarily exclude events like this. In addition to events that had precipitation occurring after the assumed time here, this would also catch events where the delay between flooding and that condition being reported long enough to push the causal precipitation before the analysis window.

The next, and most important step for building an AI/ML model is to develop a "negative" case dataset. In order to create a model that effectively discriminates between flooding and non-flooding cases, there must be a robust set of events where flooding did not occur alongside the database of events where flooding did occur. This will allow the model to learn parameters that separate the flooding events from non-flooding events.

However, this negative event dataset must be created with care. It would be simple to make one out of sunny days - but these sunny days would not be useful to the model in learning what non-flooding rain events look like. Conversely, naively choosing points along roadways on rainy days based solely on the fact that the road is not in the MoDOT closure database risks the inclusion of events where flooding occurred, but was for whatever reason not reported, or the flooding was not reported in time to close the roadway, and thus make the database. This would teach the model that flooding events were actually not flooding events, which degrades model performance.

The data assembled to teach the AI as part of this project would be useful in the creation of a negative dataset. The QPE fields can be combined with road shapefiles to find road segments that received non-negligible rainfall. The closure database can be used as an easy way to get rid of positives. Then, output from the FLASH model (peak streamflow) and the flash flood guidance ratio product can be used to ensure that none of the potential negative events are probable flooding cases, despite not appearing in the MoDOT closure database.

Chapter 5: Conclusions

This project created a precipitation and hydrologic dataset for the purposes of future development of an AI/ML model. This model will predict the probability of flooding of sufficient magnitude to require the closure of a roadway. The predictors for this model include QPEs from the MRMS system, as well as output from the FLASH hydrologic model. In addition to these precipitation-related predictors, additional static fields representing flooding-related factors, such as permeability, land use, land slope, and more were acquired. A mask was developed and applied to these fields to size the datasets down to only cover Missouri, to limit the size of the files and speed up the training and analysis processes.

Preliminary analyses of these datasets were performed to understand the interactions between the precipitation data and the closure database. Some important limitations were identified. One of the most important limitations that will need to be addressed in future research is the higher temporal resolution of the precipitation dataset. With hourly accumulations (and every two-minute rates), these data require more precise timing information than is available with the closure records. Flash flooding, the type of flooding that is most dangerous to a roadway system and a model like this would be most useful in predicting, happens on a time scale of hours to minutes. This time scale is well within the daily observation period of the closure database. The second important limitation is the lack of a negative events database, which would teach an AI/ML model the characteristics of rain events that did not result in flooding sufficient to close a road.

While these initial limitations mean that the data could not instantly be used to train an AI/ML model, the work performed this year has set up the solutions to the issues discovered. By having precipitation data tied to the closure database, future research can devise a programmatic set of rules to estimate the time of flooding for events in the MoDOT database, as well as assemble a robust set of negative events for the model from which to learn the properties of non-flooding events. The work performed during this project period has not only developed the datasets necessary to use the MoDOT flood closure database for machine learning, but it also identified the necessary analyses using the data assembled therein to complete that training work.

Chapter 6: References

NOAA, 2019: Operations and Services Water Resources Services Program, NWSPD 10-9, Definitions and General Terminology. National Weather Service Manual 10-950, 5 pp., www.nws.noaa.gov/directives/sym/pd01009050curr.pdf.

Ashley, S. T., and W. S. Ashley, 2008: Flood Fatalities in the United States. *J. Appl. Meteor. Climatol.*, **47**, 805–818, https://doi.org/10.1175/2007JAMC1611.1.

Gourley, J. J., and Coauthors, 2017: The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States. *Bull. Amer. Meteor. Soc.*, **98**, 361–372, https://doi.org/10.1175/BAMS-D-15-00247.1.

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.