Dam Analysis Modernization of Tools, Applications, Guidance and Standardization (DAM-TAGS) Project

Final Report

(June 1, 2021 – Oct 31, 2024)

Project Summary

In mid-20th century America, extensive drought, coupled with uninformed land management, produced a calamity of epic proportions. Valuable topsoil was literally gone with the wind during the infamous Dust Bowl. Once rains returned, massive flood events washed thousands of tons of soil from the unprotected remaining land surface. In response to these conditions, federal initiatives, led by the newly formed Soil Conservation Service, provided funding and expertise to construct over 12,000 earthen flood control structures nationwide.

In a testament to the skill of those original fabricators, virtually all of these structures are still functional and meeting their intended use, well after the end of their 50-year design life. However, although their basic function has remained the same, in many cases, the areas that these dams protect have changed drastically. Population growth and urban sprawl of the 21st century have replaced pastures and cropland with subdivisions and communities. The increased hazard potential associated with these developments accentuates the need to properly monitor, maintain, or even rehabilitate, the dams in question.

This need is well known to the principal investigators on this project; the Oklahoma Water Resource Center (OWRC) at Oklahoma State University and the USDA-ARS Agroclimate & Hydraulic Engineering Research Unit. Both are located in Oklahoma, which has over 2,100 earthen dams within its borders, more than any other state in the Union. Perhaps this proximity to the issue helped them recognize the enormity of this task requires a broad-based effort. One that goes beyond a traditional approach dependent upon technical expertise within the main research area but instead reaches out to disparate sources for multidisciplinary solutions.

Consequently, this project assembled a team of scientific researchers, engineers, data managers, and computer programmers. In addition, they joined their efforts to the Partnerships for Data Innovations (PDI), an intra-agency initiative focused on developing technological workflows and tools for use by USDA-ARS scientists from a variety of research areas. This collaboration effectively doubled the size of the team as they worked together to provide working solutions for not only improving dam analysis, monitoring and rehabilitation, but also cross-cutting solutions benefiting research~~ers~~ in cotton production, plant pathogens, soil science, and cattle fever ticks, among many others. Below is a partial listing of the solutions created by the DAM-TAGS team in their work with PDI:

- Cloud-based platform for methods development and collaboration
- Three chapters added to National Engineering Handbook
- Cloud environment; including workflows for PostgreSQL database management, UAV image processing, and IoT hub
- Pathway for data ingestion and transfer to knowledge graph for data organization and analysis
- Automated integration of weather information into dataset from both online repositories as well as continuous monitoring sensors
- Procedures for cloud-based data ingestion, sharing, and visualization tools
- Improved models for small watershed reservoir/lake water level predictions
- Software to improve efficiency of laboratory processes for handling soil and water samples
- GIS-based platform which transforms field observations to online maps and dashboards, including draft version for dam inspections
- Multiple outlets for training, education, and overall communication of dam monitoring information

The benefit of the PDI collaboration was two-fold. First, the project efforts provided real-world exposure for the DAM-TAGS team in working with scientists from myriad research areas. The team learned practical skills in communications regarding research data management, regardless of discipline, and developed pathways for successfully transferring data between multiple formats and environments. Second, several of the completed PDI tools, although built for specific research areas, could easily be modified for use in other disciplines. And, since the development process includes troubleshooting issues pertaining to the application itself, much preliminary work for these tools has already been completed. In effect, the runway was cleared for updating these data flows and tools with revisions and modifications specific to the needs for dam rehabilitation, maintenance, and monitoring.

Unfortunately, the promise of this effort was not fully realized. Multiple factors, including programmatic and personnel shifts within PDI, and upheaval in the federal system, hampered project goals. Not the least of these limitations was the unprecedented cessation of the DAM-TAGS funding prior to the final year of the project. This severely restricted the team's ability to complete these conversions as originally intended. Despite these roadblocks, the DAM-TAGS team was able to complete an impressive array of tasks, all documented in the following pages.

Finally, it is the belief of the entire team that these accomplishments demonstrate the utility of the novel approach adopted by the project. Indeed, they are a small sample of the potential solutions that these tools and innovations could eventually provide. It is their hope that extended funding will be made available for the opportunity to further explore and refine them for widespread adoption.

Project Administration

The Oklahoma Water Resources Center (Water Center, or OWRC) provided technical and outreach/communications support for the USDA-ARS-led DAM-TAGS project and associated scientists. Key management accomplishments included the hire and management of nine (9) full-time employees and seven (7) undergraduate student workers, project coordination via routine team meetings and regularly scheduled meetings with USDA-ARS Agroclimate & Hydraulic Engineering Research Unit (AHERU) Research Leader, and fiscal management and oversight of project funds.

Personnel Management

Project requirements continue to be routinely examined with USDA-ARS to determine personnel needs and ensure needed expertise and knowledge weare available to meet project goals and objectives. Through the project, nine (9) full time Team Members were recruited and hired to fill six (6) full time positions required to meet project needs including:

- Alan Gunderson, Senior Software Developer
- Weiping Li, GIS Specialist III/IV
- Shengfang Ma, Senior Software Developer
- Sravya Nallamaddi, Senior Software Developer
- Tim Propst, Science Coordinator/Technical Writer
- Lei Qiao, Computer Modeler
- Yan Song, Senior Software Developer
- Kyle Stoltz, Senior Software Developer
- Mark Witte, Senior Software Developer

In addition to the six full-time team members, OSU students (Ethan Boyer, Sara Howard, Adrian Toquothty, Patrick Joyce, Lily Schneberger Sawatzky and Gloria Evans) provided support for project activities including data management, coding, communications, and field work, while receiving valuable training for their future careers.

Additionally, two graduate students were partially supported by the project. Sylvester Thompson, under the direction of Dr. Jaime Schussler (Assistant Professor, OSU School of Civil & Environmental Engineering), is conducting research on smart stormwater sensing at the AHERU lab. Adeyinka Ogunbajo, under the supervision of Dr. Jeff Sadler (Assistant Professor, OSU Biosystems and Agricultural Engineering Department), is working to develop a framework for integrating diverse hydrologic variables into a machine learning (ML) rainfall-runoff model to evaluate the impact of storm events on reservoirs. DAM-TAGS project funding ended before these students were able to complete their research. Dr. Jaime Schussler is using her startup funds to support the completion of Sylvester Thompson's research. Dr. Kevin Wagner is using funding from his Thomas E. Berry Professorship to support the completion of Adeyinka Ogunbajo's research.

Team Meeting Coordination

The Water Center Director and USDA-ARS AHERU Research Leader held regular meetings to review progress, address challenges, present needs, and share collaboration opportunities. Approximately every other month, the entire team joined the discussion. Additionally, special topic meetings were ~~also~~ called as needed.

Fiscal Management

Fiscal management of the agreement, including budget modifications and revisions, was provided by the OSU Division of Agriculture and Natural Resources Sponsored Programs office and the Water Center.

Project Management

Significant progress was made by the team in supporting USDA-ARS and NRCS objectives of modernizing data management, dam inspection, dam monitoring, and dam rehabilitation. The following summarizes key activities and accomplishments for each objective outlined in the Cooperative Agreement.

Writing and revising guidance documents

Online protocols.io Platform Documents Data Management Workflow

The USDA-ARS Partnerships for Data Innovations (PDI) team has created a protocols.io workspace on the platform. One of its uses is housing a collection of developmental protocols that will be compiled into a handbook to document the workflows by which PDI obtains, processes, ingests, analyzes, and shares all data that is part of their environment. This handbook helps standardize the flow of data through the PDI environment, improving both the efficiency and efficacy of PDI operations.

Last year, membership in the ARS workspace grew to 300 researchers. This included scientists from the Long-Term Agricultural Research (LTAR) group, who used it to publish a series of 40 standard methods for soil, air, and water sampling. In addition, the protocols.io executive team initiated the FedRAMP approval process, which will help ensure the continued use of this resource within the federal system.

One of the major impacts of the platform will be increased interoperability among disparate data streams, encouraging unparalleled collaboration among researchers from various disciplines.

Design Chapters in Development for NRCS National Engineering Handbook

In response to comments from NRCS reviewers on a second draft submitted last year, OWRC and ARS staff are updating example problems for the new Stepped Spillway design chapter. Once these changes are completed and approved, this resource will be ready for publication.

An update of Design Note No. 6, “Riprap Lined Plunge Pool for Cantilever Outlet”, was submitted last year and received extensive comments from NRCS reviewers. A comprehensive revision of this draft was initiated in response to this feedback and will be resubmitted in early 2025.

The team also submitted a revised Technical Release No. 70, “Hydraulic Proportioning of Two-Way Covered Baffle Inlet Riser,” in Summer 2024 and is currently awaiting feedback.

These resources will provide clear, comprehensive guidance, with up-to-date information, for the design process of various dam components. They will greatly impact the efficiency and effectiveness with which engineers and contractors are able to develop and build these structures.

Developing a cloud-based database network for reservoir ~~&~~and levee monitoring

General Cloud Administration

For the duration of the project, Dr. Kyle Stoltz has acted as the PDI cloud engineer, architect, and administrator. These roles required the building, maintaining, and securing of all PDI cloud resources in the Azure Commercial Environment. This entailed continual maintenance activities, as well as administrative responsibilities, e.g., account creation and management, for both Azure resources and ArcGIS Enterprise. Dr. Stoltz also performed complete updates of the PDI ArcGIS Enterprise software suite, which included new functionality upgrades for ArcGIS Knowledge, Apps, and other tools.

Growing the PDI Cloud Environment

The Fleet Management System (FMS) was developed within the PDI Enterprise to enable the collection, processing, and management of Unmanned Aerial Vehicle (UAV) imagery while also capturing UAV flight information as metadata for the images. FMS includes a Survey123 form that allows pilots to submit flight information to the cloud, which is later attached to the processed UAV products. A standardized workflow was developed to streamline the upload, processing, and management of UAV images. This enables pilots to upload raw UAV images to the PDI cloud, select those of interest, and have them automatically processed in SiteScan. The final processed products—including Orthomosaics, Digital Terrain Models (DTMs), Digital Surface Models (DSMs), Point Clouds, and 3D meshes— are then migrated to the PDI Enterprise system. This process ensures the standardization of processed UAV imagery across pilots and guarantees that users and researchers can efficiently access the data. Hopefully, additional funding will be made available to implement these established pathways into a UAV-based dam monitoring network as originally envisioned.

Soil Carbon Database

PDI helped CRP (Conservation Reserve Program) scientists initiate a data management solution that integrated data collection, submission, transfer, and storage components associated with Soil Carbon Measurement ~~&~~and Evaluation (SCME) research. Project efforts focused on developing an Azure cloud-based PostgreSQL database to meet these needs. In addition, Survey123-based data ingestion tools were created for both the Forestry and Wetland groups to help streamline their annual field collection processes. These apps replace paper forms and manual data entry workflows, greatly improving field work efficiency and data analysis capabilities.

Unfortunately, multiple factors, including loss of personnel and program shifts within PDI, delayed the field apps development such that they were not ready in time to meet the seasonal schedule for the year 1 or 2 soil samplings. Many of the scientists lost interest, resulting in a loss of participation that exacerbated the slow progress. In year 3, the scientists simply chose not to utilize the apps. As a result, the complete data catalogue was collected using non-standardized methodology, and comprises multiple, independent datasets. These included many new measured variables and alternate naming conventions that were not included in the data dictionary originally developed by the PDI team, in effect bringing the single, interoperable database creation effort to a standstill.

However, recent efforts to revive this task have been bolstered by a proposed collaboration with NAL that will focus on using the NALT concept space as a controlled vocabulary to help harmonize the various data collections. This effort will take advantage of preliminary groundwork for a soils research data schema that has already been laid (see NSF Soil Organic Carbon Knowledge Graph section in this Final Report.). Additional testing and extensive revisions have also been made as part of the process to integrate the database with outputs from Esri Survey123 forms (please see the Survey123 Data Transfer Tool section in this Final Report).

Hopefully, the completion of these efforts will be realized through extension of the DAM-TAGS project goals and funding, resulting in a comprehensive data management platform to assist CRP researchers and scientists. In addition, the pathways developed by this work, especially the integrations with Survey123 forms and data transferring APIs, will serve as templates to be scaled or transformed for use in additional research by PDI, ARS, and USDA.

PDI ~~Used~~ Utilization of Esri Knowledge and NALT to Update Database Infrastructure and Provide Cutting Edge Technology to End Users

PDI worked to integrate Esri's Knowledge Graph database system with the National Agricultural Library Terms (NALT) concept space. This project's intent was to lay a foundation for data upload, analysis, and retrieval unlike any other within ARS. The Knowledge environment includes graphical data analysis tools that have industry-changing potential. NALT provides a framework for developing ontologies and metadata that maximize machine learning capabilities and data retrieval.

The past year saw numerous changes in the NALT reconciliation workflow. Recent improvements identified a pathway that allows multiple concept definitions to be appended to a single NALT entry. This means definitions can be uploaded directly from accepted glossaries to the NALT concept space, removing the need, and intensive labor required, for subject matter experts to craft a single definition. In response, through a fellowship program with NAL during Summer 2024, two University of Maryland students searched the web for USDA publications, and other established documents (i.e., professional society resources), for accepted definitions. They catalogued the contents from approximately fifty (50) glossaries that will be utilized to broaden the scope of NALT. This effort is ongoing and has now doubled the number of glossaries included in this repository.

The NALT resource includes a unique, machine-readable Universal Resource Identifier (URI) for every concept in the platform. Creating a more robust NALT resource streamlines the process by which researchers can label their data with the appropriate URI. Once uploaded into the Esri Knowledge-based PDI data environment, these URI will be the connection points that make the entire resource interoperable. First steps for this process were developed, also in Summer 2024, as part of an NSF-funded project investigating the efficacy of the voluntary carbon credit market on increasing soil organic carbon levels. As part of this collaboration between University of Texas-Arlington (UTA) computer scientists and USDA soil scientists, PDI worked to integrate the project's data dictionary with NALT. The resulting pathway was a pioneering effort that combined ontological semantic language processing with review by the data domain experts to obtain a comprehensive overview of the data. These efforts will greatly inform the workflow and configuration of the PDI Esri Knowledge-based data environment.

Additional build-out of the PDI Esri Knowledge resource was the subject of an effort utilizing data from the PDI Pollinator Data Hub project during Summer 2024. With the ultimate goal of creating an online resource for end users seeking information on pollinator habitat improvement indices, this pilot project walked through every aspect involved in creating a data portal based on a knowledge graph. Resources left unused from other PDI efforts were redirected to enlist Esri experts to establish the needed applications and pathways, with the added directive of training the PDI tech team on their design and upkeep along the way. In addition to the portal, this task included various functionalities including data ingestion, knowledge graph creation, query formulation, and metadata documentation. A basic scope of work was created in collaboration with the Esri team and efforts to delineate the specific endpoints required and their associated timelines.

The Pollinator Knowledge Graph, and associated end user portal, was completed in August 2024. Although it is a useful resource, unforeseen difficulties in obtaining and formatting the foundational data for the knowledge graph reduced the time and resources available for the NALT integration piece, one of the project's goals. However, the lessons learned from this pilot project are being applied to the current effort underway to create an Esri Knowledge-based knowledge graph of the soil organic carbon data mentioned above.

COTMAN (COTton MANagement)

This large (over 85,000 lines of code) C++ computer program was developed for use by cotton researchers and cotton farmers. Initially written in 1997-1999, and last updated approximately 20 years ago, the resource needed major updates. PDI partnered with the COTMAN team to make the necessary improvements through a series of four, stepped software versions. The original plan called for COTMAN Version 3.1 to fix browser compatibility issues and convert it to a PDI-compatible database engine; 3.2 to replace very old, specialized handheld field data collection devices (that are currently far beyond their end-of-life) with Apple iPad data entry using the Esri Survey~~r~~123 data collection tool; 3.3 ~~would~~ to integrate temperature and weather data available on the PDI Azure cloud into COTMAN; and 3.4 ~~would~~ to move the standalone COTMAN Microsoft Windows application to the PDI Azure cloud environment.

The team successfully met the goals designated for version 3.1~~;~~ - the browser compatibility program issue was fixed, and the software is now working with PostgreSQL (the PDI compatible database engine~~.~~). However, the group went back to using Microsoft Access for current software delivery, as COTMAN users, for the most part, already have that software on their computers and installation of COTMAN with a local PostgreSQL database proved too complicated. As an alternative, the team also considered using an SQLite database, but COTMAN uses ODBC (Object Database Connectivity) for database interaction and the lowest-cost ODBC connector for SQLite is $175~~.00~~. This is cost-prohibitive to most COTMAN users, so SQLite was not a viable option.

To accommodate the continued use of Microsoft Access, COTMAN was updated to run on both Windows 10 and 11. The data layer of the application was generalized so it is a simple matter to change among several database platforms, namely Microsoft Access and PostgreSQL. In addition, the team found issues with the reporting, export, and import functions of the application; these were corrected and are all now fully operational. These improvements were made available through several updates to the 3.1 branch of the COTMAN software. (The final version, 3.2.0, released on 10-10-2024, included all the transformations listed below.)

As work progressed, it became apparent that additional course corrections to the original plan were needed. For instance, the team determined that most COTMAN users utilize laptop computers in the field and often do not have internet connectivity. Furthermore, there is no funding or interest among users to switch to iPads or have instantaneous cloud-based performance. These findings removed the need to develop a version of the software for Esri's Survey123 (v3.2) or move it to the Azure environment (v3.4).

As a result, the effort to integrate automatic weather updates into the software, originally slated as COTMAN 3.3, was re-designated as COTMAN 3.2, and moved up in priority. A PDI student worker from Colorado State University (who continued as a term employee after graduation) was tasked with developing an extension to the COTMAN software that could pull weather data into the program.

A major accomplishment of the COTMAN project in 2024 was the connection of this extension to the Environmental Data Hub, another resource developed by PDI as part of a separate effort. Integration with this data repository allows the COTMAN application to obtain daily weather data (high and low temperatures) for each cotton field directly from a nearby weather station. Previously, this could only be accomplished via manual data entry, a task that was not only tedious but also prone to error.

Instead, the extension upgrade allows the COTMAN software to automatically pull the previous day's temperatures from the weather station and preload them into the program on the day of field data collection. The weather information is then included in the Bollman calculations that cotton farmers utilize to make management decisions regarding irrigation rates and pesticide application. This improves the accuracy of the calculations, increases the efficiency of the farmer's operations, and helps preserve water resources and reduce pollution from unnecessary pesticide application.

In addition to the benefits of this effort to the cotton industry, the creation of a workflow that successfully integrates weather data from the PDI Cloud into a desktop computer program could have far-reaching effects. The team is hopeful that additional funding will be made available to develop applications for its use in dam monitoring efforts.

Developing an application programming interface (API) to transfer data collected to collaborators ~~&~~and stakeholders

AMS Plant Variety API and Dashboard

The PDI team worked with USDA Agriculture Marketing Service (AMS) to create a web application for their Plant Variety Protection Office (PVPO). AMS researchers, plant breeders, farmers, research scientists, and the public can use the PVPO Online Dashboard (POD) to access information on all plant varieties in the PVPO database with issued, or expired, Plant Variety Protection Certificates from the year 2000 forward. The dashboard, hosted on the PDI Azure cloud, allows users to search the PVPO database for plant variety names, characteristics, and worldwide areas of adaptation. Use the https://pvpo.scinet.usda.gov/ link to visit.

The plant variety data are provided and updated monthly by USDA-AMS in CSV format through GitHub. Personnel changes during the last year transferred responsibility of the PVPO data to a new member of the AMS team. PDI helped during the transition by setting up a new PostgreSQL database, transferring data from CSV files into the database, and fixing issues during loading. Documented scripts and steps for future database configuration were shared in the Year 3 Annual Report.

Establishing the framework for a dam monitoring network including software development for data collection, management, and transfer

PDI IoT Hub Advancements

After developing the first version of the PDI IoT Hub, Dr. Stoltz added processes that push data not only to an ArcGIS GeoEvent server, but also to the PDI PostgreSQL server and to cold storage for archiving. With the raw data located in the PDI PostgreSQL server, unique QA/QC can be performed, as well as leveraging it for machine learning algorithms. An email alert system was implemented that actively monitors connected devices and incoming data and will send out automated alerts if anomalous activity is detected.

This project demonstrates the ability to obtain, catalogue, and monitor incoming data from continuous monitoring sensors; a key feature in the comprehensive dam monitoring network that was one of the DAM-TAGS project's goals. Hopefully, additional funding will become available to allow the team to fully implement this capability as originally envisioned.

Small Watershed Reservoir/Lake Water Level Predictions

The progress of this work since 2021 was chronicled in the three annual project reports. The most recent progress includes training of deep learning models (LSTM, BiLSTM, and BiGRU) and many rounds of intensive revisions. A final version is represented in a published manuscript about this effort:

> *Lei Qiao, Daniel Livsey, Jarrett Wise, Kem Kadavy, Sherry Hunt, Kevin Wagner. Predicting flood stages in watersheds with different scales using hourly rainfall dataset: A high-volume rainfall features empowered machine learning approach, Science of The Total Environment, Volume 950, 2024, 175231, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2024.175231.*

The methods, applications and predictive accuracies of the machine learning models (e.g., random forest, support vector machine, and XGBoost) can be found in the manuscript or from the information provided in the DAM-TAGS Year 3 Report. These results were very promising and gave accuracy at the highest level the best model can reach. These findings were also shared with experts and users in the field at the 2024 Oklahoma Governor's Water Conference.

LIMS (Laboratory Information Management System)

LIMS is a general term for a comprehensive software program that helps laboratories document and manage their overall operations. These can significantly improve (by an estimated 5-9x) the efficiency of laboratory testing workflows. The USDA-ARS Pasture Systems & Watershed Management Research Unit in University Park, PA, reached out to PDI for assistance in implementing a LIMS solution for their chemistry lab, which conducts extensive soil and water sample analyses for scientists from ARS, other USDA agencies, and collaborating partners. PDI worked with Principal Investigator and ARS Research Chemist Kyle Elkin to complete this task.

NOTE: As previously stated in the Year 3 annual report, the original vision of this effort included using the output as a template for creating an agency-wide LIMS program that would use the Internet of

Things (IoT) to connect equipment between ARS laboratories. Such a platform would offer multiple advantages, like increased interoperability between labs, shared updates, standardized workflows, etc. However, this goal had to be dropped when the team discovered that many of the devices in question utilize older software, or operating systems, that are either impossible, or extremely difficult, to adapt for online usage. Thus, there is no discussion of the proposed IoT solution in this Final Report.

The open-source SENAITE system was selected as the base LIMS for the project. This system catalogs samples as they enter the laboratory, splits them into testing racks of up to 100 tests, and sends the racks to the various testing machines within the lab. Once the specified tests are run, results are sent back to the LIMS from the machines, where they are compiled and made available for review by lab personnel. Upon approval, the LIMS is used to deliver results to the clients.

The team customized this software to export sample batches out of the LIMS system to several chemistry testing systems used in the ARS lab. They worked with an outside contractor to build the first interfaces for the ICP OES testing system. In 2024, a second interface for the ThermoFisherScientifc Gallery 862 was also completed. This discrete analyzer is used to measure multiple indices in both water and soil samples, including ammonium (NH_4N) and vanadium nitrate (NO_3N) levels in soil.

Although the Gallery 862 can process bulk racks of samples at once, considerable additional processing is required to transform these test results into applicable information. Previously, this required the USDA chemistry labs to manually transfer the data into Microsoft Excel spreadsheets for thousands of water and soil samples each day. The new LIMS from PDI has automated this workflow, saving chemists multiple hours per week. As the testing of soil samples is critical to understanding the soil composition of dams, the efficiency of the new LIMS system could have direct implications for dam monitoring and rehabilitation efforts. The team would welcome the opportunity to put this solution to work in a reimagined DAM-TAGS project should funding become available.

Forage Data Hub (FDH) Establishes Centralized Data Repository for Pastures and Grasslands

Co-PIs and collaborators from approximately fifty (50) institutions are assembling a historical collection of forage crop and animal livestock data into an online, cloud-based resource. The hub includes three hierarchical levels (annual forage; perennial pasture; and integrated systems) by USDA-NRCS Major Land Resource Area (MLRA). The database includes annual vs Diverse Perennial Circular Systems (DPCS), such as long-term grazing systems research data (soil, plant, and animal). This effort 1) combines existing databases/networks to develop a framework for integrating multiple data sources to address systems (GxExMxP) research; 2) advances data visualization tools through statistical inferences and decision support tool development; and 3) develops a cloud-based computing network for scientists to integrate into diverse research programs across multiple spatial and temporal scales (“off ramp”).

PDI technical personnel worked with colleagues from University of Texas-Arlington (UTA) to create this resource. The team completed a viable data model and data ingestion workflow which is being used to inform the configuration utilized in constructing the user interface. The UTA team is building out the data ingestion side of the resource using Django software. PDI continues to work closely with them to seamlessly integrate the output side of the Esri Knowledge hub that they are creating, which will allow users to retrieve data and review various analyses.

This was a novel approach, as this is one of the first attempts to utilize the new Esri Knowledge-based infrastructure. Ingested data are not simply stored in the cloud, but are also joined into the PDI knowledge graph, "powering" this resource by the analytics only these structures make possible.

The PDI team continues to finalize an earlier draft of the FDH knowledge graph and has developed connections to useful supplemental data sources (e.g., SSURGO soils, USEPA Ecoregion maps, etc.) using the Spatial Join functionality of ArcGIS Pro. They also demonstrated the utility of the Esri Knowledge Data Interoperability extension for dataset management and manipulations. Unfortunately, the unprecedented changes in the federal system in general and in the DAM-TAGS project specifically, resulted in losses of funding and personnel that severely hampered the completion of this task. The remaining members of the PDI and DAM-TAGS teams are attempting to salvage as much of the work as they can, given the extraordinary circumstances. Hopefully, the lessons learned will be directly applicable to the FDH build, and by extension, to every other data ingestion project undertaken by PDI.

Developing GIS based tools for dam monitoring and inspection

Survey123 Data Transfer Tool

ArcGIS Survey123 software from Esri allows the user to create digital ingestion tools (i.e., online forms, phone apps) that store data in the ArcGIS OnLine (AGOL) cloud-based platform. PDI has utilized this tool for a wide variety of projects. Data collected using Survey123 can be saved and downloaded in multiple formats, including CSV, Excel, Shapefile, File Geodatabase, GeoJson, and several others, in a local computer. Unfortunately, relational Azure PostgreSQL databases, which PDI utilizes for analysis and processing purposes, are not one of the default options for saving S~~s~~urvey123 data and require a complicated, time-consuming conversion process. This project sought to develop a customized command, or hyperlink, that would simplify data ingestion from Survey123 into PostgreSQL.

The PDI team working on the CRP Soil Carbon Measurement and Evaluation project (SCME) undertook the development of this data transfer tool. They wanted to streamline the integration process for data from Survey123 instruments utilized in their study into the PDI infrastructure. Two different approaches to achieve this goal were considered: one was to use the Esri-contracted FME server that is embedded in the Esri Knowledge Data Interoperability extension, and the other was to use Python script to create a menu command, or hyperlink, for downloading data with a specific survey ID. Since detailed instructions on the use of the Esri FME server option were not available until after the Pollinator Pilot project (please see report section above), PDI team members focused on the second option and began developing the appropriate Python application to add a hyperlink to any Survey123 form.

This seemed feasible as Python script is already utilized within the AGOL environment for connecting to AGOL accounts, downloading feature maps, and saving them to a local PC. The team set out to determine whether a similar Python solution could be developed for fetching survey data. Unfortunately, the cessation of the DAM-TAGS project, as well as program shifts within PDI, precluded them from completing this task.

Should funding become available to extend the vision of the DAM-TAGS project, it is likely that a Survey123 tool would be one of the main goals (see Dam Inspection Survey Tool section of this report.) The development of a Survey123 Data Transfer tool should parallel that effort in order to take full advantage of the dam monitoring data it collects. However, the foundational work done by the team on the current task identified three specific needs that must be met to adequately test the utility of the proposed Python transfer tool.

First, the team will require access to published survey data for testing. It would be best for this to be a pre-existing dataset, so developers do not have to delay their work until a proposed Survey123 tool is operational. Second, the script development team members must have Survey123 "Creator" and "Publisher" roles associated with their AGOL profiles to successfully implement their programming changes. Lastly, a "duplicate" of the Survey123 database and tables must be configured within the PDI PostgreSQL environment to allow the team to adequately test the capability of the Python script to transfer Survey123 data between a local PC, PostgreSQL database, and the AGOL environment.

In addition, PDI is gaining momentum in utilizing the Esri Data Interoperability extension, so a thorough examination of the suitability of this tool for Survey123 data transfer could also be a part of future work. The team is hopeful that they will be able to explore both ~~of these~~ possibilities in the near future.

Cloud-based Safety Inspection Tools Facilitate Creation of Inspection Reports and Track Hazardous Finding Trends

PDI worked with the USDA-ARS Safety, Health, and Environmental Management Branch (SHEMB) to develop a suite of tools for use by all REE safety officers and their proxies while conducting safety inspections. This included a Survey123-based form to document specific location information, record photos, note the regulating standard, and complete a written description of hazard findings that can then be uploaded to the cloud. It replaced a pen and paper methodology, helped standardize reports and reduced the need for post-inspection data entry. A dashboard was also created that allows inspectors to review their findings, document follow-up efforts, and generate official reports for distribution to responsible persons. Program leaders utilize the dashboard to monitor reports agency-wide, providing them the ability to detect national trends and inform their decision-making process (applicable to goal #7 below). A StoryMap User Guide is available as a reference as inspectors become familiar with the new tools.

An update from the SHEMB collaborator indicated, "The tool has improved ARS-wide safety inspection programs, with estimated time savings of months of analysis at the Branch level (HQ). All ARS Locations are impacted by expedited and better organized inspection reports." These efforts can serve as a template for developing similar tools to be utilized in dam monitoring and inspection efforts.

Rangeland Analysis Platform (RAP)

This innovative online tool, Rangeland Analysis Platform (rangelands.app), was first developed by the University of Montana in partnership with Working Lands for Wildlife of the USDA's Natural Resources Conservation Service, the Bureau of Land Management, and the USDA ARS. The application allows users to analyze more than three decades of United States rangeland vegetation data with an interactive mapping tool.

According to reports, "researchers from ARS and the University of Montana used RAP technology to help land managers assess vegetation trends on public lands and to assist private ranchers in maximizing their grass production. Other ARS and collaborator improvements to RAP have resulted in new vegetation models that allow land managers to forecast location-specific wildfire risk across the Great Basin." (USDA-ARS Now Stewarding the Rangeland Analysis Platform : USDA ARS)

RAP is based on the Google Earth engine and utilizes satellite imagery to create long-term datasets on vegetation cover and production. Overall management and hosting of the tool was transferred to USDA-ARS, with PDI playing a critical role in making this transition, including updating Angular, the web application framework on which RAP is built, from version 8 to version 13. The experience and knowledge gained during this process expanded the team's toolkit, broadening the scope of resources that they can utilize to create useful data tools for researchers.

Dam Inspection Survey Tool

Early in the DAM-TAGS project, PDI helped create a draft Survey123 solution for documenting dam inspections to replace existing NRCS pen and paper methodology. During the process, questions regarding the specific details of how this new tool would be implemented were raised that required review and approval from NRCS. Unfortunately, at that exact time, there was significant turnover within the NRCS administrative personnel group, and these questions were never fully answered, halting development of the tool.

With the advent of new hires by NRCS, Dr. Sherry Hunt (Research Leader, USDA-ARS AHERU) indicated that it was time to revive this task. The existing tool files were assigned to a new owner within PDI, as the original creator had left the organization. The new technical lead completed a review of the file content and functionality and found all was in good order. He was also able to connect the survey to the US Army Corps of Engineers National Inventory of Dams (NID). By way of a drop-down menu, users select the dam of interest, and the survey pulls directly from the NID to autopopulate fields with the appropriate information. This improvement saves time and reduces the potential for errors in manual data entry.

PDI was planning to reconvene a working group of NRCS and ARS experts to review the tool, offer input on the suitability of the potential improvements, provide suggestions for any additional changes, and finally, to reach consensus on the implementation questions that stalled the initial project. Unfortunately, the cessation of the DAM-TAGS funding, as well as ~~the mass firings~~ personnel losses in the federal system, have resulted in the stoppage of this effort. Hopefully, additional funding will be made available to move this work forward.

Developing and delivering training curriculum in various formats

Cattle Fever Tick Eradication Program (CFTEP) and PDI Use of Technological Improvements to Streamline Tasks and Improve Information Flow

PDI and partners in ARS, APHIS-Vet Services, and the Texas Animal Health Commission, developed multiple data input and management tools to maximize CFTEP efforts. For instance, pen and paper methodology for reporting new tick quarantine areas was replaced with the Esri FieldMaps application. This upgrade decreases the time required for creating accurate maps from weeks to mere minutes, ensuring that administration and policy makers have up-to-date information at their fingertips. PDI is continuing to work with technology vendors to develop a durable Bluetooth-enabled counter for auto populating Survey123 fields; a tool that will further streamline CFTEP data collection, with broad application potential beyond this program alone. Both of these examples could be revised and integrated into an overall dam monitoring procedure if additional funds are made available to continue the efforts cut short by termination of the DAM-TAGS project.

ARS and PDI are Development of ~~ing~~ Web-based WinDAM and SITES Training to Improve Information Flow and Curriculum

Team efforts are underway to create transcripts from videos taken of the *Using SITES and WinDAM to Hydraulically Proportion 378-Ponds and TR-60/402-Dams* training held at the Stillwater AHERU location. Subject to review by ARS and NRCS technical experts, the transcripts, videos and presented materials will be integrated into a comprehensive training curriculum for these software programs. Once the experts determine the presented information and implementation steps are accurately and effectively communicated, the goal is to share the curriculum online. Although loss of the DAM-TAGS funding severely impacted the timeline, the team is continuing to pursue this goal. Additional financial support would be a significant boost to finalizing this curriculum as an online resource. As originally envisioned, this would make the information more easily accessible to a broader audience, while lessening the costs for travel and time required by in-person training.

Developing a decision support informatics platform

DAM-TAGS Web Apps Development

This PDI project focused directly on the development of web-based applications that will aid hydraulic engineers in the design of rock chutes, plunge basins ~~&~~and stepped spillways. This effort is an extension of the DAM-TAGS project to develop new and improved software for data collection, processing, management, and transfer. It supplements the work of DAM-TAGS personnel who have been updating and developing chapters for the NRCS National Engineering Handbook. These documents include various design processes that require step-by-step completion of a series of equations involving multiple hydraulic parameters. These workflows are tedious and repetitive. Although NRCS does have some Excel spreadsheets with built-in formulas available to help designers complete the necessary calculations more quickly, there is still room for improvement.

Current efforts are developing web-based applications to help streamline the hydraulic design process. These web apps will be made available online for more widespread use, avoiding the need for emailing spreadsheets back and forth to various users. State ~~or~~ and federal transportation, dam safety, and/or consulting engineers, for example, might find this suite of tools quite helpful for their work. The result will increase efficiency, availability, and standardization, not only across NRCS, but wherever needed by hydraulic designers.

The first web application addressed by the team is focused on rock chutes design. The app mirrors the existing Excel sheet with three webpages: a design sheet, plan sheet, and instruction sheet. These frontend components are coded in HTML, while C# and Python codes were utilized for backend development. Functional integration between the two is nearly complete, while final aesthetic adjustments are currently being addressed.

The hydraulic computations for the inlet channel, chute, and outlet channel were completed by various means. The bisection method was used for one dimensional equation, while the Newton method was used for 2D/3D equations. Comparison of different optimization methods for submerged chute conditions included Sequential Least Square Programming, Constrained Optimization by Linear Approximation (COBYLA), and other methods.

Both the front-end and back-end of the Rock Chute web application were developed, integrated, and published using Visual Studio Razor ASP.net. After the equations for calculating all variables were developed in Python and converted into C#, the functions were merged and tested in ASP.net. User input parameters were passed from front-end pages into corresponding back-end functions and outputs passed back to display in proper front-end boxes, with careful testing of the input and output for each variable. Some special coding work was done to allow for the Tailwater value to either be input by the user or calculated by the program.

After all outputs are correct and in place, validation functions for each user input parameter will be added so that messages about variable ranges can be provided to users. The application will also have PDF conversion functionality to provide printable and savable output of all three major pages (instructions, design sheet and plan sheet), as well as technical and user documentation. Finally, the coding team will work with the PDI Cloud Engineer to house the Rock Chute web application in the Azure cloud environment. The lessons learned during the development of this tool will guide the ~~formation~~ development of the plunge basin and stepped spillways web apps.

Esri Knowledge Training

As mentioned above, additional training on the Esri Knowledge platform was part of the Pollinator Pilot project. PDI stipulated that Esri include a knowledge transfer component as a part of this effort. During the project, separate “tech-only” sessions were held to demonstrate the behind-the-scenes workflows utilized by the Esri experts to develop the knowledge graph. Esri provided a PDF resource to summarize foundational aspects and a PDI employee documented highlights of the training sessions on protocols.io.

Balance App ~~Provides~~ for Inexpensive, Automated Sample Weighing ~~System to~~and Streamlined Lab Processing

Based on QR label technology, the PDI team created software that allows samples to be scanned, weighed, and logged in with a simple step, reducing processing time by 70% or more. This increases productivity and data accuracy, while also helping free up scientist and technician time. The PDI team designed this package to use relatively inexpensive, off-the-shelf devices, integrated with their own app and software developed in-house, to keep costs down. The apps are approved for use on government devices, and training videos are available. This serves as a blueprint to help establish a successful training program for educational materials developed under DAM-TAGS workflows.

PDI received funds from additional sources that were used to purchase QR label printers and readers for distribution to ARS researchers interested in implementing the software. A Storymap User Guide was developed and introductory training sessions held with a few research units. However, personnel and institutional changes within PDI resulted in a loss of momentum on this rollout.

Despite this, efforts to expand the software to include Sartorius brand scales were completed during Summer 2024, so it is now compatible with models from that manufacturer, as well as Mettler-Toledo and OHAUS. The team also created a workspace on protocols.io to provide a single source for downloading the app and associated labeling software, along with step-by-step installation and operation instructions. These resources stand ready whenever the PDI administration identifies a pathway to initiate a renewed distribution program.

~~The~~ USDA-ARS Website Enhancements to Provide Greater~~s~~ Information for Partners and Stakeholders ~~While~~ and Highlight~~ing~~ Collaborative Efforts

USDA-ARS locations in Oklahoma underwent a reorganization during the project period, resulting in the creation of the Oklahoma and Central Plains Agricultural Research Center. This consolidation brought all four ARS locations in Oklahoma under a single administrative umbrella. As part of these changes, climatology research efforts from El Reno, OK, were re~~-~~aligned with hydraulic engineering research in Stillwater, OK, into a single Agroclimate & Hydraulic Engineering Research Unit (AHERU), under the supervision of Dr. Sherry Hunt.

DAM-TAGS personnel promoted the new AHERU last year with the introduction of a new webpage (https://www.ars.usda.gov/plains-area/el-reno-ok/ocparc/agroclimate-and-hydraulics-research-unit/.) This resource includes links for outside parties to request tours of the Stillwater or El Reno lab locations, plus outreach materials for stakeholders, educators, and the public. It also features research highlights, history, and a list of ongoing projects.

OWRC Communications Department Outreach

During the reporting period of March 1, 2021 – October 31, 2024, the Oklahoma Water Resources Center Communications team posted on social media (Facebook and Twitter) about the DAM-TAGS partnership 67 times with a total reach of 22,391 people. By comparison, in 2022, education information about our research updates and new utilization of technology was only posted on social media 11 times for a total reach of over 4,252 people.

For Dam Safety Awareness Day in May 2023, the OWRC, in partnership with ARS, launched a social media campaign that consisted of educational information about dam research, updates, and new utilization of technology including 7 posts, reaching 6,428 people. The DAM-TAGS research collaboration has appeared in two TV segments aired on OETA-TV and subsequent YouTube channels with a viewership of 64,000 and YouTube views of 1,647. There have been five editorial articles written on the projects, people, strategies and goals of DAM-TAGS with a total readership of over 1,300 people. Webpages with project objectives, updates, and contact information are hosted on the websites of both the OWRC (https://water.okstate.edu/our-work/dam-analysis.html; 283 visits) and the AHERU (https://www.ars.usda.gov/research/project/?accnNo=439722). Research updates were also sent out several times in the Water Center's quarterly newsletter with a subscriber base of over 1,700 people.

With the hiring of a new Communicaitons and Outreach Specialist in 2023, there were significant updates to the OWRC communications plan which was adapted to reach a larger and more diverse audience. In particular, a dedicated focus to continuous improvement on social media reach was emphasized, resulting in a significant increase in the second quarter of FY2025. Additional funding for continued communication efforts would allow the true potential of the most recent efforts to be seen.

Appendix A: Exploring Smart Stormwater as a Solution to Oklahoma Water Management

Submitted by Jaime Schussler

Summary

The USDA Small Watershed Programs have helped create earthen dams, which are critical infrastructure to manage stormwater runoff and prevent downstream flooding. However, more than half of these dams have surpassed their design service life. In this time, climate change has induced more intense and frequent events and land use changes have landlocked many of these dams. Increased runoff and sedimentation have put a higher demand on the statically designed infrastructure. "Smart" stormwater systems provide a potential solution to dynamically adjust to the climate and anthropogenic-induced changes. Smart stormwater systems use Real-Time Control (RTC) to adjust spillway conditions based on a series of sensors and actuators. This technology provides an opportunity to optimize detention systems; however, it is an emerging field. Our study aims to evaluate the feasibility of retrofitting earthen dams in Oklahoma with smart stormwater systems through a systematic literature review, field-data collection, and large-scale testing. In January 2024, we initiated a systematic literature review. Our preliminary review established that there is a need for controlled- and field- scale studies to help address several of the technical and socio-economic barriers that exist for smart stormwater implementation. As we complete our literature review, we will be working toward establishing a field-campaign for data collection and developing collaborations as we work toward an experimental testing regimen.

Statement of Critical Need for Oklahoma

Since the mid 1900's, more than 2,000 earthen dams have been constructed in Oklahoma under the USDA Small Watershed Programs (OCC, 2024). These dams provide storage volume and flow attenuation for stormwater runoff during wet-weather events. These dams are not only critical infrastructure that provide downstream protection against flooding and erosion, but also provide ecosystem services. The discharge from the dam impoundments is governed by a static engineered principal spillway that is most typically a pipe-drop drainage structure. If flows exceed the outlet capacity, then an earthen auxiliary spillway will be activated.

More than half of the earthen dams in Oklahoma have surpassed their design service life of 50 years (DOI, 1987). In the elapsed time, we have started to experience more intense and frequent storm events due to climate change (Wuebbles et al., 2017). Additionally, we have intensified our land use, through urbanization and production, which has landlocked our impoundments in many cases. The combination of climate and land use changes has yielded higher stormwater runoff and sediment loads. Increased sedimentation within the impoundment consumes some of the available storage volume. These changes increase the likelihood of the auxiliary spillway being activated for a storm event where it historically was not needed. Spillway-related drivers, such as spillway deficiency, erosion, pipe, riser or chute failures are major contributors to dam incidents and failures. According to the ASDSO Dam Incident Database (1900-2023) spillway-related primary drivers were 22% of the recorded incidents. Of the spillway-related incidents, 19% resulted in dam failure. In severe cases of inadequate spillways or intensified wet-weather events, dams may also have an increased risk of overtopping. Overtopping contributed to another 30% of total recorded incidents and 45% of those resulted in failure (ASDSO, 2024). There is a critical need to rehabilitate and retrofit our dam spillways to be more adaptive to the climate and anthropogenic changes we are experiencing in these small watersheds.

Smart Stormwater Systems

In recent years, “smart” stormwater systems have been researched to dynamically adjust to real-time storm and flow events. These stormwater systems involve pairing water quality and quantity sensors, within a single impoundment or across watershed, with actuators on an outlet to enable real-time control (RTC) (Kerkez et al., 2016, Shishegar et al., 2021). RTC allows the outlets to dynamically adjust to sensor-read and analyzed data and may even incorporate forecasted precipitation events to be more adaptive to climate and anthropogenic- induced changes (Mullapudi et al., 2016). If a storm is forecasted, the dynamic outlets can then release all or part of the detained impoundment to create more storage volume for the projected runoff. Conversely, dynamic outlets can be closed to maintain a desired water level in dry periods or create longer detention times for gravitational settling of pollutants prior to discharging downstream.

Problem Statement and Objectives

RTC applied to stormwater systems is an emerging field. The available technology has primarily been implemented on small stormwater detention systems, such as detention ponds in a neighborhood development. Our study aims to evaluate the feasibility of retrofitting earthen dams with RTC systems to improve stormwater management in small watersheds. We will accomplish this aim through three primary objectives:

1) Establish the state of the practice through a systematic literature review
2) Use field-collected spillway flow information from an earthen dam impoundment in Stillwater, OK to inform a large-scale testing regimen
3) Develop a large-scale testing model and methodology to report on water quality and quantity implications after RTC implementation

Progress

Our project started in January 2024, and we have been working toward a systematic literature review following PRISMA guidelines (Page et al., 2020) to establish the current state of the practice and evaluate critical research needs for RTC systems. We have used the Scopus database to search titles, abstracts, and keywords for the key terms "real-time control " OR "smart stormwater" OR "automated outlet." We used initial search criteria to consider research articles available in the English language authored from 2010 – 2024. This search yielded 6,877 documents. The search was then further refined by adding another search field with the Boolean operator “AND” with keywords "stormwater" OR "storm water.” The additional operation resulted in 108 documents that were used as the basis for our literature review. This second operation revealed that RTC and automation is becoming more relevant and implemented across other infrastructure and systems, but that the technology is still developing within the stormwater field. Finally, we did manual screening to meet the criteria for inclusion. The inclusion criteria were that the article was a) primarily authored around RTC in stormwater applications, b) an original research study, c) not about sewer systems, d) not a review paper or book. The manual screening resulted in 69 articles, which are included in the appendix. Our screening process is summarized in Figure 1.

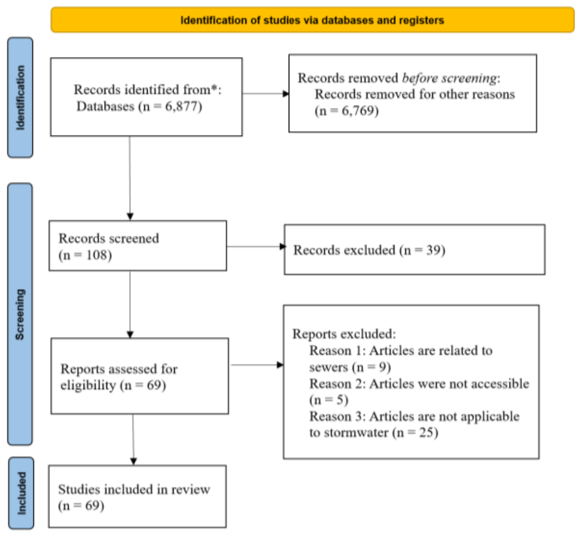

Figure A-0-1. Literature search and selection criteria

Preliminary Findings from the Literature Review

Figure 2 shows the annual publication frequency of Stormwater RTC articles. The first publications from our search period (2010-2024) were published in 2013. Published studies accelerated in 2020-2022 but slightly dropped in 2023. The drop could be due to the research-publication lag that resulted from COVID pandemic interrupting regular research operations. The publication rate has indicated that there is still much to learn about RTC stormwater systems before implementation.

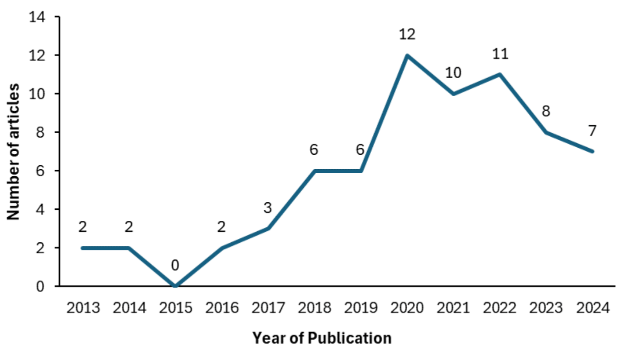

Figure A-2. Publication date of articles included in review

Next, we analyzed study locations to determine if there were similar regions or climates to Oklahoma, to determine the feasibility and applicability of RTC stormwater systems in Oklahoma. Figure 3 shows publication numbers by country.

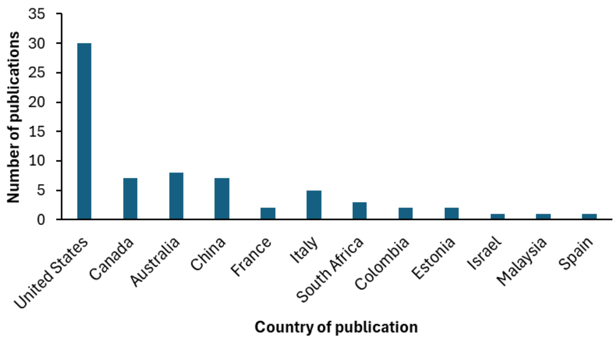

Figure 3. Publication by country of study

Preliminary analysis indicates that many of the studies are in regions with sensitive or culturally important waterbodies (e.g., Great Lakes, Salt Lake). Our ongoing literature review intends to evaluate and organize studies into three major categories: a) computer-based models, b) small-scale laboratory studies, and c) field-scale implementation and monitoring efforts. Additional topics of interest include system (e.g. single detention pond, watershed), catchment size, land use, water parameters of interest (e.g., flow rate, water quality constituents), and control algorithm.

Research Gaps

From preliminary analysis, a significant portion of the literature is focused on modeling studies, largely conducted in the Stormwater Management Model (SWMM). The studies have evaluated various algorithms for control of outlet structures (e.g., water level, inflow rate, forecast data); however, there is a significant gap in experimental studies and field-scale monitoring. Sweetapple et al. (2023) conducted a review to establish potential barriers to adoption. The study identified technical challenges (e.g., availability and reliability of technologies, technological and physical limitation, decision making, uncertainty and security) and socioeconomics (e.g., trust and lack of confidence, resistance to change, cost, and lack of guidance) as barriers to implementation. We believe that several of these barriers, such as reliability of technologies, understanding limitations, and decision-making, can be improved through controlled-, large-scale testing. Large-scale testing also offers a safe environment to experiment and prove implementation-readiness while building confidence in new technologies.

Next Steps

~~We expect to finish our literature review and establish a plan to collect field data for small impoundments in Stillwater, OK (e.g. SCS Stillwater Creek Site 24 or 26) (National Inventory of Dam, 2021) by the end of August 2024. Field data will include inflow rate, discharge through primary and~~

auxiliary spillway, and sediment-related water quality (e.g. turbidity and TSS). This information will be used to develop a large-scale model to test RTC technologies. The Digital Water Lab at the University of Michigan has initiated field scale efforts, primarily in Ann Arbor, Michigan (Kerkez, 2024). They have created an open-source database with sensor hardware and firmware (digitalwaterlab.org/build), which will be the basis of the RTC used in our experimentation. We will also work toward building partnerships to help with building and modifying the system architecture. A natural partnership would include the Oklahoma Mesonet and Dr. Ning Wang from OSU's Department of Biosystems and Agricultural Engineering, who has expertise in intelligent sensing and control technologies. With these challenges in mind, Dr. Schussler has partnered with researchers to investigate smart stormwater solutions in the Great Plains through extramural grant opportunities. Specifically, she participated on two proposals submitted to the EPA and NSF which include:

- Great Plains Center for Green, Advanced Stormwater Solutions, EPA Stormwater Centers of Excellence, *Awarded- Pending Contract (Schussler, Co-PI)*
    - The overall objective of the GPC4GrASS is to advance resilience of stormwater management in the Great Plains to reduce impacts associated with extreme events, while also creating benefits for communities across the region. One of the research subtasks is to deploy smart stormwater control on GI and stormwater ponds to investigate the feasibility of watershed-based operation of smart stormwater infrastructure to optimize flooding and stormwater quality improvements in the context of the extreme precipitation and droughts common to the Great Plains.

- Smart Stormwater in the Great Plains: Use-Inspired Technology for Resilient Stormwater Management, NSF EPSCoR RII Track-2, FEC, *Under Review (Schussler, Senior Personnel)*
    - The overarching goal of the proposed work is to build capacity across the Great Plains region to advance smart stormwater management technologies as an adaptive tool to climate extremes and address barriers that prevent stakeholders from adopting smart stormwater infrastructure. The proposed research would engage community and industry partners on use-inspired research and educational and workforce development initiatives related to smart stormwater. Schussler was proposed as the task lead for multiple demonstration sites in Nebraska, Kansas, and Oklahoma. The demonstration sites would serve as test beds for water quantity and quality sensing and control, provide performance data to inform models, and serve as venues to support the educational and workforce development initiatives. The proposed work incorporates social scientists and economists to assess socio-economic barriers to implementation of smart stormwater systems in the regions. Additionally, the education and workforce development goals would create materials on smart stormwater that were regionally applicable.

Formatted: List Paragraph, Bulleted + Level: 1 + Aligned at: 0.25" + Indent at: 0.5"
Formatted: Font: Italic
Formatted: List Paragraph, Bulleted + Level: 2 + Aligned at: 0.75" + Indent at: 1"
Formatted: List Paragraph, Bulleted + Level: 1 + Aligned at: 0.25" + Indent at: 0.5"
Formatted: Font: Italic
Formatted: List Paragraph, Bulleted + Level: 2 + Aligned at: 0.75" + Indent at: 1"

References

Association of State Dam Safety Officials (ASDSO). (2024). Dam Failure Incident Database.

Kerkez, B. (2024). Digital Water Lab at U-M. *Available from https://www.digitalwaterlab.org/*

Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., Bedig, A., Kertesz, R., Braun, T., Cadwalader, O. and A. Poresky. 2016. "Smarter stormwater systems." *Environmental Science & Technology*, 50 (14), 7267-7273 doi.org/10.1021/acs.est.5b05870.

Mullapudi, A., Wong, B. P. and B. Kerkez. 2017. "Emerging investigators series: building a theory for smart stormwater systems." *Environmental Science: Water Research &Technology*, 3(1), pp.66-77. doi.org/10.1039/C6EW00211K

Oklahoma Conservation Commission (OCC). (2024). Oklahoma's Upstream Flood Control Program. *Available from https://conservation.ok.gov/flood-control-program*

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj*, 372.

Shishegar, S., Duchesne, S., Pelletier, G., and R. Ghorbani. "A smart predictive framework for system-level stormwater management optimization." *Journal of Environmental Management*, Volume 278, Part 1. 2021. doi.org/10.1016/j.jenvman.2020.111505.

Sweetapple, C., Webber, J., Hastings, A., & Melville-Shreeve, P. (2023). Realising smarter stormwater management: A review of the barriers and a roadmap for real world application. *Water Research*, 120505.

United States Army Corps of Engineers (USACE). (2023). SCS Stillwater Creek Site- 24 and 26. *National Inventory of Dams.*

United States Department of the Interior, Bureau of Reclamation (DOI). (1987). Design of Small Dams. *A Water Resources Technical Publication*

Wuebbles, D. J., Easterling, D. R., Hayhoe, K., Knutson, T., Kopp, R. E., Kossin, J. P., Kunkel, K. E., LeGrande, A. N., Mears, C., Sweet, W. V., Taylor, P. C., Vose, R. S., and M. F. Wehner. 2017. "2017: Our globally changing climate." *In Climate Science Special Report: Fourth National Climate Assessment*, Volume I, pp. 35-72. doi.org/10.7930/J08S4N35.

Table A-1. Literature Included in Review

Author	Title	Publication Year
Gaborit E.; Muschalla D.; Vallet B.; Vanrolleghem P.A.; Anctil F.	Improving the performance of stormwater detention basins by real-time control using rainfall forecasts	2013
Campisano A.; Cabot Ple J.; Muschalla D.; Pleau M.; Vanrolleghem P.A.	Potential and limitations of modern equipment for real time control of urban wastewater systems	2013
Vallet B.; Muschalla D.; Lessard P.; Vanrolleghem P.A.	A new dynamic water quality model for stormwater basins as a tool for urban runoff management: Concept and validation	2014
Muschalla D.; Vallet B.; Anctil F.; Lessard P.; Pelletier G.; Vanrolleghem P.A.	Ecohydraulic-driven real-time control of stormwater basins	2014
Gaborit E.; Anctil F.; Pelletier G.; Vanrolleghem P.A.	Exploring forecast-based management strategies for stormwater detention ponds	2016
Riaño-Briceño G.; Barreiro-Gomez J.; Ramirez-Jaime A.; Quijano N.; Ocampo-Martinez C.	MatSWMM - An open-source toolbox for designing real-time control of urban drainage systems	2016
Mullapudi A.; Wong B.P.; Kerkez B.	Emerging investigators series: Building a theory for smart stormwater systems	2017
Rohrer A.R.; Armitage N.P.	Improving the viability of stormwater harvesting through rudimentary real time control	2017
Fisher-Jeffes L.; Carden K.; Armitage N.P.; Winter K.	Stormwater harvesting: Improving water security in South Africa's urban areas	2017
Zhang P.; Cai Y.; Wang J.	A simulation-based real-time control system for reducing urban runoff pollution through a stormwater storage tank	2018
Xu W.D.; Fletcher T.D.; Duncan H.P.; Bergmann D.J.; Breman J.; Burns M.J.	Improving the multi-objective performance of rainwater harvesting systems using Real-Time Control technology	2018
Bilodeau K.; Pelletier G.; Duchesne S.	Real-time control of stormwater detention basins as an adaptation measure in mid-size cities	2018
Wong B.P.; Kerkez B.	Real-Time Control of Urban Headwater Catchments Through Linear Feedback: Performance, Analysis, and Site Selection	2018
Mullapudi A.; Bartos M.; Wong B.; Kerkez B.	Shaping streamflow using a real-time stormwater control network	2018

Parolari A.J.; Pelrine S.; Bartlett M.S.	Stochastic water balance dynamics of passive and controlled stormwater basins	2018
Shishegar S.; Duchesne S.; Pelletier G.	An integrated optimization and rule-based approach for predictive real time control of urban stormwater management systems	2019
Ly D.K.; Maruéjouls T.; Binet G.; Bertrand-Krajewski J.-L.	Application of stormwater mass–volume curve prediction for water quality-based real-time control in sewer systems	2019
Di Matteo M.; Liang R.; Maier H.R.; Thyer M.A.; Simpson A.R.; Dandy G.C.; Ernst B.	Controlling rainwater storage as a system: An opportunity to reduce urban flood peaks for rare, long duration storms	2019
Sharior S.; McDonald W.; Parolari A.J.	Improved reliability of stormwater detention basin performance through water quality data-informed real-time control	2019
Sadler J.M.; Goodall J.L.; Behl M.; Morsy M.M.; Culver T.B.; Bowes B.D.	Leveraging open-source software and parallel computing for model predictive control of urban drainage systems using EPA-SWMM5	2019
Persaud P.P.; Akin A.A.; Kerkez B.; McCarthy D.T.; Hathaway J.M.	Real time control schemes for improving water quality from bioretention cells	2019
Li J.	A data-driven improved fuzzy logic control optimization-simulation tool for reducing flooding volume at downstream urban drainage systems	2020
Troutman S.C.; Love N.G.; Kerkez B.	Balancing water quality and flows in combined sewer systems using real-time control	2020
Altobelli M.; Cipolla S.S.; Maglionico M.	Combined application of real-time control and green technologies to urban drainage systems	2020
Mullapudi A.; Lewis M.J.; Gruden C.L.; Kerkez B.	Deep reinforcement learning for the real time control of stormwater systems	2020
Saliba S.M.; Bowes B.D.; Adams S.; Beling P.A.; Goodall J.L.	Deep reinforcement learning with uncertain data for real-time stormwater system control and flood mitigation	2020
Sun C.; Parellada B.; Feng J.; Puig V.; Cembrano G.	Factors influencing the stormwater quality model of sewer networks and a case study of Louis Fargue urban catchment in Bordeaux, France	2020
Maiolo M.; Palermo S.A.; Brusco A.C.; Pirouz B.; Turco M.; Vinci A.; Spezzano G.; Piro P.	On the use of a real-time control approach for urban stormwater management	2020

Kändler N.; Annus I.; Vassiljev A.; Puust R.	Peak flow reduction from small catchments using smart inlets	2020
Shen P.; Deletic A.; Bratieres K.; McCarthy D.T.	Real time control of biofilters delivers stormwater suitable for harvesting and reuse	2020
Xu W.D.; Fletcher T.D.; Burns M.J.; Cherqui F.	Real Time Control of Rainwater Harvesting Systems: The Benefits of Increasing Rainfall Forecast Window	2020
Ibrahim Y.A., Ph.d., P.e., M. asce	Real-Time Control Algorithm for Enhancing Operation of Network of Stormwater Management Facilities	2020
Schmitt Z.K.; Hodges C.C.; Dymond R.L.	Simulation and assessment of long-term stormwater basin performance under real-time control retrofits	2020
Shishegar S.; Duchesne S.; Pelletier G.; Ghorbani R.	A smart predictive framework for system-level stormwater management optimization	2021
Naughton J.; Sharior S.; Parolari A.; Strifling D.; McDonald W.	Barriers to Real-Time Control of Stormwater Systems	2021
Oberaschermsn M.; Rauch W.; Sitzenfrei R.	Efficient integration of IoT-based micro storages to improve urban drainage performance through advanced control strategies	2021
Bowes B.D.; Tavakoli A.; Wang C.; Heydarian A.; Behl M.; Beling P.A.; Goodall J.L.	Flood mitigation in coastal urban catchments using real-time stormwater infrastructure control and reinforcement learning	2021
Wang X.; Tian W.; Liao Z.	Offline Optimization of Sluice Control Rules in the Urban Water System for Flooding Mitigation	2021
Liang R.; Thyer M.A.; Maier H.R.; Dandy G.C.; Di Matteo M.	Optimising the design and real-time operation of systems of distributed stormwater storages to reduce urban flooding at the catchment scale	2021
Bartos M.; Kerkez B.	Pipedream: An interactive digital twin model for natural and urban drainage systems	2021
Alrabie N.A.; Mohamat-Yusuff F.; Hashim R.; Zulkeflee Z.; Amal M.N.A.; Arshad A.; Zulkifli S.Z.; Wijaya A.R.; Masood N.; Sani M.S.A.	Preliminary evaluation of heavy metal contamination and source identification in kuala lumpur smart stormwater pond sediments using pb isotopic signature	2021
Mason B.E.; Mullapudi A.; Kerkez B.	StormReactor: An open-source Python package for the integrated modeling of urban water quality and water balance	2021

Akin A.A.; Hathaway J.M.; Khojandi A.	Turbidity informed real-time control of a dry extended detention basin: a case study	2021
Liang R.; Maier H.R.; Thyer M.A.; Dandy G.C.; Tan Y.; Chhay M.; Sau T.; Lam V.	Calibration-free approach to reactive real-time control of stormwater storages	2022
Parker E.A.; Grant S.B.; Sahin A.; Vrugt J.A.; Brand M.W.	Can Smart Stormwater Systems Outsmart the Weather? Stormwater Capture with Real-Time Control in Southern California	2022
Kändler N.; Annus I.; Vassiljev A.	Controlling peak runoff from plots by coupling street storage with distributed real time control	2022
Xu W.D.; Burns M.J.; Cherqui F.; Smith-Miles K.; Fletcher T.D.	Coordinated Control Can Deliver Synergies Across Multiple Rainwater Storages	2022
Obando G.; Quijano N.; Ocampo-Martinez C.	Decentralized Control for Urban Drainage Systems Using Replicator Dynamics	2022
Shetty N.; Wang M.; Elliott R.; Culligan P.	Examining How a Smart Rainwater Harvesting System Connected to a Green Roof Can Improve Urban Stormwater Management	2022
Gomes Júnior M.N.; Giacomoni M.H.; Taha A.F.; Mendiondo E.M.	Flood Risk Mitigation and Valve Control in Stormwater Systems: State-Space Modeling, Control Algorithms, and Case Studies	2022
Altami S.A.; Salman B.	Implementation of IoT-Based Sensor Systems for Smart Stormwater Management	2022
Snir O.; Friedler E.; Ostfeld A.	Optimizing the Control of Decentralized Rainwater Harvesting Systems for Reducing Urban Drainage Flows	2022
Xu W.D.; Burns M.J.; Cherqui F.; Duchesne S.; Pelletier G.; Fletcher T.D.	Real-time controlled rainwater harvesting systems can improve the performance of stormwater networks	2022
Bowes B.D.; Wang C.; Ercan M.B.; Culver T.B.; Beling P.A.; Goodall J.L.	Reinforcement learning-based real-time control of coastal urban stormwater systems to mitigate flooding and improve water quality	2022
Mogano M.M.; Okedi J.	Assessing the benefits of real-time control to enhance rainwater harvesting at a building in Cape Town, South Africa	2023
Jung Y.E.; Jeong M.M.; Jun H.; Smith T.	Contemplation of Improvement Efforts to Manage Combined Sewer Overflows †	2023
Li J.; Burian S.J.	Evaluating real-time control of stormwater drainage network and green stormwater	2023

	infrastructure for enhancing flooding resilience under future rainfall projections	
Sun L.; Xia J.; She D.; Guo Q.; Su Y.; Wang W.	Integrated intra-storm predictive analysis and real-time control for urban stormwater storage to reduce flooding risk in cities	2023
Oh J.; Bartos M.	Model predictive control of stormwater basins coupled with real-time data assimilation enhances flood and pollution control under uncertainty	2023
Rimer S.P.; Mullapudi A.; Troutman S.C.; Ewing G.; Bowes B.D.; Akin A.A.; Sadler J.; Kertesz R.; McDonnell B.; Montestruque L.; Hathaway J.; Goodall J.L.; Norton J.; Kerkez B.	A simulation sandbox for the development and evaluation of stormwater control algorithms	2023
Zhou H.; Li R.; Liu H.; Ni G.	Real-time control enhanced blue-green infrastructure towards torrential events: A smart predictive solution	2023
Técher D.	Real-time control technology for enhancing biofiltration performances and ecosystem functioning of decentralized bioretention cells	2023
Kim M.-G.; Bartos M.	A digital twin model for contaminant fate and transport in urban and natural drainage networks with online state estimation	2024
Sharifi A.; Tarlani Beris A.; Sharifzadeh Javidi A.; Nouri M.; Gholizadeh Lonbar A.; Ahmadi M.	Application of artificial intelligence in digital twin models for stormwater infrastructure systems in smart cities	2024
Wang S.; Wang J.; Xue C.; Qiu R.; Sun S.; Yang Z.; Qiao Y.	Assessing the effectiveness of a residential-scale detention tank operated in a multi-objective approach using SWMM	2024
Lin C.-Y.; Yang Y.-C.E.; Moazeni F.	Flood Risks of Cyber-Physical Attacks in a Smart Storm Water System	2024
Altobelli M.; Evangelisti M.; Maglionico M.	Multi-Objective Performance of Detention Basins and Rainwater Harvesting Systems Using Real-Time Controls with Rainfall Forecasts	2024
Persaud P.P.; Hathaway J.M.; Kerkez B.; McCarthy D.T.	Real-Time Control and Bioretention: Implications for Hydrology	2024
Chen Y.; Wang C.; Huang H.; Lei X.; Wang H.; Jiang S.; Wang Z.	Real-time model predictive control of urban drainage system in coastal areas	2024

Appendix B: A Framework for Integrating Diverse Hydrologic Variables for Modeling Impact of Storm Events on Reservoirs, Year 3 DAM-TAGS Progress Report for Jeff Sadler

Summary

This year, our we focused on the first steps of integrating diverse hydrologic data into a machine learning (ML) rainfall-runoff model. We selected two study sites in collaboration with the Arkansas-Red River Forecast Center: Kiamichi River at Antlers, OK, and Washita River at Paul's Valley, OK. We established a baseline ML model using a Long Short-term Memory (LSTM) neural network, with simulated runoff data from the Sacramento Soil Moisture Accounting (SAC-SMA) model as input. Initial results showed the ML model performing at a similar level to the SAC-SMA model. Presentations of our work were made at the Oklahoma Governor's Water Conference and the Oklahoma Chapter of the American Society of Biological and Agricultural Engineers. Next steps include tuning the ML model and incorporating additional rainfall data to improve predictive accuracy.

Statement of critical regional or state water problem

Hydrologic information and predictions are becoming increasingly more important, however, the ability needed to fully leverage growing data streams still needs to be developed. As climate changes, the extremes of the hydrologic cycle are projected to intensify, making floods more likely (Fischer and Knutti 2015). At the same time, expanding human development has made the consequences of flooding events more consequential. For example, many small flood control dams, though originally built in rural settings, are now surrounded by residential and/or commercial development. Current operational hydrologic models do not provide usable operational predictions for most small flood control reservoirs including the more than 30,000 reservoirs in Oklahoma. Given the importance of information related to these structures, large amounts of data are currently being collected and even more data collection is planned. These data include a diversity of variables and collection methods including in situ, remote sensing, and unmanned autonomous systems (UAS). These data will be collected in a variety of spatial and temporal resolutions. Conventional, physics-based hydrologic models cannot easily incorporate such diverse data. Recently developed machine learning (ML) models can learn from large amounts of hydrologic data (Kratzert et. al 2019), however, the data have typically been continuous in time and lumped in space. There is a critical need, therefore, to develop approaches for utilizing the growing and planned streams of diverse data to improve hydrologic understanding and forecasting.

Approach and Progress

Our long-term goal is to create a framework for integrating a diverse set of hydrologic data sources into a ML rainfall runoff model. In line with this objective, our efforts to date have focused on three main tasks: 1) selecting study locations for model development and evaluation, 2) setting up a baseline model including collecting model inputs and outputs, and 3) assessing the baseline model's accuracy with a process-based model as a predictor.

Selecting Study Locations: Collaboration with Arkansas-Red River Forecast Center

One of our main tasks this year has been to select study locations for which we will develop and evaluate our ML modeling and data integration approaches. To train the ML model, we needed locations that have a relatively large amount of streamflow observation data. There are several locations in Oklahoma that have long-term streamflow records that we could have used for our modeling experiments. To have a greater impact on decision makers, we have collaborated with National Weather

Service Arkansas-Red River Forecast Center (AR-RFC) for selecting two pilot locations: the Kiamichi River at Antlers, OK (ANT) and the Washita River at Paul's Valley, OK (PV) (Figure 1). At these locations the AR-RFC forecasts consistently need manual adjustments. By focusing on these two locations, we will be able to develop and evaluate our models, and we will hopefully provide an improved forecast tool for the AR-RFC and thus improved information for decision makers that rely on those forecasts.

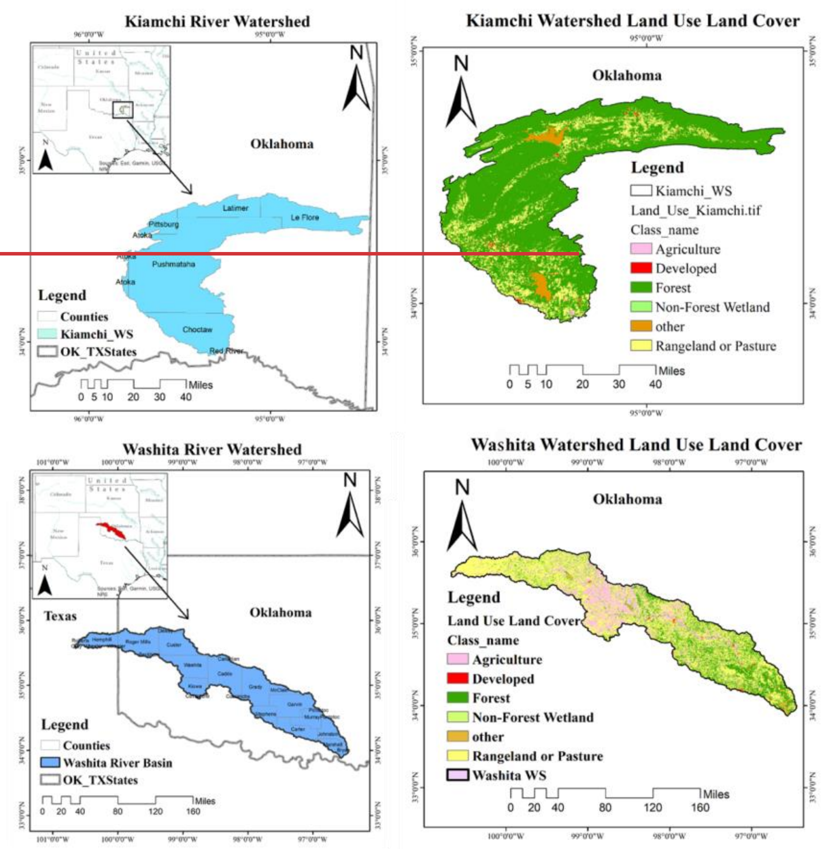

Figure 0-3. Initial Study Locations - Kiamichi River at Antlers (ANT) and the Washita River at Paul's Valley (PV)

Setting Up the Baseline ML Model

ML Model Architecture

In our initial modeling, we have evaluated a particular recurrent neural network (RNN) architecture: Long Short-term Memory (LSTM). RNNs pass information from one sequence member (e.g., time step) to the next, making them especially appropriate for environmental variables where temporal autocorrelation is common. LSTMs provide additional benefits through using "cell-states" and "gates" which allow information to accumulate and then be used for later predictions. This is important in natural systems where there can be a delay between input (e.g., rainfall) and response (e.g., runoff). LSTMs have been shown to be successful at modeling many water-related variables including streamflow (Kratzert et al. 2019), water level (Bowes et al. 2019), and water temperature (Read et al. 2019).

Input Data

In the first stage of our modeling, we have only used one model input - the simulated runoff from a process-based model. This simulation data has been provided by the AR-RFC which uses the Sacramento Soil Moisture Accounting (SAC-SMA) model. It can be useful to use process-based model output for input to a ML model because it can allow the ML model to learn the rainfall-runoff patterns without needing observations. With process-based model predictions as input, an ML model may learn to correct any bias in the process-based model. Furthermore, process-based model predictions are becoming more widespread and accessible at many locations throughout Oklahoma and the US. For example, the National Water Model produces process-based model outputs at over 1 million locations in the US. In Oklahoma, the OK-HAWQS system can be used to generate streamflow predictions at many locations in Oklahoma. Thus, we believe experimenting with using process-based model outputs as inputs to an ML model is a useful first step in our modeling process. As we move forward, we will add other data sources as model input.

Output Data

The output (or target) data of our model are streamflow estimations. We use streamflow observation data to train the ML models to predict streamflow. The streamflow observations come from the USGS which operates stream gauges at both of our study sites.

Assessing Baseline Accuracy and Evaluating Initial Approach

To assess our modeling approach, we first established the baseline accuracy of the SAC-SMA model. Although the SAC-SMA model accurately follows the streamflow trends, it often inaccurately predicts the peak flows (as is common in many rainfall-runoff models) (Figures 2 and 3). The Nash Sutcliffe Efficiency (NSE) values were 0.7 for ANT and 0.59 for PV. The ML model will seek to improve upon this performance.

In our initial modeling results, the LSTM using the SAC-SMA model as input is performing at the same level as the SAC-SMA itself (Table 1). In other words, the ML model, as currently configured, is not effectively correcting the bias of the process-based model. This is our baseline accuracy. As we continue to adjust the LSTM configurations and, later, add more input data, we anticipate that the ML model will begin to improve upon the process-based model outputs.

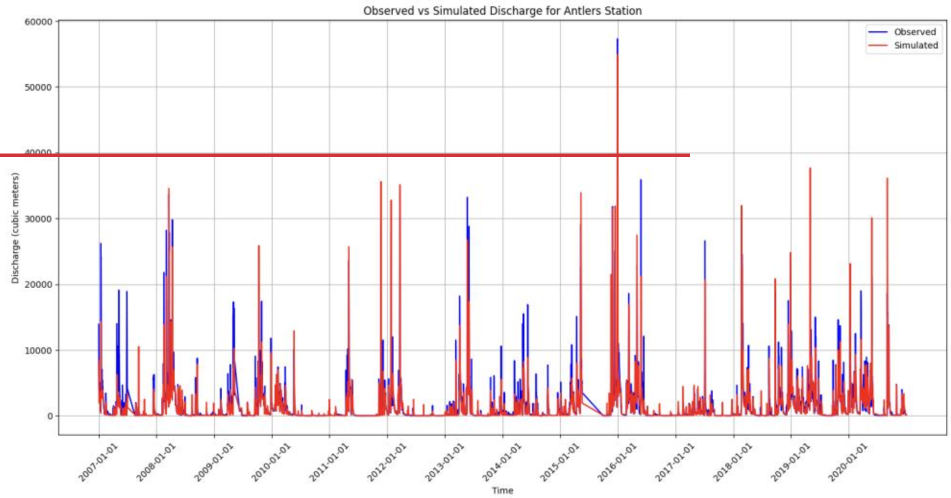

Figure 0-4. Streamflow observations and process-based model outputs (used for ML inputs) at the Kiamichi River at Antlers OK

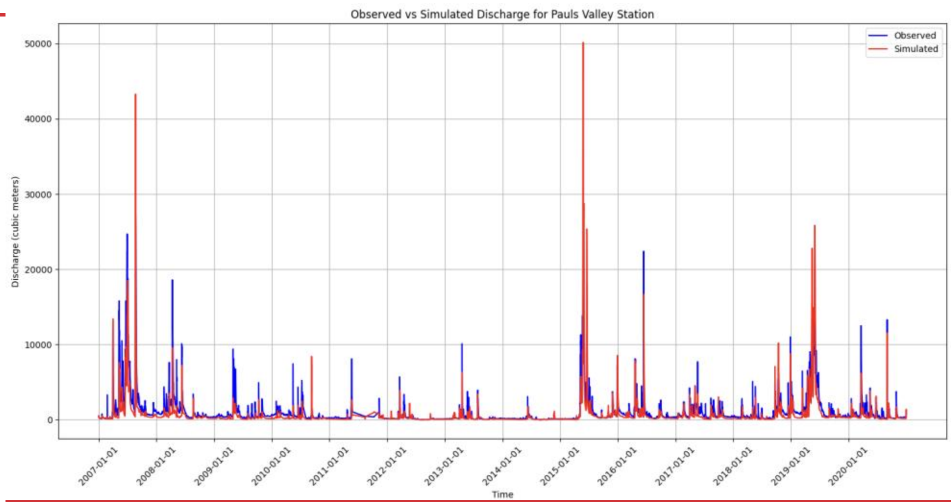

Figure 0-5. Streamflow observations and process-based model outputs (used for ML inputs) at the Washita River at Paul's Valley OK

Table 0-2. Baseline metrics (NSE) for LSTM and SAC-SMA models

	Antlers	Paul's Valley
SAC-SMA	0.7	0.59
LSTM	0.7	0.58

Project Outputs

In this year the student working on this project has presented this work in two research meetings: The Oklahoma Governor's Water Conference and The Oklahoma Chapter of American Society of Biological and Agricultural Engineers.

Next Steps

For our next steps, we will seek to improve the ML model in two main ways. First, we will tune the current configuration of the model to see if we can improve model performance. With ML models, the configuration (e.g., the number of previous time steps are used to make a prediction, the complexity of the model) can have a large impact on model performance. Second, we will look at the effect of different rainfall inputs on the model performance. As the major driver of changes in streamflow, we believe that including rainfall information (not just the process-based model outputs) will improve model performance.

References

Bowes, Benjamin D, Jeffrey M Sadler, Mohamed M Morsy, Madhur Behl, and Jonathan L Goodall. 2019. "Forecasting Groundwater Table in a Flood Prone Coastal City with Long Short-Term Memory and Recurrent Neural Networks." Water 11 (5): 1098.

Fischer, Erich, and Reto Knutti. 2015. "Anthropogenic Contribution to Global Occurrence of Heavy-Precipitation and High-Temperature Extremes." Nature Climate Change 5 (April). https://doi.org/10.1038/NCLIMATE2617.

Kratzert, Frederik, Daniel Klotz, Guy Shalev, Günter Klambauer, Sepp Hochreiter, and Grey Nearing. 2019. "Towards Learning Universal, Regional, and Local Hydrological Behaviors via Machine Learning Applied to Large-Sample Datasets." Hydrology and Earth System Sciences 23 (December): 5089–5110. https://doi.org/10.5194/hess-23-5089-2019.

Read, Jordan S., Xiaowei Jia, Jared Willard, Alison P. Appling, Jacob A. Zwart, Samantha K. Oliver, Anuj Karpatne, et al. 2019. "Process-Guided Deep Learning Predictions of Lake Water Temperature." Water Resources Research 55 (11): 9173–90. https://doi.org/https://doi.org/10.1029/2019WR024922.

Background and Introduction

Hydrologic information and predictions are becoming increasingly important, however, the ability needed to fully leverage growing data streams still needs to be developed. Our long-term goal is to create a framework for integrating a diverse set of hydrologic data sources into a machine learning (ML)

Formatted: First Paragraph

rainfall-runoff model. In line with this objective, our efforts this year have focused on understanding the importance of using spatially distributed rainfall data.

Rationale

In several instances, machine learning (ML) models have been shown to be more accurate than process-based models at predicting streamflow (cite). As is the case in all streamflow prediction models, ML models use several weather-related inputs including rainfall data, which is obviously an important predictor to changes in streamflow. In most documented cases, however, the rainfall data ML models have used has been spatially-aggregated - for the entire watershed, the rainfall data is assumed to be spatially-uniform. This is, of course, far from the truth since rainfall totals can vary widely across a watershed and the location of rainfall on a watershed will have a large impact on how the streamflow changes.

Approach

We have trained two versions of a ML model to predict streamflow at the Pauls Valley USGS station on the Washita River (USGS site 07328500) in south-central Oklahoma. The Washita River watershed at this location is very large and long - over 13k square km in area and over 300 km in length. The model we trained was a type of recurrent neural network called Long Shorterm Memory (LSTM). For both models, we used used rainfall values from 10 Oklahoma Mesonet weather stations (mesonet.org) across the length and width of the watershed. With the first version of the model, we aggregated the rainfall data geographically so that there was only one average rainfall value for each day. With the second version of the model, we disaggregated the rainfall data geographically so that there were 10 rainfall values for each day (one for each weather station).

Findings

We found that using the data from each individual weather station provided much more accurate predictions than using the spatially averaged rainfall data. We used Nash-Sutcliffe Efficiency (NSE) to measure the model accuracy, where a perfect model would have an NSE of 1, a model predicting a constant average streamflow value would have an NSE of 0, and a model worse than the average value would have a negative NSE. When we used the data from the 10 individual weather stations, the NSE was 0.35. When we used the spatially averaged data the NSE was -1.85.

Conclusion

Our findings highlight the critical role of spatially distributed rainfall data in improving streamflow predictions using ML models. The flexibility of ML allowed us to leverage data from multiple weather stations, a capability that process-based models typically lack. This demonstrates the exciting potential of ML in integrating multiple data sources to enhance hydrologic predictions.