Semi-automated workflow for multi-basin, multi-scenario flood risk mapping, and impact assessment.

Introduction

Floods are one of the most devastating natural hazards, causing massive loss of life and property. Although it is challenging to predict where flooding will occur, identifying highly vulnerable areas allows for proactive planning to mitigate potential damage. The growing concerns over climate change increase the likelihood of more frequent and intense flood events (Milly et al., 2002). Recognizing the need to protect communities from the severe impacts of flooding, it is crucial to develop accurate and reliable flood maps for effective flood mitigation and management strategies (Merwade et al., 2008). Flood inundation maps are invaluable assets for identifying areas susceptible to flooding and enhancing public awareness of the risks associated with living in flood-prone regions (National Research Council, 2009).

Various approaches have been developed and utilized to assess flood risk across specific areas, ranging from intricate hydraulic and hydrodynamic models to simpler multi-criteria analysis (MCA) techniques. However, (FAU, 2020) approached and approved the cascade 2001 routing model (SFWMD, 2001) for mapping flood-vulnerable areas across Florida from the Florida Division of Emergency Management (FDEM) and the Federal Emergency Management Agency (FEMA) to aid in making suitable watershed masterplans for various communities. The major advantage of using this model is that it gives a flood level indicating a potential flood below that surface, and it is consistent with the FEMA map (Hindle, 2021) in terms of flood extent. The use of this cascade model has been successfully implemented (Hindle, 2021; Rojas,2020; Hewett, 2021; Rodrigues,2021) to map the flood inundation across the watershed level to the local level. This Cascade 2001 model simulates the water from one basin to another based on their connection (SWFMD,2001).

Commented [HS1]: Add 1 or 2 reference for cascade 2001 model. The reference list can be found from the software manual: \\engsynws01.eng.fau.edu\Project_mastercopy\Software\Cascade 2001\Cascade\docs

Commented [HS2R1]:

Automated post-flood mapping using multispectral imagery has been extensively utilized to determine the extent of flooding and assess damage (Sun et al., 2012; Karamvasis & Karathanassi, 2021). However, only a few studies have automated the process of flood risk mapping using LFP Python packages (Sosa et al., 2019). As Centre for Water Risk Reduction Resiliency (CWR3) is actively engaged in flood mapping across Florida’s communities to provide watershed management master plans and reduce flood insurance costs, there is a need for accurate, consistent, and precise mapping for impact assessment. For this an automated process is essential, as the manual approach involves repetitive and time-consuming tasks, so we aim to automate flood risk mapping and impact assessment process except hydrological modeling.

FAU's flood inundation modeling and mapping methodology demands several datasets and multi-scenario modeling, necessitating meticulous preparation of Cascade inputs. This involves a series of intricate operations, including:

- Inputs preparation for cascade 2001
- Hydrological Simulation in Cascade 2001 routing model
- Vulnerability and Impact calculations
- Visualization of inputs and outputs through map

When executed manually, these processes became cumbersome and error-prone, particularly when working with numerous Hydrologic unit code (HUC) boundaries, each requiring independent GIS geoprocessing. Moreover, the comprehensive modeling of inundation scenarios under diverse rainfall events, sea level rise, and king tide conditions can escalate to a staggering 48 distinct scenarios, magnifying the potential for inconsistencies and time consumption. To avoid these challenges, a compelling approach will be programming the repetitive and manual aspects of data preparation and map generation, ensuring swift, consistent, and reliable results.

Motivated by the limitations inherent in existing flood map generation methods, such as the tedium and potential reproducibility issues associated with the methodology, this research pursues the development of a Python-based standalone script for automated data preparation and flood map creation. Our primary objective is to explore the feasibility of automating flood map generation within the existing framework of the FAU (2020) screening tool. This entails automating the programming of currently manual, tedious, and repetitive tasks associated with Cascade parameter generation, including those for Mean rainfall values for different scenarios, Soil storage values for different scenarios, Time of concentration calculations, Minimum and maximum elevation extraction from DEMs for specific watersheds and sub-watersheds. By automating these critical processes, we aim to achieve a streamlined and cost-effective workflow for flood map generation, particularly beneficial for comprehensive mitigation planning requiring complex, multi-scenario analyses. For instance, this script could enable efficient modeling and mapping across 48 distinct rainfall, king tide, and sea level rise scenarios in coastal regions. It is important to emphasize that this research does not seek to validate or replace the existing FAU (2020) screening tool. Instead, our focus lies on refining and automating specific processes within that framework to address critical bottlenecks and enhance the tool's overall functionality and efficiency.

Methods

Flood maps need to be created for better planning and understanding of the vulnerable regions within a watershed at the local level. FAU (2020) designed a framework to map the vulnerable areas using a probabilistic approach within a sub-watershed level using a HUC to address this. This methodology (Figure 1) requires several datasets ranging from elevation, groundwater table elevation, soil storage, different rainfall/storm scenarios, land cover, impervious mask, water mask, and stormwater infrastructure.

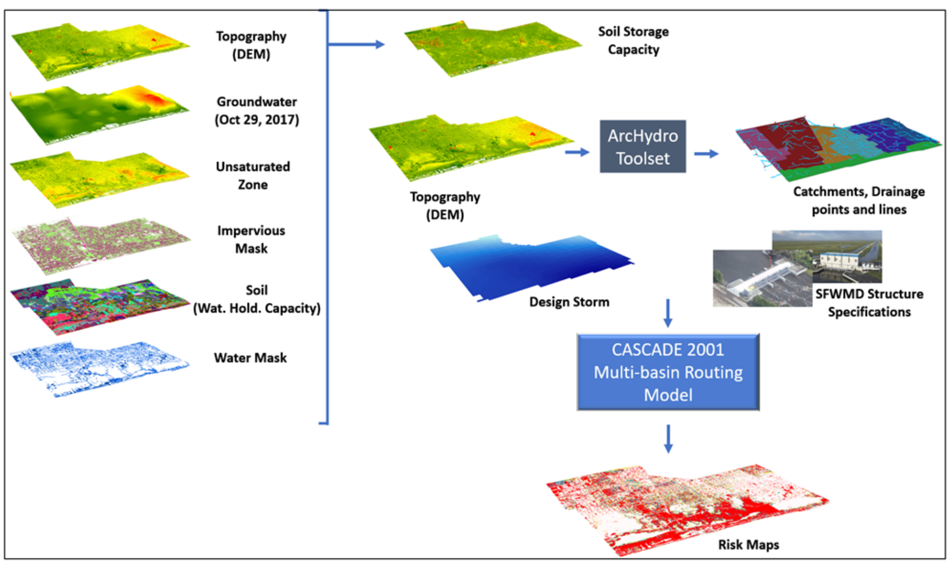

Figure 1. Screening tool approach developed by (FAU, 2020) for creating flood risk maps.

Flood inundation modeling for numerous HUC boundaries currently involves a cumbersome workflow of individual GIS operations, parameter extraction, separate Cascade model runs, and manual flood map creation for different scenarios. This iterative process, while thorough, is time-consuming and can be prone to human error. We aim to automate most of the workflow using Python to overcome these limitations. Scripting routine GIS tasks like ArcHydro, raster calculations, and zonal statistics can dramatically reduce processing time for multiple HUCs, potentially completing analyses within a day. Furthermore, automating Cascade parameter derivation and integrating it with flood inundation calculations eliminates the need for separate model runs, further streamlining the process. This comprehensive automation not only boosts efficiency but also minimizes human error and maximizes flexibility, allowing for quick adjustments to input datasets and map generation through code revisions. Automating this flood inundation modeling workflow, except Cascade modeling, offers a faster, more reliable, and cost-effective approach for assessing flood risk across diverse scenarios and geographical scales.

The detailed workflow created here in this study is based on the above methodology.

Input preparations:

The input preparation has been divided into different sections. As discussed earlier, FAU methodology has been derived for multi-basin watersheds. The first section (Section 1.1) deals with sub-set datasets for all the basins that formed a watershed, which is defined by the HUC boundary. Input_preparation.ipython file can be easily implemented into the ArcGIS environment to achieve this task, which takes a list of input datasets and a shapefile for a watershed boundary, as shown in the table below, and creates a number of subdirectories inside the directory HUCs based on the HUC id which is then used to stores subsets of raster datasets per HUC boundary.

The second section processes ArcHydro tools to obtain the initial stage and time of concentration for the basins with some intermediate files such as drainage lines, drainage points, etc. This step requires the processing of the first step, or else required files, mainly DEM, in the proper folder structure. The following flow charts illustrate the processing order with required inputs.

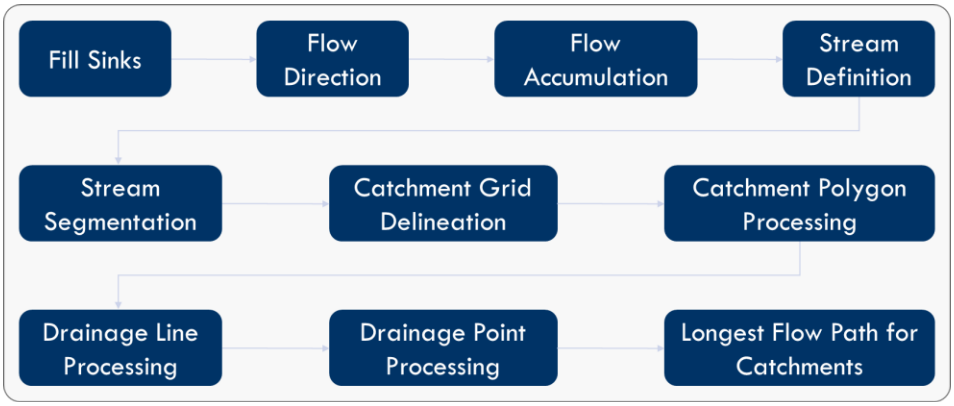

Figure: ArcHydro preprocessing workflow

Following the ArcHydro process, computations of unsaturated zones and soil storage are begun based on the given scenarios. The depth of the unsaturated zone was calculated by subtracting the (water table elevation **minus** sea level) raster from the digital elevation model (DEM), yielding a spatial representation of the unsaturated soil profile. Soil storage capacity was subsequently calculated using the following formula, incorporating key factors influencing water retention:

Commented [AM3]: We need to redefine this based on slr defination may be good to show full formula.

Soil Storage (inches) = Unsaturated Zone (feet) * Water Holding Capacity (feet) * Water Mask * Impervious Mask * 12

For making soil storage non-negative, a conditional function was applied to make positive soil storage values to ensure physical plausibility, as negative storage is conceptually invalid.

The final section of the input preparation stage is statically summarizing required input parameters for the Cascade 2001 into the Excel table for further modeling or simulating hydrological models based on the different design storms. This section yields an Excel file containing the information with the headings below.

HUC ID	Maximum Elevation	Minimum Elevation	Design Rainfall	Soil Storage Capacity (SSC)

The design rainfall column can be different as it depends upon the design storm events in the parent folder or in the data list mentioned in the first section. Similarly, soil storage capacity (SSC) can have more than one column if we model for multiple scenarios because value of SSC differs based on scenarios as it is affected by the rise in sea level, king tide and others.

Hydrological simulation in Cascade 2001

The South Florida Water Management District (SFWMD) developed a model known as Cascade 2001, designed to simulate water flow through multiple basins (SFWMD, 2001). The model utilizes several key input parameters derived from DEM processing, including the initial stage, ground storage, concentration time, DEM elevations, and a defined rainfall scenario. It also considers various stormwater infrastructures, such as pipes, pumps, and spillways. These inputs collectively represent the actual environmental conditions of the region, ensuring that the flooding dynamics of the watershed are accurately reflected in the model's results. The model can handle more than one basin in a catchment, simultaneously mimicking its hydrological conditions. By employing this model, the flood level for each sub-watershed boundary is determined. Any surface area below this specified level is assumed to be flooded, aiding in flood risk assessment and planning. We use this hydrological model because of its advantages, such as simplicity, integration with GIS-derived products, and applicability. Similarly, several researchers, including (Bloetscher et al., 2021, FAU, 2020; Rojah, 2020; Hindle, 2021; Hewett, 2021; and Rodrigues, 2021), have successfully used this model to map inundation across different watersheds in Florida.

Risk and Impact Assessment:

A maximum headwater height is obtained after the hydrological simulation in Cascade 2001, which can be converted into the Z-score raster, which is later classified in terms of the probability of flood inundation. The following equation is used to calculate the z-score using the topography and a constant value, which is often known as the combined uncertainty of LiDAR and Cascade:

Z-score = (Headwater Height - DEM)/0.46.

The z-score calculation is automated in Section 3.1 of the Risk_impact_assesment.ipython , which takes an Excel table with the following headings and yields a raster file for each scenario and each HUCs.

HUCs	Head-Water Height	Scenario

Section 3.2 of the above-mentioned ipython file deals with the calculation of the composite score. It takes a feature class/ shapefile of the parcels/tiers and inundation raster and creates a feature

class/ shapefile with the scoring. The details of the scoring and implemented methods can be found here.

The final section of stage 3 calculates the area of different types of lands such as state, federal, Indian, and/or with and without open spaces under the probabilistically flooded zone.

Mapping and Visualizations

While dealing with many datasets, creating maps to visualize them sometimes requires extensive time. On the other hand, it is frustrating if it requires multiple minor revisions. The Mapping_visualizations.ipython automates the creation of maps using pre-stored ArcGIS project files to visualize the raw datasets and the processed outputs. The script assumes that the naming conventions have been followed by the data management guide provided on the GitHub project.

Results and Discussion

In this section, along with time comparison, we can demonstrate the case study of one specific area to show that it works with the time taken by python script.

In Result sections, we will show one case study or results of automation and based on manual process how this is better.

Discussions

Discussion can be shorter just to explain what the benefits of automations was, and how useful will be this. Similarly, we can discuss it limitations such as not fully developed automated due to intricacy of Cascade 2001 model. And suggests to make surrogate modeling of Cascade 2001 in further research for fully automations.

The outcomes of this investigation reveal a notable coherence in both results and mapping products. A case study conducted in Hialeah underscores the consistency observed between the automated approach and manual methods. Notably, the time efficiency of automation is strikingly incomparable to that of human personnel. Completing inundation modeling and mapping tasks within a single day, facilitated by the automated process, starkly contrasts the approximate two-week timeframe required for manual execution when conducted on a rapid scale.

Commented [HS4]: Providing a table to compare the time needed for each sub task in the data processing between the automated method and manual approach.

Find from other similar article on how to develop and use an evaluation matrix to assess the improvement of software efficiency and automation.

Overall, it takes up to 16.4 (2 days of 8 hour works daily) hours to make a cascade input parameter if all the procedure runs smoothly and efficiently without any obstruction. But there can arise a mistake and software issues any time, which might delay the performance of humans. Therefore, to reduce the time consumption for making all these inputs for a cascade model, we wrote a script that can do all these tasks at 1-2 hours within a day. This saved lots of time.

Also, map making for each HUC for every layer requires time, we need to change scale, legend, north arrow, name of map etc, which consumes more time. So, we also developed a script to automate all these maps making which has become a life savior for we are working on this watershed masterplan project. Previously if we had to refine or update maps, it would take significant time, but with this automation, we can update or refine the maps features in no time, which is innovative in this research.

Conclusions

In conclusion, implementing automation in our research has yielded significant time and effort savings, particularly in manual tasks. The efficiency achieved by this approach surpasses human capabilities, enabling expeditious completion of work while ensuring consistency with the original template. An additional advantage is the facilitation of revisions or updates without the need for manual intervention, enhancing the overall efficiency of the research process. It is important to acknowledge the limitations of our approach, notably the reliance on ArcGIS Pro, notebook, and the manual execution required for Cascade Routing modeling. Future research endeavors can focus on automating or surrogating the routing model to further enhance the routing model to enhance further the comprehensive automation of flood vulnerability mapping for counties or cities. While our research may not introduce novel concepts, it is a valuable resource for practitioners in similar projects. The simplified workflow eliminates the need for cumbersome and repetitive geoprocessing tools, contributing to the overall efficacy and accessibility of the research methodology. This tool or automation can be very helpful for coastal and flood risk planners to map the vulnerabilities across the areas.

Acknowledgments

We acknowledge all the Professors, Research Associates and Research Assistants involved in watershed masterplan management phase II.

References

FAU (Florida Atlantic University). (2020). Watershed Master Planning Template. FAU Center for Water Resiliency and Risk Reduction, Boca Raton, FL

Milly, P. C. D., Wetherald, R. T., Dunne, K. A., & Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. Nature, 415(6871), 514-517.

Karamvasis, K., & Karathanassi, V. (2021). FLOMPY: An Open-Source Toolbox for Floodwater Mapping Using Sentinel-1 Intensity Time Series. *Water*, *13*(21), 2943.

Sun, W., Mckeown, D. M., & Messinger, D. W. (2012, October). An automated approach to flood mapping. In *Earth Resources and Environmental Remote Sensing/GIS Applications III* (Vol. 8538, pp. 261-268). SPIE.

Marco, J. B. (1994). Flood risk mapping. In Coping with floods (pp. 353-373). Dordrecht: Springer Netherlands.

ESRI. (n.d.-a). Extract values to points (spatial analyst). Extract Values to Points (Spatial Analyst)-ArcGIS Pro | Documentation. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/extract-values-to-points.html

Sosa, J., Sampson, C., Smith, A., Neal, J., & Bates, P. (2020). A toolbox to quickly prepare flood inundation models for LISFLOOD-FP simulations. Environmental Modelling & Software, 123, 104561.

SFWMD (South Florida Water Management District). (2001). User's Guide for the Routing Model CASCADE 2001 Version 1.0. Technical Publication REG-003, SFWMD, West Palm Beach, FL

NOAA (National Oceanic and Atmospheric Administration). (2010). Mapping Inundation Uncertainty. NOAA Coastal Services Center, Charleston, SC.

Bloetscher, F., Rojas, G., Abbate, A., Hindle, T., Huber, J., Jones, R., ... & Zhang, C. (2021). A Framework for a Subwatershed-Scale Screening Tool to Support Development

Commented [HS5]: Can you find several recent relevant articles from the same journal you plan to submit?

Please add this reference and cite it:

https://www.fau.edu/engineering/research/cwr3/pdf/framework-subwatershed-scale-resiliency-solutions-flood-protection-priority.pdf

of Resiliency Solutions and Flood Protection Priority Areas in a Low-Lying Coastal Community. *Journal of Geoscience and Environment Protection*, *9*(10), 180-205.

Rodrigues, S. (2021). *Flood Analysis in the Pensacola Basin, Florida* (Doctoral dissertation, Florida Atlantic University).

Rojas, G. (2020). *Establishing a Screening Tool to Support Development and Prioritization of Watershed Based Flood Protection Plans* (Doctoral dissertation, Florida Atlantic University).

Hewett, M. (2021). *Flood Analysis in Nassau County, Florida* (Doctoral dissertation, Florida Atlantic University).

Hindle, T. (2021). *An Examination of Downscaling a Flood Risk Screening Tool at the Watershed, Subwatershed, and Municipallevels* (Doctoral dissertation, Florida Atlantic University).

Merwade, V., Olivera, F., Arabi, M., & Edleman, S. (2008). Uncertainty in flood inundation mapping: current issues and future directions. *Journal of hydrologic engineering*, *13*(7), 608-620.

National Research Council. (2009). Mapping the Zone: Improving Flood Map Accuracy. Washington, DC: The National Academies Press.