An examination of extreme floods, effects on land-use change and seasonality in the lower St. Johns River Basin, Florida using HSPF and statistical methods

Samantha Kovalenko1 · Christopher James Brown1 · Cigdem Akan2 · Alexandra Schonning1

Abstract
As population growth and urbanization are steadily rising, the need for dependable flood estimation techniques is crucial. This study evaluates extreme flood events in select sub-basins of the Lower St. Johns River in Florida, USA. The study considers the effect of urbanization on the natural hydrologic processes and flood magnitudes in the watershed. Additionally, the effects of varying seasonality into the hydrologic modeling procedure are also investigated. This research focuses on determining the 10-, 25-, 50-, and 100-year return frequency flood flows in Julington Creek, Ortega River, and Pablo Creek of the Lower St. Johns River Basin in Florida, USA. The major findings of this research indicate that by implementing a range of flood estimation methods one can better describe the inherent uncertainty with traditional estimates. Also, the research showed that varying seasonality in the hydrologic modeling procedure does not result in vast differences in the resulting flood estimates. However, various land-use scenarios may produce simulated flood flows of greater magnitude—especially when a more urbanized land-use scenario is modeled.

Introduction
Extreme flood estimation techniques are a crucial necessity for all communities. In a world where urbanization, sea-level rise, and climate change are prevalent, advancements in flood estimation techniques must occur to better protect humanity. When accurate flood estimation is implemented, economic and community wellbeing are secured. In the United States of America, flooding impacts the state of Florida most severely and it is estimated that Florida suffered $2.5 billion in flood-related losses from 1990 to 2003 (Brody et al. 2007). Therefore, the need for accurate flood estimation techniques is evident. The research presented herein was conducted to better assess flood flow estimates for portions of the Lower St. Johns River in Florida, USA. Traditionally, flood estimates have relied upon one estimating methodology per study resulting in only a “best estimate” of various extreme storm events. This research intends to demonstrate that developing a wider range of estimates with multiple methods provides a more realistic approach and provides a much-needed way to assess estimate uncertainties. The worked discussed herein focuses upon six sub-basins namely Julington Creek, Durbin Creek, Big Davis Creek, Black Creek, Ortega River, and Pablo Creek. Seasonality of flow events in Northeastern Florida is analyzed by studying the effects of projected urbanization in the watershed on flood estimates. This paper presents flood estimates developed using the model code HSPF as well as multiple statistical methods suggested by the literature.

Flood estimates for this study were developed by utilizing multiple methodologies including statistical modeling and numerical modeling. In addition, published estimates from flood insurance studies were also compared and included in the compilation. The resulting matrix of estimates provides a fairly wide range of reasonable flood flow estimates for each of the sub-basins investigated. Ultimately, the flood flow...
estimates themselves were combined at the different return frequencies to look at the underlying distribution. First, a major part of the overall effort was modification of existing numerical models. Later sections of this paper discuss the statistical estimates.

The research of Kovalenko (2020) encompasses the modification of existing hydrologic models that were initially developed by the St. Johns River Water Management District (SJRWMD). Several water management districts collaborated in 2012 to develop the St. Johns River Water Supply Impact Study (WSIS). The WSIS was developed to simulate the effects of withdrawing water from the St. Johns River and The Ocklawaha River (SJRWMD 2012a, b). The Hydrologic Simulation Program—FORTRAN (HSPF) and Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) software was used in the SJRWMD’s research. These models were developed for every basin and sub-basin along the St. Johns River. Readers are referred to the SJRWMD report for further details on the massive model development effort.

Using the existing hydrologic models developed by the SJRWMD, Kovalenko (2020) developed a new methodology to utilize the models for flood estimation. The technique incorporated the consideration of land-use change, varying precipitation frequency values, antecedent moisture conditions, and two rainfall distributions in the development of the 10-, 25-, 50-, and 100-year return frequency flood estimates in Northeast Florida. Limited research was previously conducted regarding the true impacts of land-use change in the flood estimation technique. Therefore, this research extends the original work by further investigating the impacts of future land-use conditions as well as seasonality.

**Study location**

This research focuses on the Black Creek, Julington Creek, Durbin Creek, Big Davis Creek, Ortega River, and Pablo Creek sub-basins of the Lower St. Johns River Basin located in Northeast Florida. The Lower St. Johns River Basin is shown in Fig. 1. These sub-basins of interest are encompassed within Clay County, St. Johns County and Duval County. All the sub-basins presented in Fig. 2, 3C, D, H, and I are pertinent to this research. These locations represent the physical boundaries of the models utilized in this research. Clay County encompasses 1668 square kilometers, including 111 square kilometers of water (FEMA 2014). St. Johns County comprises an area of 1577 square kilometers with approximately 109 square kilometers of Atlantic Ocean shoreline (FEMA 2011). Duval county consists of 2378 square kilometers and includes 373 square kilometers of water area (FEMA 2013). Nearly level to gently sloping terrain with poorly drained and sandy soils can be found in all three counties (FEMA 2011, 2013, 2014).
As previously mentioned, the sub-basins of interest in this research are part of the Lower St. Johns River Basin. The Lower St. Johns River basin is the northernmost basin along the St. Johns River. These sub-basins have been previously assigned a watershed unit number by the SJRWMD. The SJRWMD created model runs for each of the sub-basins. Table 1 presents all the sub-basins located in the Lower St. Johns River Basin along with the watershed unit number.
assigned by the SJRWMD and its model area in square kilometers. For the research program, the watershed unit numbers 3C, 3D, 3H, and 3I are of most interest and these are highlighted on Table 1 below. It is important to note the Julington Creek (watershed unit number 3H) contains two additional locations of interest that were assessed in this research—Durbin Creek and Big Davis Creek. These two additional sub-basins were modeled as individual reaches within the Julington Creek hydrologic model. Additionally, the Intracoastal Waterway model (watershed unit number 3I) contains the Pablo Creek sub-basin, which is the last location of interest for this research.

Methods

Hydrologic model background

The Hydrologic Simulation Program—FORTRAN (HSPF) and Better Assessment Science Integrating Point and Non-point Sources (BASINS) software was used in this research study. The original HSPF models incorporate either physical or empirical model input parameters (SJRWMD 2012a, b). The physical parameters are watershed areas, land use, precipitation, evaporation, slope, roughness, infiltration, and system hydraulics (SJRWMD 2012a, b), among others.

The models were originally calibrated by the SJRWMD using specific streamflow gages per sub-basin. Table 2 depicts which United States Geological Survey (USGS) streamflow gages were used to calibrate each sub-basin of interest. The Black Creek sub-basin was calibrated with the North Fork gage and South Fork gage. The calibration using the North Fork gage was described as overall very good and the calibration using the South Fork gage was described as overall good (SJRWMD 2012a, b). The 02246318 (Kirwin Rd.) Ortega River gage calibration was unsatisfactory, and 02246300 (103rd St.) Ortega River gage calibration was overall good (SJRWMD 2012a, b). The Big Davis Creek gage 02246150 calibration was adequate and lastly, the Pablo Creek gage 02246828 calibration was reasonable (SJRWMD 2012a, b). During the calibration process, the observed gage data was directly compared to “synthetic” gage data produced from the model output. SJRWMD (2012a, b) determined that the calibration performance results for the Nash–Sutcliffe statistic (NSE) of the WSIS watersheds show that the calibrated model is rated “good” or “very good for 30 of 39 flows and “unsatisfactory” for only 2 of 39 flows.

One of the major requirements of the HSPF model is the incorporation of existing meteorological data. The SJRWMD maintains point rain gauge and Doppler radar rainfall datasets (SJRWMD 2012a, b). The Doppler radar data set provided only 13 years of rainfall data, whereas the National Weather Service (NWS) stations provided data dating back to the early 1900s. Due to the need to run long term simulations, the NWS rain gauge data was selected for the WSIS model. The model scenario simulations run from 1975 through 2008 (SJRWMD 2012a, b). The SJRWMD implemented a Theissen polygon network to establish the area of influence for the NWS rain gauges. The rain gauges of interest will be discussed in further detail in the upcoming sections of this manuscript. The temporal resolution for the model variables is daily and hourly for the precipitation, temperature, and ET.

As previously explained, the locations of interest are modeled within four separate watersheds in HSPF. Figure 3 shows a comprehensive hierarchy of the model locations of interest. The reaches are the main points of interest and are the locations where flood estimates were developed in this research. Table 3 depicts modeled area ($km^2$) of each location of interest.

Statistical analysis

The statistical analysis conducted included multiple methods as suggested by literature including the Log-Pearson Type III (LP3) approach and two flavors of the Power Law (PL) distribution approach namely a linear fit and a non-linear fit. The LP3 distribution was selected because it is the official methodology for analyzing hydrometeorological data used by the United States Army Corps of Engineers (USACE 1994). There is also an abundance of sources providing supporting evidence that the LP3 is a reliable distribution for flood frequency analysis. Kumar (2019), Saf et al. (2007), and Kidson and Richards (2005) all provided substantial evidence of the reliability of the LP3 distribution. The PL distribution was selected because of its increase in popularity for predicting extreme natural events (Alipour et al. 2016). Kidson and Richards (2005) state that it is arguably one of the simplest probability distributions. They (Kidson and Richards, 2005) demonstrated the successful implementation of the PL for flood estimation in Thailand. Alipour et al. (2016) demonstrated the successful use of the PL in region 3 (entire Florida, almost entire Alabama, Georgia, South Carolina, North Carolina, and part of Virginia and Mississippi) and region 8 (Mississippi, Louisiana, Arkansas, and Tennessee) of the United States.

The research team used these methods and applied them to both real and synthetic data (e.g. model-derived datasets). Real flood data was obtained from the stream flow gages that were used to calibrate the models. Synthetic gage streamflow data was obtained from the HSPF model output (if synthetic gage data was available). In addition, synthetic flow data was also obtained from the flow output of each reach of interest outlined in Table 3 thus providing a large matrix of statistical flood estimates.
Hydrologic modeling

The original SJRWMD HSPF models were modified to produce the 10-, 25-, 50-, and 100-year flows in the sub-basins of interest. The general procedure associated with these modifications involves (1) selection of land-use scenario, (2) determination of target date for rainfall simulation, (3) determination of precipitation frequency value for simulation, (4) addition of antecedent moisture conditions, and (5) consideration of rainfall distribution during simulation. The proposed procedure itself is novel and could be adopted in other watersheds of interest along the St. Johns River where HSPF models are already available.

The numerical model simulations investigated the following scenarios and variables:

- Rainfall amount for different frequencies and uncertainty threshold (median value and 90% confidence interval amount);
- Rainfall distribution (e.g. constant rainfall intensity or more natural distribution);
- Antecedent moisture conditions;
- Flow starting condition in each sub-basin; and,
- Land-use boundary (e.g. either 1995 or projected 2030 land-use).

To produce the 10-, 25-, 50-, and 100-year flood flows, a 10-, 25-, 50-, and 100-year frequency precipitation (24-h duration) event was simulated. The simulated flood flows are considered to be rainfall driven. The process of simulating any given frequency precipitation event involves identifying the appropriate amount of rainfall to simulate on a specific date, referred to as a target date in this research. Before identifying the precipitation frequency values to simulate in the model, the target dates were determined.

The first step in selecting the appropriate target dates is to reference the output data generated from the original models. To do this, the original models obtained from the SJRWMD were simulated. The simulated flow data output values were obtained in graphical form. Next, the 50th percentile flow (±15%) was identified from the flood frequency curve, which depicts exceedance probability vs. flow. The 50th percentile flow is a standard baseline for various flood frequency analysis procedures (Malamud and Turcotte 2006). Once the 50th percentile flow was obtained, the period of record was observed. Ten (10) target dates were selected from the period of record on which the 50th percentile flow occurred. The ten (10) selected target dates occurred in varying months of the year to account for the varying rainfall conditions to consider average seasonality. This process was repeated for every sub-basin of interest. Originally, the possible seasonality of the flood flows was not investigated in Kovalenko (2020), but the additional work completed this year rectified that oversight. The modeling of seasonality is discussed in “Discussion” section.

Now that there is understanding on which specific dates the 10-, 25-, 50-, and 100-year rainfall events were simulated, a discussion on specifically how much rainfall to simulate is presented. Using rain gauges previously used in the WSIS study, precipitation amounts were chosen using both the median values and 90% values from the National Oceanic Atmospheric Administration (NOAA) Atlas 14 (NOAA 2017).

Table 4 depicts the median and 90th percentile 24-h 10-, 25-, 50-, and 100-year rainfall events at each sub-basins’ corresponding rainfall gage.

The SJRWMD defines the antecedent soil moisture conditions as an indicator of watershed wetness and soil storage availability before a storm (SJRWM 1985). It is known that these conditions have a significant effect on runoff volume and runoff rate. Three levels of antecedent moisture conditions (AMC) exist: AMC-I for dry, AMC-II for normal, AMC-III for wet conditions (SJRWM 1985). Table 5 depicts the seasonal rainfall limits for these three AMCs. According to Schiariiti (n.d.), AMC II is considered for modeling purposes because it is essentially the average moisture condition. According to the SJRWMD’s Technical Publication SJ90-3 (1990), the rainy season is in northeast Florida lasts from June to October and the dry season lasts from November to May (1990). The simplest method that Kovalenko (2020) discovered to simulate antecedent moisture conditions (AMC) in the existing HSPF models was to simulate it as daily rainfall. As depicted in Table 5, the AMC was simulated over the course of five days. Therefore, to summarize the application of AMC, 5.3 cm of rainfall over the course of five days were simulated before each target date.

When applied on an hourly basis, the rainfall data described in Table 5 represents a uniform distribution—each hour will receive the same amount of rainfall. However, in reality, rainfall occurs in varying temporal distributions. Therefore, an additional rainfall distribution was also applied in a separate scenario. According to Suphunvorranop (1985), the Soil Conservation Service (SCS) has developed four types of rainfall distributions—Types I, II, III (representative of different climates in the United States), and Type II Modified (representative of Florida specifically). These synthetic rainfall distributions occur over a 24-h time period. The SCS Modified Type II rainfall distribution, obtained from the Suphunvorranop (1985), was also modeled to determine the effects on the flood magnitude predictions when different rainfall distributions are applied. It was assumed that the combination of extreme precipitation, median starting flows, and wet antecedent moisture conditions would produce extreme flows in the various streams simulated.
Results

Existing flood insurance studies

Existing flood estimates were obtained from the Federal Emergency Management Agency’s (FEMA) Flood Insurance Studies (FIS). The FIS contain flood-risk data that aids in the development of flood insurance rates for communities’ efforts of safe floodplain management (FEMA 2014). Between Clay County, St. Johns County, and Duval County, three different methodologies were implemented for flood-risk analysis. The FEMA FIS estimates were adjusted accordingly to represent the modeled drainage basin area (Kovalenko 2020). In some cases, the HSPF modeled drainage basin area did not match the FEMA FIS drainage basin area from which the results were obtained. Therefore, a simple extrapolation was conducted to standardize the FEMA FIS estimates’ drainage basin area to match the drainage basin area of the HSPF modeled locations for more accurate comparison.

Clay County implemented the Magnitude and Frequency of Flood Discharges in Northeast Florida from the Technical Publication SJ-86–2 (Rao 1986) to conduct their hydrologic analyses and obtain their flood estimates. The methodology involves the determination of different return periods (T). Given an annual exceedance probability of a maximum event, the return period is defined as $T = 1/P$, where $P$ is the annual exceedance probability of a maximum event (Rao, 1986). The technique consists of applying an appropriate probability distribution and fitting it to sample data, where the sample data consists of observed annual peak flows (Rao 1986). From there, probability distributions such as the two-parameter Gumbel distribution and the five-parameter Wakeby distribution were implemented (Rao 1986). The Log-Pearson Type III distribution was also recognized and used as it was recommended by the United States Water Resource Council (Rao 1986).

St. Johns County Flood Insurance Study is a compilation of previously printed FIS reports (FEMA 2011). The 2003 countywide analyses were conducted by the United States Army Corps of Engineers using the HEC-1 computer program. The methodology was deemed appropriate for the characteristic drainage basin conditions; however, it was determined that the limited history of stream gage records prevented effective statistical analysis (FEMA 2011). The HEC-1 models incorporated the Natural Resources Conservation Service (NRCS) unit hydrograph and kinematic wave routing methods, sub-basin runoff curve numbers, lag times, stream cross sections, and Manning’s “$n$” roughness factors (FEMA 2011). United States Geological Survey (USGS) topographic maps, field inspection, and aerial photos were utilized. Additionally, the modified SCS Type II rainfall distribution was implemented into the model. Overall, only Durbin Creek and Sixmile Creek were adequately calibrated due to a lack of sufficient stream gage data.

Duval County utilized the Environmental Protection Agency Stormwater Management Model (EPA SWMM5) versions 12 to 14 to conduct hydrologic analyses (FEMA 2013). The model applied precipitation across hydrologic units and performed hydrologic calculations, which account for hydrologic unit geometry, land use, and soil characteristics. The computed surface runoff hydrographs were routed to the dynamic hydraulic model.

These FIS reports were used as a reputable baseline for reasonable flood estimate values to expect for results obtained through the statistical analysis and hydrologic modeling approaches of Kovalenko (2020).

Initial Flood Estimates.

The purpose of this section is to provide the reader with an understanding of the variety of flood estimates and scenarios that could be obtained using the method outlined in Kovalenko (2020) and to compare it to the current research. Tables 6 and 7 depict 25- and 100-year return frequency flood estimates, respectively.

Research findings and conclusions

The research of Kovalenko (2020) produced mostly reliable flood frequency estimates in the sub-basins of interest. The resulting peak flood flows in each model run typically occurred one to two days after the Target Date that was modified. However, the resulting peak flow occurred directly on the Target Date in several model runs. It is believed that when the peak flow occurred on the Target Date, the daily timestep of the model was potentially smoothing the peak to some degree. Since the original HSPF models were created by various individuals at the SJRWMD, there is a possibility that there are slight differences in modeling approaches between sub-basins. The Black Creek sub-basin naturally produces higher flow rates compared to the other basins that were studied. Because the Black Creek sub-basin contains naturally higher flow rates, the simulation of additional precipitation frequency values resulted in the model crashing consistently. Attempts to resolve these problems in HSPF were not successful. Luckily, other watersheds did not exhibit similar issues. Like the observation made by Gebremariam et al. (2014), source-code modification for the HSPF model was challenging primarily because of lack of documentation related to code structure and subroutines. The most complete set of flood estimates obtained in the Black Creek sub-basin were the 10- to 50-year flood estimates using the 1995 land-use condition, median precipitation frequency values, uniform rainfall distribution, and antecedent...
moisture conditions. The HSPF model run flood estimates were generally lower than the FEMA FIS estimates.

The model runs in the Julington Creek, Durbin Creek, and Big Davis Creek sub-basins produced satisfactory results. In this area, land-use variation, variation of precipitation frequency values, and variation of rainfall distribution were assessed. Overall, the 2030 land-use paired with the 90th percentile precipitation frequency, uniform rainfall distribution, and antecedent moisture conditions produced the highest flood flows; however, these relatively high flood flow values were still lower than the adjusted FEMA FIS estimates for those locations. Although the 2030 land-use runs were successfully simulated in these basins and produced relatively high flood flows, because of the complications involved in running the 2030 land-use runs in the Black Creek sub-basin, the 1995 land-use became the default for the remainder of the sub-basins. It is also important to note that the mouth of Julington Creek is more strongly tidally influenced and could be influenced by storm surge, which makes hydrologic modeling in those locations more challenging.

The Ortega River and Pablo Creek sub-basin HSPF models ran well. The 1995 land-use, 90th percentile precipitation frequency values, uniform distribution, and antecedent moisture conditions were considered in the model runs. Ortega River and Pablo Creek are also more heavily tidally influenced at the mouth and could experience a greater deal of storm surge, which poses additional challenges to the hydrologic modeling process. Additionally, it was previously observed that the FEMA FIS drainage areas were usually larger than the drainage areas in the HSPF models.

A possible reason for the difference in area is the consideration of tidal influence. The hydrologic HSPF models do not consider the tidal areas of each sub-basin. The tidal areas of the project locations were instead modeled in the hydrodynamic portion of the model, which is not part of the research.

In a case study conducted by Ninov et al. (2008), the results of their HSPF modeling for flood assessment yielded flood flows that were 130% higher than the historical flood flows, while the modeled annual, seasonal, and low flows were approximately 25–33% less than the observed respectively. This case study provides a perspective on the variety of results that can be obtained with the use of the HSPF model. This is a valuable perspective to consider when developing a flood estimation methodology. The research of Ninov et al. (2008) produced flood flows that were too high while the results of the HSPF flood modeling presented in Kovalenko (2020) appeared lower compared to FEMA FIS estimates. The results and certainty of the HSPF model simulations cannot be concretely proven and serve only as estimates.

Overall, the Log-Pearson Type III (LP3) statistical computations were successful. There trends found in the data were conclusive and expected. This was the expected outcome of the LP3 results as it is a common method for flood frequency estimation. The Power Law (PL) statistical computations were mostly successful. The PL derived flood estimates were typically much higher than the LP3 results and the HSPF modeled results. The PL derived flood estimates were even higher than the adjusted FEMA FIS estimates at times. The PL distributions produced more reasonable estimates for the 10- and 25-year flood estimates. The PL distribution seemed to massively overestimate the 50- and 100-year flood flows. The PL was selected because of the praise it received in various research studies for being a simple and effective method. However, there was certainly a caveat that the PL performs best with a larger data set (Kidson and Richards, 2005). The difference in computing the PL distribution regression coefficients using the Linear Regression model and Nonlinear Regression model also produced varying results. It was evident that the data sets were better suited for one method over the other in certain cases.

The use of the FEMA FIS estimates proved to be an asset as the estimates were derived by qualified professionals. There is more certainty in comparing the estimated developed herein to the FEMA FIS estimates. However, it has also been established that the FEMA FIS estimates were all obtained using varying methods. This raises a question regarding the consistency of the FEMA FIS estimates. It was evident that different methodologies can at times produce widely varying flood flows. As previously established, the FEMA FIS estimates are either based on hydrologic modeling or statistical estimates. Research conducted by Okoli et al. (2019), compared statistical and hydrological methods for the estimation of design floods based on 10,000 years of synthetically generated weather and discharge data. Although their hydrologic modeling did not reflect any real applications and was intended as a baseline for discussion for comparison of results, their ultimate findings suggest that more than one flood estimate should be obtained and the maximum value (within reason) should be selected to minimize the likelihood of underestimating the design flood (Okoli, 2019). As part of the research conducted, the various flood estimates were compiled together irrespective of method used to develop them. The estimates at a particular return frequency were then evaluated to discern their underlying statistical distribution. First, a normal distribution was assumed. This worked well in Ortega River Reach 3 and Julington Creek Reach 2 and 5 for the 10-year through 50-year flood events, but overall was not satisfactory since the lower tail of the distribution can lead to negative flow rate estimates for floods which is physically impossible. Alternatively, it was assumed that the underlying distribution of the flood estimates was log-normal, but this led to a narrow band that did not fit the outliers well. Ultimately, a bounded normal distribution assumption worked well as a
first iteration to more realistically portray the range of values. For example, for the Durbin Creek Reach 1, 100-year flood estimates shown in Table 7, a mean of 116.83 m^3s^{-1} and a sample standard deviation of 61.75 were calculated and the estimate range was bounded by 51 m^3s^{-1} as a minimum and 203 m^3s^{-1} as a maximum. Using a Monte-Carlo approach, this distribution was simulated 10,000 times and the 90% value (~178 m^3s^{-1}) was selected as an improved “best estimate” but with a better understanding of the considerable uncertainty. It should be noted that this improved best estimate exceeded the recommended flood insurance value.

Therefore, this work has demonstrated the possibility for obtaining a wide range of flood estimates depending on the estimation method selected. Kovalenko (2020) concluded that the establishment of extreme flood estimates based on one methodology is outdated and involves higher risk in the development of planning measures for flood protection. Therefore, this work advocates for consideration of various techniques when developing flood estimates. As presented in Kovalenko (2020), consideration of seasonality was not incorporated when developing the HSPF model runs and the final results of all methods assessed indicated that considering the 2030 land-use condition could increase the magnitude of the simulated flood flows and perhaps provide better numerical estimates in each sub-basin.

**Evaluation of land-use change**

The hydrologic modeling technique implemented in this research is in accordance with the hydrologic modeling technique developed by Kovalenko (2020). Table 8 provides an overview of the parameters included in the hydrologic models developed in new effort. The 2030 land-use condition was used for all model runs in this research. The 90th percentile confidence interval precipitation point value was simulated at each sub-basin of interest. The precipitation frequency values were obtained using the methodology described in Kovalenko (2020) and summarized earlier in this manuscript. The Type II Antecedent Moisture Condition (AMC) was incorporated into model runs. Lastly, the uniform rainfall distribution was applied when incorporating the precipitation frequency values and the AMC. According to research conducted by Kovalenko (2020), modeling a uniform rainfall distribution versus varying temporal distributions does not have a significant effect on the resulting flood flows in the locations of interest.

Contrary to the earlier research, this research considers the effects of seasonality when simulating flood flows. For this set of simulations, the team chose to simulate ten (10) Target Dates from varying months and developed an average flood flow magnitude from all ten (10) estimates. This research considers five (5) Target Dates from the rainy season and five (5) Target Dates from the dry season in each sub-basin. The flood flows obtained from simulating rainfall in the wet season are compared to the flood flows obtained from simulating rainfall in the dry season in this research.

**Discussion**

Tables 9, 10, 11 show the 25-year and 100-year flood estimates at Ortega River and Pablo Creek. Figs. 4, 5, 6, 7, 8 represent a comparison of the 1995 land-use HSPF model runs, 2030 land-use HSPF model runs, and the adjusted FEMA FIS estimates. A graphical representation of the flood estimates at Black Creek was not created because of the lack of data. As discussed previously, the modeling capabilities in the Black Creek sub-basin were limited due problems encountered modifying the source code to produce the flood estimates desired. Therefore, only the 10-year flood estimates (2030 land-use and seasonality included) were obtained at Black Creek. The adjusted FEMA FIS estimates discussed previously were included in the tables and graphs.
for the reader for comparison purposes. After all, the FEMA FIS estimates are highly revered and professionally derived estimates. The comparison graphs shown in Figs. 4, 5, 6, 7, 8, depict that the 2030 land-use condition paired with the 90th percentile precipitation values, and antecedent moisture conditions produce higher flood flows than the 1995 land-use condition and the FEMA FIS estimates in Julington Creek Reach 2 and Ortega River Reach 3; while all other locations
depict that the FEMA FIS estimates produce higher flows than the other approaches. Generally, the flood estimates increased in magnitude from the FEMA FIS estimate, to the 1995 land-use HSPF model-derived estimates, to the 2030 land-use scenario model-derived estimates. However, in the sub-basins where the highest flood flow estimate was that of the FEMA FIS estimate, there are indications that that sub-basin is likely more tidally influenced. Therefore, the HSPF models may not be the most suitable route for flood estimation in those locations since the hydrodynamic portion of the model is not represented.

Overall, the results of this research point towards a noticeable difference between consideration of the 1995 land-use condition and the 2030 land-use condition when implementing Kovalenko’s (2020) flood estimation technique using the SJRWMD hydrologic models. Consideration of seasonality in selecting Target Dates for precipitation simulation in the hydrologic models also resulted in various flood flows. It was observed that simulating the 2030 land-use condition 10-, 25-, 50-, and 100-year return frequency floods in the dry season produced lower flood flows than simulating in the wet season which was expected. The analysis considered a range of historical and extreme storm events including both precipitation and flood flows, which were built into the HSPF model runs.

Besides the more in-depth evaluation of the effect of urbanization and land-use change, this additional work compares the effects of wet and dry seasonal precipitation

### Table 1 Sub-basins belonging to the Lower St. Johns River Basin and their model area (km²)

| Watershed Unit Number | Sub-basin Name       | Model Area (km²) |
|-----------------------|----------------------|------------------|
| 3A                    | Crescent Lake        | 1542             |
| 3B                    | Etonia Creek         | 924              |
| 3C                    | Black Creek          | 1316             |
| 3D                    | Ortega River         | 271              |
| 3E                    | Trout River          | 248              |
| 3F                    | Deep Creek           | 358              |
| 3G                    | Sixmile Creek        | 331              |
| 3H                    | Julington Creek      | 252              |
| 3I                    | Intracoastal Waterway| 268              |
| 3J                    | South Main Stem      | 997              |
| 3K                    | North Main Stem      | 630              |

### Table 2 Rainfall gages used to calibrate the original HSPF model runs (SJRWMD 2012a, b)

| Sub-basin          | Gage name               | Gage ID   | Notes                        |
|--------------------|-------------------------|-----------|------------------------------|
| Black Creek        | North Fork near Middleburg | 02246000 |                              |
|                    | South Fork Penney Farms  | 02245500 |                              |
| Julington Creek    | Big Davis Creek at Bayard| 02246150 |                              |
| Durbin Creek       | Big Davis Creek at Bayard| 02246150 |                              |
| Big Davis Creek    | Big Davis Creek at Bayard| 02246150 |                              |
| Ortega River       | Ortega 103rd St. Bridge  | 02246300 | Discontinued in 2003         |
|                    | Ortega at Kirwin Rd     | 02246318 | Replacement gage             |
| Pablo Creek        | Pablo Creek             | 02246828 | Discontinued in 2002         |
frequency induced rainfall events to obtain flood estimates. All remaining sub-basins that were examined (Black Creek, Ortega River, and Pablo Creek) exhibited an increase in flood flows when the precipitation frequency events were modeled during the wet season.

Sahu et al. (2017) deduced that the most accurate forecast of flood prediction should incorporate seasonal streamflow predictions utilizing hydrologic models which are calibrated for individual river basins. In the case of the methodology discussed in this paper, the pre-existing SJRWMD HSPF models of the St. Johns River Basins have been calibrated sufficiently such that their use should be valid. With the incorporation of seasonality and assurance of sufficient calibrations per each sub-basin, individuals who chose to implement the flood estimation technique presented herein should be producing reasonable flood estimates for most sub-basins. Additionally, Alipour et al. (2016) determined that incorporating seasonality into the hydrologic modeling procedure can improve the magnitude of flood estimates. This is especially true in the results obtained in this research. The hydrologic models which were simulated during the wet season produced higher flood flows. Although the wet season scenario produced higher simulated flood flows, the consideration of seasonality did not truly yield dramatic differences. However, according to others, the incorporation of seasonality into hydrologic flood modeling is widely recommended. Although the result of these sub-basins may have not indicated the greatest variability between wet and dry season derived flood estimates, the inclusion of seasonality in flood estimation is encouraged based on other relevant sources.

As urban expansion increases with the growth of the human population, land-use change is inevitable in humanity’s future. The results of this research indicate that the consideration of a more urbanized land-use has the potential to produce higher flood flows. This research demonstrates

| Table 3 Model area of each reach associated with each sub-basin of interest (SJRWMD 2012a, b) |
|-----------------------------------------------|
| WSIS Sub-basin Name | Reach of interest ID | Reach of interest name | Model area (km²) |
| Black Creek | 3 | South Fork | 114 |
| | 6 | Down North Fork | 15 |
| | 12 | Black Out | 43 |
| Ortega River | 3 | Ortega Riv Dns | 49 |
| Julington Creek | 1 | Durbin Creek | 104 |
| | 4 | Julington Creek | 25 |
| | 6 | Big Davis Creek | 4 |
| Intracoastal Waterway | 8 | Pablo Creek Down S | 10 |

Table 4 Rainfall frequency values (cm) applied on each target date based on appropriate rainfall gages

| Sub-basin | Rainfall Gage | 24-h 90th Percentile Rainfall for each recurrence interval (years) | 24-h Median Rainfall for each recurrence interval (years) |
|-----------|---------------|---------------------------------------------------------------|---------------------------------------------------------------|
| Black Creek | Jax Airport | 20.5 | 27.4 | 32.8 | 39.6 | 17.9 | 22.5 | 26.4 | 31.0 |
| | Glen St. Mary | 18.9 | 24.3 | 28.4 | 33.3 | 16.0 | 19.5 | 22.6 | 25.9 |
| | Starke | 18.1 | 23.2 | 26.9 | 31.8 | 14.8 | 18.1 | 21.0 | 24.2 |
| | Federal Point | 19.0 | 25.0 | 29.5 | 35.1 | 16.1 | 20.1 | 23.5 | 27.4 |
| | Palatka | 18.4 | 23.9 | 27.9 | 33.0 | 15.7 | 19.3 | 22.4 | 25.9 |
| | Average | 19.0 | 24.7 | 29.2 | 34.5 | 16.1 | 19.9 | 23.2 | 26.9 |
| Julington Creek | Jax Beach | 22.2 | 29.5 | 35.1 | 41.7 | 18.6 | 23.4 | 27.4 | 32.0 |
| | St. Augustine | 20.8 | 28.4 | 34.0 | 41.1 | 17.7 | 22.6 | 26.9 | 31.5 |
| | Average | 21.5 | 29.0 | 34.5 | 41.4 | 18.2 | 23.0 | 27.2 | 31.8 |
| Durbin Creek | Jax Beach | 22.2 | 29.5 | 35.1 | 41.7 | 18.6 | 23.4 | 27.4 | 32.0 |
| | St. Augustine | 20.8 | 28.4 | 34.0 | 41.1 | 17.7 | 22.6 | 26.9 | 31.5 |
| | Average | 21.5 | 29.0 | 34.5 | 41.4 | 18.2 | 23.0 | 27.2 | 31.8 |
| Big Davis Creek | Jax Beach | 22.2 | 29.5 | 35.1 | 41.7 | 18.6 | 23.4 | 27.4 | 32.0 |
| | St. Augustine | 20.8 | 28.4 | 34.0 | 41.1 | 17.7 | 22.6 | 26.9 | 31.5 |
| | Average | 21.5 | 29.0 | 34.5 | 41.4 | 18.2 | 23.0 | 27.2 | 31.8 |
| Ortega River | Jax Airport | 20.5 | 27.4 | 32.8 | 39.6 | 17.9 | 22.5 | 26.4 | 31.0 |
| Pablo Creek | Jax Beach | 22.2 | 29.5 | 35.1 | 41.7 | 18.6 | 23.4 | 27.4 | 32.0 |
the 1995 land-use condition in comparison to the 2030 land-use condition. Black Creek, Julington Creek, and Durbin Creek experienced a difference of up to the second order of magnitude between the flood estimates derived from the 1995 land-use and the 2030 land-use scenario models. Choi and Deal (2008), confirmed in their research that land-use change involved with urban growth can result in higher surface flows and that low flows tend to decrease. Baloch et al. (2014) not only considered land-use change but also climate change in their predictive hydrologic modeling efforts.

### Table 6 All 25-year return frequency results (Kovalenko 2020) in cubic meters per second

| Sub-Basin     | Method          | Reach Location | HSPF 1995, 90th Percentile Precip., AMC | LP3 PL Linear Regression | PL Nonlinear Regression | FEMA FIS | HSPF 1995, Median Precip., AMC |
|---------------|-----------------|----------------|-----------------------------------------|--------------------------|------------------------|----------|----------------------------------|
| Black Creek   | Reach 3         | NA             | 234                                     | 328                      | 272                    | 348      | 251                              |
| South Fork Gage | NA             | 237            | 364                                     | 230                      | NA                     | NA       | NA                               |
| Reach 6       | NA             | 328            | 553                                     | 344                      | 342                    | 265      | NA                               |
| North Fork Gage | NA             | 325            | 581                                     | 324                      | NA                     | NA       | NA                               |
| Reach 12      | NA             | 405            | 600                                     | 481                      | 856                    | 573      | NA                               |
| Julington Creek | Reach 2      | 69             | 63                                      | 95                       | 83                     | 93       | 48                               |
|               | Reach 5        | 23             | 27                                      | 40                       | 31                     | 72       | 15                               |
| Durbin Creek  | Reach 1        | 44             | 40                                      | 69                       | 49                     | 109      | 33                               |
| Big Davis Creek | Reach 6       | 20             | 26                                      | 43                       | 31                     | 17       | 12                               |
|               | Synth Gage     | 16             | 17                                      | 29                       | 19                     | 44       | 11                               |
|               | Real Gage      | NA             | 26                                      | 42                       | 30                     | NA       | NA                               |
| Ortega River  | Reach 3        | 135            | 194                                     | 253                      | 213                    | 116      | NA                               |
|               | 103rd St. Synth Gage | 88     | 90                                      | 126                      | 91                     | 70       | NA                               |
|               | Kirwin Real    | 70             | 86                                      | 78                       | 119                    | NA       | NA                               |
|               | Kirwin Rd. Synth Gage | 132   | 101                                     | 124                      | 110                    | NA       | NA                               |
| Pablo Creek   | Reach 8        | 76             | 75                                      | 109                      | 94                     | 167      | NA                               |
|               | Pablo Real Gage| NA             | 37                                      | 68                       | 47                     | NA       | NA                               |

### Table 7 All 100-year return frequency results (Kovalenko 2020) in cubic meters per second

| Sub-Basin      | Method          | Reach Location | HSPF 1995, 90th Percentile Precip., AMC | LP3 PL Linear Regression | PL Nonlinear Regression | FEMA FIS | HSPF 1995, Median Precip., AMC |
|----------------|-----------------|----------------|-----------------------------------------|--------------------------|------------------------|----------|----------------------------------|
| Black Creek    | Reach 3         | NA             | 333                                     | 718                      | 515                    | 531      | NA                               |
| South Fork Gage | NA             | 346            | 1,086                                   | 452                      | NA                     | NA       | NA                               |
| Reach 6        | NA             | 447            | 1,568                                   | 654                      | 480                    | NA       | NA                               |
| North Fork Gage | NA             | 466            | 1,791                                   | 652                      | NA                     | NA       | NA                               |
| Reach 12       | NA             | 458            | 1,166                                   | 787                      | 1,268                  | NA       | NA                               |
| Julington Creek | Reach 2        | 132            | 87                                      | 230                      | 180                    | 125      | 82                               |
|               | Reach 5        | 41             | 43                                      | 106                      | 68                     | 96       | 26                               |
| Durbin Creek   | Reach 1        | 86             | 56                                      | 203                      | 145                    | 160      | 51                               |
| Big Davis Creek | Reach 6       | 29             | 37                                      | 117                      | 62                     | 32       | 23                               |
|               | Synth Gage     | 31             | 24                                      | 88                       | 40                     | 63       | 20                               |
|               | Real Gage      | NA             | 39                                      | 122                      | 65                     | NA       | NA                               |
| Ortega River   | Reach 3        | 212            | 238                                     | 446                      | 329                    | 175      | NA                               |
|               | 103rd St. Synth Gage | 121  | 148                                     | 371                      | 203                    | 106      | NA                               |
|               | Kirwin Real    | NA             | 90                                      | 166                      | 144                    | 182      | NA                               |
|               | Kirwin Rd. Synth Gage | 189  | 134                                     | 245                      | 205                    | NA       | NA                               |
| Pablo Creek    | Reach 8        | 98             | 107                                     | 150                      | 203                    | 249      | NA                               |
|               | Pablo Real Gage| NA             | 49                                      | 196                      | 102                    | 200      | NA                               |
Although climate change was not considered in this research, it is important to note that not only did Baloch et al. (2014) observe an increase in flood flows when climate change was considered, but they also confirmed that the effects of land-use change resulted in a compounding effect on flood flows. Their research confirmed that the magnitude, frequency, and duration of flood flows are significantly modified as urbanization increases. Therefore, it is further evident that land use is a crucial component of accurate extreme flood estimation for the future.

As was mentioned earlier in this paper, the underlying distribution for each range of flood flow estimates was of great interest in this research effort. Using the NIST/SEMATECH e-Handbook of Statistical Methods (2012), the Shapiro–Wilk test was conducted on all results obtained by all methods from this research regarding Reach 3 of the Ortega River sub-basin. The Shapiro-Wilk test revealed that the 10-year, 25-year, and 50-year flood frequency estimates were normally distributed because the determined P-values were greater than alpha (0.05). However, the Shapiro-Wilk test did not indicate a normal distribution of the 100-year flood estimates of all methods applied at Ortega Reach 3 because the P-values were less than alpha (0.05). For the best understanding of the results at all sub-basins, testing for normality should be conducted. Although only Reach 3 at Ortega River was tested for normality, it is evident that these results indicate variance between testing methods and between flood magnitude. In the case of Ortega River Reach 3, the 100-year flood estimate was not normally distributed and that is expected because there is greater uncertainty involved in flood estimates of larger magnitude. As was determined previously and confirmed with the normality tests utilized, a truncated or bounded normal distribution appears to be a reasonable underlying distribution to use when evaluating and using a range of flood flow estimates instead of just the one, best estimate. It was suggested earlier that using this approach combined with a Monte Carlo simulation, it is possible to perhaps derive a superior “best estimate”. For this work, the team adopted the 90% from the Monte Carlo simulation which is consistent with many engineering reliability studies.

## Conclusions

This research assessed the significance of seasonality and land-use change when developing hydrologic flood estimating models in HSPF for the 10-, 25-, 50-, and 100-year flood...
frequency estimates at Black Creek, Julington Creek, Durbin Creek, Big Davis Creek, Ortega River, and Pablo Creek. Although the model runs developed in the wet season versus the dry season did not vary significantly in the sub-basins involved in this research, consideration of seasonality is still recommended based on other available literature (Alipour et al. 2016). The 2030 land-use scenario models resulted in higher flood flows than the 1995 land-use scenario models. Each sub-basin varied in regard to the significance that land-use variation played in the model runs. The 2030 land-use scenario produced higher flood flows than the 1995 land-use scenario in all locations except for Pablo Creek, where the two produced almost identical results. In all locations except for Ortega Reach 3 and Julington Creek Reach 2, the adjusted FEMA FIS estimate produced the greatest flood estimate. The methodology developed by Kovalenko (2020) has its limitations. The locations of the flood estimates developed in this research are closely tied to the tidal regions of the St. Johns River. Since the HSPF model does not directly consider the impacts of tidal behavior, a level of uncertainty is present regarding the 10-, 25-, 50-, and 100-year flood flows developed using this methodology. Since this research only assessed sub-basins in the Lower St. Johns River Basin, there is much opportunity remaining to assess the remaining in-land sub-basins of the St. Johns River that are not heavily influenced by the tides. Another opportunity for further research involves the assessment of the hydrodynamic activity impacts to flood estimation in the tidal regions of the Lower St. Johns River Basin. Overall, the HSPF models developed by the SJRWMD serve as a valuable tool for developing flood estimates in precise locations along the various sub-basins of the St. Johns River and it is recommended to incorporate seasonality and land-use change for most accurate development of flood estimates. To better account for the uncertainties associated with traditional methods of flood estimation, this research demonstrates that the superior flood estimate may possibly be obtained through the implementation of multiple estimation methods paired with a Monte Carlo simulation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12665-021-10038-5.

Acknowledgements Samantha Kovalenko would like to acknowledge the extensive support of the St. Johns River Water Management staff for their valuable input and for making their existing HSPF models available for use in this research effort.

Authors’ Contributions SK completed the data analysis, writing, and revisions. Dr. Brown assisted with data analysis, review, and editing. Drs. Akan and Schonning assisted with editing and review.

Funding Funding was provided by Dr. Chris Brown and by the UNF Environmental Center.

Availability of data and material Supplementary data has been uploaded.

Code availability The HSPF code for each sub-basin is available from the St. Johns River Water Management District (SJRWM).

Declarations

Conflicts of interest There are no known conflicts of interest.

Ethics approval All work presented herein was completed by the research team listed. Portions of the work was published in the Master’s Thesis of Kovalenko (2020) and in Kovalenko et al. (2020) and have been cited in the paper as such. The Master’s Thesis work is summarized and discussed thoroughly herein.

Consent to participate Consent is granted to share this research with the journal audience.

Consent for publication All necessary approvals are granted for manuscript publication.

References

Alipour M, Rezakhani A, Shamsai A (2016) Seasonal fractal-scaling of floods in two U.S. water resources regions. J Hydrol 540:232–239. https://doi.org/10.1016/j.jhydrol.2016.06.016

Baloç M, Ames D, Tanik A (2014) Hydrologic impacts of climate and land-use change on Namnam Stream in Koycegiz Watershed, Turkey. Int J Environ Sci Technol 12:1481–1494. https://doi.org/10.1007/s13762-014-0527-x

Brody SD, Zahran S, Maghelal P, Grover H, Highfield WE (2007) The rising costs of floods: examining the impact of planning and development decisions on property damage in Florida. J Am Plann Assoc 73(3):330–345. https://doi.org/10.1080/01944360708977981

Choi W, Deal B (2008) Assessing hydrological impact of potential land use change through hydrological and land use change modeling for the Kishwaukee River basin (USA). J Environ Manage 88(4):1119–1130. https://doi.org/10.1016/j.jenvman.2007.06.001

Federal Emergency Management Agency (2013) Flood insurance study: Duval County, Florida. (12031CV001A). https://fris.nc.gov/fris/

Federal Emergency Management Agency (2014) Flood insurance study: Clay County, Florida. (12019CV001A). https://fris.nc.gov/fris/

Federal Emergency Management Agency (2011) Flood insurance study: St. Johns County, Florida. 12109CV000C). https://fris.nc.gov/fris_/hardfiles/FL/FIS/12109CV000C.pdf

Gebremariam SY, Martin JF, Demarchi C, Bosch NS, Confesor R, Ludsin SA (2014) A comprehensive approach to evaluating watershed models for predicting river flow regimes critical to downstream ecosystem services. Environ Model Softw 61:121–134. https://doi.org/10.1016/j.envsoft.2014.07.004

Kidson R, Richards K (2005) Flood frequency analysis: assumptions and alternatives. Prog Phys Geogr. https://doi.org/10.1191/0309133305pp454ra

Kovalenko S, Brown C, Akan C, Schonning A (2020) An examination of extreme flood events and resiliency of the Lower St. Johns River, Florida USA Using Multiple Methods. In: 5th international electronic conference on water sciences. Doi: https://doi.org/10.3390/ECWS-5-08280.
Kovalenko S (2020) Determination of extreme flood events in the Black Creek, Julington Creek, Durbin Creek, Big Davis Creek, Ortega River, and Pablo Creek Sub-Basins of the Lower St. Johns River Basin, Florida, USA. M.S. Thesis. University of North Florida.

Kumar R (2019) Flood frequency analysis of the Rapti River basin using log pearson type-iii and Gumbel Extreme Value-1 methods. J Geol Soc India 94(5):480–484. https://doi.org/10.1007/s12594-019-1344-0

Malamud B, Turcotte D (2006) The applicability of power-law frequency statistics to floods. J Hydrol 322(2006). https://www.researchgate.net/publication/222655854_The_applicability_of_power-law_frequency_statistics_to_floods

NIST/SEMATECH e-Handbook of Statistical Methods (2012) Engineering statistics handbook. https://doi.org/10.18434/M32189

National Oceanic and Atmospheric Administration Atlas 14 (2017) https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_cont.html

Okoli K, Mazzoleni M, Breinl K, Baldassarre G (2019) A systematic comparison of statistical and hydrological methods for design flood estimation. Hydrology Res 50(6): 1665–16678. https://iwaponline.com/hr/article/50/6/1665/69899/A-systematic-comparison-of-statistical-and

Rao D (1986) Technical Publication SJ 86-2. Magnitude and frequency of flood discharges in Northeast Florida. https://www.sjrwm.com/documents/technical-reports/

Saf B, Dikbas F, Yasar M (2007) Determination of regional frequency distributions of floods in West Mediterranean river basins in Turkey. Fresenius Environ Bull 16:1300–1308

Sahin N, Robertson A, Boer R, Behera S (2017) Probabilistic seasonal streamflow forecasts of the Citarum River, Indonesia, based on general circulation models. Stoch Environ Res Risk Assess. https://doi.org/10.1007/s00477-016-1297-4

Schiariti, P., P.E., CPESC (n.d.). Basic hydrology—runoff curve numbers. http://njscdea.ncdea.org/

St. Johns River Water Management District (1985) A guide to SCS runoff procedures. (SJ85-5). https://www.sjrwm.com/documents/technical-reports

St. Johns River Water Management District (1990) Rainfall analysis for Northeast Florida Part V: frequency analysis of wet season and dry season rainfall. (Project No. 10 200 02)

St. Johns River Water Management District (2012) St. Johns River Water Supply Impact Study (SJ20120-1). https://www.sjrwm.com/documents/water-supply/#wsis-final-report

St. Johns River Water Management District (2012). Appendix M of St. Johns River water supply impact study (SJ20120-1). https://www.sjrwm.com/documents/water-supply/#wsis-final-report

Suphunvorranop T (1985) A guide to SCS runoff procedures. United States Department of Water Resources Technical Publication No. 85-5.

US Army Corps of Engineers (1994). Engineering and DESIGN: FLOOD RUNOFF ANALYSIs (Engineer Manual 110-2-1417).

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.