Response to Anonymous Referee 1

1. This manuscript introduces a coupled model to jointly simulate groundwater levels and surface flows to assess compound flooding in South Florida. The two models that are loosely coupled are FLO-2D and MODFLOW-2005.

The modelling framework is applied for three past storm events to assess flood extent and depth and the role of the different flooding drivers. Validation opportunities are limited due to missing flood inundation information, but ancillary data is collected and used to assess the model performance.

Overall, I think the manuscript is very interesting, timely, and generally well written. I have some general and several specific comments listed below which I think should be taken into consideration to improve the analysis/presentation.

Authors' comment: We thank Reviewer #1 for the positive feedback.

Action taken: The authors would like to thank the reviewer for providing thoughtful comments. Please find our responses to each comment below including how we plan (following the NHESS review process) to adjust the manuscript.

General comments:

2. I think the authors need to be clearer in how they define compound flooding for the purpose of their analysis, i.e. is a rainfall event that coincides with a high water table considered a compound event or do tides also have to contribute to escalate the flood depth/area to make it a compound event?

From the introduction it appears that the key focus is on the combination of all three flooding drivers, rainfall, ground water, and coastal water levels, but the latter barely play a role in the events that were analyzed. I am not saying that the combination of rainfall and high water table could not be considered a compound event, but there was a bit of a mismatch of what I read in the introduction/methods and what was shown in the results.

I think this can be fixed by changing the narrative and does not require additional analysis.

Authors' comment: We thank Reviewer #1 for the suggestion. This paper aims to present an integrated modeling framework capable of simulating surface-subsurface water interactions. For the purpose of this analysis, compound flooding is defined as the interaction of overland flow and groundwater emergence to the surface. For this reason, the manuscript title has been modified to " Compound flood modelling framework for surface-subsurface water interactions" and the narrative was modified to be clearer and improve the readability of the manuscript.

Action taken: The narrative has been clarified in the following sections:
Abstract (lines 19-31)

This paper presents a physics-based, loosely-coupled modelling framework using FLO-2D and MODFLOW-2005 that is capable of simulating surface-subsurface water interactions. FLO-2D, responsible for the surface hydrology and infiltration processes, transfers the infiltration volume as recharge to MODFLOW-2005 until the soil absorption capacity is exceeded, while MODFLOW-2005 returns exchange flow to the surface when the groundwater heads are higher than the surface depth. Three events characterized by short-duration intense precipitation, average tide levels and unusually high water table levels are used to assess the relevance of groundwater flooding in the Arch Creek Basin, a locality in North Miami particularly prone to flooding conditions. Due to limitations in water level observations, the model was calibrated based on properties that have experienced repetitive flooding losses, and validated using image-based volunteer geographic information (VGI). Results suggest that groundwater-induced flooding is localized, and high groundwater heads influence pluvial flooding, as the shallow water table undermines the soil infiltration capacity. Understanding groundwater flood risk is of particular interest to low-elevation coastal karst environments as the sudden emergence of the water table at ground surface can result in social disruption, adverse effects to essential services and damage of infrastructure. Further research should assess the exacerbated impacts of high tides and sea level rise on water tables under current and future climate projections.

Introduction (lines 95-106)

This paper presents a physics-based, loosely-coupled modelling framework using FLO-2D and MODFLOW-2005 that is capable of simulating surface-subsurface water interactions. FLO-2D, responsible for the surface hydrology and infiltration processes, transfers the infiltration volume as recharge to MODFLOW-2005 until the soil absorption capacity is exceeded, while MODFLOW-2005 returns exchange flow to the surface when the groundwater heads are higher than the surface depth. Three events characterized by short-duration intense precipitation, average tide levels and unusually high water table levels are used to assess the relevance of groundwater flooding in the Arch Creek Basin, a locality in North Miami particularly prone to flooding conditions. Due to limitations in water level observations, the model was calibrated based on properties that have experienced repetitive flooding losses, and validated using image-based volunteer geographic information (VGI). Results suggest that groundwater-induced flooding is localized, and high groundwater heads influence pluvial flooding, as the shallow water table undermines the soil infiltration capacity. Understanding groundwater flood risk is of particular interest to low-elevation coastal karst environments as the sudden emergence of the water table at ground surface can result in social disruption, adverse effects to essential services and damage of infrastructure. Further research should assess the exacerbated impacts of high tides and sea level rise on water tables under current and future climate projections.

Methodology (lines 370-379)

Three flood events characterized by similar high intensity rainfall, tide levels, and unusually high-water table levels with different response times were selected to compare the surface-subsurface model results (Fig. 8). Tropical Storm Leslie (2-4 October 2000) was responsible for one of the most severe events of North Miami in recent history in terms of flooding and property damages, with an accumulated rainfall of 454 mm over 65 hours and an estimated return period of 50 years (Franklin et al., 2001). Similarly, Tropical Storm Andrea (6-8 June 2013) was a short-lived storm that formed in the Gulf of Mexico which produced very heavy precipitation across Broward and MDC (Beven II, 2013), and a total rainfall of 317 mm in the Arch Creek Basin. The 25 May 2020 event is categorized as a 25-year storm with a total daily rainfall depth of 263 mm, producing localized rainfall in the North Biscayne Bay watershed, specifically in the Arch Creek Basin, due to antecedent rainfall conditions since mid-April 2020.

Results (lines 383 – 435)

- **5.1 Calibrated coupled surface-subsurface model**

Simulating surface-subsurface water physical processes through physics-based flood modelling frameworks is relevant and meaningful to better assess the severity of groundwater-induced flooding in low elevation coastal environments characterized by porous permeable soil. Fig. 9 illustrates the simulated maximum inundation depths corresponding to the magnitudes of Tropical Storm Leslie, Tropical Storm Andrea, and the 25 May 2020 storm. Tide levels per se do not pose significant threats to infrastructure as the coastal waters remain within the channels. Fig. 10 illustrates the emergence of the groundwater heads to the surface as a result of the increase in the water table. The simulation proves reasonable in terms of maximum flood depth and extent due to the similarities in the hydrologic conditions, being Tropical Storm Leslie the most severe of all three storms. FEMA's records on properties subject to frequent flooding were used as a calibration approach to verify a match between the model results with flood observations. Although the available records do not specify the observed inundation depths, an agreement between the property locations and maximum water levels may offer sufficient evidence that the model provides reasonable results (Fig. 11). The calibrated results and display of the water table timeseries in selected locations for Tropical Storm Leslie are shown in Fig. 12-13.

- **5.2 Identification of flooding hotspots**

The groundwater flood maps for Tropical Storm Leslie (37.17%), Tropical Storm Andrea (13.87%) and the May 2020 event (20.82%) are showed in Fig. 10. The simulation demonstrates that slight variations in the water table depth (Fig. 8) can exacerbate groundwater emergence extent, resulting in ≈ 10 cm across the Arch Creek Basin. Interestingly heavy precipitations scenarios with very high water tables over extended periods of time (May 2020 event) are more likely to trigger groundwater induced flooding compared to very high precipitation with high water table levels (Tropical Storm Andrea). Fig. 11 presents reasonable results between the reported claims and localized flooding, indicating that the housing infrastructure in these neighborhoods are likely to experience additional flood losses at some point in the future. The simulated storm events illustrate that most of the properties experienced moderate to high flood depths (> 0.5 meters) in predefined locations. Although rainfall-runoff is the primary source of flooding in the urbanized Arch Creek Basin, abnormally high groundwater levels triggered groundwater-induced flooding near historic waterways and zones below the County's land elevation flood criteria, with flood depths ≈ 1 meter (Fig. 12a – 12b). The groundwater plots illustrate the effect of tidal and groundwater boundary conditions on the behavior of the simulated water table, in turn demonstrating the importance of both variables in the modeling set-up and influence in subsurface dynamics, as a cyclic high-low pattern characterizes the tide fluctuations of the Biscayne Bay (Fig. 12b – 12e) compared to the defined water heads behavior from well G-852 in the western boundary of the domain (Fig. 12a, 12f). In terms of residential damage, Tropical Storm Leslie and Tropical Storm Andrea may be considered the costliest events in the Arch Creek Basin as both account for 60% of the reported claims (25 and 17 respectively) (Table 2).

Sources of uncertainty in the coupled numerical model could be reduced by increasing the model's resolution and incorporating storm-water infrastructure features (i.e., French drains). For example, the increase of the water table levels could challenge the ability of the storm drain system to convey water towards the Bay, resulting in prolonged flooding conditions, or anti-flood pump stations may alleviate the impacts of flooding by draining water from the streets and swales back to the ocean. Nevertheless, the repetitive loss records only reflect a small percentage of the damaged infrastructure and cannot be generalized at the Basin scale as the property owners may not meet the criteria to file the claim. Therefore, the presented modelling results fall more on the conservative side and might overestimate the real flooding conditions.

- **5.3 Validation using crowdsourced data from Tropical Storm Andrea**

A limited number of real-time crowdsourced flooding observations in the Arch Creek Basin were available for Tropical Storm Andrea (Fig. 13). The visual comparison indicates a spatial agreement between the maximum flood depth of the coupled simulation and the interpreted depth of the crowdsourced data (Table 3). Fig. 12a associates high flow depths (> 0.5 meters) with several properties that have experienced regular flooding conditions, while the crowdsourced photograph displays an estimated inundation depth of 0.20 meters. Despite the model's overestimation, this comparison can be seen as an effective form of validation considering the changes in land use associated with the Arch Creek flow (Fig. 2) and low topographic elevation (Fig. 3b). Regarding Fig. 13b, the US Post Office exhibits chronic flooding in the parking lot. The coupled model exhibits a reasonable level of accuracy in terms of flood depth validation results. Fig. 13c displays stagnant flood water accumulated post-event in a portion of the NE 14 Ave. The results suggest that the rise of the water table do not influence the inundation depth and extents in any of these locations. Despite the limitations on the amount of collected crowdsourced data in the study area, a larger georeferenced dataset including the date and time could improve the reliability of VGI data to validate hydrodynamic models.

Conclusions (line 471-475)

Surface-subsurface water interactions are increasing in coastal cities due to multiple factors related to climate change. The Arch Creek Basin in North Miami, which served as a vital flow corridor that connected the Everglades to the Biscayne Bay, is an appropriate location to study the influence of high water tables in flood conditions. Results corroborate that groundwater-induced flooding is localized; thus, becoming an underlying condition that must be considered in low elevation coastal karst environments where the water table dynamics are subject to swift fluctuations caused by rainfall events.

Conclusions (line 484-489)

The quality and accuracy of flood hazard mapping in urban areas are strictly related to the model spatial resolution considering that the vertical datum and built-up environment influence flow propagation dynamics. A 20-meters grid resolution was selected to balance the computational demands with a certain level of precision without compromising the quality of the simulation. However, the investigation of higher and coarser resolutions in surface-subsurface modelling studies might yield insights into the estimation of inundated areas and time performance at different scales.

3. I am not familiar with the exact models that were used for the analysis and hence I sometimes had a hard time understanding in the Methods section which parts of the modelling system were actually developed by the authors and which parts were already implemented. Much of it read like material I would expect in the technical manuals of the models. I would keep such information at a minimum and only provide the information

necessary to understand the new aspects of the framework going beyond what was already there, and rather refer to the technical documentation for details of how these models work.

Authors' comment: We agree with Reviewer #1 that the methodology section contained redundant information about both modelling frameworks and the coupling methodology is not properly explain and requires editing

Action taken: The methodology sections (4.1 to 4.4) has been fully restructured as requested (lines 208 - 369)

4. The figure captions are all very short and often don't contain enough information to fully understand what is actually shown; this is particularly important since many of the figures are stitched together and contain a lot of information. See also specific comments below regarding figures/legends.

Authors' comment: We agree that the figure captions are not self-explanatory and some figures require editing

Action taken: Most figures and captions have been improved as requested

Specific comments:

5. Line 19- "Physics-based" might be better

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

6. Line 22- "returns"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

7. Line 24- "that as" doesn't really work here I think

Authors' comment: We agree

Action taken: The text has been deleted as the narrative has changed

8. Line 25 "result"

Authors' comment: We agree

Action taken: The text has been deleted as the narrative has changed

9. Line 27 "damage of"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

10. Line 44 remove "field"?

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

11. Line 50 there is a recent paper by Gori et al. (https://doi.org/10.1029/2019WR026788) which could be added to the list

Authors' comment: We agree

Action taken: The citation has been added to the list as requested

12. Line 73 "in that study"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

13. Line 165-171 Do I understand correctly that total still water levels (tide + surge) are considered? I was confused when I read later in the discussion/conclusion that storm surges were not included in the analysis. If still water levels were used but from "calm" periods where surge component was minimal that should be made clear.

Authors' comment: We thank Reviewer #1 for the valuable suggestion. Tide levels were considered as part of the coastal boundary conditions. The three selected events are characterized by 'regular' tide level conditions that have little impact on the coast. Therefore, the surge component is minimal and does not exacerbate flood risk conditions in the study area.

Action taken: This concept has been clarified as requested in the following sections:
Introduction: Lines 100-104

For the purpose of this analysis, three events characterized by short-lived heavy precipitation, regular tide levels and unusually high-water tables were selected to demonstrate the importance of simulating surface-subsurface water interactions in urbanized karst coasts, as high groundwater heads may exacerbate flooding conditions. In the context of this paper, compound flooding is defined as the interaction of overland flow and groundwater emergence, while surge levels are normal and have a minimal influence in the inundation beyond the coast.

Methodology: Lines 324-326

Rainfall and tides were considered for the hydrologic forcing, setting the precipitation over the grid system and tide levels in the easternmost cells to represent the Biscayne Bay's coastal conditions. Both time series are structured on a one-hour basis and are presented in the following section.

Methodology: Lines 371-372

Three flood events characterized by similar high intensity rainfall, tide levels, and unusually high-water table levels with different response times were selected to compare the surface-subsurface model results (Fig. 8).

Results: Lines 387-388

Tide levels per se do not pose significant threats to infrastructure as the coastal waters remain within the channels.

14. Line 199 would drainage systems be considered as sinks or explicitly modelled?

Authors' comment: We thank Reviewer #1 for bringing this up.

The FLO-2D components are explicitly modeled, the interaction between the grid cell (surface model) and the components are physically represented and modeled.

For example, the storm drain is simulated as a fully integrated system on a computational timestep basis. The FLO-2D model moves around blocks of water on a discretized grid system. Grid elements assigned as inlets/outfalls connect the surface layer with the closed conduit storm drain system. A comparison of the grid element water surface elevation with the pressure head from the closed conduit system node in a given cell determines the direction of the flow exchanged between the two systems. Discharge from FLO-2D surface layer to FLO-2D storm drain layer is based on the inlet geometry and water surface depth, return flow is only allowed when Pressure head is greater than water surface elevation. The system is a physical system that represents the real interaction between surface and storm drain layers.

It should be noted that the storm drain component (mentioned in the methodology section) serves to only highlight the model versatility to simulate flooding conditions in urban environments and was never considered in this study.

Further research should incorporate additional urban features such as the storm drain system and pumps to improve the model's flow propagation dynamics.

Action taken: Section 4.1 was shorted to comply with comment 1.3 (lines 208-216). In addition, we acknowledged the limitations of the study and underlined potential sources of uncertainties, such as the importance of increasing the model resolution and include additional urban features (i.e., storm drain system and pumps) to improve the model's flow propagation dynamics (lines 327-328, 415-421).

Lines 208-216

FLO-2D is a physically-based volume conservation model that combines hydrology and hydraulics to simulate the propagation of water dynamics in urban, riverine, and coastal environments for flood hazard mapping, floodplain delineation, flood vulnerability assessments and mitigation planning (O'Brien et al., 1993). The flood routing model applies the dynamic wave approximation to the momentum equation to calculate the average flow velocity across the square grid system one direction at a time in eight potential flow directions over the floodplain. Hydrological processes are represented as rainfall data over the computational domain or as input hydrographs that can be specified in the channel, floodplain, or along the coasts. Various attributes (elevations, roughness coefficient), components (channel, infiltration, storm drain) and features (streets, hydraulic structures) can be incorporated into the FLO-2D model to produce more refined simulations (O'Brien, 2011). Details are described elsewhere (Annis and Nardi, 2019; Grimaldi et al., 2013; Peña et al., 2021; Peña and Nardi, 2018).

Lines 327-328

The inclusion of the storm drain system, French drains, surface water control structures and pump stations in the modelling framework is beyond the scope of this study.

Lines 415-421

Sources of uncertainty in the coupled numerical model could be reduced by increasing the model's resolution and incorporating storm-water infrastructure features (i.e., French drains). For example, the increase of the water table levels could challenge the ability of the storm drain system to convey water towards the Bay, resulting in prolonged flooding conditions, or anti-flood pump stations may alleviate the impacts of flooding by draining water from the streets and swales back to the ocean. Nevertheless, the repetitive loss records only reflect a small percentage of the damaged infrastructure and cannot be generalized at the Basin scale as the property owners may not meet the criteria to file the claim. Therefore, the presented modelling results fall more on the conservative side and might overestimate the real flooding conditions.

15. Line 234 I would change to "solvers for matrix equations" to avoid repetition

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

16. Line 329 "G&A" has not been defined; this is also an example where I wasn't sure if the authors had added new options/functionality or just selected one of different existing options already available in the model code

Authors' comment: We agree

Action taken: The term has been defined as "Green & Ampt method" and is now consistent throughout the manuscript. This observation has been previously addressed in comment 1.3

17. Line 378 should it be "canal bed"?

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

18. Line 382 "what is the CHD package feature"?

Authors' comment: We agree

Action taken: The term has been defined as requested (Line 356-357).

19. Line 397 "responsible for"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

20. Line 393-416 This is where I started wondering what the actual role of the coastal water level
will be in the analysis (after a lot was said about it before) since it is not mentioned at all, other than that the events had low storm surge levels.

Was there a particular reason to not select any events where there was at least some storm surge to actually see the effect?

Authors' comment: We thank Reviewer #1 for bringing this up. As mentioned previously (comment 1.13), the selected events have a minimal influence in the flooding conditions across the domain, as the coastal boundary only increase the water levels of the Biscayne Bay and waterways.

The combination of heavy precipitation and unusually high water table levels with high storm surge was not found in the timeseries records.

Although some exceptions presented heavy precipitation and high storm surge with above regular water table levels, these events did not experience groundwater-induced flooding. Therefore, the selection of those will not be of interest for this manuscript as we are studying the influence of surface-subsurface water interactions.

Action taken: This observation has been previously addressed on comments 1.2 and 1.13

21. Line 427 "karts environments"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

22. Line 435 I was confused here since it says Fig. 10 shows results for Leslie, but it also has results for other events.

Authors' comment: We agree

Action taken: The results section, Figures and associates statements have been improved.

23. Line 437 How is "chronic flooding" defined here? It's a term often used when analyzing high tide flooding but that is different, I think to what the authors refer to here. Please clarify.

Authors' comment: We thank Reviewer #1 for pointing this out.

Action taken: Concepts related to "chronic flooding" have been removed from the manuscript to avoid potential misunderstandings.

24. Line 441 should be "characterizes" I think

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

25. Line 453 In Fig. 10b the water table actually goes above the terrain for all events at some point, should the reference be to "(Fig. 10c-f)"?

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

26. Line 453 "consistent agreement" sounds a bit strange, maybe reword

Authors' comment: We agree

Action taken: The term "consistent agreement" has been changed to "reasonable results"

27. Line 455 do you mean SRL? I don't understand how any of the results presented here would show the effect of SLR (assuming that it stands for sea-level rise).

Authors' comment: We thank Reviewer #1 for pointing out this concern. We apologize for the confusion that was due to similarities with the abbreviation of sea-level rise (SLR).

Action taken: To avoid confusion, we define "SRL" as "Severe Repetitive Loss" throughout the manuscript.

28. Line 457 "account for 60%"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

29. Line 478 I think a better way to start the sentence is to use "Despite…" or something similar

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

30. Line 512 the paper by Serafin et al. (https://doi.org/10.5194/nhess-19-1415-2019, 2019) could be added to the list

Authors' comment: We agree

Action taken: The citation has been added to the list as requested

31. Line 515 "on record"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

32. Line 517 here the authors use again "SLR" and I think this time it stands for sea-level rise, but it has not been defined anywhere.

Authors' comment: We agree

Action taken: This observation has been previously addressed on comments 1.27

33. Line 522 see comment above about using still water levels vs predicted tides as boundary conditions and the choice of picking three events where surge component was small

Authors' comment: We thank Reviewer #1 for pointing this out.

Action taken: This observation has been previously addressed on comments 1.2, 1.13 and 1.20

34. Line 546 "coupled"

Authors' comment: We agree

Action taken: The sentence has been restructured as requested

35. Figure 6: maybe consider switching "Stress Period I" and "Stress Period II" text as you start at the bottom with DT1 and T1 but they are linked to period II.

Authors' comment: The figure required additional editing. We apologize for the confusion due to mistakes in the text.

Action taken: Figure 6 (now Figure 5) has been improved as requested

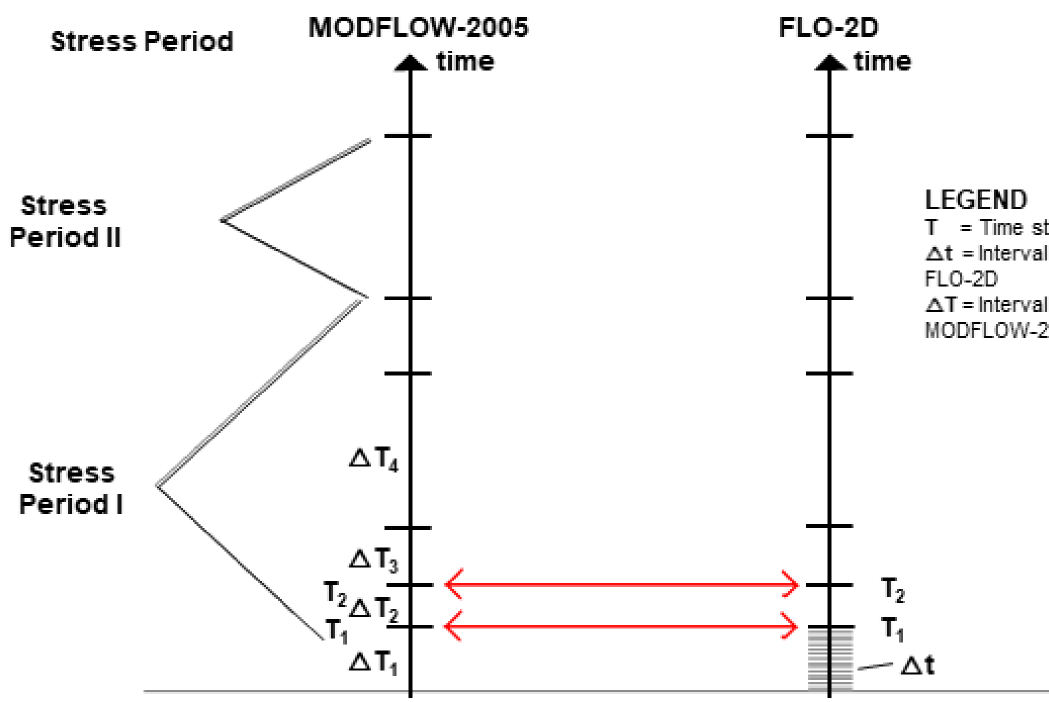

Figure 1. Time-step synchronization of FLO-2D and MODFLOW-2005.

36. Figure 7: in the top what does "statistics" refer to? I didn't see anything about that in the text. In the caption it should be "base hydraulic model"

Authors' comment: The figure required additional editing because the current description can lead to confusion.

Action taken: The terms "observations / model / statistics" were removed, and the figure has been improved to better communicate the coupling framework between FLO-2D and MODFLOW-2005 to produce a compound inundation scenario.

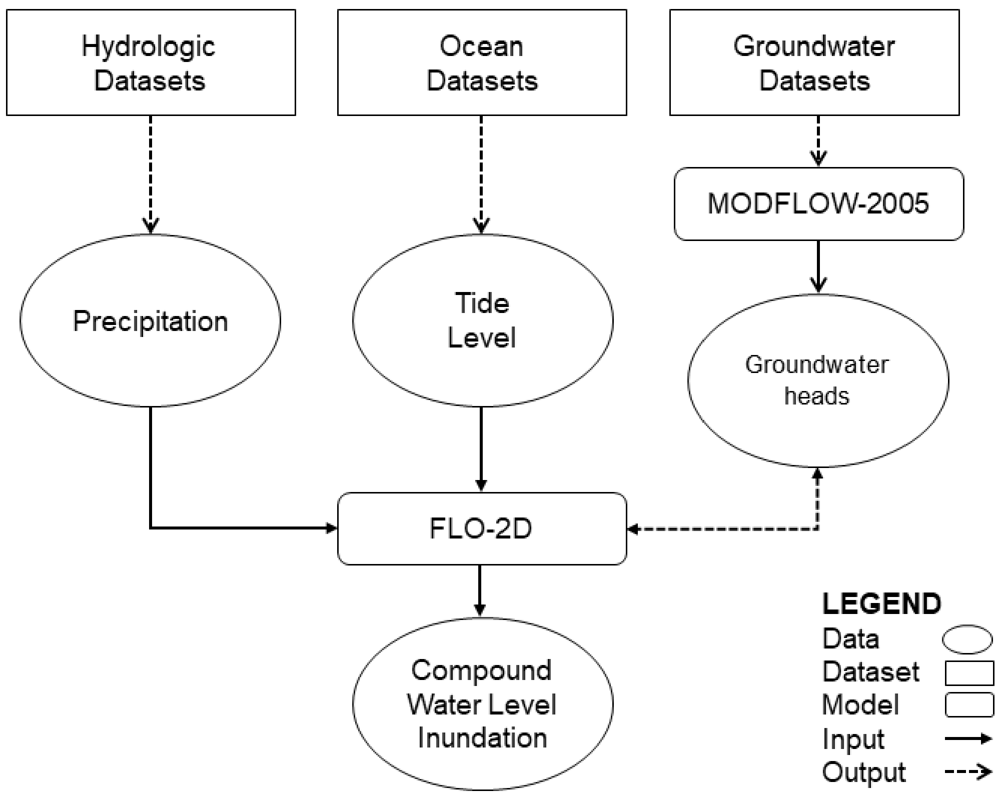

Figure 2. Flowchart representing the CF simulation using FLO-2D as the base hydraulic model. The hydrologic, ocean, and groundwater datasets were obtained through observations. The surface hydrology was incorporated as rainfall and coastal boundary conditions in FLO-2D. The groundwater heads were calculated in MODFLOW-2005 and transferred in an iterative manner to FLO-2D every time a MODFLOW-2005 time step is reached (Fig. 6). Adapted from Santiago-Collazo et al. (2019)

37. Figure 9: need to mention in the caption what the insets are and refer to the later figure where they are used.

Authors' comment: The figure required additional editing as it fails to convey meaningful information.

Action taken: The figure has been improved as requested and the insets were included (now figures 11 and 12)

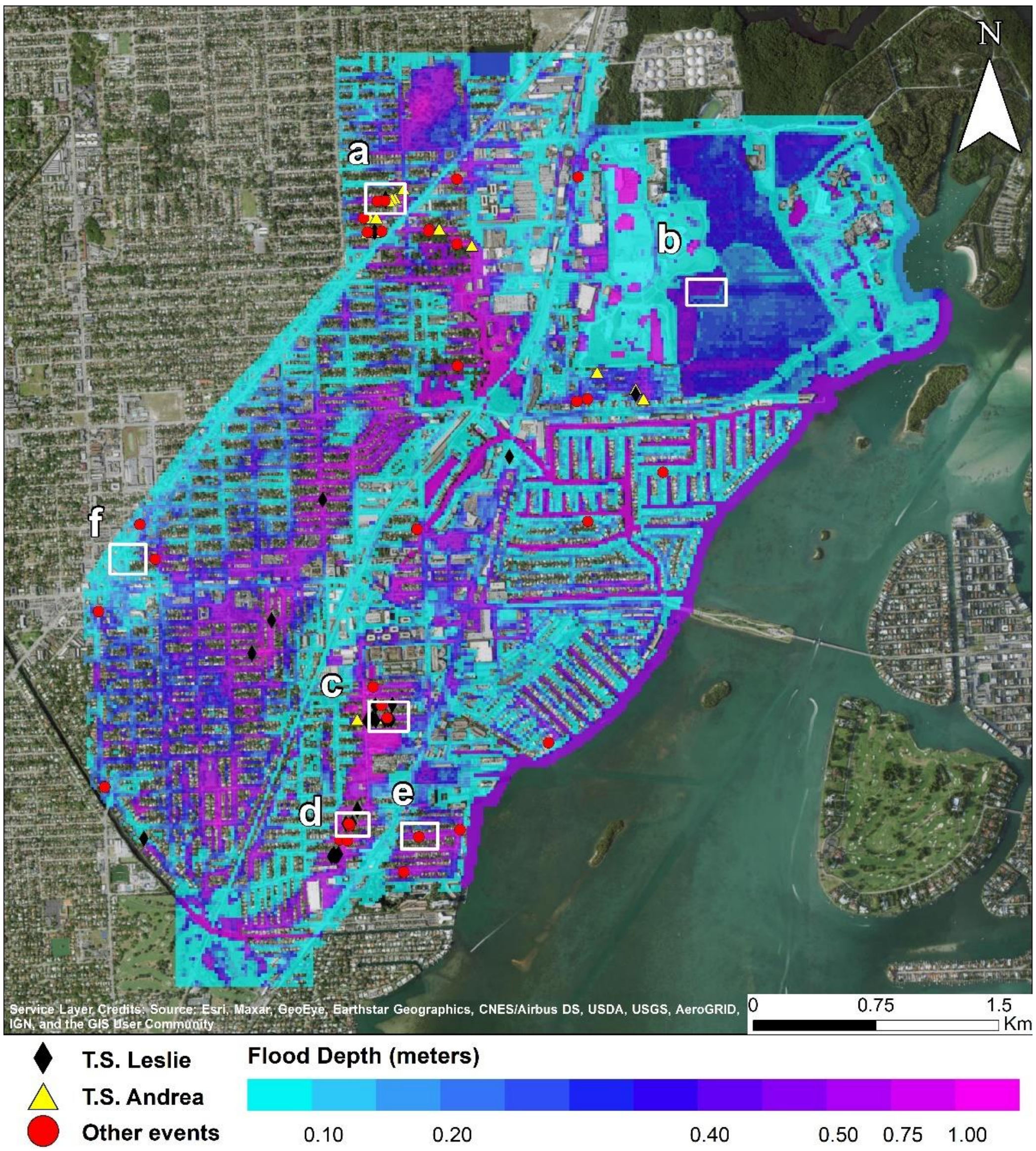

Figure 3. Distribution of maximum flood depths for Tropical Storm Leslie. The markers indicate repetitive loss properties caused by Tropical Storm Leslie (black), Tropical Storm Andrea (yellow) or other storm events (red). Maximum flood depths at six sample locations (white) are presented in Fig. 11.

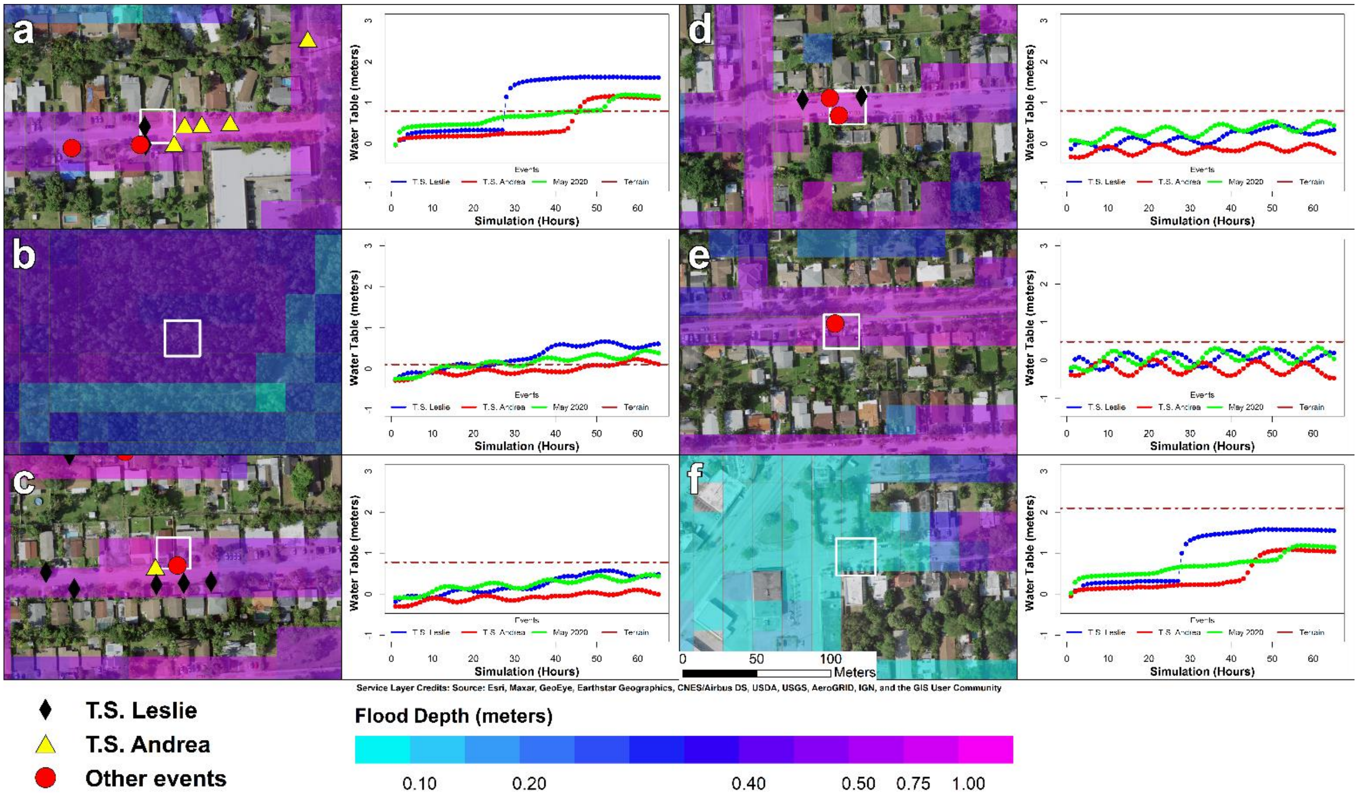

Figure 4. Six sample locations (Fig. 10) are selected to observe the maximum flood depths for Tropical Storm Leslie (left). The markers display repetitive loss properties that have been affected by Tropical Storm Leslie (black), Tropical Storm Andrea (yellow), and other storm events (red). The water table timeseries (left) display the behavior of the groundwater heads during Tropical Storm Leslie (blue line), Tropical Storm Andrea (red line) and the 25 May 2020 event (green line) at a specific location (white). Results demonstrate that the simulated water table (right pane) exceeded the surface elevation (brown line) on two locations leading to groundwater-induced flooding (a-b) while the rest are driven by pluvial flooding (c-d-e-f).

38. Figure 10: This figure has a lot of information and is a little bit hard to read with 1,2,3 and a,b,c showing up multiple times.

Maybe consider splitting the top part and the bottom part into separate figures. It's not clear what the markers (diamond, triangle, circle) represent here.

Similarly, the VGA Image and Area of Interest markers in the legend are confusing (and maybe not needed). The legend and associated text in the water table plots are way too small and impossible to read. Finally, what is meant with "other events"?

I assume that relates to flood claims from events that were not Leslie or Andrea? At this point the reader has no clue about this information being even shown in the figure and it's not mentioned in the text or caption at all; it's only mentioned later when talking about Fig. 11, so maybe it shouldn't be shown in Fig. 10 to keep the reader focused on what matters.

Authors' comment: We thank Reviewer #1 for the suggestion. The figure required major improvement as it failed to convey meaningful information.

Action taken: Two figures were created from Figure 10 (now Figures 11-12) as requested to only display relevant markers pertinent to the results section. The readability of both figures has now been improved.

39. Figure 11: why is the color yellow mentioned in the caption, are the black diamonds and red circles not representing SRL info? Please make sure that figures are 100% understandable when looking at them and reading the caption (one should not have to read the main text to understand the content of a figure).

Authors' comment: We thank Reviewer #1 for the suggestion.

Action taken: Figure 11 was removed from the manuscript as it failed to convey meaningful information. All Figures and captions were improved throughout the manuscript as requested.