Dear Editor and Referee #2,

We thank you for the time spent to evaluate and review our manuscript and for the review comments and discussions. Answers to the issues raised by Referee #2 are provided below, and have been addressed in a revised manuscript.

Here, we will repeat the reviewer's comments (italic font) and response directly below (standard font). After each response, the associated changes applied to the revised manuscript are included in a table box (coloured in red).

Sincerely,
G. Kesserwani and M. Shirvani

" The paper "Flood-pedestrian simulator for modelling human response dynamics during flood-induced evacuation: Hillsborough stadium case study" aims at showing the effect of human-body characteristics and in-model behavioural rules when included in an ABM integrated framework for flood evacuation modelling. The model is largely based on the previous version developed by the authors, and it is tested on a synthetic case study and on a real-world case. The topic of this study is definitely highly relevant and timely. However, I think that the study could be further strengthened by assessing the effect of the evacuation time and characteristics of the inflow hydrograph on the evacuation process. These factors could have a higher impact than the human-body characteristics considered by the authors for flood risk mitigation purposes."

Thank you for highlighting the timeliness and relevance of the topic investigated. Before we respond to the comments, we clarify that the new version of the simulator is augmented by empirically-derived behavioural rules to particularly: account for the realistic heterogeneity in pedestrian agents' characteristics (human age, gender and body mass) and behaviours (mobility states, risk perception thresholds); and, their back interaction on the floodwater dynamics. In the context of microscopic evacuation modelling in small urban areas, both the heterogeneity and the back interaction factors could play a determinant role to assess during-flood risk to people and the potential changes in their spatial and temporal evacuation patterns. In this context, the scope of this paper has been to explore the impact of pedestrian characteristics and behaviours on evacuation timing and risk assessment, and not the evacuating timing on general risk assessment.

In the revised introduction (3^{rd} paragraph, Sect. 1), the scope of study has been clarified, while also acknowledging the importance of evacuation time and characteristics of the inflow hydrograph as the reviewer suggested:
… These ABM-based evacuation simulation tools were developed to inform emergency plans for severe flood types, such as in the immediate aftermath of a dam-break or a tsunami wave (e.g. Lumbroso et al., 2021). The focus of these tools is mainly on estimating the loss of life, pinpointing bottlenecks and high-risk areas, and assessing how flood warnings of an impending flash flood could reduce the number of casualties and injuries. For this type of risk analysis, individuals' microscopic decisions and actions are considered insignificant in influencing the overall simulation outcomes due to the scale and speed of floodwater flow. However, for the most common flood types in urban areas, e.g. surface water due to extreme rainfall, less attention has been given to model the microscopic responses, down to the scale of the moving individuals, in and around flooded urban hubs (Ramsbottom et al. 2006). In this context, Bernardini et al. 2021 imported outputs of a flood model into a commercially available evacuation modelling tool, called MassMotion, to analyse flood risk differences in microscale and macroscale modeling with and without including pedestrians' microscopic evacuation behaviour. They concluded that incorporating pedestrians' microscopic evacuation behaviour in microscale modelling could significantly influence the spatial and temporal changes in flood risk to people, i.e. up to 15 % in absolute terms, when compared to macroscale modelling. Their findings also suggest the need to further incorporate non-homogeneous characteristics of people in a more flexible microscale modelling framework, which may result in additional differences to the analysis of flood risk to people.

Having clarified that, a point-by-point reply to each comment raised is provided below it, together with an explanation of the associated changes made to the revised manuscript.

- *" It has been recently shown that the timing in which the evacuation is issued is crucial for reducing flood risk (Alonso et al., 2020). However, in section 2.3.1 the authors state that "When the floodwater starts to propagate over the walkable area, simulation time (t) of 0 min, the pedestrian agents start the evacuation …" In a no-flooding situation, agents are randomly moving based on their behaviours and on their daily routines. Why an agent should start moving exactly when the water starts entering the building and not before or after? Are there supporting evidence to justify such an assumption? Then, when flooding occurs, there can be two extreme situations. On the one hand, if the agent is doing something else it may not notice the flooding until the pedestrian agent goes on a flooded area. On the other hand, the agent could be informed earlier about a coming flood and start to evacuate earlier. These two scenarios could have dramatic consequences. I suggest including more modes (table 6) accounting for different evaluation timing. This study could show the role of human-body characteristics and in-model behavioural rules in reducing flood risk when evacuation is issued late. "*

To evaluate the relevance of newly-added pedestrian agent's human-body characteristics and in-model behavioural rules, we used Take a Previous model and Add Something (TAPAS) approach, as justified in the initially submitted preprint (Sect. 2.3). To do that, we reconsidered the same synthetic test case of the shopping centre (Sect. 2.3.1) under the same flooding condition as investigated before in the previous version of the simulator (Shirvani et al., 2020). This test case is designed to evaluate the simulator in capturing (spatially and temporarily) the two-way dynamic interactions between the pedestrians and the floodwater flow. Therefore, the evacuation was set to happen at the onset of flooding to produce the most critical scenario that would cause the maximum interactions between moving pedestrians and flowing floodwater. This assumption is based on the fact that pedestrians are mainly reliant on their own visual detection of the upcoming hazard and/or an immediate announcement in such a small urban hub. In addition, the selected evacuation timing allows us to analyse the relevance of the added characteristics and rules across a full range of flood risk HR states. This means that further exploring pre-flood or post-flood evacuation scenarios, while feasible, won't allow us to assess the potential impacts neither from the added rules and characteristics nor from the during-flood interactions.

- *" My second concern relates to the shape of the hydrograph considered in the synthetic experiment. I understand that using a hydrograph with a high flood peak would lead to significantly bigger HR and the consequent loss of life. However, is it not the scope of this model to represent worst-case scenarios to improve flood risk management and reduce loss of people? Also, not necessarily using a shorter hydrograph can lead to loss of life. I invite the authors to run different scenarios keeping the same volume of the input hydrograph but changing the timing of the peak. The timing of the peak is a crucial factor in any flood risk management application and I do not understand why its influence was not included in this study."*

Although not reported in this paper, we previously explored four different hydrographs, as suggested, based on keeping the same volume but changing the timing of the peak (Shirvani et al. 2021). Here, we chose the most suited hydrograph (timing and peak) that generates a full range of flood risk HR states to conduct the aforementioned evaluation for the new version of the flood-pedestrian simulator. Supported by the HR analysis in Shirvani et al.(2021), this choice for the hydrograph is representative of flood types that frequently occur in urban areas, hence for evaluating the simulator with a focus on individual-scale risk analysis while considering their two-way interaction with the floodwater flow and realistic human-body characteristics. As we already clarified in the initially submitted preprint (see the 3rd paragraph in Sect. 2.3.1), choosing higher peak or shorter timing would lead to bigger HR values associated with loss of life or injury, hence are not suited to test the relevance of this simulator.

In the revised version of the paper, clarification on the choice of hydrograph has been provided in the 3rd paragraph of Sect. 2.3.1 to clarify:
The flooding inflow was generated based on an inflow hydrograph of a discharge, Q (m^3/s) propagating over a duration of 7.5 min, and peaking to 160 m^3/s at 3.75 min (Fig. 3). The hydrograph was produced based on the Norwich inundation case study, and because it results in a range for the HR that is inclusive of all the ranges listed in Table 1, i.e. HR < 7 (Shirvani et al., 2021). Deploying a hydrograph with shorter duration or a bigger peak would lead to significantly bigger HR, which is indicative of potential loss of life or injury where a person can take very limited actions to carry on moving to the emergency exit (hence it is outside the scope of this study). When the floodwater starts to propagate over the walkable area, simulation time (t) of 0 min, the pedestrian agents start the evacuation and the simulation terminates when all the pedestrian agents have evacuated the walkable area.

- *" Section 2 is a very large and dense section. There are many headings and sub-headings and I found myself lost with a need to scroll up and down. Would not be better to move 2.3.1 and 2.3.2 in a new section 3 called synthetic case study used to test the model and then introduce section 4 (now section 3) on the real-life experiment?"*

Section 2 is focused on the simulator, with Sect. 2.3 dedicated to assessing the relevance of the newly added agent rules and characteristics. As part of Sect. 2.3, the synthetic case study (Sect. 2.3.1 and Sect. 2.3.2) is used to serve the aim of assessing the relevance of these new rules and characteristics within the simulator. Therefore, it makes sense to keep it with the section describing the simulator and its new rules and characteristics.