We sincerely thank Referee #2 for their constructive, helpful, and thorough review that has improved this paper. Below, we address each referee comment. Each referee comment is in **BOLD** and our response directly below. Page/line numbers in referee comments refer to the original submission.

1. **The use of drones for natural hazards damages evaluation is a well-known topic. It is important to point out that there is a special issue published on NHESS dedicated to UAV and natural hazards, I think that authors can find several interesting suggestions considering the published revision paper or the others. Another revision paper has been published by Gomez, C. and Purdie, H.: UAV- based Photogrammetry and Geocomputing for Hazards and Disaster Risk Monitoring – A Review, Geoenvironmental Disasters, 3, 1–11, 2016.**
    - We thank the referee for recommending these publications. We performed an extensive review of the literature and referenced what we believe to be the most relevant and applicable papers to our study.

2. **Page 3 line 20: the use of nadiral acquisition (both by drones or planes) can be critical in urbanized areas. In Giordan et al 2018, (see comment on chapter 4.2) the effect of damages caused by a flood was defined using a mixed approach based on drone and terrestrial acquisitions.**
    - We thank the referee for providing this context. In Section 4.2 of the original submission, we acknowledged that the addition of terrestrial images to drone-based aerial imagery has been shown to improve the 3D textured mesh model in urban study areas, and referred to Wu et al. (2018) as a demonstrative study. The scope of our study was to investigate the sole use of drone-based aerial imagery, and this imagery set was not nadir. As noted in Section 2.3, the 828 images were collected at an average pitch angle of 7 degrees off nadir.
    - To provide more detail about the obliquity of the imagery set, we propose the following modification to the last sentence of Section 2.3: "A total of 828 oblique images were captured. The median image pitch angle was 7.35 ° off nadir (3.55 ° interquartile range), with a minimum and maximum of 1.22 ° and 11.83 °, respectively".

3. **Page 4 line 15 "The RTK/PPK image georeferencing capabilities of the drone replaced the need for ground control points (GCPs), which are not practical to distribute and survey in an emergency mapping context." This is not correct. The RTK/PPk correction improves the accuracy of images acquisition points. The number and the needs of GCPs depend on the required accuracy of the SFM results. During the mission planning, it is possible to have an estimation of final accuracy and decide if GCPs are required or not. For fast acquisitions, often GCPs are not required, but for a pre-event acquisition, the required accuracy should be high, and I do not think that it is possible to avoid GCPs.**
    - We agree with the referee that GCPs can and should be used for pre-disaster mapping to maximize the geospatial accuracy of the data. We retain our position that GCPs are not practical to use for emergency (i.e., post-disaster) mapping. We used PPK corrections instead of GCPs because of the applicability of this georeferencing method to both pre- and post-disaster mapping.

- To communicate the important point the referee has made, we propose the following modification of this paragraph in Section 2.2: “A senseFly eBee Plus drone with real-time kinematic (RTK)/post-processed kinematic (PPK) functionality and senseFly Sensor Optimised for Drone Applications (SODA) red-green-blue (RGB) 20-megapixel camera were used to collect imagery. The RTK/PPK image georeferencing capabilities of the drone replaced the need for ground control points (GCPs), which are not practical to distribute and survey in an emergency (i.e., post-disaster) mapping context. It is important to note that, for pre-disaster mapping, GCPs should be used to maximize geospatial accuracy. Hugenholtz et al. (2016) demonstrate the improvement in DSM vertical accuracy when using a non-RTK senseFly eBee with GCPs compared to an RTK-enabled senseFly eBee without GCPs. For this pre-disaster mapping exercise in downtown Victoria, we chose to use RTK/PPK image georeferencing because this method is also applicable to post-disaster mapping. One of our objectives was to assess the geospatial accuracy of the pre-disaster data, which has implications for the use of RTK/PPK-enabled drones for post-disaster mapping and change detection applications.”

4. **Page 6 line 10: In my experience, this is not the correct way to operate. The first step is the check of the right alignment of surveys. This can be done in particular using large plane areas (like car parking). Then you can compare buildings or other structures. The validation of the right position of DTM is mandatory to be sure that all used DSM are correct form the geodetic point of view. In an exercise like the one presented by authors, they could easily use as a sequence of Ground checkpoints to assure the accuracy of the obtained DSM. These checkpoints can be acquired using natural or artificial elements like (manholes) also after the UAV acquisition. To be rigorous, authors should present a more detailed study of the accuracy of the obtained DSM. This is a crucial point because the accuracy of the DSMs comparison is a function of the accuracy of used DSMs.**
    - We thank the referee for making this critical suggestion. We agree with this comment and propose a modification to the vertical accuracy assessment. We propose to replace the LiDAR checkpoints with 47 ground-surveyed (total station) checkpoints located on sewer manhole covers throughout the study area. These ground checkpoints follow the guidelines for vertical checkpoints as outlined in the 2015 ASPRS Positional Accuracy Standards for Digital Geospatial Data (ASPRS, 2015).
    - The proposed modification of the vertical accuracy assessment will be reflected in changes to Section 2.5, Section 3.1, Section 4.1, Figure 1, Table 1, and Table 2.

5. **Chapter 2.6 it is not clear which is the goal of this chapter. Using nadiral images for facades is not correct, and this is not a novelty. Authors should clarify better if the final goal is the identification of damages comparing the geometry of roofs or the study of facades.**
    - To clarify our goal, we added the following text to the beginning of Section 2.6: “In addition to geospatial accuracy, we wanted to assess the quality of building representation in the drone-derived 3D data. This assessment would have implications on the usability of the 3D data for identifying damages to building roofs and facades”.
    - As described in our response to Comment #2, the imagery set was not nadir.

6. **Authors should present the metadata of 2013 LiDAR before using it as a benchmark like the number of acquired points per meters, the accuracy of the survey, and the density of DSM point cloud. In particular, the DSM density is an important data. If the LiDAR density is not adequate, how authors can be sure that they comparing two points acquired in the same position or they are comparing a surveyed point and an artifact?**
    - Please refer to our response to Comment #4, where we propose to use 47 ground checkpoints instead of the LiDAR checkpoints.

7. **Chapter 4.1 the presented "key lesson 1" is quite trivial. Authors presented obvious data for people familiar with LiDAR and drones DSM. Several critical issues are quite evident in this chapter: the most critical point is the a priori definition of LiDAR resolution and accuracy using García-Quijano et al. (2008). The final resolution of LiDAR surveys is a function of many parameters, like the point density, the flight velocity, the post processing accuracy, and many others. In this paper, authors never mentioned the characteristics of the available LiDAR survey. Another important element is that without the acquisition of checkpoints, authors are not able to define the accuracy of their UAV DSM. I think that this lack of information cannot be accepted in a scientific paper.**
    - We thank the referee for identifying the improvement that should be made to our reference for piloted LiDAR vertical accuracy. We propose to remove García-Quijano et al. (2008) as the reference in Section 4.1 and Table 2. Instead, we propose to use the 47 ground checkpoints to calculate the $RMSE_z$ of the LiDAR DSM generated using the piloted LiDAR data from our study area. We will use the $RMSE_z$ of the LiDAR DSM to modify Section 4.1 and Table 2. Additionally, we will add the LiDAR metadata to Section 4.1.

8. **Page 10, line 5. The presence of differences in the geometry of several houses in the studied area could be useful for better development of the DSM comparison methodology. Using the comparison of DSM an images to check the first results, authors can be able to distinguish damages from building modifications. An improvement of the presented approach and the definition of an effective methodology for the recognition of damages can be an essential add value for this work, and it can also reduce the need of a continuous update of the DSM, which can generate a strong improvement of cost with a limited benefit. The only real result presented in chapter 4.1 is the difference between the results obtained by pix4d using the "rapid" and "full" point cloud. In my opinion, this cannot be considered an adequate result.**
    - If we are understanding this comment correctly, then the referee is suggesting we develop a methodology to distinguish between damaged buildings and modified buildings in the DoD. This is an excellent suggestion that would indeed contribute to reduced costs and time associated with continuously updating a pre-disaster DSM. If our interpretation of the referee's comment is correct, then we believe we do not have sufficient data for the proposed analysis. As shown in Figure 1b, the changes during the 5 years between the LiDAR (2013) and drone (2018) data acquisitions include new construction, structure removal, and parking lot excavation. The buildings that underwent new construction and structure removal could serve as examples of "modified buildings" in the referee's proposed analysis. However, our data lack

examples of "destroyed buildings". Therefore, we believe we cannot perform what we interpret as the proposed analysis. However, we propose the following addition to the end of Section 4.1: "To reduce costs and time associated with continuously updating a pre-disaster DSM, future research should focus on developing methodologies to distinguish between construction-modified and disaster-damaged buildings in a DoD".

9. **Chapter 4.2 the presented "key lesson2" is focused on an interesting point. The nadiral acquisition of an urbanized area is not enough for the correct reconstruction of facades. Giordan et al. (Giordan, D., Notti, D., Villa, A., Zucca, F., Calò, F., Pepe, A., Dutto, F., Pari, P., Baldo, M., and Allasia, P.: Low cost, multiscale and multi-sensor application for flooded areas mapping, Nat. Hazards Earth Syst. Sci., 18, 1493-1516, 2018) published a multi-scale approach aimed to detect and measure damages on facades. The approach is different, but the topic is important for a correct estimation of damages. One of the problems of this article is the organization. If the authors want to analyze facades, they have to introduce this topic in advance and propose a possible methodology. The publication of a sequence of well-known limitations cannot be considered sufficient for an international scientific journal like NHESS.**
    - We strongly disagree with the referee's comment. We believe our research is novel because this is the first government-approved drone mapping mission over a major Canadian city. This was a multi-stakeholder effort that included the municipal emergency management office, federal aviation authority, and air traffic control. In their review paper concerning RPAS for natural hazards monitoring and management, Giordan et al. (2018) recommend that future research should "propose faster and automated approaches. In particular during emergencies, the time required for RPAS data set processing is an important element that should be carefully considered". Giordan et al. (2018) also recommend that, "In the following years, it would be desirable to witness the transfer of best practices in the use of RPASs be then from the research community to government agencies (or private companies) involved in the prevention and reduction of impacts of natural hazards. The scientific community should contribute to the definition of standard methodologies that can be assumed by civil protection agencies for the management of emergencies".
    - Consistent with the recommendations of Giordan et al. (2018), we present and evaluate a legal and plausible scenario. This is evidenced by our description of the multi-stakeholder coordination (Section 1.2), our use of the only legally approved drone for urban overflight in Canada to date (Section 1.2), our gridded flight plan for efficiency (as opposed to circular flights around individual buildings) (Section 1, Section 2.2), our use of PPK image georeferencing (as opposed to GCPs) (Section 2.2), and our examination of "rapid" image processing (Section 2.4, Section 4.1). By constraining our study to comply with the legal and logistical practicalities of pre- and post-disaster mapping in a major Canadian city, we believe our results have implications on the usability of the regulatory-approved drone for assisting in rescue and damage assessment activities. Specifically, the results inform the federal aviation authority (Transport Canada) of the limitations of this drone and camera configuration, and we suggest an equally safe alternative for legal approval (Section 4.2, Section 5). We also provide evidence-based lessons/best practices for practitioners such as emergency management offices. These

best practices pertain to drone hardware (e.g., tilting cameras for 3D mapping [Section 4.2] and RTK/PPK georeferencing for change detection applications [Section 4.1]), data collection (e.g., takeoff and landing locations [Section 2.3] and up-to-date DSMs [Section 4.1]), and data processing (e.g., "rapid" processing for sub-meter building collapse detection [Section 4.1]).

- Additionally, by revising the accuracy assessment as recommended by the referee, we believe we provide a more rigorous analysis of the drone and LiDAR DSM accuracies.

**References:**

American Society for Photogrammetry and Remote Sensing (ASPRS): ASPRS Positional Accuracy Standards for Digital Geospatial Data, Photogramm. Eng. Remote Sens., 81(3), 1–26, doi:10.14358/PERS.81.3.A1-A26, 2015.

García-Quijano, M. J., Jensen, J. R., Hodgson, M. E., Hadley, B. C., Gladden, J. B. and Lapine, L. A.: Significance of Altitude and Posting Density on Lidar-derived Elevation Accuracy on Hazardous Waste Sites, Photogramm. Eng. Remote Sens., 74(9), 1137–1146, doi:10.14358/PERS.74.9.1137, 2008.

Giordan, D., Hayakawa, Y. S., Nex, F. and Tarolli, P.: Review article: the use of remotely piloted aircraft systems (RPASs) for natural hazards monitoring and management, Nat. Hazards Earth Syst. Sci., 18(4), 1079–1096, doi:10.5194/nhess-18-1079-2018, 2018.

Wu, B., Xie, L., Hu, H., Zhu, Q. and Yau, E.: Integration of aerial oblique imagery and terrestrial imagery for optimized 3D modeling in urban areas, ISPRS J. Photogramm. Remote Sens., 139, 119–132, doi:10.1016/j.isprsjprs.2018.03.004, 2018.