Gridded Flood Depth Estimates from Satellite Derived Inundations

Seth Bryant[1,2], Heather McGrath[3], Mathieu Boudreault [4]

[1] GFZ German Research Centre for Geosciences, Section 4.4. Hydrology, Potsdam, Germany
[2] Institute for Environmental Sciences and Geography, University of Potsdam, Germany
5   [3] Canada Centre for Mapping and Earth Observation, Natural Resources Canada, Ottawa, Canada
[4] Department of Mathematics, Université du Québec à Montréal, Montréal, Canada

Correspondence to: Seth Bryant (seth.bryant@gfz-potsdam.de)

Abstract. Canada's RADARSAT missions improve the potential to study past flood events; however, existing tools to derive
10   flood depths from this remote-sensing data do not correct for errors, leading to poor estimates. To provide more accurate
gridded depth estimates of historical flooding, a new tool is proposed that integrates Height Above Nearest Drainage and Cost
Allocation algorithms. This tool is tested against two trusted, hydraulically derived, gridded depths of recent floods in Canada.
This validation shows the proposed tool outperforms existing tools and can provide more accurate estimates from minimal
data without the need for complex physics-based models or expert judgement. With improvements in remote-sensing data, the
15   tool proposed here can provide flood researchers and emergency managers accurate depths in near-real time.

1 Introduction

Flooding has become the costliest natural disaster in Canada, with economic losses estimated around $2.5 billion per year
(Office of the Parliamentary Budget Officer, 2016). To mitigate this flood risk, large investments in infrastructure and planning
have been made by the federal government in the past decade (National Disaster Mitigation Program, 2017; Government of
20   Canada, 2021); however, accuracy and the absence of data on historical flooding remains challenging for the models
underpinning these investments (McGrath et al., 2015; Bryant et al., 2021). While new satellite missions have improved
capabilities for mapping inundation extents, data on maximum flood depth, which is commonly found to be the most significant
indicator of building damage following European floods (Mohor et al., 2020; Laudan et al., 2017; Merz et al., 2010), remains
scarce. The absence of such depth data in Canada limits the utility of flooding research, ultimately leading to less informed
25   flood management decisions.

Relying on microwave pulses that can reflect the ground surface at night and through clouds, Synthetic Aperture Radar (SAR)
instruments have become an important tool for measuring flood inundation at large scales (Shen et al., 2019). For example,
the recently launched, three-satellite, RADARSAT Constellation Mission provides regular medium-resolution (30-100 m)
SAR observations across Canada but can capture high-resolution (1-3 m) observations when requested for flood disasters
30   (Canadian Space Agency, 2021b). A common approach for identifying flooded areas from SAR observations employs a

threshold to the measured backscatter values to classify water covered areas based on their surface roughness (Benoudjit, 2019). For example, the 'Floods in Canada' (FiC) project calculates a backscatter threshold from historical inundation data to classify open-water flooding; before identifying adjacent areas with flooded vegetation using a second threshold (Natural Resources Canada, 2020).

35 While SAR measurements are advantageous for identifying inundated areas remotely (compared to optical measurements), the signal technology has limitations. Shen et al. (2019) identify three common error sources challenging SAR-derived inundation algorithms: 1) smooth dry-surfaces that return similar signals to inundated surfaces or rough water surfaces that return uncharacteristically rough signals; 2) georeferencing of SAR images (which often relies on ancillary terrain data); and 3) inundated areas near dense obstructions that scatter returning signals. These errors lead to less accurate inundation predictions

40 in areas with: dense urban infrastructure, dense vegetation, floating debris (e.g., ice), waves/rapids, steep river-banks perpendicular to the instrument (Natural Resources Canada, 2020; Cian et al., 2018), or recent construction/earthwork.

Simplified conceptual or '0D' inundation models provide an alternate means for estimating inundation efficiently; however, some calibration data is required to simulate specific events (Teng et al., 2017). Height Above Nearest Drainage (HAND) methods are a raster-based class of 0D models that leverage a digital elevation model (DEM) and drainage network information

45 to implement a three-phase routine for identifying flooded regions: 1) generate a hydraulically conditioned DEM; 2) calculate the height of each cell above the drainage network (HAND value); then 3) map all cells below some HAND value threshold, typically derived from observations or rating curves (Rodda, 2005; Rennó et al., 2008; Donchyts et al., 2016). In the U.S., HAND techniques have been coupled with the National Weather Model to produce uncalibrated continent-scale 10 m resolution inundation predictions (Liu et al., 2018) that were later shown to capture 19-25% of inundated areas accurately

50 (Johnson et al., 2019).

Advancements in remote sensing and terrain analysis have improved the availability and accuracy of historical inundation data; however, the corresponding (higher-dimensional) gridded depth data, desired by flood vulnerability research, has proven more illusive. Cian et al. (2018) provides a review of methods to derive gridded depths from remote sensing data, starting with work that manually overlaid LANDSAT imagery on terrain contours to estimate reservoir volumes (Gupta and Banerji, 1985).

55 This class of 'inundation-polygon terrain-overlay' methods seeks to first construct a water surface by identifying and projecting the land-water interface or 'shoreline', before subtracting the DEM from this water surface to yield gridded depths. When such methods are employed with DEMs that omit or ignore bathymetry, depth values are underestimated for waterbodies. A second, more relevant, error source is introduced by inaccuracies in the inundation extents or polygons. In areas with significant topography (e.g., steep riverbanks) small errors in inundation polygons can yield large errors in shoreline elevations (Nguyen

60 et al., 2016).

Within the inundation-polygon terrain-overlay class of depth estimation methods, a distinction can be made between those methods assuming a flat-water shoreline (or constant elevation) and those that are agnostic or allow shoreline elevation

heterogeneity to propagate into the water surface estimates. Flat-water shoreline methods are well-suited to floods with near-zero water surface gradients and high segmentation between inundation polygons (Cian et al., 2018; Gupta and Banerji, 1985).

65 Shoreline agnostic methods are better suited to handle floods with higher water surface gradients and more continuous inundation polygons; however, these methods are more sensitive to errors in shoreline location (Cohen et al., 2018; Brown et al., 2016; Nguyen et al., 2016). A second distinction can be made between those requiring expert input for parameterization (Cian et al., 2018; Brown et al., 2016; Nguyen et al., 2016) and those that are automated or only require simple data inputs (Cohen et al., 2018).

70 Early attempts to estimate gridded-flood depths from SAR-derived inundations implemented largely manual workflows. For a 2014 flood of a low-lying area in the U.K., Brown et al. (2016) selected shoreline elevation points (from LiDAR data) from which the flooded water surface was interpolated and finally subtracted from the elevation model to obtain gridded depths. This method was compared to LiDAR measurements of the flood surface, yielding a root-mean-square-error (RMSE) of 15 cm for overlapping inundation cells. Rather than use the SAR-derived inundations directly, Nguyen et al. (2016) used them to

75 identify a best-fitting parameterization of a simplified 2D hydrodynamic model for a low-lying floodplain in Vietnam. No RMSE was reported, and the method failed to predict inundation in "several small [sub] areas".

Cian et al. (2018) developed a semi-automated flat-water method employing statistical analysis of the raw shoreline elevations to identify the value where 5 percent of adjacent sorted values differ by less than 10 cm. Manual correction was used for inundation polygons whose raw shoreline elevation values failed to yield a conforming elevation. These results were compared

80 against hydrodynamic modelling derived depths, yielding a RMSE between 55 cm and 79 cm for overlapping inundation cells.

Leveraging user supplied flood path transect lines and boundary masks, Scorzini et al. (2018) developed the 'RAPIDE' tool. This shoreline agnostic tool was tested against a hydrodynamic model simulation of a flood in Italy using the simulated inundation for the tool input (rather than satellite derived inundations), yielding a RMSE between 38 and 79 cm for overlapping inundation cells.

85 Using a fully automated open source algorithm, Cohen et al. (2018) developed the Flood Water Depth Estimation Tool (FwDET) version 1.0 in the proprietary ArcGIS platform using a raster-based shoreline agonistic 'nearest boundary cell elevation' routine to interpolate shoreline elevations onto the interior inundated region. Version 2.0 replaced the interpolation routine with a more efficient 'cost allocation' routine better suited for inundations with incomplete boundaries (e.g., coastal flooding) (Cohen et al., 2019). Cohen et al. (2019) tested this tool against hydrodynamic model results for two flood-prone

90 regions in the U.S. using the simulated inundation for the tool input (rather than satellite derived inundations) and found average errors of 0.18 and 0.31 m, with some errors exceeding 1.5 m. Version 2.0 was later ported to Google Earth Engine (Peter et al., 2020). Cohen et al. (2019) reports on a second tool 'FwDET-QGIS', similar to Version 1.0 but for the QGIS open source platform, with the GRASS 'r.grow.distance' function (GRASS GIS manual, 2021) providing the nearest boundary cell elevation routine (Raney and Cohen, 2019). With the exception of a two-fold decrease in run-time (compared to Version 2.0),

95   no comparison was reported for FwDET-QGIS (Cohen et al., 2019). Aside from a low-pass filter applied to depth results, the FwDET tools propagate all shoreline errors into depth estimates.

All the aforementioned depth estimation methods report accuracy by comparing against some trusted source for all overlapping grid cells, rather than comparing depth estimates at asset locations (e.g., buildings) — the metric generally sought by flood vulnerability research. Further, no study with a fully automated method reported accuracy against satellite derived inundations.

100   In this context, our study pursues the following objectives: 1) present a fully automated shoreline agnostic tool that provides moderate error correction of shorelines; and 2) test this tool and FwDET-QGIS on two recent floods in Canada using typical Canada-wide datasets.

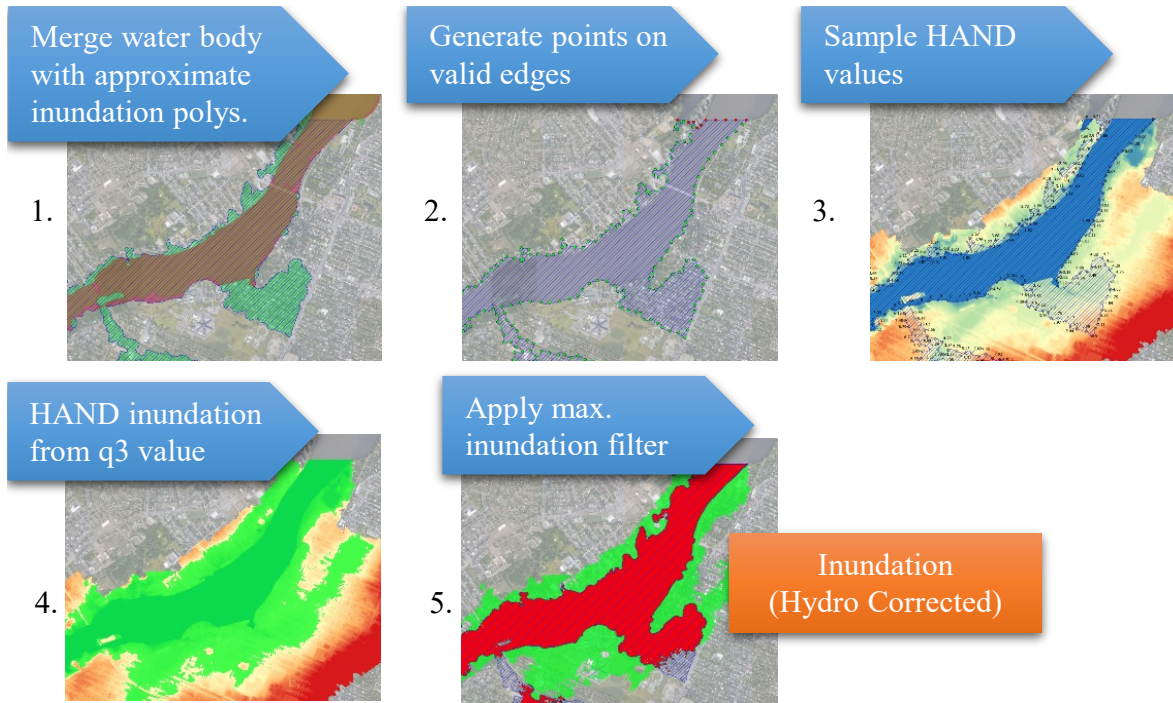
2 Methods

This study develops the novel Rolling HAND Inundation Corrected Depth Estimator (RICorDE) Tool for generating gridded
105   depth estimates of past flood events from approximate inundation polygons and a DEM. This tool allows for moderately varying water-land interface 'shoreline' values and does not require expert input. To demonstrate the accuracy of RICorDE, depth estimates are generated for two historical flooding events in Canada using publicly available datasets. These satellite derived depth grids are then compared against 'trusted' depths simulated by others with more sophisticated site-calibrated hydrodynamic models. To provide a comparison, the FwDET-QGIS tool is also tested against the same datasets.

110 **2.1 RICorDE v1.0**

RICorDE produces gridded water depth estimates by incorporating a HAND sub-model and cost distancing algorithms to extrapolate edge values into the inundated region. Built for estimating depths from an approximate polygon produced by the 'Floods in Canada' (FiC) project, RICorDE draws all input data from Canadian web-hosted sources, only requiring the user to specify the period and area of interest. However, users can supply similarly formatted input data from alternate sources.
115   Following data downloading, input and pre-processing, RICorDE uses the WhiteboxTools 'ElevationAboveStream' tool (Lindsay, 2014) to generate the HAND values raster from user-supplied permanent waterbody polygons and the DEM. Using these inputs, the central depth estimating algorithm of RICorDE has three phases: 1) hydraulically correcting the approximate inundation polygon to remove egregious over-predictions; 2) interpolating rolling HAND values for the flooded domain and corresponding water surface levels (WSL); and finally 3) subtracting the water level grid from the DEM. The remainder of
120   this section provides additional detail on these three phases of the algorithm.

The first phase to hydraulically correct the approximate raw inundation polygon is summarized in Fig. 1. To address under-predictions in the raw polygon (areas falsely shown as dry), user supplied polygons denoting permanent water bodies are used to fill erroneously dry areas (Fig. 1 Panel 1). To address over-predictions (areas falsely shown as wet) the newly corrected inundation polygon is used to generate shoreline HAND values along valid edges (Fig. 1 Panel 2 and 3). From these samples,

125 the upper or third quartile (value between the median and maximum) is calculated and used to generate a HAND inundation polygon by polygonizing a mask of all lesser HAND raster values (Fig. 1 Panel 4). Finally, this new 'maximum inundation' polygon is used to clip all external values from the approximate inundation (Fig. 1 Panel 5). In this way, egregious over- and under- predictions in the raw inundation polygon are corrected before proceeding with the algorithm.

130 **Figure 1: Process diagram of the first phase of RICorDE's depth algorithm for generating the hydraulically corrected inundation polygon showing five basic steps; where 'q3' is the third quartile of the sampled HAND values. Basemap imagery from © Maxar Technologies.**

The second phase of RICorDE's depth algorithm is summarized in Fig. 2 and develops WSLs from the hydraulically corrected inundation of the previous phase. The second phase begins by generating an interior surface of HAND values that best

135 represents the flooding in each grid cell ('Rolling HAND Grid') from which the final WSL is generated by mosaicking the HAND derived WSLs corresponding to each cell. This process begins with a second sampling of the shoreline HAND values using the hydraulically corrected inundation of the previous phase (Fig. 2 Panel 1). These second-generation values are again filtered using statistics from the initial first-generation HAND sampling: where a lower bound is forced from the first quartile and an upper bound is forced from the third quartile (Fig. 2 Panel 2). This second aspatial filtering of HAND values is required

140 to address sampling errors that arise from small spatial shifts introduced in the polygonization process of the hydraulic correction in the first phase (Fig. 2 Panel 4), while preserving the inundated area. From these corrected values, continuous shoreline HAND values are generated by first interpolating using an Inverse Distance Squared Weighting algorithm (GRASS Development Team, 2017) as shown in Fig. 2 Panel 3, before masking out the interior region (Fig. 2 Panel 4). Interior HAND

values are then generated using the WhiteboxTools 'CostAllocation' algorithm (Lindsay, 2014) as shown in Fig. 2 Panel 5.
This algorithm is similar to the one used by FwDET version 2.0 and our testing showed this produced more hydraulically reasonable results than the Inverse Distance Squared Weighting from previous steps. The final Rolling HAND Grid is obtained by smoothing with a low-pass filter (GRASS Development Team, 2017) iteration loop until a HAND value gradient less than or equal to 0.1 m/m is achieved (Fig. 2 Panel 6). With this smoothing, the algorithm balances the flooding surface implied by the hydraulically corrected inundation and a flooding surface (of constant height) derived from the DEM. By smoothing HAND values, rather than WSLs, the algorithm generates results with a bias towards the vertical profile of the flow path (e.g., WSL in a flooded river channel), rather than a flat-water surface.

Once the 'Rolling HAND Grid' is obtained, the second phase continues by generating a HAND inundation raster for each unique value within the rolling grid, following the same procedure described in phase 1 (Fig. 2 Panel 7). For each of these inundations, a corresponding WSL raster is generated by first masking out interior regions on the DEM, then applying the 'CostAllocation' algorithm to grow the shoreline values into the interior, and finally masking out the exterior (Fig. 2 Panel 8). This is similar to the FwDET version 2.0 routine; however, here more realistic inundation regions obtained from the HAND model are used (rather than approximate inundation polygons). Because these inundations are derived from the DEM itself, DEM artifacts do not lead to internal inconsistencies within the WSL rasters. Phase 2 concludes by mosaicking values from the WSL set according to the positions of the corresponding HAND value found in the rolling HAND grid, then applying a final low-pass filter (Fig. 2 Panel 9).

In the third and final phase of RICorDE's depth algorithm, the DEM is subtracted from the WSL Mosaic to obtain the raw depth raster. From this raster, depths less than or equal to zero are removed, and the remaining raster is clipped to match the extents of the hydraulically corrected inundation (from the first phase).

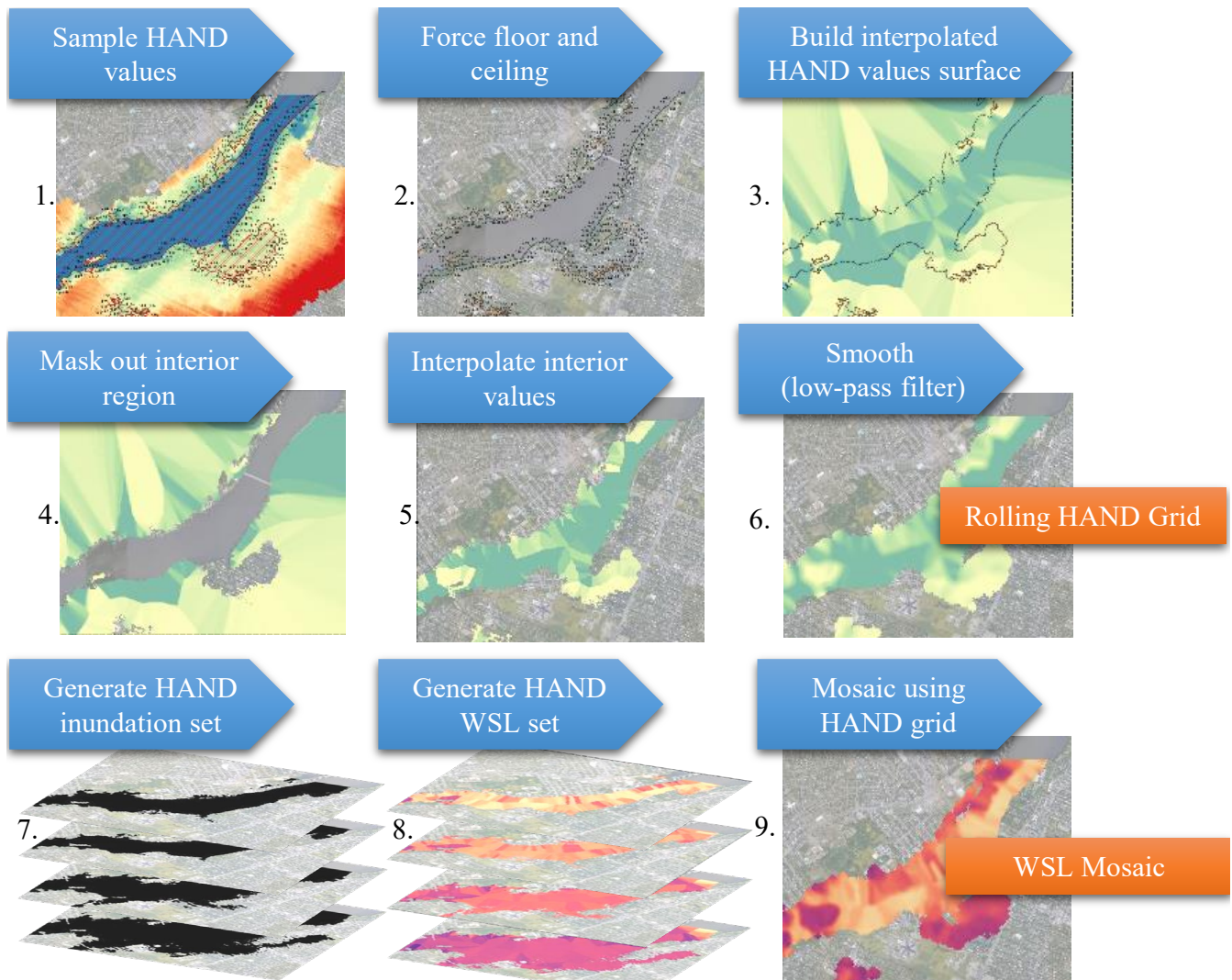



Figure 2: Process diagram for the second phase of RICorDE's depth algorithm for generating the Rolling HAND Grid and the
Water Surface Level (WSL) Mosaic. Basemap imagery from © Maxar Technologies.

2.2 Evaluation

To evaluate the performance of RICorDE v1.0's novel depth estimation algorithm, approximate SAR-derived inundation
polygons produced by the Floods in Canada project (FiC) for a 2018 flood in New Brunswick and a 2017 flood in Quebec,
Canada are used to generate gridded depth estimates. Additional gridded depth estimates are generated using FwDET-QGIS
(Cohen et al., 2019) for comparison. These simulated depth grids are then tested against corresponding 'trusted' grids obtained
from hydrodynamic modelling done by others. To test performance, metrics based on the depth sampled at building locations
are used.

2.2.1 Study Flood Events

175   Two recent spring fluvial flooding events in Canada were selected for evaluation.

Rivière des Prairies at Montreal, Quebec — May 2017 Flood: The Rivière des Prairies is a deltaic channel dividing Laval from Montréal, two cities in Southern Quebec at the confluence of the regulated St. Lawrence River and Ottawa River. The confluence has a drainage area of roughly 240,000 km^2 and 150,000 km^2 for the St. Lawrence and Ottawa rivers respectively. Above-average precipitation in April and early May 2017, combined with snowmelt, contributed to the highest flow on record

180   for the Ottawa River and regulators discharging the maximum allowable flow to the St. Lawrence River (Teufel et al., 2019). On May 8th, 2017 the water level peaked near the mouth of the Rivière des Prairies at an elevation of 24.8 m, the highest in the 26-year record, while the discharge peaked at 3,310 m^3/sec (Environment and Climate Change Canada, 2021).

Following the 2017 event, the Communauté métropolitaine de Montréal (CMM) generated a maximum WSL map using a two-dimensional hydrodynamic model of the Rivière des Prairies. This model was built in the H2D2 platform (INRS-ETE, 2022)

185   using LiDAR data obtained by CMM in April 2016 and calibrated using high water marks collected by CMM on April 29th 2019 (Edwards, 2022). These measurements showed the 2019 event reached levels within 5 cm of the 2017 peak (Communauté Métropolitaine de Montréal, 2019). Levees and temporary flood mitigation structures were not included in the model (i.e., some areas that remained dry in 2017 are shown as flooded on CMM's map). To correct for this, levee-protected areas identified through visual inspection, as shown in Fig. S1, were removed for this analysis.

190   *Saint John River at Fredericton, New Brunswick — May 2018 Flood:* The Saint John River is a regulated river that drains roughly 55,000 km^2 of mostly forested regions within Maine (U.S.), Quebec (Canada), and New Brunswick (Canada) before reaching the Bay of Fundy at Saint John in New Brunswick (Newton and Burrell, 2016). The lower reaches of the river, which are broad and shallow, form the flood-prone New Brunswick Lowlands, home to the provincial capital of Fredericton and numerous recorded flood disasters, most notably in 1973, 2008 (McGrath et al., 2015), and 2018 (Environment and Local

195   Government, 2021b). Triggered by rapid snowmelt and rain, on May 1st, 2018 the Saint John River peaked with a discharge of 6,070 m^3/sec and a WSL of 8.2 m at Grand Falls and Fredericton respectively, the fourth highest WSL in the 85-year record (Environment and Climate Change Canada, 2021). An estimated 12,947 buildings were damaged by the flood requiring $CAD 80 million in government relief (Hrabluk, 2019).

Following the 2018 event, the Province of New Brunswick (PoNB) generated a maximum WSL map through hydrodynamic

200   modelling (Environment and Local Government, 2021a). The model used was a one-dimensional HEC-RAS (version 4.1.0) parameterization of the Saint John river system with 83 cross-sections cut from a digital terrain model developed by the Province from various LiDAR missions flown from 2015 to 2018 (GeoNB, 2022). PoNB adapted this model to match the maximum instantaneous water level observations from May 1st, 2018, at the 8 available gauge stations, rounding WSL values to the nearest decimeter (Boisvert, 2021). PoNB validated their result against aerial imagery of the 2018 peak inundation and

205 high water marks surveyed during the flood peak of the following year, which had a gauge reading at Fredericton within 5 cm of the 2018 level (Environment and Climate Change Canada, 2021).

2.2.2 Tool Inputs

Three publicly available datasets produced by Natural Resources Canada provided the primary inputs for the depth estimation tools used in this evaluation: 1) DEMs were sourced from the High-Resolution Digital Elevation Model (HRDEM); 2)
210 approximate inundation polygons were sourced from the Floods in Canada project (FiC); and 3) permanent waterbody polygons were obtained from the National Hydro Network project (NHN). The remainder of this section summarizes these initiatives.

Since 2011, Natural Resources Canada has been modernizing Canada's terrain data through the HRDEM Mosaic project, which seeks to provide continuous high-resolution elevation data for the whole country. For the southern regions considered
215 here, this effort involves collecting historical LiDAR data from partners and supporting new acquisitions. From this collection of point-clouds, HRDEM algorithms generate a Triangular Irregular Network (TIN) from 'ground' and 'water' classified points. This TIN is then rasterized, mosaicked, and hosted as a Web Coverage Service (WCS) with a resolution of 1 or 2 m (Government of Canada and Natural Resources Canada, 2020). As of July 2021, HRDEM covers nearly 500,000 km^2 of Canada. For the HRDEM areas used here, source LiDAR data matches that described above for Rivière des Prairies and the
220 Saint John River hydrodynamic models.

NHN is a national database of inland waters, providing vector data of waterbodies, watercourses, reservoirs, man-made obstructions (e.g. dams), etc. from the best available provincial and federal collections of data, at 1:50,000 scale or better (Government of Canada, 2004).

In December 2007, the Canadian Space Agency launched the RADARSAT-2 mission which carries a 15 m Synthetic Aperture
225 Radar (SAR) antenna (Canadian Space Agency, 2021a). Expanding disaster monitoring capabilities, the RADARSAT Constellation Mission was later launched in June 2019 (Canadian Space Agency, 2021b). A primary user of these missions is the Emergency Geomatics Services team at Natural Resources Canada. During major flood events, both optical and SAR satellite imagery are used in conjunction with ancillary data layers to generate near-real time mapping of flood events to support emergency response activities as part of the FiC project (Natural Resources Canada, 2020). A multi-step process is employed
230 to map open water and flooded vegetation, which includes supervised machine learning classification and threshold-based region growing. The FiC data repository (https://open.canada.ca/data/en/dataset/08b810c2-7c81-40f1-adb1-c32c8a2c9f50) contains active flood extents (current to past 72 hours) and archived extents dating back to 2011.

For the Saint John River event, the FiC scene from May 3rd, 2018, was selected. For the Rivière des Prairies event, the FiC scene from May 9th, 2017, was selected. Corresponding metadata states that both scenes have a 'moderate' confidence level
235 and are SAR derived. Maps summarizing the tool data inputs used for both events are provided in Fig. S1 and Fig. S2.

2.2.3 Performance Metrics

To test the performance of the depth-estimate tools, similar studies often compare per-cell depth values (Nguyen et al., 2016; Brown et al., 2016; Cian et al., 2018; Cohen et al., 2018; Scorzini et al., 2018), biasing the performance towards areas without assets (i.e., focusing on the whole domain rather than building locations). In regions with heterogeneous asset densities, like in the two areas of this study where development has occurred along riverbanks, this per-cell performance reporting strategy is less useful for flood vulnerability researchers who are interested in the exposure of assets (e.g., buildings) not open floodplains. Further, this per-cell reporting often obscures the performance of binary wet/dry predictions by only calculating metrics for overlapping cells (i.e., where both the validation and estimated grids indicate flooding).

To address these challenges, this study focuses on metrics based on building locations and reports performance for inundation (wet vs. dry) predictions separate from depth value predictions. This separation allows for a more robust evaluation of each tool as inundation predictions are more sensitive to the input inundation polygon, while the depth value predictions are more closely related to algorithm performance.

For both inundation and depth metrics, depth values are first sampled from the grids/rasters at asset locations obtained from centroids of the 'CanadianBuildingFootprints' project layers (Microsoft, 2019). For each tool, the inundation performance is then calculated against values sampled from the trusted grid. To report on the performance of the depth predictions, samples with zero value are discarded. To calculate difference and correlation metrics, first the trusted and simulated depth sets (with zeros removed) are paired. For the correlation analysis, missing pair values are discarded. For the difference analysis, missing pair values are replaced with zeros before subtracting each trusted depth from its paired simulated depth. Because each depth raster differs in extents (i.e., which cells are predicted wet/dry), this paired analysis results in different size datasets within the same trial for each tool comparison.

3 Results and Discussion

Maps of the resulting gridded-depth simulations for the two study floods are provided in Fig. S3 and S4. Raw asset samples for the gridded-depth simulations generated by RICorDE and FwDET-QGIS, and corresponding trusted values, for the two flood events are summarized in Table 1. Computation time for all simulations, executed on a single core of an Intel i7-10700 (2.90 GHz), are provided in minutes. This table also summarizes four performance metrics by comparing simulated values against the trusted values at each asset: 1) root mean square error (RMSE); 2) mean of all difference values (simulated – trusted); 3) count of difference values (simulated – trusted) exceeding 2 m; and 4) the correlation coefficient. This table shows that RICorDE outperforms FwDET-QGIS for seven of the eight metrics and trials considered. The remaining underperforming metric, the 'difference mean' of the Saint John River event, is discussed below. Of particular interest is the improvement provided by RICorDE in reducing the number of predictions with large errors, as shown by the 'difference > 2m (count)' metric.

Table 1: Summary of the trusted and simulated depth grids for the two study floods.

depth grid code	study flood event trial	method	run time (mins)	count	max (m)	RMSE (cm)	difference mean (m)	difference > 2m (count)	r value
				raw values			**performance**		
SJ_PoNB	2018 Saint John River	trusted		1045	6.0				
SJ_FQ	2018 Saint John River	FwDET-QGIS	3	1108	29.6	194.2	0.12	114	0.202
SJ_RIC	2018 Saint John River	RICorDE	90	569	6.6	79.0	-0.31	11	0.582
dP_CMM	2017 Rivière Des Prairies	trusted		542	1.3				
dP_FQ	2017 Rivière Des Prairies	FwDET-QGIS	1	1741	12.2	60.6	0.16	19	0.118
dP_RIC	2017 Rivière Des Prairies	RICorDE	30	776	3.1	50.9	0.10	5	0.147

To show the accuracy of inundation predictions (i.e., wet vs. dry) from the two tools using the FiC approximate polygons, Fig. 3 shows the portion of over- (tool predicts the asset is wet in error), under- (tool predicts the asset is dry in error), and accurate-inundation predictions. This figure shows that both tools yielded less accurate *inundation* predictions for the Des Prairies trial (than for the Saint John), in contrast to the more accurate *depth* value predictions shown for this trial in Table 1.

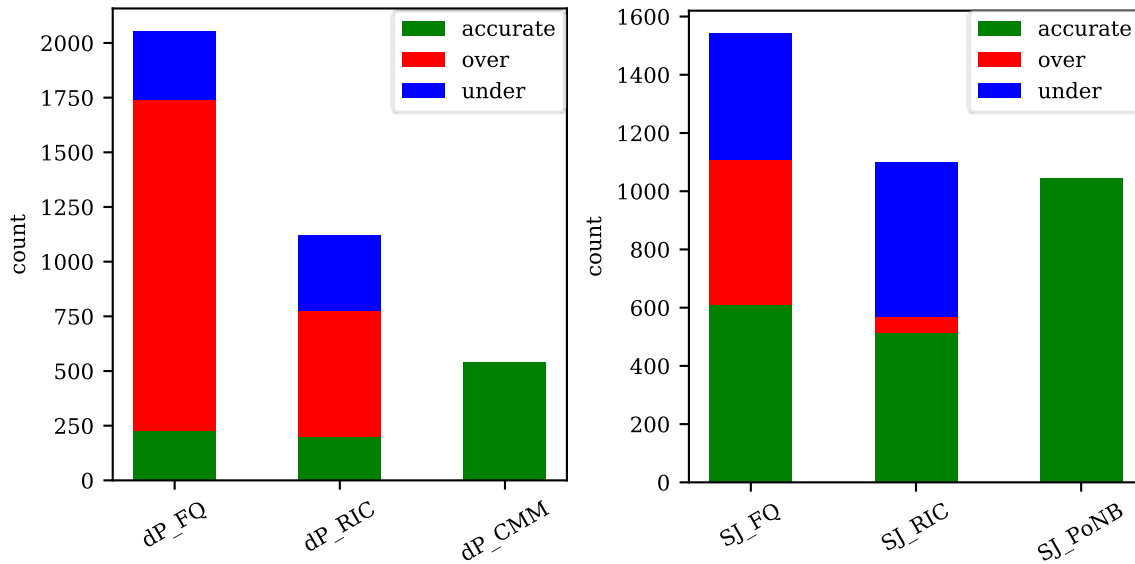



Figure 3: Wet/dry asset sample performance for the two study floods showing: [left] 2017 Rivière Des Prairies performance and [right] 2018 Saint John River performance. See Table 1 for additional legend.

Because FwDET-QGIS maps onto the raw inundation approximation (i.e., no hydraulic correction), values for these simulations ('FQ') reflect the accuracy of the FiC polygons themselves, rather than some underlying algorithm. From this, Fig.

3 shows that the two FiC scenes investigated both over- and under-predict inundation of assets; however, over-prediction is more severe. This is especially true for the Rivière des Prairies event (Fig. 3 'dP_FQ'), where the FiC polygons predicted less than half of the total assets accurately (228 of 542). This could be attributed to erroneous treatment of levees in the trusted data 'dP_CMM' (rather than errors in the FiC polygons); however, comparing the two trials suggests this possible artifact is less relevant than the inaccuracies inherent in the FiC polygons. Alternatively, the performance improvement between the 2017 des Prairies and the 2018 Saint John trials could also be attributed to advancements in the FiC project made during the year between the two floods.

From Fig. 3, the advantage of RICorDE's hydraulically corrected inundation can be seen in the relative decrease of 'over' predictions (compared to FwDET-QGIS). However, RICorDE performs slightly less well with 'under' prediction errors (predicts dry when the trusted shows wet). This 'dry bias' could be attributed to some over-correction of the inundation during the second phase of the depth algorithm (Fig. 2 Panel 2). This bias is also reflected in the lower (more negative) 'difference mean' values shown on Table 1 (relative to FwDET-QGIS), leading this metric to underperform for the Saint John Event despite the lower RMSE (0.657 vs. 1.915 m). If instead the *absolute* difference values are examined, the mean for RICorDE would outperform by 23 cm (0.643 vs. 0.876 m). In other words, the combined effect of slightly better under predictions and much worse over predictions somewhat balance, pulling the difference mean of the FwDET-QGIS Saint John trial closer to zero, despite having more erroneous predictions.

To demonstrate the performance of the simulated depth values for the 2018 Saint John River trial, sample values, difference values, and correlation plots are provided in Fig. 4.

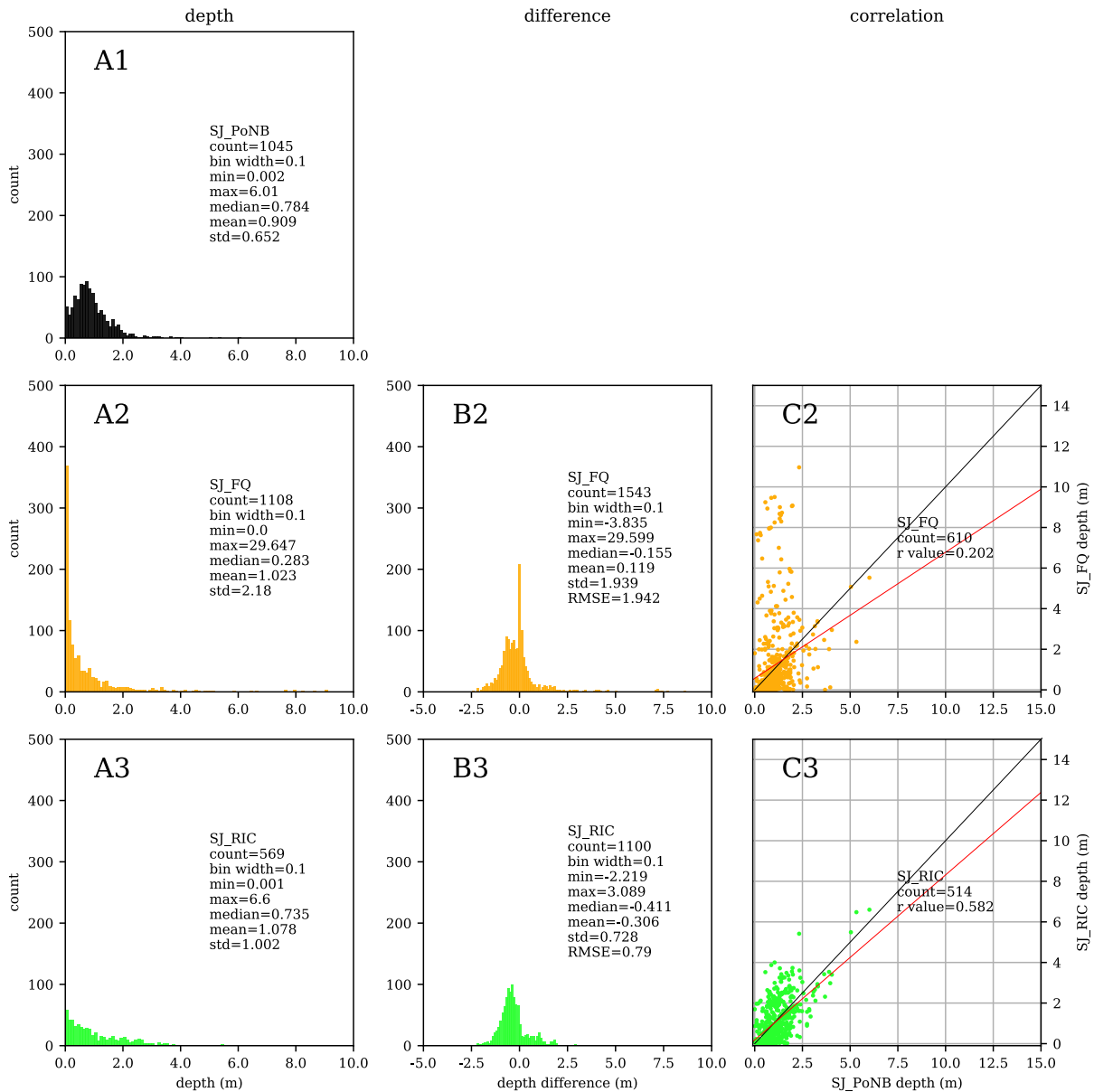



Figure 4: Performance metrics matrix plot with rows for: [3] RICorDE simulated, [2] FwDET simulated, and the [1] trusted depth grid of the 2018 Saint John River flood, showing values at building locations with columns for: [A] depth; [B] difference (simulated – trusted); and [C] linear correlation of simulated- (y-axis) against trusted-depths (x-axis). Values exceeding 10 m are hidden from row 2 panels for clarity. Common statistical metrics for the plotted data are shown within each plot along with the 'Depth Grid Code' from Table 1. Coloration is applied for convenience when cross-comparing with other figures.

The somewhat normally distributed depths of the trusted dataset are shown in Fig. 4 Panel A1 with a mean of 0.91 m and a max of 6.01 m. In contrast to this, Fig. 4 Panel A2 and A3 show a more exponential shape for the simulated results. While FwDET-QGIS does not have a built-in zero-filter, the 181 'zero depth' values were removed prior to this performance analysis

as previously noted. Similarly, the ten FwDET-QGIS simulations with depth values exceeding 10 m are hidden on the figure (these outliers are discussed further below). In contrast to this, the maximum depth estimated by RICorDE is 6.6 m. The disparities in inundation predictions between raster pairs, shown in Fig. 3, is also evident on Fig. 4 in the different sizes of each paired data set, shown as 'count' values on each panel. For example, the 'count=610' shown on Fig. 4 Panel C2 is equivalent to the height of the green bar shown on Fig. 3 'SJ_FQ'. The aforementioned 'dry bias' of RICorDE is also replicated in the depth values shown on Fig. 4 Panel B3.

Fig. 5 presents comparable plots for the 2017 Rivière Des Prairies flood showing similar performance. Interestingly, the shape of the trusted depth values histogram for this trial (Fig. 5 Panel A1) differs from that in the Saint John trial (Fig. 4 Panel A1). This could be a result of differences in topography, development patterns, flood behavior, levee performance, or the hydrodynamic modelling methods used (by others). While insufficient information was available to evaluate this further, these disparities point to the importance of incorporating multiple heterogeneous trials when evaluating tools like RICorDE.

For the Des Prairies trial, Fig. 5 again shows RICorDE outperforming FwDET-QGIS; with even the mean difference metric being more favorable by 5 cm. However, the comparable and relatively favorable performance of both tools in predicting depths should be weighed against the poor inundation performance shown in Fig. 3 for this trial. This suggests the FiC polygon used in this trial led to many false predictions; however, where the FiC polygon was accurate, RICorDE yielded better depth estimates.

All the performance metrics discussed above are sensitive to the treatment of zero values and paired values (with one dry or missing value). The treatment used here (and described above) was selected to provide broad and clear metrics; however, alternate treatments would also be reasonable. For example, the difference analysis could have discarded any paired values with a missing value, rather than setting the missing value to zero. Many of these alternate metrics were explored by the study-team — all yielded similarly favorable results for the performance of RICorDE.

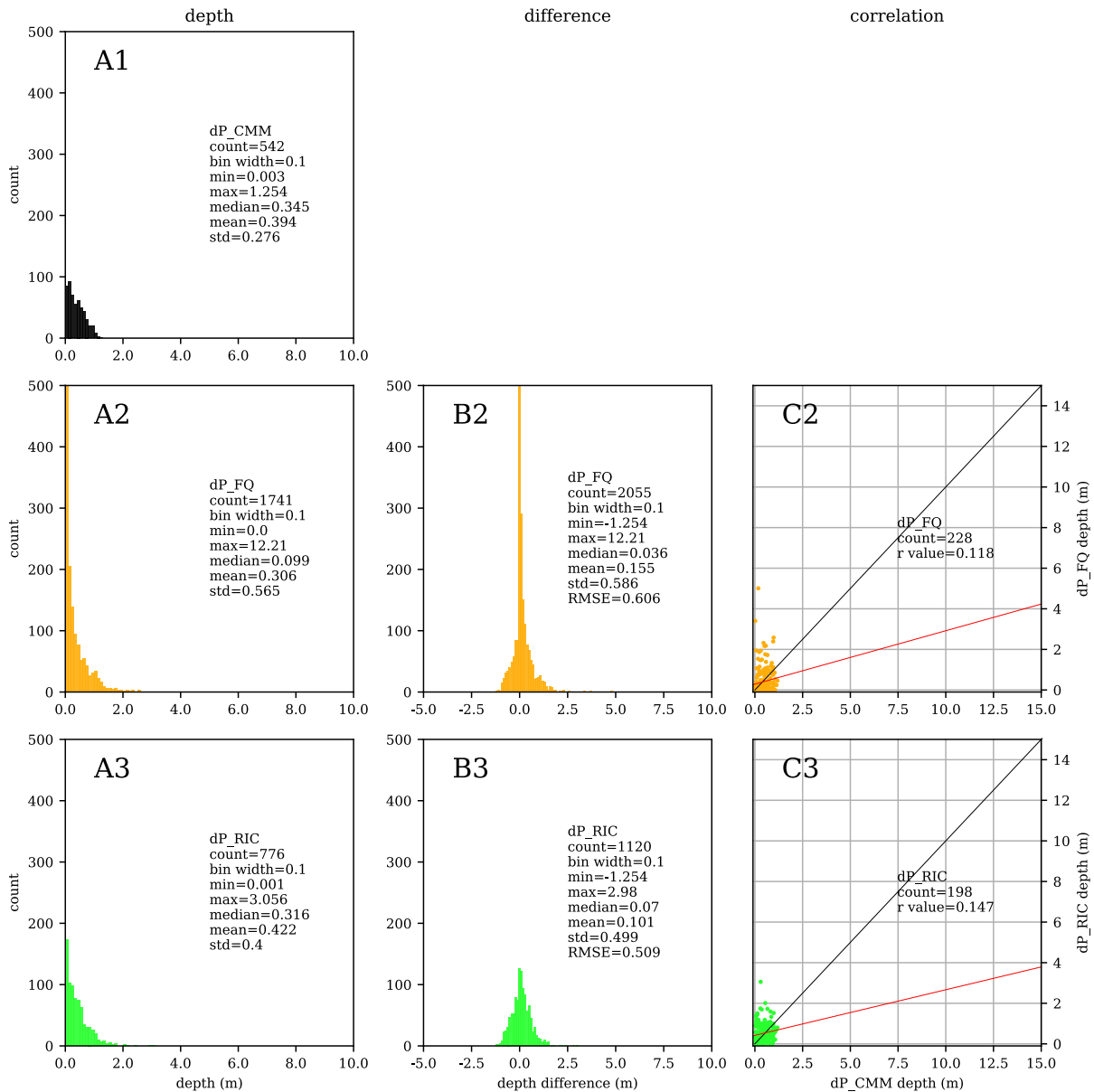



Figure 5: Performance metrics for the 2017 Rivière Des Prairies flood as in Figure 4.

To further investigate the performance of the two tools, Fig. 6 provides an overview and two comparable maps of the 2018
Saint John River study event. This shows a portion of the river where the FiC polygon erroneously identified inundation up
the riverbank to an elevation of 60 m (lower left of Fig. 6 Panel A and B). FwDET-QGIS interpolated these shoreline values
directly onto the interior, until values propagated from the opposing bank were encountered by the routine (Fig. 6 Panel B
black arrows) yielding depth errors in excess of 20 m for assets that should have been dry (according to the validation data).

Contrary to this, RICorDE's hydro-correction clipped this inundation to achieve shoreline elevations between 7 and 10 m before interpolating these onto the interior. This mechanism contributed to the overall more accurate predictions of RICorDE discussed above but was especially relevant in reducing the count of large errors indicated by the 'difference >2 m' metric in Table 1.

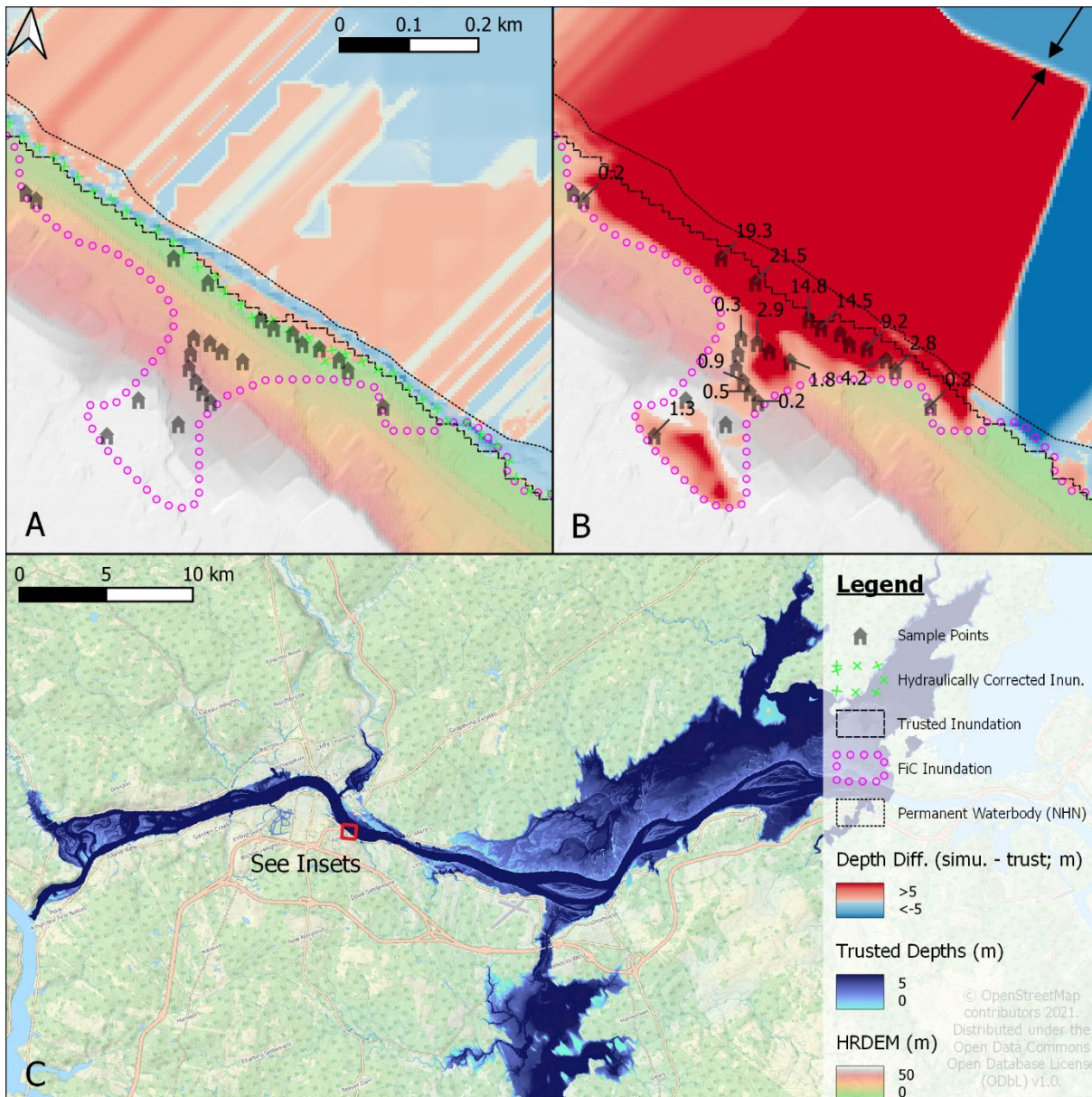



Figure 6: Gridded depths for the Saint John River, May 2018 flood event showing: [C] trusted depth estimates for the full study area; [A] RICorDE simulations minus trusted depth results with hydraulically corrected inundation; and [B] FwDET-QGIS minus trusted. All sample points with depth values are labelled accordingly.

4 Conclusions

This study developed the Rolling HAND Inundation Corrected Depth Estimator (RICorDE) Tool for predicting depths from approximate inundation polygons. Similar to previous tools, like FwDET-QGIS, RICorDE provides an efficient method that does not assume a flat-water surface and leverages remote sensing data and does not require hydrodynamic modelling expertise or difficult to obtain calibration data (e.g., bathymetry, high-water marks). Unlike previous tools, RICorDE incorporates some error correction of approximate inundation polygons and is structured around a HAND sub-model to facilitate more realistic water surfaces. These enhancements come at the cost of algorithm complexity and longer runtimes. This work enhances the utility of satellite derived data for studying flood events; thereby improving society's ability to plan and prepare for flood disasters.

To test the performance of RICorDE, two recent flood events were examined. Depth estimates were generated for these events from public data and approximate inundation polygons from the Floods in Canada project (FiC) using RICorDE and the popular FwDET-QGIS tool, before comparing against trusted depth grids. The presented results suggest the novel RICorDE's algorithm outperforms FwDET-QGIS in depth predictions for the two study floods investigated with RICorDE having a RMSE of 79 and 51 cm for the two trials. Inundation performance was mixed, with FwDET-QGIS having slightly fewer 'under' predictions (tool predicts the asset is dry in error) while RICorDE had far fewer 'over' predictions (tool predicts the asset is wet in error), suggesting a slight 'dry bias' for RICorDE. RICorDE substantially outperformed FwDET-QGIS in the treatment of outliers, returning an order-of-magnitude fewer predictions with errors exceeding 2 m for one trial. These trials demonstrated that both algorithms remain limited by inaccuracies in satellite-derived inundations, with both tools predicting less than half of asset inundations accurately for one trial. Future work should consider improving the underlying satellite derived inundations, improving the run times and computational efficiency of RICorDE, testing against different flood hazards (e.g., ice jam flooding), integrating uncertainty quantification, and porting the tool to a more user-friendly environment.

Where RICorDE is used to predict depths at building locations from FiC polygons, the two trials performed here suggest reasonable depth estimates can be obtained for those buildings that were truly inundated; however, the two FiC polygons examined from 2017 and 2018 were unable to reliably predict this inundation. While the objective of this study was not to test the accuracy of the FiC project, improvements were found between the 2017 and 2018 FiC polygons. This limited observation suggests new satellite missions and internal advancements may bring more useful FiC polygons. With such inundation polygons, RICorDE provides a more accurate means of studying historical floods remotely and at scale.

Code Availability

Code is available by request from the authors.

Data Availability

The Saint John River 2018 maximum WSL validation data is hosted by GeoNB (http://www.snb.ca/geonb1/e/DC/floodraahf.asp) under the 'GeoNB Open Data License'. The Rivière des Prairies 2017
375 maximum WSL trusted data was provided by Brent Edwards (brent.edwards@cmm.qc.ca) of the Communauté métropolitaine de Montréal under license to the project team. Datasets used for the depth tools are publicly available at the locations mentioned in the text.

Author Contribution

Motivation for the work stemmed from discussions between the three authors about developing gridded depths from the Floods
380 in Canada project and well-known tools. All authors contributed to collecting and locating the input and trusted data. SB developed the code, performed the simulations and analysis, and prepared the manuscript with contributions from MB and HM.

Competing interests

The authors declare that they have no conflict of interest.

385 **Acknowledgements**

The authors thank Jasmin Boisvert and Reid McLean of the New Brunswick Department of Environment and Local Government for their insight into the 2018 flooding of the Saint John River. We are similarly grateful for the efforts of Brent Edwards and Nicolas Milot of the Communauté métropolitaine de Montréal for their assistance in obtaining data on the 2017 floods there. Charles Papasodoro at the Canada Centre for Mapping and Earth Observation provided valuable guidance on
390 integrating HRDEM into the project.

References

Benoudjit, A.: Operational mapping of the flood extent and depth from SAR images, Ph.D., University of Surrey, Surrey Space Centre, 2019.

Boisvert, J.: Frederickton Flood Simulations. Personal Communication., 2021.

395 Brown, K. M., Hambidge, C. H., and Brownett, J. M.: Progress in operational flood mapping using satellite synthetic aperture radar (SAR) and airborne light detection and ranging (LiDAR) data, PROG PHYS GEOG, 40, 196–214, https://doi.org/10.1177/0309133316633570, 2016.

Bryant, S., Davies, E., Sol, D., and Davis, S.: The Progression of Flood Risk in Southern Alberta Since the 2013 Flood [accepted], J. Flood Risk Manag., https://doi.org/10.13140/RG.2.2.26504.14086, 2021.

Canadian Space Agency: https://www.asc-csa.gc.ca/eng/satellites/radarsat2/default.asp, last access: 10 September 2021a.

Canadian Space Agency: https://www.asc-csa.gc.ca/eng/satellites/radarsat/default.asp, last access: 19 August 2021b.

Cian, F., Marconcini, M., Ceccato, P., and Giupponi, C.: Flood depth estimation by means of high-resolution SAR images and lidar data, Nat. Hazards Earth Syst. Sci., 18, 3063–3084, https://doi.org/10.5194/nhess-18-3063-2018, 2018.

GRASS GIS manual: https://grass.osgeo.org/grass78/manuals/r.grow.distance.html, last access: 20 August 2021.

Cohen, S., Brakenridge, G. R., Kettner, A., Bates, B., Nelson, J., McDonald, R., Huang, Y.-F., Munasinghe, D., and Zhang, J.: Estimating Floodwater Depths from Flood Inundation Maps and Topography, J Am Water Resour Assoc, 54, 847–858, https://doi.org/10.1111/1752-1688.12609, 2018.

Cohen, S., Raney, A., Munasinghe, D., Loftis, J. D., Molthan, A., Bell, J., Rogers, L., Galantowicz, J., Brakenridge, G. R., Kettner, A. J., Huang, Y.-F., and Tsang, Y.-P.: The Floodwater Depth Estimation Tool (FwDET v2.0) for improved remote sensing analysis of coastal flooding, Nat. Hazards Earth Syst. Sci., 19, 2053–2065, https://doi.org/10.5194/nhess-19-2053-2019, 2019.

Communauté Métropolitaine de Montréal: RIVIÈRE DES PRAIRIES AMONT - ÉVALUATION DES COTES RELATIVES AUX PLUS HAUTES EAUX CONNUES - Final Report, 2019.

Donchyts, G., Winsemius, H., Schellekens, J., Erickson, T., Gao, H., Savenije, H., and van de Giesen, N.: Global 30m Height Above the Nearest Drainage, 2016.

Environment and Climate Change Canada: https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html, last access: 23 August 2021.

Edwards, B.: Map of 2017 flood event. Personal Communication., 2022.

Environment and Local Government: https://www2.gnb.ca/content/gnb/en/departments/elg/news/news_release.2020.08.0446.html, last access: 23 August 2021a.

Environment and Local Government: https://www.elgegl.gnb.ca/0001/en/Flood/Details/360, last access: 20 December 2021b.

GeoNB: https://geonb.snb.ca/image/rest/services/Elevation/DEM_Raw_MNE_Brut/ImageServer, last access: 14 January 2022.

Government of Canada: National Hydro Network Data Production Catalogue, 2004.

Government of Canada: https://www.infrastructure.gc.ca/dmaf-faac/index-eng.html, last access: 9 September 2021.

National Disaster Mitigation Program: https://www.publicsafety.gc.ca/cnt/mrgnc-mngmnt/dsstr-prvntn-mtgtn/ndmp/index-en.aspx, last access: 17 February 2017.

Government of Canada and Natural Resources Canada: High Resolution Digital Elevation Model Mosaic (HRDEM Mosaic) – CanElevation Series – Product Specifications, 2020.

430 GRASS Development Team: Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.2, Open Source Geospatial Foundation, 2017.

Gupta, R. and Banerji, S.: Monitoring of reservoir volume using LANDSAT data, Journal of Hydrology, 77, 159–170, 1985.

Hrabluk, L.: New Brunswick Underwater: The 2018 Saint John River Flood, MACINTYRE & PURCELL, 1-77276-115-X, 2019.

435 INRS-ETE: http://www.gre-ehn.ete.inrs.ca/H2D2/contenu_description_en/introduction-to-h2d2, last access: 14 January 2022.

Johnson, J. M., Munasinghe, D., Eyelade, D., and Cohen, S.: An integrated evaluation of the National Water Model (NWM)– Height Above Nearest Drainage (HAND) flood mapping methodology, Nat. Hazards Earth Syst. Sci., 19, 2405–2420, https://doi.org/10.5194/nhess-19-2405-2019, 2019.

Laudan, J., Rözer, V., Sieg, T., Vogel, K., and Thieken, A. H.: Damage assessment in Braunsbach 2016: data collection and 440 analysis for an improved understanding of damaging processes during flash floods, Nat. Hazards Earth Syst. Sci., 17, 2163–2179, https://doi.org/10.5194/nhess-17-2163-2017, 2017.

Lindsay, J.: The whitebox geospatial analysis tools project and open-access GIS, in: Proceedings of the GIS Research UK 22nd Annual Conference, The University of Glasgow, 16–18, 2014.

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., and Wang, S.: A CyberGIS Integration and Computation Framework 445 for High-Resolution Continental-Scale Flood Inundation Mapping, J Am Water Resour Assoc, 54, 770–784, https://doi.org/10.1111/1752-1688.12660, 2018.

McGrath, H., Stefanakis, E., and Nastev, M.: Sensitivity analysis of flood damage estimates: A case study in Fredericton, New Brunswick, Int. J. Disaster Risk Reduct., 14, 379–387, https://doi.org/10.1016/j.ijdrr.2015.09.003, 2015.

Merz, B., Kreibich, H., Schwarze, R., and Thieken, A.: Review article "Assessment of economic flood damage," Nat. Hazards 450 Earth Syst. Sci., 10, 1697–1724, https://doi.org/10.5194/nhess-10-1697-2010, 2010.

Microsoft: microsoft/CanadianBuildingFootprints, https://github.com/microsoft/CanadianBuildingFootprints, 2019.

Mohor, G. S., Hudson, P., and Thieken, A. H.: A Comparison of Factors Driving Flood Losses in Households Affected by Different Flood Types, Water Resour. Res., 56, https://doi.org/10.1029/2019WR025943, 2020.

Natural Resources Canada: Floods in Canada Product Specifications, https://data.eodms-sgdot.nrcan-455 rncan.gc.ca/public/EGS/EGS_FGP_Geodatabases/Flood_Inondation/EGS_FloodExtent_ProductGuide.pdf, 19 March 2020.

Newton, B. and Burrell, B. C.: The April–May 2008 flood event in the Saint John River Basin: Causes, assessment and damages, Can. Water Resour. J., 41, 118–128, https://doi.org/10.1080/07011784.2015.1009950, 2016.

Nguyen, N. Y., Ichikawa, Y., and Ishidaira, H.: Estimation of inundation depth using flood extent information and hydrodynamic simulations, Hydrol. Res. Lett., 10, 39–44, https://doi.org/10.3178/hrl.10.39, 2016.

460 Office of the Parliamentary Budget Officer: Estimate of the Average Annual Cost for Disaster Financial Assistance Arrangements due to Weather Events, Ottawa, Canada, 2016.

Peter, B. G., Cohen, S., Lucey, R., Munasinghe, D., Raney, A., and Brakenridge, G. R.: Google Earth Engine Implementation of the Floodwater Depth Estimation Tool (FwDET-GEE) for Rapid and Large Scale Flood Analysis, IEEE Geosci. Remote Sensing Lett., 1–5, https://doi.org/10.1109/LGRS.2020.3031190, 2020.

465   Raney, A. and Cohen, S.: Floodwater Depth Estimation Tool, Python, csdms-contrib, https://github.com/csdms-contrib/fwdet/blob/master/fwdet/FwDET_QGIS_plugin.py, 2019.

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia, Remote Sens. Environ., 112, 3469–3481, https://doi.org/10.1016/j.rse.2008.03.018, 2008.

470   Rodda, H. J. E.: The Development and Application of a Flood Risk Model for the Czech Republic, Nat Hazards, 36, 207–220, https://doi.org/10.1007/s11069-004-4549-4, 2005.

Scorzini, A., Radice, A., and Molinari, D.: A New Tool to Estimate Inundation Depths by Spatial Interpolation (RAPIDE): Design, Application and Impact on Quantitative Assessment of Flood Damages, Water, 10, 1805, https://doi.org/10.3390/w10121805, 2018.

475   Shen, X., Wang, D., Mao, K., Anagnostou, E., and Hong, Y.: Inundation Extent Mapping by Synthetic Aperture Radar: A Review, Remote Sens., 11, 879, https://doi.org/10.3390/rs11070879, 2019.

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F. W., Dutta, D., and Kim, S.: Flood inundation modelling: A review of methods, recent advances and uncertainty analysis, Environ. Model. Softw., 90, 201–216, https://doi.org/10.1016/j.envsoft.2017.01.006, 2017.

480   Teufel, B., Sushama, L., Huziy, O., Diro, G. T., Jeong, D. I., Winger, K., Garnaud, C., de Elia, R., Zwiers, F. W., Matthews, H. D., and Nguyen, V.-T.-V.: Investigation of the mechanisms leading to the 2017 Montreal flood, Clim. Dyn., 52, 4193–4206, https://doi.org/10.1007/s00382-018-4375-0, 2019.