Multi-Event Machine Learning for Annual Flood Susceptibility Prediction across Canada

McGrath, Heather*[1]

*[1]Natural Resources Canada, Ottawa, ON Canada

Keywords: machine learning, flood mapping, flood susceptibility, temporal mapping

Abstract

Machine learning for flood susceptibility mapping (FSM) has traditionally relied on narrowly scoped events and temporally constrained datasets, limiting the generalizability and long-term utility of predictive models. We present a multi-event, multi-temporal modelling framework that leverages discrete flood occurrences from 2005 to 2023 to train a unified model capable of inference across an extended temporal horizon. Each flood event is treated as a spatio-temporal marker, enabling the model to learn evolving driver–event relationships and underlying temporal trends. Dynamic inputs (e.g., climate data, land use/land cover) are integrated with static geophysical features (e.g., digital terrain model and derivatives) to capture both transient and persistent influences on flood susceptibility. An XGBoost model is trained, tested, and validated using a 70/15/15 split, achieving an overall accuracy of 0.945, with true positive and true negative rates of 0.95 and 0.94, respectively. Precision scores for wet (flood-prone) and dry (non-flood-prone) classes are 0.94 and 0.95. Generated yearly national FSM maps from 2000 to 2023 were evaluated against published flood event datasets. Validation using national flood records, climate variability bulletins, and spatio-temporal analyses of year-to-year raster correlations confirms that years with elevated predicted susceptibility correspond to observed flood events. In addition, a weighted wetness score identified the years with both widespread and extreme flood-prone conditions, highlighting the model's ability to capture multi-scale temporal dynamics. These results demonstrate that multi-event, multi-temporal modelling enhances the temporal reach and robustness of geospatial flood prediction, providing a foundation for long-term monitoring, trend analysis, and policy-relevant scenario planning.

1. Introduction

Until recently, floods were typically modelled using physically-based hydrological and hydraulic methods — approaches that demand extensive data inputs, require careful calibration, and depend on experienced analysts to set up, run, and interpret the results (Canada, 2023; GOV.UK, 2023; Pechlivandis et al., 2013). Over the past decade, however, the application of machine learning (ML) to support flood mapping has expanded substantially (Pourzangbar et al., 2025).

ML-based flood susceptibility mapping (FSM) diverges from physically based hydrological and hydraulic models in conceptual approach but also in the nature and precision of outputs, producing a gradient of flood susceptibility instead of explicit metrics like inundation extent, water depth, or flow velocity. While hydrological models simulate the physical processes governing runoff generation and flow propagation, ML models learn statistical relationships from data and conditioning variables. As a result, ML derived FSMs may differ in spatial patterns and accuracy from process-based simulations and their suitability varies depending on their intended application.

Many FSM studies rely on data from a single flood event or a limited temporal window, constraining their ability to generalize beyond the calibration period, estimate flood likelihoods over multiple years, or detect long-term trends in flood dynamics. This limited temporal scope further reduces the robustness of FSM approaches, particularly under changing climatic and land-use conditions.

In FSM, an event represents a discrete occurrence that combines both spatial and temporal attributes. Each flood instance carries the imprint of its environmental and climatic drivers—such as precipitation intensity, antecedent soil moisture, land cover, and catchment storage capacity—that may vary significantly over time. Traditional single-event models implicitly assume that these driver–response relationships are temporally stable. However, such relationships evolve under the influence of climate change, increasing storm frequency and intensity, land-use modification, and the expansion of built infrastructure. Failing to incorporate this spatio-temporal heterogeneity risks overfitting models to specific periods and reduces their capacity for extrapolation and long-term inference.

A promising pathway to overcome these limitations is through a multi-event (ME) modelling framework, in which many flood events spanning extended temporal and spatial domains are integrated during training. By exposing ML algorithms to a diverse set of hydrological and climatic scenarios, these ME frameworks support the learning of generalized, temporally resilient patterns, thereby enhancing model robustness.

In this paper, we present a ME ML framework for FSM that integrates historical flood data and associated environmental drivers across two decades. The framework and outputs are evaluated in terms of predictive accuracy, temporal extension (including back-casting and forecasting), and the interpretability of driver–event relationships. By demonstrating the efficacy of a ME training regime, this study advances FSM mapping toward more temporally comprehensive and generalizable inference, extending beyond the constraints of conventional single-epoch approaches.

* Corresponding (presenting) author

2. Related Work

ML has become a dominant approach for FSM mapping, with numerous studies comparing algorithmic performance, feature engineering strategies, and validation methods.

For example, numerous studies have evaluated the performance of multiple classifiers for FSM. These include comparisons among Naïve Bayes Trees and Naïve Bayes (Khosravi et al., 2019); Support Vector Machines (SVM), Decision Trees (DT), Random Forest (RF), Light Gradient Boosting Machine (LightGBM), Extreme Gradient Boosing (XGBoost) (Seydi et al., 2022); as well as RF, Neural Networks (NN), SVM, MARS, (McGrath and Gohl, 2022) across hydrological and physiographic contexts to benchmark performance. Likewise, hybrid and ensemble ML frameworks are widely tested and used, (Choubin et al., 2019; Fang et al., 2022; Paryani et al., 2023; Prasad et al., 2022). Collectively, this body of literature demonstrates extensive experimentation across diverse hydrological and physiographic contexts, with each study offering localized insights and model recommendations. Beyond the flood domain (Fernández-Delgado et al., 2014) conducted a comprehensive benchmarking of hundreds of ML classifiers across spatial and non-spatial real-world problems, illustrating the breadth of algorithmic variability that also characterizes FSM research. Feature engineering and data preprocessing strategies have been examined as well in many papers. These include the selection and preprocessing of numeric and classified features, as discussed by (Al-Kindi and Alabri, 2024; Dunbar et al., 2025; Islam et al., 2025; McGrath and Gohl, 2022). As well, validation approaches vary widely, with some studies emphasizing spatial cross-validation, while others rely on random splits or temporal holdouts. A comprehensive review of parameters and methods in the Web of Science core collection found the number of methods used is quite high, and the parameters included is quite varied (Kaya and Derin, 2023). The accuracy and reliability of FSM which may be used for identifying flood-prone areas, land use planning, emergency preparedness, and/or disaster risk reduction is subject to the choice of parameters, modelling method and judgement of the expert, and all of these choices can lead to different FSM for a single region (Kaya and Derin, 2023).

Despite this rich body of work, several persistent limitations remain. One key limitation is temporal scope. Recent studies highlight the importance of explicitly accounting for temporal drift and evolving environmental conditions to improve model generalizability. (Masrur et al., 2022) emphasize the need to incorporate spatio-temporal heterogeneity within model structures to enhance interpretability and transferability across different flood events and periods. Long-term, multi-event modelling has been less common but demonstrates the feasibility and benefits of extending predictive horizons. For example, (Costache et al., 2024) examined flood potential at a basin scale over a 30-year period (1990–2020). However, they did not fully articulate strategies for temporal partitioning of training and testing data, leaving room for further methodological refinement.

Another distinctive thread in FSM literature concerns the integration of static versus dynamic predictors. Typical approaches combine static terrain and land-cover variables (e.g.: elevation, slope, curvature) with dynamic climatic factors (e.g.: rainfall, temperature). Many studies treat these inputs as temporally invariant snapshots, implicitly assuming stability in driver–response relationships. Some studies are beginning to challenge that assumption and include characteristics such as rainfall intensity, duration and volume and peak flood discharge (Islam et al., 2025).

However few FSM studies make the leap toward multi-event, multi-temporal training.

In summary, while ML-based FSM has matured in classifier benchmarking, feature-selection, and mapping accuracy, it remains under-developed in its treatment of temporal extension and long-term trend inference. The present study addresses this gap by proposing a ME modelling framework that spans two decades of flood events and associated drivers—thereby advancing FSM mapping toward greater temporal robustness and generalisability.

3. Methodology

This study develops a multi-event (ME) XGBoost ML model trained on flood events from 2005, 2006, 2008, 2011, 2013–2015, and 2017–2023.

3.1 Training Data and Label Generation

Training labels for each historical event were derived from satellite-based flood-extent products generated by Natural Resources Canada's Emergency Geomatics Services (EGS) during major flood activations. Primarily these activations occur in the spring due to the combination of snow melt, raising temperatures and heavy precipitation, however they may be activated for any major flood. These operational flood delineations integrate synthetic aperture radar (SAR) and optical imagery, providing reliable, event-specific representations of inundated areas (Canada, 2025). This dataset defines the positive ("flood") class for model training. Non-flood samples were generated from the extent of the processed area but outside the identified inundation extents. The training data were grouped by year and location.

3.2 Independent Variables

The independent variables comprise three broad categories: (i) static terrain and geological data, (ii) temporal hydroclimatic and environmental data, and (iii) geospatial reference grids.

Static Data

Static predictors capture landscape characteristics, including terrain, lithology, and hydrography. A digital terrain model (DTM) (McGrath et al., 2025) and its topographic derivatives—aspect, roughness, slope, terrain roughness index, topographic position index, and Geomorphons—were incorporated to represent flow direction, accumulation zones, and areas of potential pooling or dispersion. Stream networks were derived from the DTM using tools from Whitebox Tools (Whitebox Geospatial Inc., 2025). From streams of Strahler Order ≥5, two additional layers were created: Euclidean distance to streams and height above nearest drainage (HAND). These variables capture proximity and elevation relationships relative to drainage pathways.

A surficial geology index map (Canada, 2015) from NRCan was included to represent variations in infiltration potential and surface runoff dynamics.

Temporal Data

Although spring temperature and precipitation are often regarded as the dominant temporal controls on the freshet, hydrological conditions in the preceding fall and winter are equally critical.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-M-1-2026-21-2026 | © Author(s) 2026. CC BY 4.0 License.

Autumn precipitation influences antecedent soil moisture and groundwater recharge, raising the water table and reducing infiltration capacity prior to freeze-up. During winter, the accumulation and compaction of the snowpack determine the snow water equivalent (SWE), which governs the potential volume of meltwater. Mid-winter thaws can further modify snowpack structure—partial melting followed by refreezing increases snow and ice density, altering the thermal and hydrological response during spring melt. When soils enter winter in a saturated state, subsequent thaws or freshet events produce greater surface runoff and reduced infiltration, increasing flood potential. Thus, effective flood likelihood modelling must account for the cumulative hydroclimatic conditions of the antecedent seasons, rather than relying solely on spring meteorological variables. Thus, we have included spring, winter, and fall datasets of precipitation and temperature.

Land-use and land-cover (LULC) data, produced every five years as part of the North American Land Change Monitoring System ((NALCMS) (Natural Resources Canada, 2025a; Pasos, 2025) were also included. For this project, datasets spanning 2000 thru 2020 were used, with unique land-cover classes one-hot encoded for model ingestion.

Finally, Normalized Difference Vegetation Index (NDVI) data (Canada, 2018) were incorporated. Yearly NDVI composites are available from 1987 onward, with each band representing a Julian week. For this study, an average NDVI was computed across Julian weeks 16–23 to capture typical early-summer vegetation conditions.

Geospatial Grids

Because spatial dependencies are not explicitly modelled within the XGBoost framework, spatial reference variables were introduced to preserve geographic context. Latitude and longitude coordinates to represent north–south and east–west positional gradients. Morton and Hilbert spatial indexes were produced to encode two-dimensional spatial relationships in a form interpretable by the model.

3.3 Modelling

To construct the training dataset, the labelled data were combined with the corresponding independent variables on a per-year basis, ensuring that each temporal event was matched with the correct environmental and climatic conditions. A metadata file was compiled to describe how each variable should be accessed and pre-processed. A ‘type’ field indicates whether a variable is numeric or non-numeric. For numeric variables, the minimum and maximum values across the entire study area were recorded to enable min–max scaling, ensuring that the model could correctly interpret unseen data. Non-numeric variables—such as LULC, surficial geology, and Geomorphons—were one-hot encoded (OHE) to allow them to be used effectively in the model.

Although the model does not ingest an explicit time index during training, the temporal evolution of flood-related drivers is fully represented in the predictor variables. Each sample carries the time-specific values of weather, hydrology, land use, and other dynamic variables for its corresponding year. As such, the model learns from temporally varying conditions without requiring a sequential or recurrent architecture such as an LSTM. Because the dataset consists of independent spatial samples rather than ordered time series at fixed locations, sequence-based models are not directly applicable and would not provide additional benefit for this type of spatial classification problem.

Once the data was compiled and pre-processed, the distribution of the data was explored and feature importance (FI) tested using three different tests: Partial Mutual Information, Partial Correlation and Recursive Feature Elimination. See (Dunbar et al., 2025) for details of the testing. While each FI method found differing subsets, the complete set of features were found to produce the best results three different ML models, XGBoost, ANN and CNN.

Deep learning alternatives—including a fully connected artificial neural network (ANN) and a convolutional neural network (CNN)—were benchmarked against tree-based models in Dunbar et al., (2025) using the same labelled dataset and predictor framework applied here. As demonstrated in that work, XGBoost outperformed both the ANN and CNN across multiple metrics (log-loss, AUC, F1, and class-specific accuracy), particularly in a spatially heterogeneous environment where deep models required substantially more tuning and computational resources without corresponding gains in predictive skill. Based on those results, and to avoid redundant benchmarking on the same dataset and feature space, the present study adopts XGBoost as the preferred modelling approach and focuses on uncertainty assessment, and national-scale application rather than model-type comparison.

The XGBoost model was selected to train the model and create historic FSM for 24 years, 2000-2023. XGBoost classifier is a tree-based ensemble method that builds many decision trees sequentially (Chen and Guestrin, 2016). Each tree learns from the ‘mistakes’ of the earlier ones. The configured model in this work is setup to output a probability estimate – not just a binary classification, this is useful in a flood susceptibility problem where probability estimates may matter as much as classification accuracy. The training/test/validation dataset has 268,049 samples, providing a strong statistical foundation for the model.

The final configuration consists of an ensemble of 10 independently trained XGBoost models, each initialized with a different random seed to reduce randomness-induced bias. Each model comprises 600 decision trees with a maximum depth of 20. Although deep trees can capture complex nonlinear relationships, they risk overfitting; therefore, subsampling (subsample = 0.8) and column subsampling (colsample_bytree = 0.8) were applied to improve generalization. The learning rate was set to 0.01, and regularization parameters $\alpha = 0.5$ (L1) and $\lambda = 5$ (L2) were used to control model complexity. These regularization values were inherited from Dunbar et al. (2025) and validated in this workflow through bagging-based sensitivity tests (10 bootstrap replicates) and 5-fold cross-validated learning curves using log-loss, confirming effective control of overfitting. The model was trained using the log-loss objective with a sample size of 600 per iteration, balancing computational efficiency with model performance.

Model performance was evaluated using confusion-matrix-derived statistics, including accuracy, F1 score, true positive rate, true negative rate for both wet and dry classes, and precision.

After configuring and validating the model, it was applied to unseen data across all of Canada to generate medium-resolution flood susceptibility maps.

3.4 Validation

XGBoost model validation follows established ML evaluation practices. It was assessed using the receiver operating characteristics area under curve (ROC-AUC), precision, accuracy, F_1 score and true positive (TPR) and true negative (TNR) rates derived from the confusion matrix outputs. These metrics provide complementary perspectives on model discrimination ability, balance between precision and recall, and overall classification reliability.

To explore temporal trends, pixel-based classification outputs were summarized to quantify the proportion of dry, wet, and extremely wet conditions, both nationally and by the National Hydrographic Network (NHN) Work Units (WU). The WUs were created based on Water Survey of Canada Sub-Sub-Drainage Area (Secretariat and Secretariat, 2023) and were used to evaluate outputs on a hydrologically consistent basis across regions. Furthermore, to access the temporal variability from year to year correlation analyses were conducted, to evaluate the long-term variability in the predicted flood susceptibility patterns.

Finally, model outputs were compared with three independent national-level datasets for contextual validation. These include: (i) Canadian Disaster Database (CDD) which documents significant flood and other disasters affecting Canadians since 1990 (Government of Canada, 2018), (ii) Environment and Climate Change Canada (ECCC) Climate Trends and Variation Bulletins (CTVB) (Environment and Climate Change Canada, 2017), which summarize seasonal and long term precipitation and temperature trends and anomalies, and (iii) the Historic Flood Events (HFE) layer, a compilation of point flood records from multiple sources (Natural Resources Canada, 2025b). This comparative analysis helped corroborate spatial and temporal patterns.

4. Results

4.1 Modelling Results

The XGBoost model demonstrated strong predictive performance across all key classification metrics. The model achieved an overall accuracy of 94.5%, with nearly balanced precision and recall across both wet and dry classes, Table 1, Figure 1. Specifically, the True Positive Rate (TPR) of 0.95 indicates that the model correctly identified 95% of flood-prone (wet) areas, while the True Negative Rate (TNR) of 0.94 shows similarly high reliability in recognizing non-flood-prone (dry) regions. Low error rates, False Positive Rate (0.06) and False Negative Rate (0.05), further confirm the model's stability and robustness. Precision values of 0.94 for the wet class and 0.95 for the dry class suggest minimal class confusion, and the balanced F_1 Scores (0.945) highlight the model's ability to maintain consistent performance across categories. Overall, these results indicate that the XGBoost model effectively discriminates between wet and dry conditions, making it a reliable tool for spatial flood susceptibility mapping.

In addition to the results from Figure 1, the prediction for all pixels within historic flood events were evaluated. While the model was trained on random samples within the flood extents, this does not represent the entirety of the available data. The data for each year of a flood was masked by the flood extents, and the distribution of all values were computed. The results were explored both by year and province and combined in a violin plot, Figure 2. Since only values within flood extents are selected, the expectation is that the pixel values and distribution should be strongly predicted to be wet (closer to 1 (or 100 in our case)), and values should be above the wet/dry threshold value. As per our metrics, ~95 percent of the data should be within these ranges, which we do see in the results, Figure 2.

Metric	Value
True Positive Rate (TPR)	0.95
True Negative Rate (TNR)	0.94
False Positive Rate (FPR)	0.06
False Negative Rate (FNR)	0.05
Precision (Wet class)	0.94
Precision (Dry class)	0.95
Accuracy	0.945
F_1 Score (Wet class)	0.945
F_1 Score (Dry class)	0.945

Table 1 XGBoost Model results on test set

4.2 Pixel Based classification

For a high-level analysis, the national count of pixels classified as susceptible to flooding were counted for each year and plotted, Figure 3. From this plot we can see there is variation throughout the years. The maximum number of pixels which are classed as flooded in found 2017 and the next highest year is 2021. While we find lowest number of flood classed pixels in 2007 and 2009. Looking at the trend line from 2000 to 2023, the positive slope is indicating an increasing trend of flood susceptible pixels overall in Canada from 2020 to 2023.

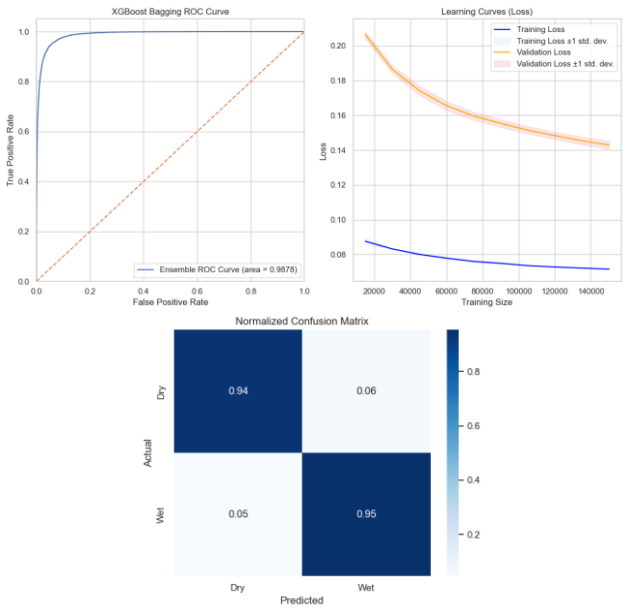

Figure 1 Model Results, (a) ROC Curve, (b) learning curves and (c) confusion matrix.

4.3 Spatio- temporal variability

To determine how the FSM predictions changed per-pixel over time, we explored spatio-temporal variability from year to year, by pairwise year-to-year correlations. Due to the large spatial extent of the study area, correlations were calculated using a windowed approach, reading the data in manageable chunks to efficiently handle memory constraints. No-data pixels were excluded, and Pearson correlation coefficients were computed

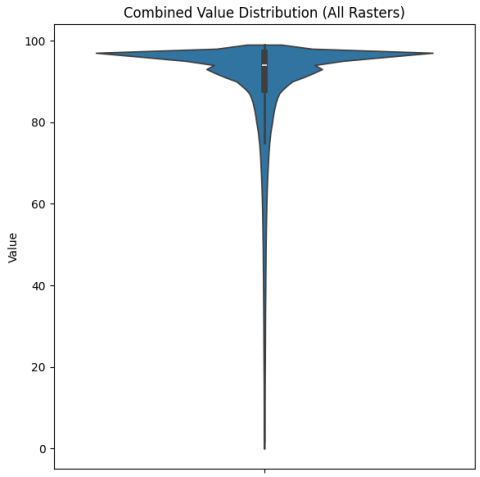

Figure 2 Distribution of pixel values in flood extent polygons, all years, locations combined.

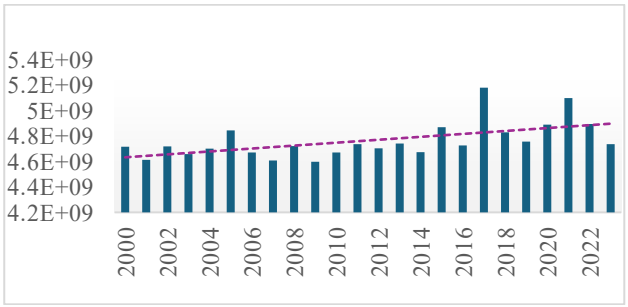

Figure 3 National Pixel Count likely susceptible to flooding.

for the overlapping valid values within each window before aggregating across the full raster extent.

A rolling three-year average was applied to the year-to-year correlation coefficients to illustrate temporal trends in flood susceptibility predictions, Figure 4. The blue line highlights periods of greater variability or stability across time, minimizing short-term fluctuations. Higher correlations indicate consistent susceptibility patterns between consecutive years, while localized drops in the 3-year rolling mean suggest temporal shifts potentially linked to climatic or land surface changes, e.g. 2004 to 2005 and 2016 to 2017.

In addition to this general national trend, analysis of FSM by WU were calculated. This approach enables systematic identification of temporal patterns in flood susceptibility and highlights WUs experiencing recurrent extreme wetness.

In this test, the pixels are divided into three categories, dry, wet, and extremely wet pixels. Where dry and wet bounds are based on the F_1 optimal threshold and extremely wet lower boundary is defined as the bottom value of a 68% confidence or prediction interval, from Figure 2. To complete this, the annual FSM were clipped to each WU and the pixels in each of these three categories were counted and saved to a csv. Then, a composite wetness score was computed by combining total wet and extreme wet pixels with a weighted formula, to rank years in terms wettest or highest FSM predictions. This analysis identified the top wettest years for each WU, illustrated in Figure 5. The years 2017 and 2023 were most frequently ranked the highest, and trends can be seen across the country.

The third spatio-temporal test involved correlation analysis within each Work Unit (WU). The correlation coefficients ranged from 0.513 to 0.999, reflecting substantial variation in year-to-year consistency. An example is presented in Figure 6 for WU 05AJ000, which encompasses a portion of the South Saskatchewan River near the town of Bow Island and the northwestern area of Medicine Hat, Alberta exhibits distinct temporal variability in modelled FSM values. Noticeable year-to-year changes in FSM values, e.g. 2005, 2013, 2017, indicate that the multi-event (ME) model captures temporally dynamic hydrological conditions rather than merely reproducing static spatial patterns.

A record flood in Alberta in 2013 resulted in millions of dollars in damage and evacuation of approximately 10,000 residents of Medicine Hat (Marks, 2023) among other cities and towns in Alberta. In the correlation matrix, 2013 shows lower correlation with preceding years, reflecting this substantial hydrologic change. Another year which stands out is 2017, previously identified (Figure 3) as one of the three wettest years on record. Following the 2013 flood, a comprehensive resiliency plan and overland flow protection strategy were implemented (Stantec, 2025) which may partly explain the reduction in flood impacts in subsequent years.

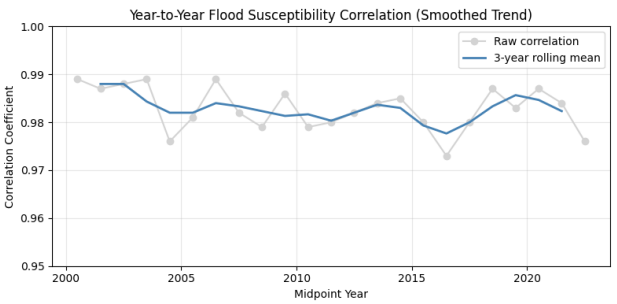

Figure 4 Smoothed (rolling mean) correlation trend, year to year, 2000 to 20203

4.4 Auxiliary Dataset Comparison

As illustrated in Figure 3, 2017 was the wettest year in the FSM yearly model results. According to the 2017 Spring CTVB, this season ranked as the third-wettest spring on record, with total precipitation 19.2% above the baseline average (Environment and Climate Change Canada, 2017). Wetter-than-average conditions were most pronounced across southern British Columbia, central Alberta, central Saskatchewan, southeastern Ontario, southern Quebec, and parts of southern and northern Nunavut. Conversely, southern Saskatchewan, Manitoba, northern Quebec, and western Northwest Territories experienced drier anomalies. Pixel-based counts of "wetter-than-average" and "drier-than-average" categories from 2000 to 2023 show regionally coherent trends that align with historic flood

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-M-1-2026-21-2026 | © Author(s) 2026. CC BY 4.0 License.

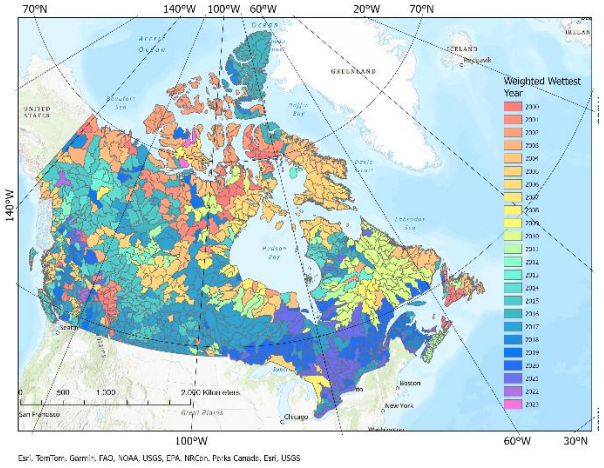

Figure 5 Weighted wettest years by National Hydrographic Network (NHN) work units (WU).

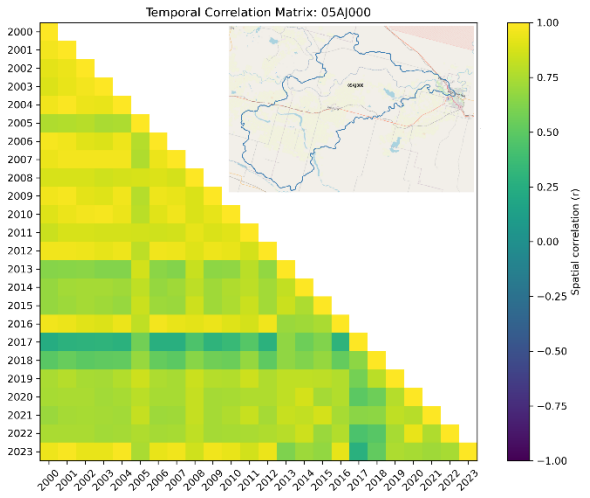

Figure 6 Temporal correlation matrix, 2000 to 2023, and location of NHN WU 05AJ000 in southern Alberta.

occurrences and broader climatic variability Figure 3. In the CVTB, 2012 was quite a wet year with precipitation being 22% above average. 2013 had temperatures 0.6° above the baseline and drier than average. 2014 has among the coldest temperatures on record in 4 of the 11 regions, and precipitation 4% above the baseline, while the west coast was quite wet.

The CDD confirms several flood events during 2017 in British Columbia, eastern Ontario, and western and southern Quebec, consistent with the spatial precipitation patterns reported by CTVB. FSM outputs were clipped to three representative regions—southern BC, southern Saskatchewan, and the National Capital Region (Ottawa–Gatineau)—covering wetter, drier, and transitional conditions respectively. In Figure 7 records from the HFE are plotted for the year 2017. There is a cluster of points in eastern Ontario and southwestern Quebec. The average pixel values within the hashed-polygon areas of FSM time-series maps 2000 to 2023 are plotted below the map. We can see the variation in the areas over the time period, with highest values in 2017 in both the NCR and southern Saskatchewan. The FSM prediction in southern BC is quite low, relative to the other sites, and not significantly higher than any other years in this location.

In Calgary 2013 there was massive flooding due to heavy rain over a melting snowpack on the Bow River and contributing areas (Water Services, 2022). Cross-section profiles of the data around downtown Calgary show the highest FSM values in the years 2005, 2013, and 2022. While there were floods in 2013, high water levels were recorded by river gauges in the area as per the National Data Archive HYDAT database (Canada, 2010) and we find records in the CDD of this event.

We find generally positive correlation of the trends of the yearly FSM predictions align with records in CDD, CTVB and HEF. Through this exploration of the model performance over several years and locations, we observe the ME framework is sensitive to evolving patterns and underlying data distributions. Suggesting the ME ML model is not merely fitting static features but is learning event-specific variations and adapting to changes over time. This is critical for applications involving non-stationarity or seasonally influenced phenomena.

Finally, though pixel-by-pixel trend analysis we can visualize wetting/drying trends from the FSM predictions across Canada, giving us a first look at how flood susceptibility may be changing over time, Figure 8.

However, there are several limitations to this approach, as it does not consider any resiliency measures that may be put in place. For example, around Medicine Hat and in Calgary after the 2013 floods, rebuilding included resiliency measures. The temporal variables of precipitation, temperature, land use land cover and NDVI may not adequately incorporate or account for these resiliencies and potentially over-estimate FSM. With the predictions at a 30m by 30m pixel resolution, they FSM values can provide an estimate to the general characteristics of the area and its susceptibility to flooding, however it is not a replacement for traditional flood hazard mapping. Instead, this would be more

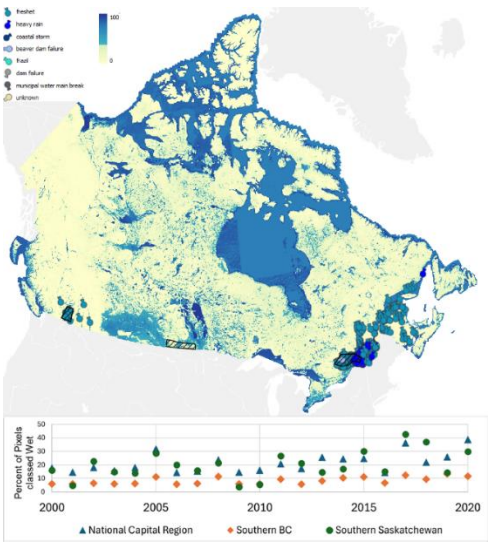

Figure 7 Historic flood events layer from 2017 overlaid on 2017 FSM map, selected sites and pixel value analysis from 2000 to 2023 for three sites, from west to east: southern BC, southern Saskatchewan and Manitoba, and the National Capital Region.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-M-1-2026-21-2026 | © Author(s) 2026. CC BY 4.0 License.

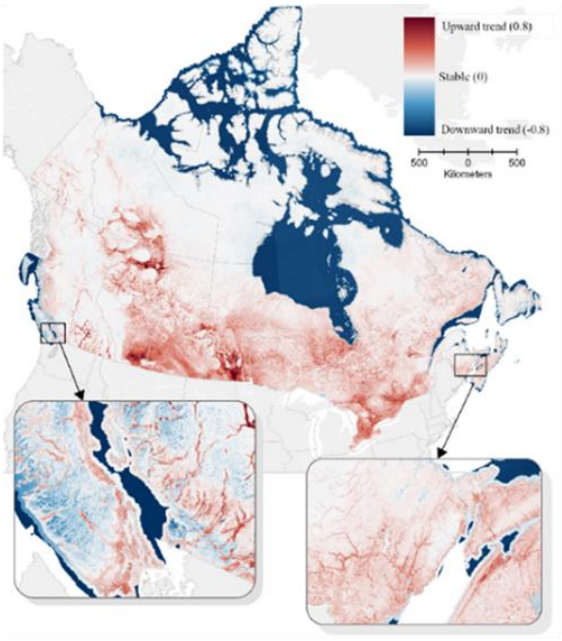

Figure 8 Flood Susceptibility trend from 2000 to 2023.

suitable for general trend analysis and identifying areas where fluctuations occur and may require further data collection and more robust methods to identify flood risk.

5. Conclusion

Flood susceptibility mapping is critical and can provide wide scale coverage, enabling planning and mitigation of flood risk. However, many studies focus on single events, small geographic regions and produce static maps that fail to capture or demonstrate temporal variability. This work introduces a novel multi-event temporal ML framework to predict annual flood susceptibility across Canada. Using datasets that integrate spring temperature and precipitation, previous winter and fall climate along with land use and land cover (LULC) normalized difference vegetation index (NDVI), topographic, lithology, hydrography, and spatial indexes, an XGBoost model is trained to capture the complex environmental drivers of spring freshet. The model achieves an overall accuracy of 0.95, true positive and true negatives rates of 0.95 and 0.94 respectively.

A key advantage of the ME approach is in its ability to aggregate data that are otherwise limited in spatial and temporal coverage across the country. By combining diverse climatic, hydrological, and landscape conditions into a unified training dataset, the model learns more generalizable patterns, enhancing its robustness under varying environmental contexts.

By producing yearly flood susceptibility predictions across diverse Canadian landscapes, this method not only advances methodological rigor, but delivers practical insights to aid flood risk management and planning and exploring temporal patterns of changing flood susceptibility. This framework represents a step towards more dynamic, data-driven flood mapping, highlighting the value of integrating ML with temporal environmental data for scalable, interpretable flood susceptibility assessments.

References

Al-Kindi, K.M., Alabri, Z., 2024. Investigating the Role of the Key Conditioning Factors in Flood Susceptibility Mapping Through Machine Learning Approaches. Earth Syst Environ 8, 63–81. https://doi.org/10.1007/s41748-023-00369-7

Canada, E. and C.C., 2010. National Water Data Archive: HYDAT [WWW Document]. URL https://www.canada.ca/en/environment-climate-change/services/water-overview/quantity/monitoring/survey/data-products-services/national-archive-hydat.html (accessed 10.2.25).

Canada, N.R., 2025. Floods in Canada - Archive - Open Government Portal [WWW Document]. URL https://open.canada.ca/data/en/dataset/74144824-206e-4cea-9fb9-72925a128189 (accessed 7.23.25).

Canada, N.R., 2023. Federal Flood Mapping Guidelines [WWW Document]. Federal Flood Mapping Guidelines. URL https://natural-resources.canada.ca/science-data/science-research/natural-hazards/federal-flood-mapping-guidelines (accessed 11.6.25).

Canada, N.R., 2015. Surficial geology index map - Open Government Portal [WWW Document]. URL https://open.canada.ca/data/en/dataset/cebc283f-bae1-4eae-a91f-a26480cd4e4a (accessed 9.16.25).

Canada, S., 2018. Corrected representation of the NDVI using historical AVHRR and VIIRS satellite images (1 km resolution) from 1987 to present - Open Government Portal [WWW Document]. Corrected representation of the NDVI using historical AVHRR and VIIRS satellite images (1 km resolution) from 1987 to present - Open Government Portal. URL https://open.canada.ca/data/en/dataset/44ced2fa-afcc-47bd-b46e-8596a25e446e (accessed 9.16.25).

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16. Association for Computing Machinery, New York, NY, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785

Choubin, B., Moradi, E., Golshan, M., Adamowski, J., Sajedi-Hosseini, F., Mosavi, A., 2019. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Science of The Total Environment 651, 2087–2096. https://doi.org/10.1016/j.scitotenv.2018.10.064

Costache, R., Pal, S.C., Pande, C.B., Islam, A.R.Md.T., Alshehri, F., Abdo, H.G., 2024. Flood mapping based on novel ensemble modeling involving the deep learning, Harris Hawk optimization algorithm and stacking based machine learning. Applied Water Science 14, 78. https://doi.org/10.1007/s13201-024-02131-4

Dunbar, K.E., McGrath, H., Khan, U.T., 2025. Enhancing flood susceptibility modelling in Canada: Integrating seasonal meteorological data, feature selection and machine learning approaches (No. EGU25-3871). Presented at the EGU25, Copernicus Meetings. https://doi.org/10.5194/egusphere-egu25-3871

Environment and Climate Change Canada, 2017. Climate Trends and Variations Bulletin – Spring 2017 [WWW Document]. Climate Trends and Variations Bulletin – Spring 2017. URL https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/trends-variations/spring-2017-bulletin.html (accessed 7.8.25).

Fang, L., Huang, J., Cai, J., Nitivattananon, V., 2022. Hybrid approach for flood susceptibility assessment in a flood-prone mountainous catchment in China. Journal of Hydrology 612, 128091. https://doi.org/10.1016/j.jhydrol.2022.128091

Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D., 2014. Do we need hundreds of classifiers to solve real world classification problems? The journal of machine learning research 15, 3133–3181.

Government of Canada, P.S.C., 2018. Canadian Disaster Database [WWW Document]. URL https://cdd.publicsafety.gc.ca/srchpg-eng.aspx (accessed 10.2.25).

GOV.UK, 2023. Hydraulic modelling: best practice (model approach) [WWW Document]. Guidance Hydraulic modelling: best practice (model approach). URL https://www.gov.uk/government/publications/river-modelling-technical-standards-and-assessment/hydraulic-modelling-best-practice-model-approach (accessed 11.6.25).

Islam, T., Zeleke, E.B., Afroz, M., Melesse, A.M., 2025. A Systematic Review of Urban Flood Susceptibility Mapping: Remote Sensing, Machine Learning, and Other Modeling Approaches. Remote Sensing 17. https://doi.org/10.3390/rs17030524

Kaya, C.M., Derin, L., 2023. Parameters and methods used in flood susceptibility mapping: a review. Journal of Water and Climate Change 14, 1935–1960. https://doi.org/10.2166/wcc.2023.035

Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.-B., Gróf, G., Ho, H.L., Hong, H., Chapi, K., Prakash, I., 2019. A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. Journal of Hydrology 573, 311–323. https://doi.org/10.1016/j.jhydrol.2019.03.073

Marks, S., 2023. “It was devastating”: Medicine Hat reflects on 2013 flood [WWW Document]. CTVNews. URL https://www.ctvnews.ca/calgary/article/it-was-devastating-medicine-hat-reflects-on-2013-flood/ (accessed 11.8.25).

Masrur, A., Yu, M., Mitra, P., Peuquet, D., Taylor, A., 2022. Interpretable machine learning for analysing heterogeneous drivers of geographic events in space-time. International Journal of Geographical Information Science 36, 692–719. https://doi.org/10.1080/13658816.2021.1965608

McGrath, H., Gohl, P.N., 2022. Accessing the Impact of Meteorological Variables on Machine Learning Flood Susceptibility Mapping. Remote Sensing 14, 1656.

McGrath, H., Papasodoro, C., Moreau, J.S., Bourgon, J.F., Fortin, M., Fuss, C., Perry, B., Tardif, P., 2025. Descriptor: Medium Resolution Digital Elevation Model From Natural Resources Canada’s CanElevation Series (MRDEM-30). IEEE Data Descriptions 2, 211–217. https://doi.org/10.1109/IEEEDATA.2025.3576318

Natural Resources Canada, 2025a. 2020 Land Cover of Canada - Open Science and Data Platform [WWW Document]. 2020 Land Cover of Canada. URL https://osdp-psdo.canada.ca/dp/en/search/metadata/NRCAN-FGP-1-ee1580ab-a23d-4f86-a09b-79763677eb47 (accessed 11.3.25).

Natural Resources Canada, 2025b. Historical Flood Events (HFE) - Open Government Portal [WWW Document]. Historical Flood Events (HFE). URL https://open.canada.ca/data/en/dataset/fe83a604-aa5a-4e46-903c-685f8b0cc33c (accessed 8.28.25).

Paryani, S., Bordbar, M., Jun, C., Panahi, M., Bateni, S.M., Neale, C.M.U., Moeini, H., Lee, S., 2023. Hybrid-based approaches for the flood susceptibility prediction of Kermanshah province, Iran. Natural Hazards 116, 837–868. https://doi.org/10.1007/s11069-022-05701-4

Pasos, M., 2025. North American Land Change Monitoring System. Commission for Environmental Cooperation. URL https://www.cec.org/north-american-land-change-monitoring-system/ (accessed 11.3.25).

Pechlivandis, I., Jackson, B.M., McIntyre, N.R., Wheater, H.S., 2013. Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. Global NEST Journal 13, 193–214. https://doi.org/10.30955/gnj.000778

Pourzangbar, A., Oberle, P., Kron, A., Franca, M.J., 2025. Analysis of the Utilization of Machine Learning to Map Flood Susceptibility. Journal of Flood Risk Management 18, e70042. https://doi.org/10.1111/jfr3.70042

Prasad, P., Loveson, V.J., Das, B., Kotha, M., 2022. Novel ensemble machine learning models in flood susceptibility mapping. Geocarto International 37, 4571–4593. https://doi.org/10.1080/10106049.2021.1892209

Secretariat, T.B. of C., Secretariat, T.B. of C., 2023. National Hydro Network - NHN - GeoBase Series - Open Government Portal [WWW Document]. URL https://open.canada.ca/data/en/dataset/a4b190fe-e090-4e6d-881e-b87956c07977 (accessed 7.28.23).

Seydi, S.T., Kanani-Sadat, Y., Hasanlou, M., Sahraei, R., Chanussot, J., Amani, M., Seydi, S.T., Kanani-Sadat, Y., Hasanlou, M., Sahraei, R., Chanussot, J., Amani, M., 2022. Comparison of Machine Learning Algorithms for Flood Susceptibility Mapping. Remote Sensing 15. https://doi.org/10.3390/rs15010192

Stantec, 2025. Medicine Hat Overland Flow Protection Strategy [WWW Document]. Medicine Hat Overland Flow Protection Strategy. URL https://www.stantec.com/en/projects/canada-projects/o/overland-flow-protection-strategy (accessed 11.8.25).

Water Services, C., 2022. Flooding in Calgary - Flood of 2013 [WWW Document]. https://www.calgary.ca. URL https://www.calgary.ca/content/www/en/home/uep/water/flood-info/flooding-history-calgary.html (accessed 5.11.22).

Whitebox Geospatial Inc., 2025. Whitebox Geospatial [WWW Document]. Whitebox Geospatial Inc. URL https://www.whiteboxgeo.com/ (accessed 9.16.25).