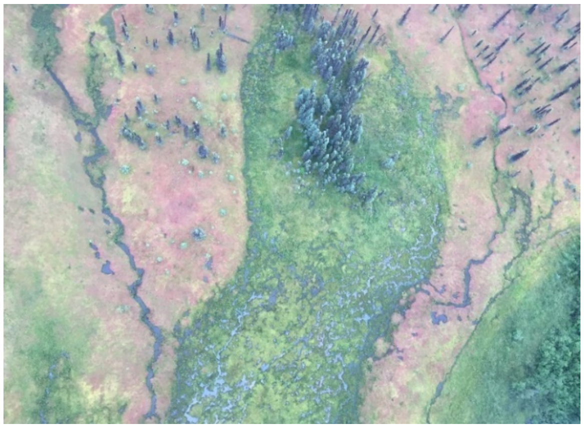
PREDICTIVE WETLAND MAPPING OF THE WILLISTON DRAINAGE BASIN: UPDATE

Ministry of Environment Project BAPID: 6538
FWCP Project ID: PEA-F19-W-2896-DC

30-June-2020



PREPARED FOR:

PREPARED BY: *DEEPA FILATOW (P.Geo, P.Ag.), GILLIAN HARVEY (MSc.), TYSON CARSWELL (MSc.), EMILY CAMERON (R.P.BIO., MSc.)with Preliminary Disturbance and Risk Analysis by EMILY CAMERON (MSc., R.P.Bio)*
BC Ministry of Environment and Climate Change Strategy

PREPARED WITH FINANCIAL SUPPORT OF THE FISH AND WILDLIFE COMPENSATION PROGRAM ON BEHALF OF ITS PROGRAM PARTNERS: BC HYDRO, THE PROVINCE OF BC, FISHERIES AND OCEANS CANADA, FIRST NATIONS AND PUBLIC STAKEHOLDERS.

Ministry of Environment & Climate Change Strategy
Environmental Sustainability & Strategic Policy Division

Knowledge Management Branch
Ecosystem Information Section
Terrestrial Ecosystems Unit

Mailing Address:
PO Box 9358 STN. PROV. GOV.
Victoria BC V8W 9M1

Citation:

Filatow, D., Harvey, G., Carswell, T., and E. Cameron. 2020. Predictive Wetland Mapping of the Williston Drainage Basin: Update. Knowledge Management Branch, BC Ministry of Environment and Climate Change Strategy, Victoria. BC.

Ministry of Environment &
Climate Change Strategy
Environmental Sustainability &
Strategic Policy Division

Knowledge Management Branch
Ecosystem Information Section
Terrestrial Ecosystems Unit

Mailing Address:
PO Box 9358 STN. PROV.
GOV.
Victoria BC V8W 9M1

TO: Chelsea Coady, FWCP Peace Region Manager
Fish & Wildlife Compensation Program (FWCP), 3333 22nd Ave., Prince George, BC V2N 1B4

FROM: Deepa Spaeth Filatow, P.Geo , P.Ag., Provincial Bioterrain Specialist

RE: Deliverables for Fish and Wildlife Compensation Program (FWCP) and Ministry of Environment and Climate Change Strategies (MoECCS) 'Letter of Agreement' and Modification #1 (July 1, 2018 – July 1, 2020).

This report and accompanying data will be publicly available on the Ecological Reports Catalogue, pending Treasury Board Review. Digital deliverables data package will be provided to FWCP Peace Region Manager and are described in in detail in the file **Metadata_datasets.xlsx**. The core deliverables include:

- Final Report **Predictive Wetland Mapping of Williston Drainage Basin_2020.pdf**
- 3-category raster map (86.7% accuracy) **3CategoryPrediction/20190926-103025_map_recl.tif**
- 10-catagory raster map (52.8% accuracy) **10CatagoryPrediction/20190627_114031_map_recl.tif**
- Field data
- Training data
- Supplementary data including wetland and training point density by mapsheet
- Links to field photographs
- Metadata

The **ArcGIS interactive mapping** application will be available through the following metadata link, pending publication. https://catalogue.data.gov.bc.ca/dataset/1d174499-72af-41d5-ad5e-19460b1e6fb1

Updated scripts are available at https://github.com/bcgov/wetlandmapR.

Sincerely,

Deepa S Filatow, P.Geo, P.Ag. deepa.filatow@gov.bc.ca

Disturbance Analysis signed off by:

Emily Cameron, R.P.Bio, MSc. emily.cameron@gov.bc.ca

EXECUTIVE SUMMARY

This mapping project builds on the data and methodology developed in Predictive Wetland Mapping of the Williston Drainage (Filatow *et al.*, 2018) to improve the understanding of abundance, distribution and connectivity of wetlands in the FWCP Peace region, British Columbia, Canada. New 3-category (water, wetland, upland), 10-category (bog, fen, swamp, marsh, low-bench floodplain, mid-bench floodplain, high-bench floodplain, shallow water wetland, upland, and water) and 46-catagory predictive wetland raster maps were produced. The 3-category map is 86.7% accurate at predicting wetland, upland and water across the FWCP Peace region, and had an internal model accuracy of 93% making it a reliable tool for use in wetland and riparian ecosystem planning and management. The 10-category map has an accuracy of 53% and should be used with caution. The 46-category raster prediction is not reliable and should not be used.

Updates to the predictive wetland and riparian ecosystem mapping incorporated several key improvements. One thousand new expert interpreted training points and 620 field observation points were added in the Parsnip sub-basin to address over capture of wetlands in disturbed areas. To better capture riparian areas in the 3-category prediction map, low-bench and mid-bench floodplain ecosystems were included in the “wetland” category, as were shallow water wetlands – a diversion from the methods applied in the first run of the model. Open source scripts in the ModelMapR package were tested to re-run the model and accuracy metrics were recalculated using independent validation data from 2017, 2018 and 2019 field observations. A comparison between the first and second run of the model was completed to quantify overall improvement between models and specific improvements in the Parsnip sub-basin. Across the FWCP Peace region, model accuracy improved by 13% for the 3-category, and 7% for the 10-category predictions. At the sub-basin level, the 3-category prediction was improved by 3% in the Finlay sub-basin and 14% in the Parsnip sub-basin where new training points were added.

The disturbance analysis method was refined, and an example case study analysis is provided as proof of concept. Disturbances intersecting both the footprints of ecologically connected wetlands and a wider 1km buffered area of influence are summarized. Eight disturbance layer types were included in the analysis, with the output displaying relative disturbance patterns in the FWCP Peace region. Example maps and watershed group summaries demonstrate how the wetland prediction map can be used with other information to inform wetland and riparian priority actions.

Guidance on the appropriate use, and limitations, of the predictive 3-category map and analysis output is provided. Continuous improvement recommendations outline potential updates to the model and analysis methods, as well as areas for further research and development. Education and outreach related to this project can leverage the Esri ArcGIS Online map application, hosted on the B.C.’s Map Hub. The findings from this report can be used to inform potential updates to FWCP priority actions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

BACKGROUND[1]

Fish and Wildlife Compensation Program (FWCP) and Ministry of Environment and Climate Change Strategies (ENV) signed a '*Letter of Agreement*' (Dec 2015 - March 31, 2018 [Phase 1] and July 2018 – July 1, 2020 [Phase 2]) to complete mapping of wetland and riparian units representing wetland extent and classification (e.g., fen, bog, marsh, swamp, flooded). The goal was to create consistent, defensible and repeatable maps with supporting documentation that improves current inventories at a fraction of the traditional inventory mapping costs. The project was designed to address riparian ecosystem goals by determining the distribution, abundance and connectivity of wetland and riparian areas as outlined in the Riparian and Wetlands Action Plan[2] , specifically:

Objective 1*: Improve the understanding of the abundance, distribution, trend and connectivity of riparian and wetland ecosystems.*

Sub-objective 1a*: Improve understanding of the abundance, distribution, trend and connectivity of riparian ecosystems.*

Action 1a-1*: Inventory the distribution, abundance, current function and connectivity of remaining riparian ecosystems.*

Rationale: *Before feasible targets can be established for riparian ecosystem restoration or enhancement, an inventory of existing habitats within the Peace Basin is required to identify potential sites and their current status.*

Sub-objective 1b*: Improve understanding of the abundance, distribution, trend and connectivity of wetland ecosystems.*

Action 1b-1*: Inventory the distribution, abundance, current function and connectivity of remaining wetland ecosystems.*

Rationale*: Before feasible targets can be established for wetland ecosystem restoration or enhancement, an inventory of existing habitats within the Peace Basin is required to identify potential sites and their current status.*

A timeline of activities and deliverables/outcomes is available in Appendix A. Trend and current function analyses are beyond the scope of this project.

---

[1] The Background section content is reproduced from the "Predictive Wetland Mapping of the Williston Drainage Basin" report from 2018, with additions to the previous and current phase summaries. http://www.env.gov.bc.ca/esd/distdata/ecosystems/TEI/WETLANDS/Williston_Wetlands_Reporting/

[2] "Peace Basin Riparian and Wetlands Action Plan - Fish and Wildlife Compensation Program" http://fwcp.ca/app/uploads/2015/07/fwcp-peace-riparian-and-wetlands-action-plan-march-31-2014.pdf. Accessed 16 Apr. 2018.

MODEL RUN 1 OVERVIEW

The first stage of the project included an investigation into the current status of ecosystem mapping in the FWCP Peace region and an evaluation of available inventory methods in two pilot areas. The key findings from the pilot projects initiated the third stage of investigation. This stage explored a machine learning modelling approach (Random Forest algorithm) as a cost-effective method to improve current wetland and riparian mapping products. The Random Forest modeling technique proved successful and was therefore used to predict wetland and riparian areas throughout the FWCP Peace region (i.e., Williston Basin).

The Random Forest model was applied using a suite of environmental predictor variables to predict wetland and riparian classes, terrestrial-uplands (hereafter referred to as "Upland" or "T"), and water. Model training data were generated for fifteen 1:20,000 map sheets along with additional random points throughout the study region. By interpreting satellite images, mapping contractors – working with ENV staff – attributed the training points using 48 map codes. The model predicted a series of classified products that included 3, 9, and 48-category surfaces. The modeled outputs (hereafter referred to as "Run 1") were evaluated using the out-of-bag error metric and accuracy was measured through comparison with field collected data from 2017. Based on model evaluation and accuracy metrics, the highest performing models were selected. The 48-cateogory model was the highest performing model. Reclassifying the 48-category model into 9 and 3 category surfaces, represented wetland and riparian ecosystems better than the raw model output. Overall, the reclassified 3-category surface was reliable in differentiating between upland, wetlands and water. The reclassified 9-class model was superior to the human delineated wetlands in the provincial Terrain Resource Information Management (TRIM) polygon dataset, representing a more consistent and reliable estimate of wetland distribution, type and area. Key model limitations included the over-capture of wetlands in disturbed areas and miss-classification errors in wetland type.

A preliminary threats analysis was employed to evaluate wetland distribution and density per watershed sub-basin and overall anthropogenic disturbance. Based on the model and threats analysis a series of recommendations were outlined to guide future mitigation actions, research, and analyses.

To conclude the first model run of the project, a comprehensive report outlining the approaches, results, and recommendations was compiled in June of 2018 (Filatow *et al.,* 2018).

MODEL RUN 2 OVERVIEW

Following the recommendations from the first report and model (hereafter referred to as Run 1), The second run of the model (hereafter referred to as Run 2) began in July 2018 with the primary goal to improve the overall model prediction by:

- reducing the over-capture of wetlands in disturbed areas,
- incorporating additional training points to increase model performance, and
- collecting additional field data in underrepresented areas of the Parsnip sub-basin.

Additional objectives included:

- expansion and further development of threat analysis explored as a case study,
- application of open-source tools to distribute project results and facilitate collaboration and engagement with project material.

Contributions to these goals included the development of an R package distributed through an open source platform (e.g., GitHub) to support transparency and encourage a continuous improvement cycle of the model. A web mapping platform (Esri ArcGIS Online) was used to demonstrate the capabilities of online platforms for facilitating the access and use of project data and results.

Key milestones and project timeline for model Run 1 and Run 2 are listed in Appendix A.

FWCP PEACE REGION[3]

The wetland project area is confined to the Williston Reservoir Drainage Basin (FWCP Peace Region) comprised of 7.2 million hectares ranging from the alpine ecosystems of the high mountain ranges to forested lowlands of the Rocky Mountain Trench. Central to the FWCP Peace is the Williston Reservoir; a 177,300-hectare designed water body formed in 1967 by the construction of the WAC Bennett Dam (Golder Associates, 2010). Elevation of the FWCP Peace region ranges from 500 to 3,000 meters and is further described by four sub-basins including the Finlay, Parsnip, Peace and Dinosaur (Figure 1). These cover four physiographic types inclusive of mountains, foothills, plateaus and the Rocky Mountain Trench, and falling within five biogeoclimatic zones. This extent spans the mapping grids 94B-F, 93I, 93J and 93M-O at 1:250,000 scale. At the 1:20,000 scale, the area covers 623 map sheets[4] of which ~150 are partial coverage along the FWCP Peace region boundary.

Factors influencing wetland and riparian ecosystem diversity within the region include topography, substrate (bedrock and soils), climate, hydrology, biota (vegetation, wildlife and other organisms) and disturbance history (Filatow *et al.*, 2018). These factors ultimately determine the abundance, distribution, current state, function and response rate (trend) of wetland and riparian ecosystems. Of these, topography and substrate are enduring features and are generally unlikely to change over hundreds of years. Conversely, hydrology, natural disturbances and biota change occur more rapidly in response to escalating disturbance due to human activity and a changing climate (Yang *et al.*, 2018).

Understanding the interactions and relationships of these factors to wetland and riparian systems is key to improving understanding, prioritizing actions and setting feasible targets (National Research Council, 1995; Euliss *et al.*, 2008). Consistent data on wetland extent and type will facilitate analysis to investigate the interrelationships that different environmental factors and disturbance have on ecosystem function, condition and trends.

---

[3] The *FWCP Peace Region* section content is reproduced from the "Predictive Wetland Mapping of the Williston Drainage Basin" report from 2018, with minor modifications to figure appearance (Filatow *et al.,* 2018). http://www.env.gov.bc.ca/esd/distdata/ecosystems/TEI/WETLANDS/Williston_Wetlands_Reporting/

[4] Provincial, geometrically corrected aerial photograph that displays ground features in their true ground position with a constant scale throughout the image. https://www2.gov.bc.ca/gov/content/data/geographic-data-services/digital-imagery/orthophotos

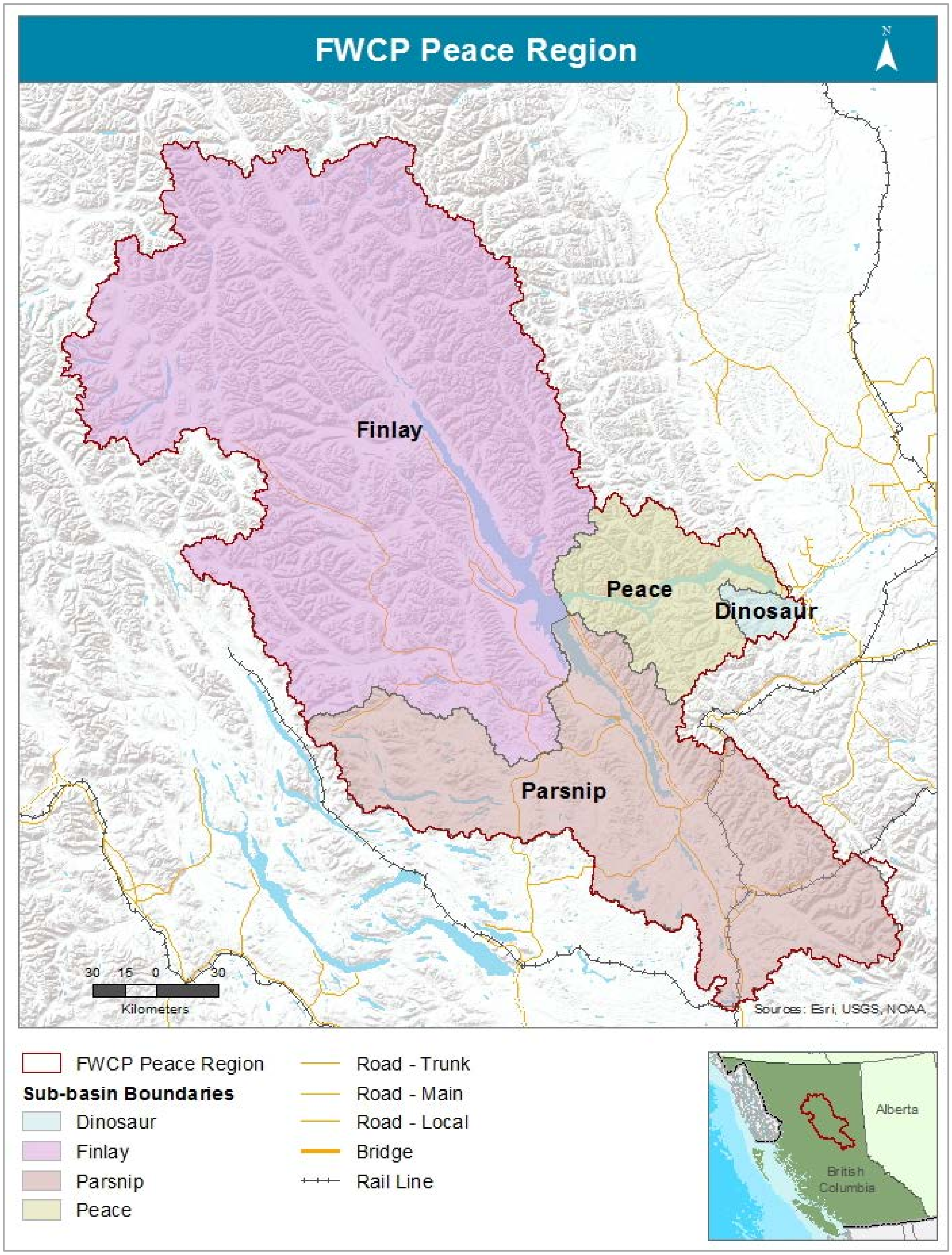


Figure 1: Overview map with topography, FWCP Peace Region and sub-basin boundaries.

METHODS

Methods for the second model run are split in to three key sections: (1) model refinement, (2) disturbance analysis case study, and (3) data and information access.

MODEL REFINEMENT

APPROACH OVERVIEW

The approach used to model wetlands is consistent with Run 1 of the project. A Random Forest model was used to estimate upland, water, riparian and wetland categories across the FWCP Peace region. Random Forest is a machine learning method often used for ecological prediction (see established use in Run 1 report; Filatow *et al.*, 2018). Unlike traditional statistical models, Random forest is a data driven model that uses recursive partitioning of data to predict the terminal node classes (e.g., bog, marsh swamp and flood categories). In contrast to traditional regression models, Random Forest does not assume predefined statistical distributions illuminating macro and micro differences in the natural environment. Random Forest is a popular modeling technique as it is generally resistant to overfitting, outliers and data gaps (Breiman, 2001a; Breiman, 2001b; Gislason *et al.,* 2006; Belgiu and Dragut, 2016). Random Forest algorithms have been successfully applied to classify ecosystems and vegetation cover across the landscape (e.g., Ayala-Izurieta *et al.*, 2017; *Corcoran et al.*, 2013; Timm and McGarigal, 2012). Specifically, the algorithm has been shown to outperform other popular nonparametric machine-learning methods (e.g., rule-based classifiers and individual decision trees) in overall accuracy when modelling wetlands (Berhane *et al.*, 2018) and complex ecosystems (Rodriguez-Galiano *et al.*, 2012). In addition, Random Forest modelling is cost effective and repeatable (Filatow et al., 2018). For additional details on the justification of the selected approach – Random Forest – see Filatow *et al.* (2018).

Random Forest models are a machine-learning algorithm developed as an improvement over the individual decision tree-based classifier (Gislason *et al.*, 2006; Prasad *et al.*, 2006; Rodriguez-Galiano *et al.*, 2012). Rather than modeling a single decision tree, a Random Forest approach generates many decision trees, and the results from each tree count as a vote towards the final prediction class based on the most popular class. As a result, the ensemble modeling approach improves overall prediction accuracy over single tree methods. The ensemble framework of many independent trees means that errors from a single tree will have less of an effect of the final prediction as the errors are less influential when combined across the 'forest' of trees.

In a supervised Random Forest classification model, each tree is built using a bootstrapped sample of training data and the overall majority vote across all trees predicts the final class. To build the model, each training point is attributed with a class (response variable) and associated values from the environmental layers (predictor variables). These data are compiled into a classification matrix. Each decision tree is built using a different random bootstrapped sample of training data from the classification matrix, while the remaining held out data are used to test the overall model performance. Overall model performance is measured via an internal estimate of error (out-of-bag error) and evaluates how well the model performs at predicting the held-out data from the training set.

Validation using field-collected observations provides an additional measure of accuracy by testing model performance for areas not known by the model. The validation data represent site visits in the field by crews not involved in the creation of the interpreted training data. This external validation dataset indicates the model ability to accurately predict the correct class as assessed on the ground and in areas not covered by training points. High validation errors for certain categories may indicate a need for the collection of additional training data points for the category. Errors could also be due to interpretation and classification errors in the training point dataset (see Step 1: Training Data Additions and Review).

Figure 2 illustrates the modelling and evaluation workflow of the wetland and riparian predictions generated in Run 2. Building upon the Random Forest model created in Run 1, Run 2 provides additional improvements with respect to:

1. Training data additions, review and data cleaning
2. Field data additions and data cleaning
3. Script development
4. Random Forest model refinement
5. Wetland prediction summaries

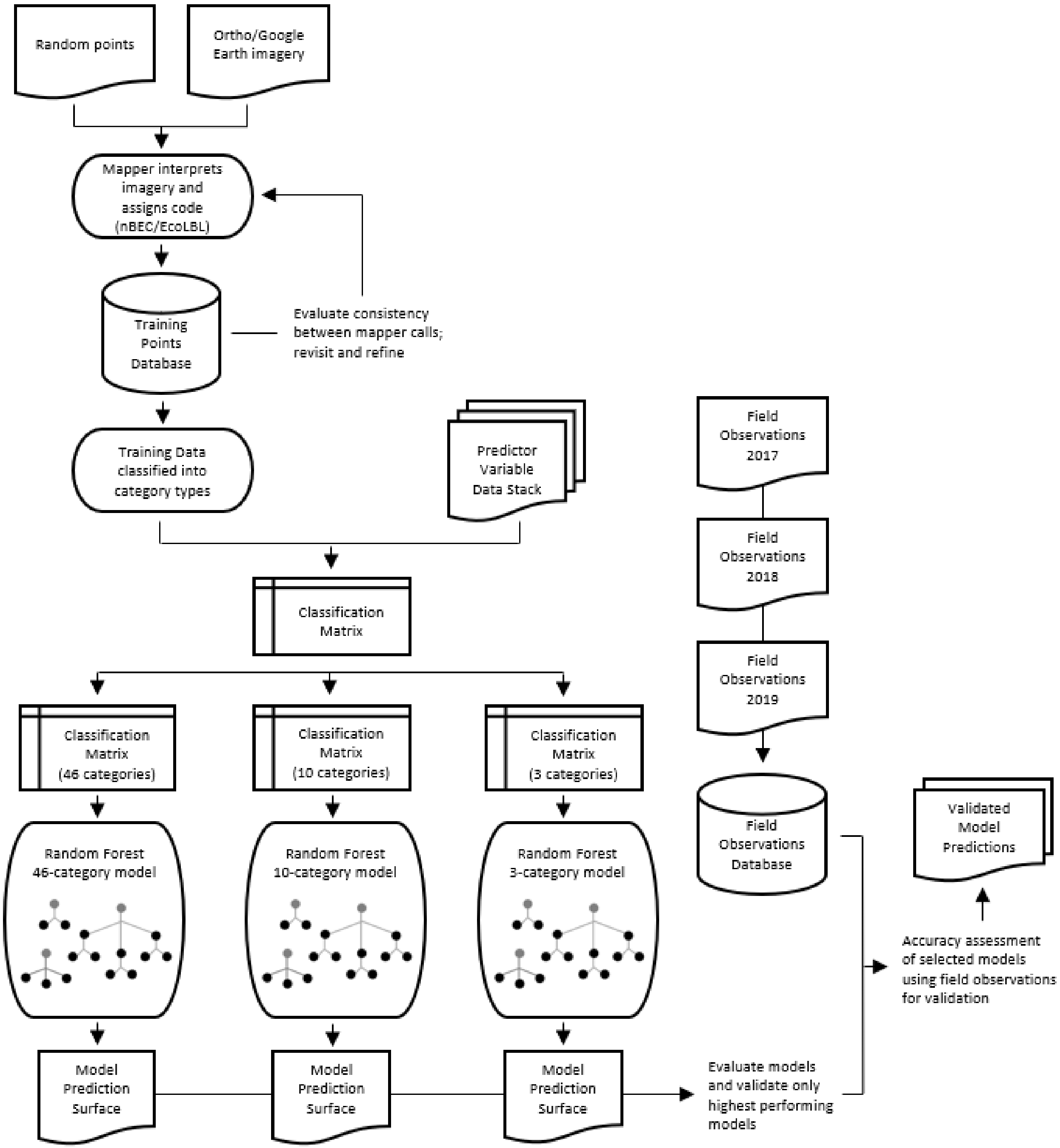


Figure 2: Diagram of Run 2 modeling process for the FWCP Peace region.

STEP 1: TRAINING DATA ADDITIONS, REVIEW AND DATA CLEANING

Training data were developed by mapping contractors who worked with ENV staff to interpret satellite imagery and assign upland, water, and riparian and wetland category codes for randomly generated points. To ensure consistency with existing mapped points, mappers who were previously trained during Run 1 were also retained for mapping during Run 2.

The Parsnip sub-basin was selected for additional training points as this area contained particularly high prediction errors identified during Run 1. Within the Parsnip sub-basin, two new map sheets were selected based on presence of timber harvest, density of wetlands, and site access for field validation.

1000 additional training points were interpreted for 1:20,000 NTS map sheets 093J067 and 093J047; each containing 500 randomly generated points (Figure 3). Identical to the Run 1 approach, contractors assigned each point a wetland group/class code or nBEC code (Mackenzie, 2004; Mackenzie, 2012) – the model response variable. Where feasible, contractors also interpreted additional attributes including hydrogeomorphic systems and subsystems, as well as anthropogenic units. Standardized codes for attribution were reduced from the previous 48 to 46 codes, by removing 'ponds' category and lumping all previously attributed ponds with 'lakes' (no size restriction) as well as eliminating the 'other' category *(Appendix C).*

Overall, a total of 13,500 training points were attributed. Of the total 13,500 training points, 37 points fell outside the FWCP Peace boundary, where the map sheet extended past the edge of the boundary. These 37 points were excluded prior to running the Random Forest model.

Contractors also revisited training data from Run 1 and corrected inaccuracies and inconsistences in interpretations from air photos. To further verify the accuracy of mapper calls, training point locations were visited in the field where independent field calls were recorded. Training and field points were compared between, and within, categories to assess overall agreement and disagreement between calls.

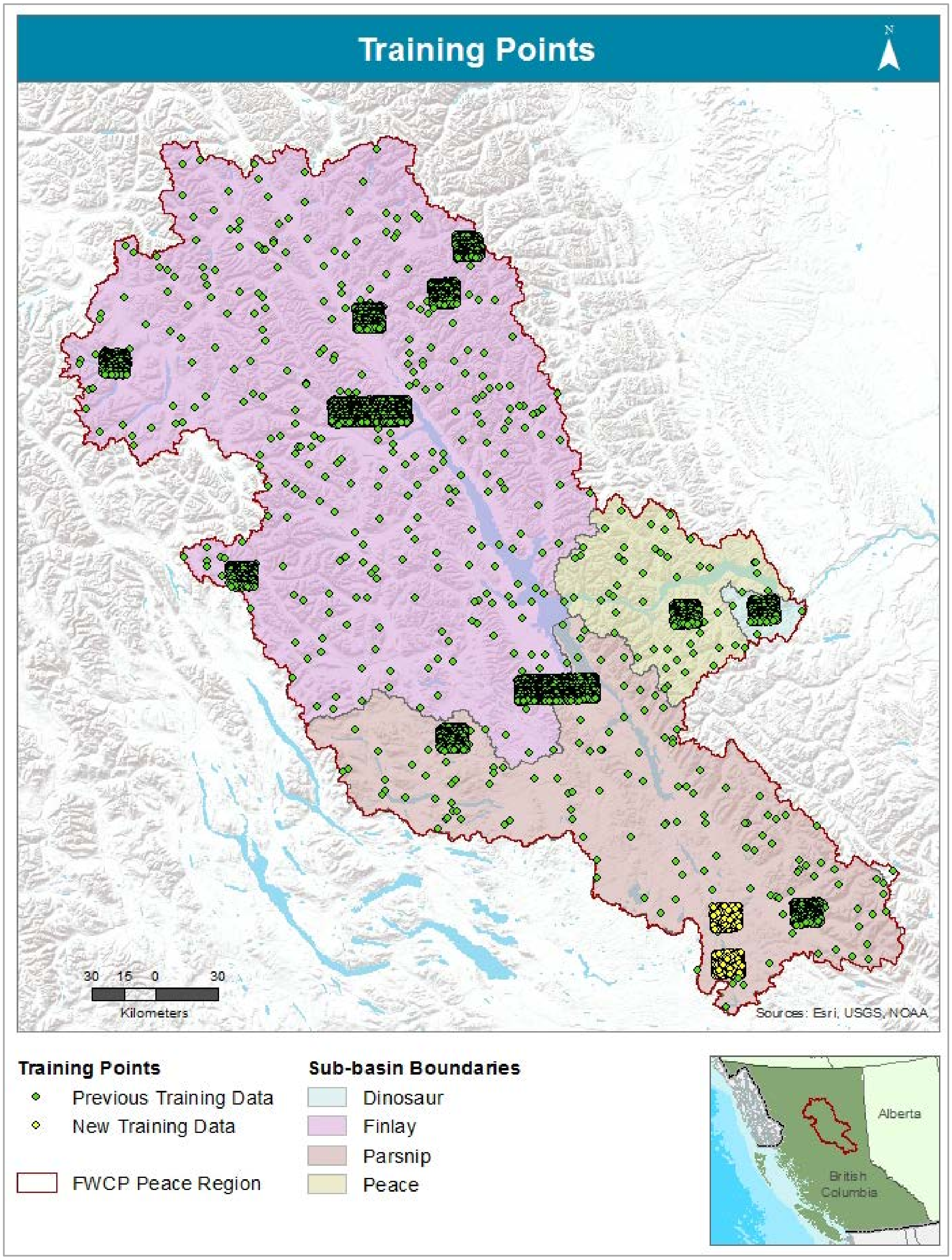


Figure 3: Model training points from previous model iterations (green) and new 2019 additions (yellow). These data were used to train the Random Forest model. Total training points 13,500.

STEP 2: FIELD DATA ADDITIONS AND DATA CLEANING

2017 field efforts were concentrated along Williston Lake within the Finley and Parsnip sub-basins. In 2018, due to access and time constraints, field work was conducted via truck-based inspections and on-site plots in a smaller geographic area within the south Parsnip sub-basin (Figure 4).

Field work priorities for 2019 focused on the Parsnip sub-basin – previously identified as regions with high model error – and in areas of disturbance. Additionally, this area exhibited high concentrations of wetlands and good access into sites for truck and helicopter-based sampling. During Run 2, additional training points were added for two map sheets in the Parsnip sub-basin. Where possible, field crews targeted training points for field visits on the ground, to allow for validation of mapper calls as well as model accuracy.

Truck-based field work was completed in 2019 over 10 days with multiple crews working simultaneously throughout the Parsnip sub-basin in up to 4 vehicles. Helicopter based sampling was conducted over two days and focused in areas inaccessible by ground vehicles.

The example map sheet 093J070 was selected as a case study region of which to further demonstrate model improvements and confusion. The map sheet covers a large connected wetland complex, containing a mosaic of different wetland types. Map sheet 093J070 was visited in the field and contained a dense sampling strategy by helicopter. As a result, this map sheet is a good area to assess the modelled results from both phases (run 1 and run2).

Over the summer of 2019, 620 new field observation points were collected. During this period, ENV field staff worked closely with First Nation partners from Prophet River First Nations and McLeod Lake Indian Band who provided local area knowledge and expertise during field sampling.

Digital field collection forms were created using Collector (Esri ArcGIS), applying domains to generate picklists and required fields following provincial standards where appropriate (British Columbia Ministry of Forests and Range and British Columbia Ministry of Environment, 2010). Digital capture was preferred over paper forms as it improved collection speed and eliminated the need to transcribe paper forms post field collection. Additionally, data collection via tablets enabled the user to place field collection points in locations physically inaccessible but viewable from a distant vantage point.

A simplified picklist of wetland types was used to improve efficiency in recordings and reduce potential input errors and confusion. Recording the 3-category call (upland, water, wetland) at each site was prioritized and allowed for general categorization of land cover type when detailed 46-category calls were not achievable. Where possible, field technicians recorded wetland categories down to the 46-category level (Appendix C). However, in many cases, site access or conditions – such as transitional zones within wetland complexes – made it difficult to establish a call at this level of detail. In addition, the 46-category level was difficult to discern from visual calls from a distance or by helicopter. As a result, most field observations could not be collected at the level of detail required for the 46-categories. A summary of the field visits in each category for the 10 and 3 category groupings can be found in Table 1 and Table 2 respectively.

Over the three field seasons, data were collected on various GPS enabled devices with varying GPS positional accuracy capabilities. Sites with dense overhead vegetation or locations where topography limited receiver access to satellite signal (e.g., valleys or low-lying areas) exhibited reduced GPS accuracy (U.S. National Coordination Office for Space-Based Positioning, December 2017). In these circumstances, additional positional error is introduced and added to the already ± 5-10 metre variable accuracy from device sensitivity.

Given the GPS positional variability, plot locations may not perfectly coincide with the physical location from which they were recorded. Due to the 25 x 25 metre resolution of the model predictions, GPS error may shift field points into an adjacent pixel. These types of shifts may be more prevalent or exacerbated along edges of wetland complexes or transitional zones where field technicians may have trouble accessing or identifying homogeneous landscape conditions from which to sample. In some cases, site access (e.g., water level) may have resulted in a point collection along a pixel edge rather than the centre. In areas where location inaccuracy was identified, the point was manually moved to a more appropriate location on the data collection device.

Geo-enabled photographs were taken to supplement the information collected at field location sites. Photographs were assigned a plot number and corresponding metadata attributes, including the date, 24-hour time of capture, plot number, and type. Image types included flora (Fl), fauna (Fa), landscape (L), soil (S), and other (O) that allow users to filter for desired types. Images from the 2018 and 2019 field efforts were hosted on the BC Data Distribution site[5] to allow remote access.

Geo-enabled photographs are embedded with coordinate location information and may also contain other metadata attributes such as aspect or elevation. 2018 and 2019 field collected photographs were captured with various GPS enabled devices, with differing locational accuracy, under various conditions. Therefore, image locations may not coincide perfectly with plot location, or point of image capture. Additionally, metadata attributes may contain systematic errors due to device calibration issues. Some devices assign a -9999 value to the aspect field and therefore, images with -9999 assigned will not accurately represent the correct capture aspect.

---

[5] BC Data Distribution Site address: http://www.env.gov.bc.ca/esd/distdata/ecosystems/TEI/WETLANDS

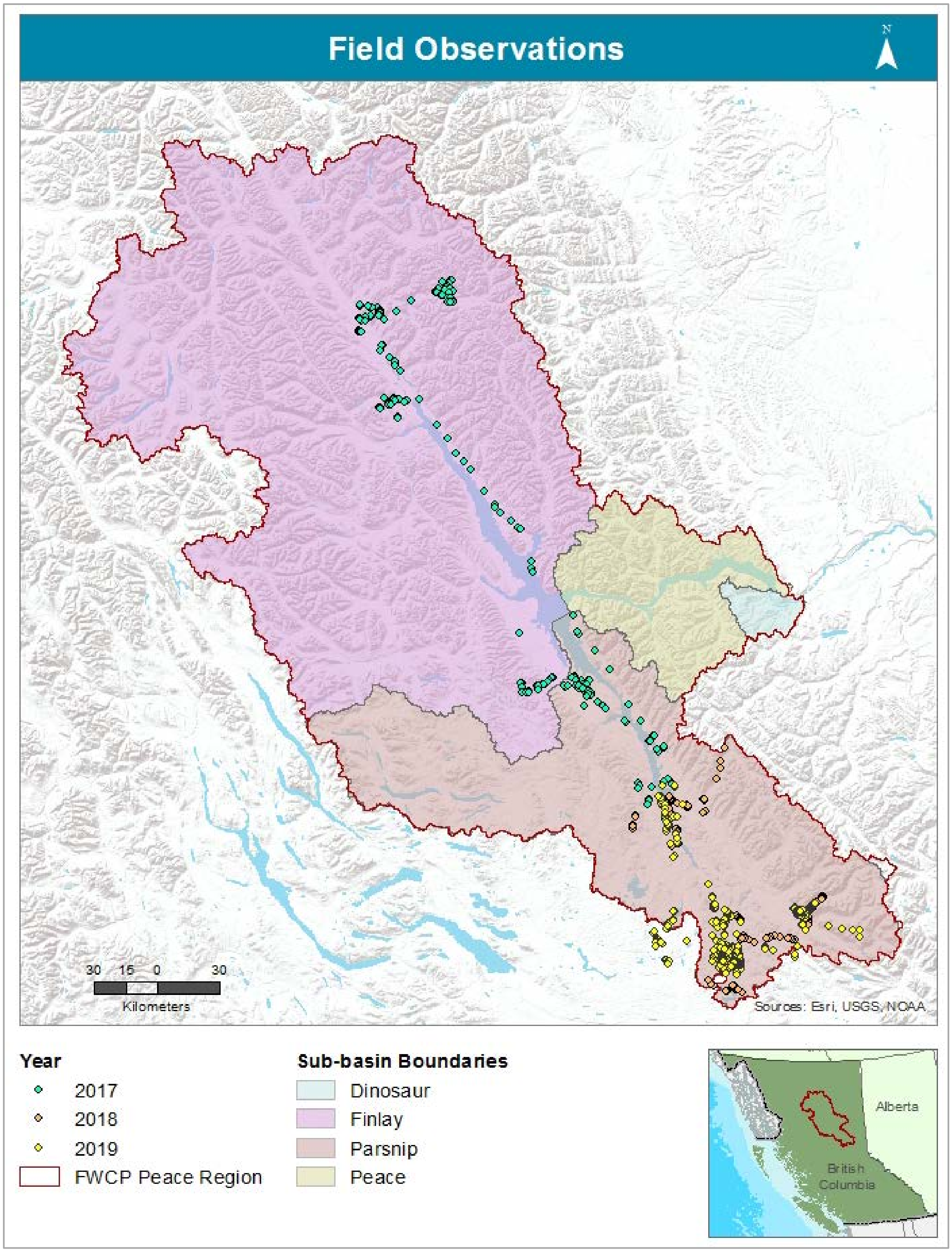


Figure 4: Field collected data points for 2017, 2018, and 2019 summer field work efforts. Only field data within the FWCP Peace Region were used to validate model predictions.

Table 1: The proportion of each wetland class observed during each field collection year. Calculations were generated from all field data collected, including observations that fall outside the FWCP Peace Region.

Field Observations by Class										
Year	**Bog (Wb)**	**Fen (Wf)**	**Marsh (Wm)**	**Swamp (Ws)**	**Upland (T)**	**High-bench Floodplain (Fh)**	**Mid-bench Floodplain (Fm)**	**Low-bench Floodplain (Fl)**	**Water (W)**	**Shallow Water (Ww)**
2017	31 12%	80 31%	30 11.6%	47 18.2%	21 8.1%	2 0.8%	9 3.5%	20 7.8%	12 4.7%	6 2.3%
2018	33 13%	45 18%	14 5%	33 13%	69 27%	5 2%	11 4%	41 16%	5 2%	0 0%
2019	90 15%	98 16.6%	39 6.6%	52 8.8%	216 36.5%	1 0.2%	8 1.4%	40 6.8%	33 5.6%	15 2.5%
	Total 154	**Total** 223	**Total** 83	**Total** 132	**Total** 306	**Total** 8	**Total** 28	**Total** 101	**Total** 50	**Total** 21

Table 2: The proportion of each category (Upland, Water and Wetland) observed during each field collection year. Calculations were generated from all field data collected, including observations that fall outside the FWCP Peace Region. High-bench flood units were grouped with upland calls, while mid- and low-bench flood units were grouped with wetland calls. Total field points 1,135 (35 observations removed during the QA process).

Field Observations by Category				
Year	**Total number**	**Upland (T)**	**Water (W)**	**Wetland (WL)**
2017	258	20 7.75%	13 5.04%	225 87.21%
2018	257	69 26.85%	5 1.95%	183 71.21%
2019	620	217 35.00%	33 5.32%	370 59.68%
		Total 306	**Total** 51	**Total** 778

STEP 3: SCRIPT DEVELOPMENT

Scripts developed for the first run of the model were updated and incorporated into an R package (R Core Team, 2019) called wetlandmapR (https://github.com/bcgov/wetlandmapR). With the package the script was organized into functions. These are reusable chunks of code that can be applied to a variety of workflows or used in future script development. Functions allow for replication of interim steps in the process.

Development of the R package also allowed for documentation, creation of a data dictionary and metadata which facilitates repeatability and transparency of the process. An example data package was assembled for inclusion in the wetlandmapR package using map sheet 093J070. This test data was used for script testing, quality control and debugging and will be valuable for future script additions improvements.

STEP 4: RANDOM FOREST MODEL REFINEMENT

The wetlandmapR scripts were used to run the Random Forest model using the 13,463 training points and the predictor variables developed in the first model run of the project. For model Run 2, all predictor variables remain the same as Run 1, as updated datasets were not available at the time of model processing (Appendix B). Where necessary, variables were resampled to a standardized grid size of 25 meters, selected to match the resolution of the Digital Elevation Model (DEM).

Predictor layers were selected for a variety of reasons and prepared using a series of preprocessing steps. For a detailed description of preprocessing steps as well as layer selection rationale and limitations, see the first project report (Filatow *et al.,* 2018). The following input layers were used:

- Provincial Digital Elevation Model (DEM) at 25 m resolution (British Columbia, 2002);
- System for Automated Geoscientific Analysis (SAGA) derived topographic indices generated from the DEM;
- Three band composite of Landsat-8 satellite imagery optimized for vegetation;
- 13-band multispectral Sentinel-2 satellite imagery;
- ClimateBC layers version 5.10 (1961-1990 climate normal; Wang *et al.*, 2012). At the time of model processing, no updated climate variables were available.

Applying the new standardized training data categories (46 categories rather than 48 used in Run 1; Appendix C) as the model response variable, several iterations of the Random Forest model were generated using the wetlandmapR scripts that called on functions from the ModelMap package (Freeman and Frescino, 2009) in R.

MODEL ACCURACY ASSESSMENT

Two metrics of model performance were used to evaluate model results: (1) an internal estimate of model error (out-of-bag [OOB] error) and (2) model validation using field sampled data.

First, model performance was determined based on an internal error metric (OOB error). The OOB error indicates how well the algorithm performed when predicting held-out (unseen) training data. A random sample of 1/3rd of the input training data (bootstrap) is used to generate a fully-grown decision tree which is then compared with held out data to calculate error. Metrics from all other trees are combined to establish overall difference (Breiman, 2001b).

Second, model prediction categories were compared to field collected data and the match up between field calls and model predictions were evaluated. By assessing field data agreement with modelled results, it becomes apparent which categories the model was less successful at predicting. Field validation results are reported in percent correct match.

In addition, both OOB errors and accuracy using field data were used to evaluate the 3, 10 and 46-category predictions. The 46 categories were aggregated into broader wetland groupings as an approach to improve model performance and prediction accuracy. Various model response variable configurations were attempted and only the best performing models were selected (Table 3). Multiple configurations for grouping the 10-categories into the 3-categories were assessed to select the best performing options. A crosswalk table of the category codes is available in Appendix C.

TRAINING AND FIELD VALIDATION COMPARISON

In addition to model accuracy, the accuracy of expert interpreted training data is a key assumption of this predictive wetland mapping methodology. It is assumed that the interpreted training points used to train the model were assigned the correct response variable classes. However, any errors in interpretation or class assignment may lead to model confusion between class types. Therefore, it is important to evaluate how well the training data were interpreted by validating training data calls against field observations.

Accuracy between mapper interpreted training data calls and field calls were assessed by evaluating how well the training data calls matched the field observations. When feasible, field technicians for the 2019 season attempted to visit as many training data sites as possible to provide an on-the-ground record of the site conditions and assign wetland classes and other categories interpreted in the training data set. To eliminate potential bias, points were collected on site without technician’s knowledge of the training data call – only training point locations were provided to field crews.

Table 3: Random Forest model was run under various category configurations to find the optimal and best performing model. Of the iterations performed, only the best performing models were selected.

Model Iterations Evaluated		
Category	**Category Description**	**Selection**
46-Category	• 46 individual categories (Appendix C)	NOT SELECTED (Appendix C)
10-Category	• **Upland:** All upland units • **Water:** All water units • **Bog** • **Fen** • **Marsh** • **Swamp** • **Shallow Water Wetlands** • **High-bench Floodplain** • **Mid-bench Floodplain** • **Low-bench Floodplain**	**SELECTED** **(Appendix C)**
3-Category	• **Upland:** All upland units including high-bench floodplain • **Water:** All water units • **Wetland:** All wetland units (bog, fen, marsh, swamp, shallow water wetlands) including mid and low-bench floodplains	**SELECTED** **(Appendix C)**
3-Category	• **Upland:** All upland units • **Water:** All water units • **Wetland:** All wetland units (bog, fen, marsh, swamp, shallow water wetlands) including high, mid, and low bench floodplains	NOT SELECTED
3-Category	• **Upland:** All upland units including high, mid, and low bench floodplains • **Water:** All water units • **Wetland:** All wetland units (bog, fen, marsh, swamp, shallow water wetlands)	NOT SELECTED
4-Category	• **Upland:** All upland units • **Water:** All water units • **Wetland:** All wetland units (bog, fen, marsh, swamp, shallow water wetlands) • **Floodplain:** high, mid, and low bench floodplain units	NOT SELECTED

STEP 5: MODEL IMPROVEMENT AND RELIABILITY ASSESSMENT

MODEL IMPROVEMENT

A qualitative approach was used to assess model improvement by visually comparing model Run 1 predictions (3 and 10-category) with those of Run 2. The improvements were demonstrated by providing side by side visual comparisons. Emphasis was placed in disturbed areas that were identified as being poorly predicted in the first run of the model.

In addition, a quantitative approach using field data was applied to calculate accuracy difference between model Run 1 and Run 2. Field data were compared against model results from model Run 1 and Run 2 to determine overall improvement in the 3 and 10-category predictions. Improvements were assessed across the entire FWCP Peace region as well as for the Parsnip and Finlay sub-basins. As some of the classification groupings changed between phases (model run 1 vs run 2), mid and low-bench floodplains and shallow water wetlands were removed prior to calculations. This ensured that only like categories were compared.

RELIABILITY ASSESSMENT

The reliability of the modelled results should be critically assessed by the end user and should be considered prior to the implementation of any management actions. The user can assess the model results through: (1) comparing model predictions against field observations, (2) assessing model predictions against training data density, as well as (3) model to model comparison. To assist with evaluating model reliability, figures were generated to summarize the spatial distribution of training and field data points across the FWCP Peace region. For each 1:20,000 scale map sheet the number of training and field data points were calculated. Training data were displayed using manual break values to account for the skewed distribution that resulted from low density random points across the entire FWCP Peace region and higher density data within selected map sheets. Field data were displayed using quantile breaks. Additionally, an online mapping application was created to facilitate model to model comparison (see *Data and Information Access: Web Mapping Application* methods section for additional details).

STEP 6: WETLAND PREDICTION SUMMARIES

Prediction results from the 3-category model were converted to vectors and summary statistics were calculated for the FWCP Peace region, sub-basins, and 1:20,000 map sheets. The FWCP Peace region boundary used for summary calculations was based on the exact raster extent of the model predictions and maintained the cell edges of the surface (i.e., polygon edges were not simplified or smoothed). The region was subset into sub-basins using a polygon feature provided by FWCP. The method for establishing the region and sub-basin boundaries were different between Run 1 and Run 2 and, as a result, some discrepancies in area calculations may be present between reports.

Model results for the 3-category prediction were converted to vector using GRASS GIS version 7.8.1 r.to.vector module. The '–s' flag was used to smooth the corners of the resulting polygons (GRASS Development Team, 2019).

Based on the converted polygons, wetland density for the FWCP Peace region was calculated for each 1:20,000 scale map sheet. Prior to calculation, lakes greater than 1000 ha were removed from the map sheet areas. Removed lakes included Williston, Tchentlo, Thutade, Carp, Chuchi, McLeod, Tatlatui, Tsayta, Germansen, Witch, Kitchener and Summit Lakes. To account for edge effects, only map sheet areas that fell within the FWCP Peace region boundary were included in the calculation.

The Random Forest model results for the 3-category prediction were summarized in tabular form where the total area and density of wetlands for each sub-basin in the FWCP Peace region were calculated. Distribution of large wetlands was assessed via a figure to establish trends in wetland size across the region.

DISTURBANCE ANALYSIS CASE STUDY

APPROACH OVERVIEW

Cumulative impacts can affect wetland structure and function both directly and indirectly. Direct impacts include disturbances that occur on the mapped wetland footprint, whereas indirect impacts may include disturbances to upland catchment areas that are hydrologically connected to wetlands. Landscape-level disturbances that can directly and indirectly impact wetland function include (but are not limited to):

- Fire perimeters
- Timber harvest
- Road construction
- Mineral and Placer leases
- Coal mining
- Mountain Pine Beetle infestation
- Spruce Beetle Infestation
- Land conversion to Agricultural and Residential/Agricultural mixed land cover
- Climate Change
- Linear right-of-ways
- Reservoirs

[6]Although timber harvest is not necessarily a permanent disturbance on the landscape in the same way that conversion to urban land cover is; it often impacts wetlands by altering evapotranspiration, water levels (Dube *et al.*, 1995) and localized ecological function, increasing sedimentation rates (Moore and Wondzel, 2005), and altering hydrologic, thermal, and chemical regimes (Mellina, 2002). Additional work has shown that timber harvest can also negatively impact invertebrate communities (Batzer *et al.*, 2000, Kreutzweiser *et al.*, 2008), change the base of food webs, and affect leaf litter decomposition and nutrient cycling (Kreutzweiser *et al.*, 2008). Road networks and densities can be used as a proxy for human activity and negative effects on a landscape (Trombulak and Frissell, 2001). Roads contribute to fragmentation and edge effects, can alter hydrology patterns, and increase invasion by exotic species. Furthermore, it is demonstrated that overall species richness in wetlands, as defined as number of different species represented in an ecological community, decreases with increased road density (Findlay and Houlahan, 2003). Mining activity (coal, placer, mineral leases and claims) changes the landscape in several ways including permanent land conversion, altered hydrology patterns, and fragmentation. These activities can release undesirable contaminants into watersheds and contribute to siltation. Successful remediation of ecosystems and wetlands after a disturbance may be difficult, as natural

---

[6] This paragraph content is reproduced from the “Predictive Wetland Mapping of the Williston Drainage Basin” report from 2018. http://www.env.gov.bc.ca/esd/distdata/ecosystems/TEI/WETLANDS/Williston_Wetlands_Reporting/

wetlands can take several decades to develop a complex level of functional interactions. Potential future mining activity can be inferred by placer or mineral claims and reserves.

A number of these disturbances have spatial data available through the BC Data Catalogue, except for climate change impacts. As climate change impacts are anticipated to cause significant changes to wetland form and function (altered hydrology patterns, water inputs etc.), their spatial impacts in this region are not yet well-documented or researched and are out of the scope for this disturbance analysis.

Given the spatial nature of a number of these disturbances, a method to determine how they might impact wetlands could be used to stratify wetlands in various ways. For example, wetlands with relatively few disturbances could be identified as candidates for both protection and monitoring as their baseline condition is presumed to be relatively undisturbed. Conversely, wetlands with significant disturbance could be identified as candidates for enhancement and remediation efforts, with the effectiveness of these efforts recorded through monitoring programs.

Multiple methods of aggregating threats and impacts to wetlands have been developed (Faber-Langendoen *et al.*, 2012; IUCN, 2012). However, aggregation of threats impacts may be difficult as 1) their severity and magnitude depend on wetland type (e.g., a bog responds differently to altered hydrology patterns than a marsh), 2) the severity of impacts to wetlands, for both direct and indirect threats, is relatively unknown and requires more research, 3) how threats interact with each other on the landscape is unknown, and 4) the map scale of the disturbance layers may be at broader or finer scales than the mapped wetlands.

The disturbance assessment method is an example case study that applies a Tier 1 desktop assessment and will benefit from additional work on the ground (Tier 2). This method evaluates relative disturbance levels and their spatial relationship to wetlands using a scalable approach that can be applied to different sub-areas within the FWCP Peace region. The disturbance assessment analysis was completed in three stages: 1) aggregating wetlands into unique groupings based on hydrological connectivity and proximity, 2) calculating disturbances, and 3) generating disturbance density summaries for multiple disturbance types to investigate watershed group level trends. Results are presented as an example analysis, and therefore future implementations of this approach should be tailored to specific wetland management questions and targets.

STEP 1: WETLAND AGGREGATION

Individual wetlands were aggregated together to account for ecologically meaningful connectivity (e.g., hydrology and seed dispersal) as wetlands are known to occur both as complexes and as individual unconnected wetlands. To group wetlands into aggregated units representing connected systems, a series of data processing steps were required. The following data sources were used:

- Freshwater Atlas Stream Network
  https://catalogue.data.gov.bc.ca/dataset/freshwater-atlas-stream-network
- Freshwater Atlas Analysis Watersheds
  https://catalogue.data.gov.bc.ca/dataset/freshwater-atlas-assessment-watersheds
- Riparian potential layer (The Nature Trust of BC, 2019)

As discussed in Step 6 of the Model Refinement methods, predicted wetlands were converted from raster surfaces to polygons. Wetland polygons that share a vertex were lumped together and assumed to be connected features. In other words, if any wetlands "touched", they were considered to be part of the same complex.

Next, wetlands were further aggregated to account for hydrological connectivity (ecologically meaningful connectivity). The Freshwater Atlas Stream Network was used to represent surface connectivity and a riparian potential layer was used as a surrogate for subsurface hydrological connectivity. These two spatial layers were used to attribute whether a wetland was connected (1) or not connected (0) to neighbouring wetland polygons. Wetlands that fell within a riparian potential polygon connected by a stream were grouped as connected units.

The Freshwater Atlas Assessment Watersheds were used to account for height of land and stream order. A third assessment of connectivity was included in the form of a distance threshold. The threshold was included to account for separation distances that would influence seed dispersal etc. For example, after a certain distance (e.g., user defined threshold of 1 km), seed dispersal would not be effective. This separation distance is applied from NatureServe methods and relevant element occurrence specifications (Cadrin and Christy, 2013) and applied in this analysis as 1 km from the nearest "connected-watershed" polygon. Note, as per all user defined threshold parameters, this separation distance should be researched and justified depending on the value or function being assessed.

This aggregation method resulted in wetlands that were grouped according to:

- Watershed ID – hydrologically connected – within 1 km
- Watershed ID – hydrologically connected – further than 1 km
- Watershed ID – not hydrologically connected

Assumptions with these groupings include:

- Spatial accuracy of all input layers are reasonable for the 25m prediction surface.
- Attribution of the input layers is accurate.
- Use of hydrologically significant study area. There are inconsistencies between administrative and watershed boundaries (height of land) in the FWCP Peace region. The

FWCP Peace boundary also does not match the drainage divide of the 25m DEM nor the analysis watersheds and watershed groups.

- To account for phenomena such as seed dispersal, 1 km separation distance is appropriate for wetlands that are likely to be hydrologically connected. This value was selected based on Element Occurrence (EO) specifications, but is not necessarily well-supported (i.e., seed dispersal vectors on waterfowl, etc.).
- Hydrological connectivity is based on surface hydrological flow. Factors such as parent material and porosity (geological factors) are not accounted for. Groundwater flow is unknown.
- Hydrological connectivity is related to the riparian potential layer.

STEP 2: DISTURBANCE CALCULATION

Disturbances were calculated on grouped wetland area (i.e., footprint) and within the adjacent landscape (i.e., 1 km buffered footprint in this example) to determine direct-footprint and indirect-landscape disturbance. Calculations summarized area of disturbance type within unique wetland complexes (groups) and associated buffers. Note that the 1 km buffer of the wetland footprint was arbitrarily chosen and should depend on the analysis question. In other words, the analysis parameters, such as the buffer distance threshold, should be tailored to the question.

Example disturbance layers from the BC Data Catalogue were used to illustrate the method. Disturbances incorporated into the example analysis included:

- Digital Road Atlas (DRA)
  https://catalogue.data.gov.bc.ca/dataset/digital-road-atlas-dra-master-partially-attributed-roads
- Consolidated Cutblocks
  https://catalogue.data.gov.bc.ca/dataset/harvested-areas-of-bc-consolidated-cutblocks-
- Fire Perimeters
  https://catalogue.data.gov.bc.ca/dataset/fire-perimeters-historical
- Pest Infestation Spruce Beetle
  https://catalogue.data.gov.bc.ca/dataset/pest-infestation-polygons
- Pest Infestation Mountain Pine Beetle
  https://catalogue.data.gov.bc.ca/dataset/pest-infestation-polygons
- Mines (mineral and coal tenures)
  https://catalogue.data.gov.bc.ca/dataset/mta-mineral-placer-and-coal-tenure-spatial-view
- BTM (agriculture, urban)

https://catalogue.data.gov.bc.ca/dataset/baseline-thematic-mapping-present-land-use-version-1-spatial-layer

With the exception of roads, linear features are not assessed (such as right-of-ways for oil and gas activities). Note that road analysis depends specifically on density, and impacts are assumed to be related only to linear density and not any kind of decay function related to road proximity. In this analysis, reservoirs and dam impacts, including operational impacts (e.g., drawdown impacts), were not assessed. Impacts related to the Williston and Dinosaur Reservoir are addressed in detail under the BC Hydro's Water Use Program, Williston Reservoir Management Plan[7], and Dinosaur Reservoir Management Plan[8] and are therefore out of scope for this disturbance analysis.

Disturbance layers were not weighted to evaluate cumulative impacts as more research is required to appropriately select weighting values. For example, roads and timber harvesting tend to co-occur in space, but each impacts wetlands differently. Consequently, the management of these impacts would be specific to each threat. However, to provide an overall disturbance summary, aggregation of disturbance by count (i.e., number of disturbance categories within each wetland group or buffer) is an acceptable means of summarizing disturbance load on landscape.

STEP 3: CALCULATION AND CLASSIFICATION OF DISTURBANCE DENSITIES

The density of disturbances, within wetland footprints or 1 km buffer, was calculated to standardize the data based on area. Fields for each layer includes area (actual), density (relative), and presence. The presence of each disturbance type was used to calculate overall disturbance load for each aggregated wetland footprint and associated buffer. Note that applications and future land-use such as coal license applications and proposed roads were not included in cumulative landscape disturbance calculations.

Display of the results is based on the FWCP Peace region. However, area of interest can vary and therefore represents a "relative extent". The results are binned by quartiles, so the project extent influences relative density distribution in the histogram, (i.e., the most disturbed wetland group in the FWCP Peace region may be different than the most disturbed wetland group in Parsnip sub-basin). Quartiles were selected as the most appropriate method to bin disturbance results as the ends of the distribution (high and low) could be easily identified and therefore, disturbance load could be assessed relative to the distribution of the data within the area of interest.

---

[7] https://www.bchydro.com/toolbar/about/sustainability/conservation/water_use_planning/northern_interior/peace_river/williston_reservoir.html

[8] https://www.bchydro.com/about/sustainability/conservation/water_use_planning/northern_interior/peace_river/dinosaur_reservoir.html

DATA AND INFORMATION ACCESS

WEB MAPPING APPLICATION

A core goal of the project was to establish an effective way for interested parties, such as the public and key stakeholders, to share and engage with project material. Web mapping applications, like those offered by ESRI's ArcGIS Online, present an opportunity to develop online tools that facilitate the sharing of spatial data and information though an engaging and interactive user experience. As a result, ArcGIS Online Story Maps was a key solution for disseminating project results.

Web mapping platforms cater to various levels of expertise and are displayed in an intuitive and familiar context for many non-expert users. In our increasingly connected and digital world, many individuals have become familiar with digital mapping platforms (e.g., Google Maps and OpenStreetMap) that are typically accessed via smart phone, tablet or computer. Web mapping applications draw on familiar functionality from these commonly used applications for navigating and exploring spatial data through an online map interface.

At its most basic level, online web mapping applications typically contain a map window with conventional zooming and scrolling capabilities. More advanced web maps may include customized tools for analyzing, visualizing or exploring map content. In some instances, web mapping applications allow for the inclusion of contextual information containing links, documents, and photographs. Thus, the user can explore all relevant content (not only spatial datasets) and build a holistic understanding of a given topic. As a result, web maps offered a unique solution for distributing project results while offering an open and accessible resource for accessing project material.

An ArcGIS Online web mapping application prototype, Williston Wetland Explorer Tool (WWET), was developed using the following data layers:

- Training data
- Field observations
- Field photographs
- FWCP Peace boundary
- 3-category model
- 10-category model
- Ancillary data sets (e.g., Consolidated Cutblocks and Digital Roads Atlas)

The application capabilities include the ability to add additional datasets as needed. Available ancillary data from BC Data Catalogue (e.g., BEC Units) are potential datasets that may be added to the application as future need arises.

A user guide, project information page, and web map were developed within the ArcGIS Online Story Map template. Each section was partitioned via separate tabs where users can navigate to access various information content. If required, additional tabs may be added.

Demonstrations of the WWET application were conducted to the professional community of Biologists at the Annual Professional Biology Conference & AGM on April 28 and 29, 2020. Further demonstrations to interested groups such as First Nations Communities or FWCP board members are also feasible.

SCRIPT AND DATA DISSEMINATION

Project data and scripts are available through open source tools and platforms to facilitate transparency, information sharing, and collaboration. These sites include the Terrestrial Ecosystem Information (TEI) Data Distribution Site, Ecological Reports Catalogue (EcoCat), and GitHub.

Spatial data in both raw and edited formats, reports, data dictionaries, model results and graphs are available through the TEI Data Distribution Site. As new datasets and information are collected, the distribution site can be updated to provide access to these data. Project report and key data outputs will also be stored on EcoCat[9].

SCRIPT PACKAGE AND GITHUB

As introduced in the *Step 3: Script Development* section, the Random Forest model script is available through the wetlandmapR GitHub repository (https://github.com/bcgov/wetlandmapR). GitHub is a free and open source web-based platform that promotes the distribution and improvement of digital assets, such as – but not limited to – data and source code sharing, as well as software development. Contributors can modify and develop upon existing resources or generate new repositories, by following well established guidelines for appropriate use (Perez-Riverol *et al.*, 2016). Notably, GitHub's user-friendly platform has reduced common barriers for many users and therefore, promotes high participation rates (Peterson, 2013), at varying intensities, across multiple disciplines.

In a scientific research context, GitHub offers a platform upon which the 'continuous improvement cycle' of scripted workflows can be hosted, improved, peer-reviewed, and shared. A foundation of version control practices is integrated into GitHub via Git Source Code Management (SCM; http://git-scm.com/; Blischak *et al.,* 2016). Users can retrieve (pull) data from an existing repository,

---

[9] EcoCat access: https://www2.gov.bc.ca/gov/content/environment/research-monitoring-reporting/libraries-publication-catalogues/ecocat

make changes and then request that the repository owner review and implement the alterations into the master script, thus integrating, recording and disseminating the new additions.

Scientific research is often conducted in institutional silos and behind paywalls or data access restrictions. The important details related to data and method application are often obscured ("black box") when datasets and scripts have public access restrictions. Open source platforms, such as GitHub, represent an emerging trend of data sharing and collaboration that is important step forward for producing repeatable and transparent scientific research, while also facilitating partnerships between users of different expertise.

Open access to data and scripts is a priority goal for this project. Input data preparation and Random Forest scripts were uploaded to GitHub where further improvements to the script were implemented. Scripts are written in R and refined to allow for an iterative workflow. GitHub's version control facilitated collaborative script and package development. The inclusion of the example data package (map sheet 093J070) aided in testing and debugging the scripts and R package.

RESULTS

The Random Forest model was produced and validated using the best available data at the time of processing. The overall success and accuracy of the predicted product is correlated with the quality of the data and the algorithm efficiency. Therefore, the results should be interpreted and approached with a level of caution, as predicted results are estimates and thus may differ from actuality. The results provide an estimate of potential wetland and riparian distribution across the FWCP Peace region from which further study and investigation may be performed.

Key results (Table 4) are summarized and discussed in detail within their relevant sections below.

Table 4: Key project results.

Key Results			
Method	**Results Section**	**Step**	**Results**
Model Prediction	*Random Forest Model Prediction & Internal Model Accuracy Assessment*	Assess multiple iterations of the 3, 10 and 46-category groupings to establish the best performing models	3-category model (low- and mid-bench floodplain units grouped with the wetland category) produced a 7% OOB error (93% internal model accuracy)
			10-category model (shallow water wetland as 10th class) produced a 11.7% OOB error (88.3% internal model accuracy)
			46-category model produced a 32.95% OOB error (67.05% internal model accuracy)
	Model Accuracy Assessment	Determine accuracy of selected models using independent validation dataset – all field collected data points from 2017-2019	3-category model (low- and mid-bench floodplain units grouped with the wetland category) produced 86.7% field validated accuracy
			10-category model (shallow water wetland as 10th class) produced 52.8% field validated accuracy
			46-category model produced 17.2% field validated accuracy
	Training and Field Validation Comparison	Compare how well the training data (desktop inventory by mapping experts) aligned with all field observations	Percentage of correct training data for the 3-category classification: upland (81%), water (89%) and wetland (98%).
			Percentage of correct training data for the 10-category classification: upland (91%), water (89%), bog (86%), fen (63%), marsh (18%), swamp (33%), high-bench floodplain (N/A), mid-bench floodplain (29%), low bench floodplain (16%), and shallow water wetland (50%)

Model Improvement and Reliability	*Model Improvement*	Compare model accuracy of the second model run (Run 2) with the first model (Run 1) using only comparable field data to assess model improvement. *Note, to quantitatively compare the accuracy of modeled products from model Run 1 against Run 2, the field data had to be subset to account for the different grouping categories applied in each run of the model	3-category model (low- and mid-bench floodplain units as well as shallow water wetlands removed from field data and model 2 prediction prior to calculation) improved by 13% across the entire FWCP Peace region, by 3% in the Finlay sub-basin and by 14% in the Parsnip sub-basin.
			10-category model (shallow water wetlands and null values were removed prior to calculation) improved by 7% across the entire FWCP Peace region and by 9% within the Parsnip sub-basin but decreased in the Finlay sub-basin by 4%.
	Reliability Assessment	Provide options for examining the reliability of the model predictions	Produced dataset that indicates number of field points per 1:20,000 scale map sheet. Higher confidence can be applied to predictions in areas with sufficient field data coverage.
			Produced dataset that indicates number of training points per 1:20,000 scale map sheet. Higher confidence can be applied to predictions where higher densities of training points are present
			Model predictions can be compared and evaluated against supplementary datasets within Williston Wetland Explorer Tool (WWET)
Wetland Prediction Summaries	*Wetland Prediction Summaries*	Summarize wetland prediction polygons to establish trends	The 3-category model predicted 366,381 hectares of wetlands (5% of the total FWCP Peace area) 255,244 hectares of water (3.5% of the total FWCP Peace area) and 6,659,503 hectares of upland (91.5% of the total FWCP Peace area).
			Wetlands are most prevalent in Parsnip sub-basin (8% of the total area) and least common in the Peace (2% of the total area)
			Large wetlands are primarily located in the Parsnip sub-basin
Disturbance Analysis Case Study	*Wetlands*	Aggregate wetlands into unique groupings based on hydrological connectivity and proximity	A number of wetlands are likely hydrologically connected and aggregate into wetland complexes
			Largest wetland complexes occurred in the Parsnip sub-basin
	Disturbance	Calculate disturbance densities for multiple example disturbance types	Disturbance count and density varied by watershed group
			Watershed groups in the Parsnip had the highest average count and density of disturbance

Data and Information Access	*Data and Information Access*	Develop a web mapping application to facilitate stakeholder engagement	A new ArcGIS Online web mapping application was created (Williston Wetland Explorer Tool – WWET) that includes project information, user guide and web map exploration tool
		Provide script and data accessibility through open source or publicly available platforms	GitHub repository generated to house example data and scripts
			Data and report are made available through the TEI Data Distribution Site and EcoCat

RANDOM FOREST MODEL PREDICTION

Several iterations of the Random Forest model were run to optimize categorization and establish the best performing model. Many different combinations of codes were attempted (Table 3). These combinations included a 46-category, 10-category, 4-category, and three variations of the 3-category grouping. Each iteration was evaluated for internal estimates of model error (OOB error) and validated using external data (field data). The more ecosystem categories the model attempted to predict the greater the OOB error. As categories were reduced by amalgamating the groupings into a smaller number of overall categories (e.g., from 46 categories to 10 or 3), the model performance improved (Table 5).

The 4-category prediction included flood units as a separate category (i.e., upland, water, wetland and flood). However, the 4-category option performed poorly, likely due to the lack of variability across predictor variables in flood areas that would allow for the algorithm to decipher flood units as separate entities. Additionally, limited training data for flood classes may have skewed model predictions away from these categories. Consequently, the 4-category grouping was not selected due to its poor accuracy metrics, however, should input data improve the performance of this category grouping may also improve.

The highest performing 3-category model grouped high-bench floodplain units with the upland category and low- and mid-bench floodplain units with the wetland category. This also effectively added riparian ecosystems to the wetland class of the 3-category prediction. The 10-category model did better than the 46-category model but did not outperform the coarsest grouping of 3-categories. The 46-category model, representing the most detailed classification, did not perform well enough to warrant further discussion in this report. Additionally, validation at the 46-category level was hindered by the limited number of field observations collected at the required degree of detail.

Table 5: Random Forest model results for each of the key iteration categories. The 3-category prediction includes high-bench floodplain units as upland features, while low and mid-bench floodplain units are grouped with wetlands.

Random Forest Model Results					
Categories	**Total field points**	**Total Null Values**	**# correct**	**% accuracy**	**% OOB error**
Upland/Water/Wetland (3 categories)	1100	0	954	86.7	7
Upland/Water/Class (10 categories)	1100	24	568	52.8	11.7
All (46 categories)	1100	762	58	17.2	32.95

MODEL ACCURACY ASSESSMENT

Model accuracy metrics were used to evaluate the overall model performance and accuracy. Two key metrics were used to measure model results: (1) an internal estimate of model error (out-of-bag [OOB] error) and (2) model validation using field sampled data. Supporting assessment products included reporting outputs from the Random Forest model.

The highest performing model was the 3-category grouping, followed by the 10-category and lastly the 46-category (Table 5). Overall, the coarser prediction (3-category) performed the best with only 7% OOB error. The 10-category had a 21% improvement in OOB error over the 46-category results but remained behind the 3-category model by 5%.

Various reporting outputs from the Random Forest model offer additional ways to evaluate each model. One useful product offered is a confusion matrix which allows for the comparison between predicted and observed samples. Key values from the confusion matrix include the values of omission (false negatives) and commission (false positives). Omission occurs when a class call was provided in the training data but was not predicted for. Commission, on the other hand, occurs when a class call was provided in the training data but was predicted incorrectly (e.g., a bog was incorrectly predicted as a fen).

For the 10-category result, errors of omission (false positives) were highest for flood units (Fh, Fm, and Fl) and marsh (Wm) units (Table 6). Errors of commission (false negatives) were greatest for marshes (Wm), while swamp (Ws), bog (Wb), and fen (Wf) also showed large commission errors. The high errors associated with swamps and marshes, along with floodplain units, indicates the model's difficulty in separating these classes based on the input data provided. Treed wetland units, such as swamps, may have increased confusion with other treed units like upland. While marshes may often be found in transitional zones around open bodies of water. Floodplains, particularly high-bench flood units, are often heavily treed and therefore, it is not unexpected that confusion may exist between floodplains and other treed units. Given the resolution of the predictor variables (25 meters), it is anticipated that there may be errors in transitional zones and where multiple wetland types form a mosaic within a single wetland complex.

The 3-category prediction demonstrated improved omission and commission errors over the 10-category prediction, with good separability between units (Table 7). Of the three units, 5.2% of upland, 13.7% of water, and 15% of wetlands were classified incorrectly.

Table 6: Random Forest confusion matrix of Upland, Water and Wetland/Riparian (10-category 'Class' level) model. Errors of omission (false negatives) and errors of commission (false positives) between predicted and observed values are provided.

		observed											
		Fh	*Fl*	*Fm*	*T*	*W*	*Wb*	*Wf*	*Wm*	*Ws*	*Ww*	*total*	Commission
predicted	*Fh*	1	0	0	0	0	0	0	0	0	0	1	0.00
	Fl	0	12	0	1	1	1	1	1	0	0	17	0.29
	Fm	0	0	0	0	0	0	0	0	0	0	0	NaN
	T	23	28	27	10512	59	155	216	21	388	18	11397	0.08
	W	3	1	0	9	294	4	8	10	7	38	374	0.21
	Wb	0	6	2	14	4	275	87	2	48	3	441	0.38
	Wf	2	30	1	66	16	155	631	10	43	20	974	0.35
	Wm	0	0	0	0	0	0	1	0	0	0	1	1.00
	Ws	0	2	1	23	10	20	12	5	98	2	173	0.43
	Ww	0	0	1	0	9	1	10	1	6	57	85	0.33
	total	29	79	32	10625	393	611	966	50	540	138	13463	
	Omission	0.97	0.85	1.00	0.01	0.25	0.55	0.35	1.00	0.82	0.59		

Table 7: Random Forest confusion matrix of internal predicted and observed units for the Upland (T) – Water (W) – Wetland/Riparian (WL) categories. Errors of omission (false negatives) and errors of commission (false positives) between predicted and observed values are provided.

			observed			
		T	*Water*	*WL*	*total*	Commission
predicted	*T*	10414	35	540	10989	0.052
	Water	7	265	35	307	0.137
	WL	233	93	1841	2167	0.150
	total	10654	393	2416	13463	
	Omission	0.023	0.326	0.238		

When compared against field observations, the 3-category prediction has a field validated accuracy of 86.7% representing good agreement between field data and modeled results (Table 5). This metric indicates that the model is 86.7% accurate in the category call (upland, water or wetland) when compared to field observations.

The 10-category model has a 36% improvement in validated accuracy over the 46-category results, with a total of 52.8% accuracy (Table 5). Of particular interest is the large difference between the 52.8% field accuracy and the 88.3% OOB accuracy metrics for this model. The large discrepancy (35.5%) between these two accuracy measurements indicates that the 10-category model did well at predicting training data (known data) but failed to correctly predict field data (unknown data) approximately half the time.

To evaluate how well the model predicted each category unit, the field observation call was compared to the modelled prediction. The results were graphed to display the percentage of all training points that were incorrectly or correctly matched with the model prediction (Figure 5 and Figure 6).

Run 2 results showed an overall improvement in accuracy for upland and water in the 10-category prediction (Figure 5). These results demonstrate the 10-category model successfully predicted upland and water more often than was incorrect in these categories. Bog, fen, and open water wetland units were generally similar in the number of correct vs incorrect matchups, while the floodplain units, marshes and swamps were categories in which the model struggled to accurately predict.

The 3-category matchups generally coincide with the field data samples approximately half the time. The balance of error and correct classification percentages are more equal in the 3-category prediction compared to the 10-category results. This trend is relatively consistent across the three categories (upland, water, wetland).

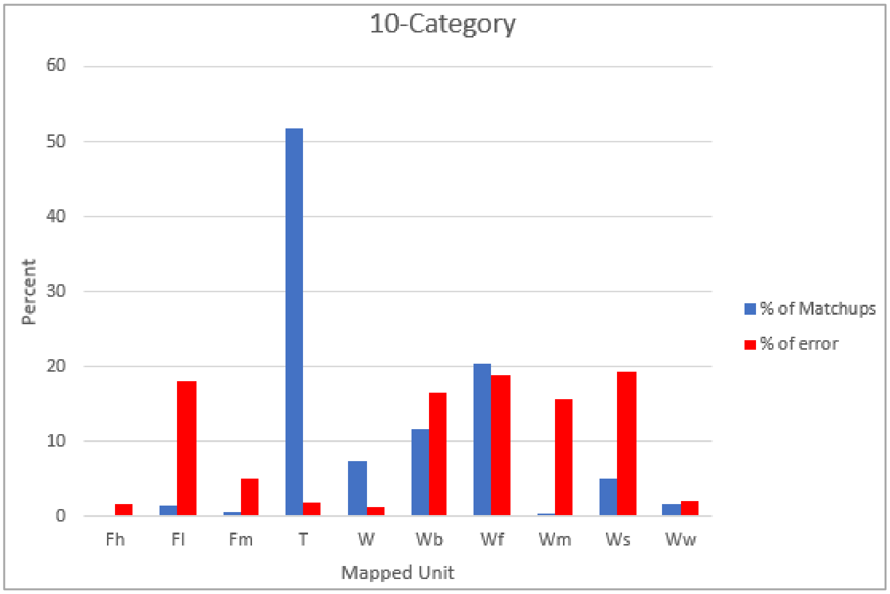


Figure 5: Percent error and correct matchups between 10-category prediction and field observations.

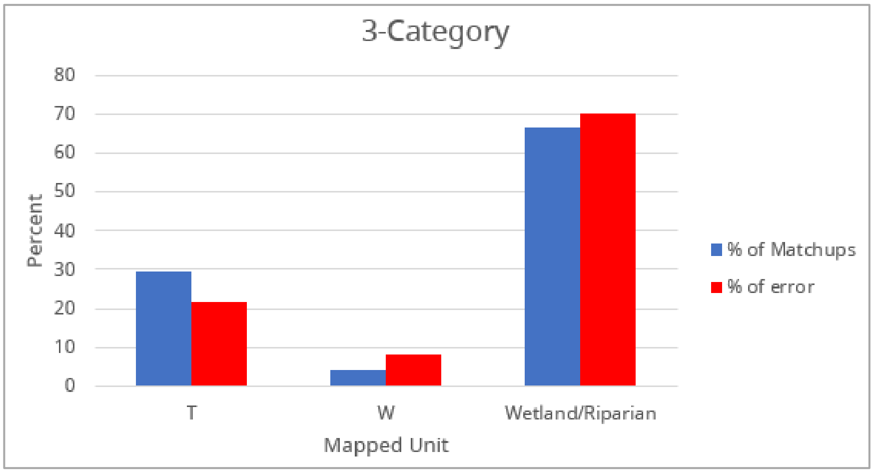


Figure 6: Percent error and correct matchups between 3-category prediction and field observations.

TRAINING AND FIELD VALIDATION COMPARISON

Training points were targeted for field validation and used to evaluate how accurately mappers interpreted wetland calls for 3 and 10 categories. There are 13,463 training points through out the FWCP Peace region and of these, 392 were visited in the field. The remainder of the 1,100 field observations (708) were opportunistically selected during field work (Figure 7). For the 3-category grouping, 392 training points were validated in the field. Two of these sites did not have the level detail required to validate the 10-category grouping. The 10-category grouping was, thus, validated using 390 points.

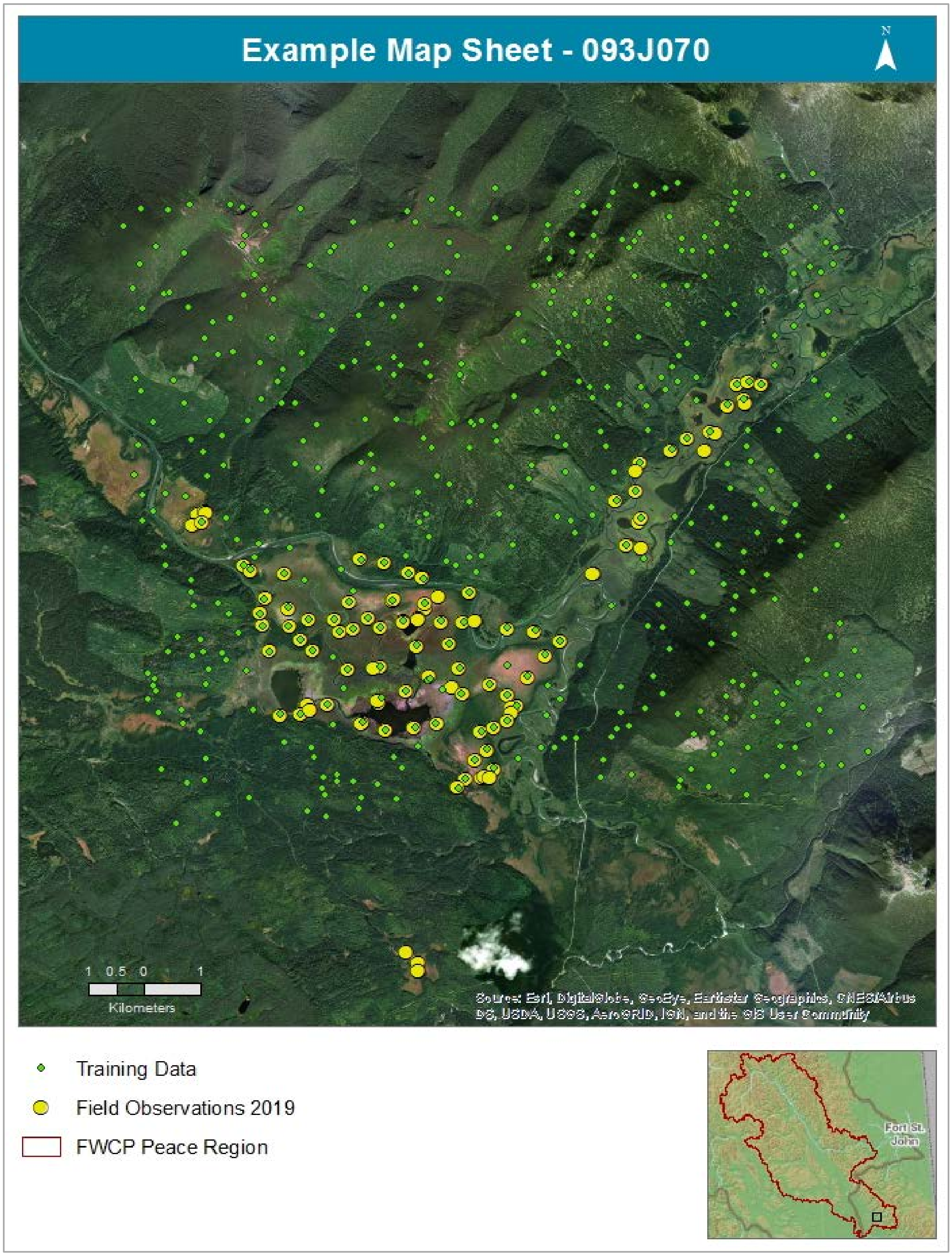
As established in the first model run, mappers responsible for generating training data were most confident in interpreting water categories (Filatow *et al.*, 2018). In model Run 2, water scored 89% correct for both the 3 and 10-category predictions. The upland class in both the 3-category (81%) and 10-category (91%) groupings were also interpreted reliably by mappers. Wetlands scored 98% correct for the 3-category model (Table 8). Of the 10-category wetland classes, bogs sored 86%, however all other wetland categories scored below 65% (Table 9). Shallow water wetland, floodplain units, and marshes were poorly assigned by mappers.

Table 8: Training data validation using field observations to identify the total number (and percentages) of correct calls versus incorrect calls for the 3-category grouping.

Training data compared to Field observations (3-Category)			
	Training Data Correct	**Training Data Incorrect**	**Total Count**
Upland (count)	169	39	208
Upland (%)	81%	19%	-
Water (count)	17	2	19
Water (%)	89%	11%	-
Wetland (count)	161	4	165
Wetland (%)	98%	2%	-
TOTAL (count)	**347**	**45**	**392**

Table 9: Training data validation using field observations to identify the total number (and percentages) of correct calls versus incorrect calls for the 10-category grouping.

Training data compared to Field observations (10-Category)			
	Training Data Correct	**Training Data Incorrect**	**Total Count**
Upland (count)	167	16	183
Upland (%)	91%	9%	-
Water (count)	17	2	19
Water (%)	89%	11%	-
High-bench Floodplain (count)	-	-	-
High-bench Floodplain (%)	-	-	-
Mid-bench Floodplain (count)	2	5	7
Mid-bench Floodplain (%)	29%	71%	-
Low-bench Floodplain (count)	3	16	19
Low-bench Floodplain (%)	16%	84%	-
Bog (count)	36	6	42
Bog (%)	86%	14%	-
Fen (count)	31	18	49
Fen (%)	63%	37%	-
Marsh (count)	4	18	22
Marsh (%)	18%	82%	-
Swamp (count)	15	30	45
Swamp (%)	33%	67%	-
Shallow Water Wetland (count)	2	2	4
Shallow Water Wetland (%)	50%	50%	-
TOTAL (count)	**277**	**113**	**390**



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

Figure 7: Training data points and 2019 field collected data points for the map sheet 093J070.

MODEL IMPROVEMENT AND RELIABILITY ASSESSMENT

MODEL IMPROVEMENT

Model results from Run 1 were compared with those produced in Run 2. Based on accuracy metrics alone, Run 2 models performed similarly to those models produced in Run 1. Due to the performance of the models, there was no need to reclassify the raw model output to more general categories such as the technique applied in Run 1. Additionally, the models produced in Run 2 exhibited marked improvements in areas previously overestimated in Run 1.

One key observation from Run 1 included the considerable over prediction of wetlands in cutblocks and cleared land, especially in low lying areas (Filatow *et al.*, 2018). Regions impacted by timber harvest, particularly in the Parsnip sub-basin, were often predicted as bogs. This model confusion was significant and covered large connected areas, often in close proximity to legitimate wetland complexes. With the addition of 1000 new training points, Run 2 results showed clear improvements in these areas, especially within the Parsnip sub-basin (Figure 8). Where previously cutblocks were classified as bogs, they are generally assigned correctly as upland or fen in both the 3 and 10-category models (Figure 8 and Figure 9).

The example map sheet, 093J070, was initially modelled in Run 1 as dominated by bogs and containing a scatterings of low-bench floodplain units along the river system (Figure 10). The first model exhibits a homogenous representation of the wetland complex and does not reflect the mosaic of wetland types that were observed during field sampling. In Run 2, the overestimation of bogs was replaced by the addition of fens around water features as well as swamps throughout treed sections. A notable limitation was the inability of either model to accurately predict marshes. Marsh categories were prevalent throughout the map sheet, but were not detected, likely due to their smaller footprint (less than 25x25 meters) and potential confusion with water. Additionally, although swamps were predicted, it is reasonable that there would be confusion between swamps and floodplain units surrounding the river network.

To quantify improvement between model runs, the number correctly predicted field points were calculated for each model. In order to maintain accurate comparison between each run of the model (Run 1 vs. Run 2), categories that were changed or modified between runs were removed prior to calculation. In the 3- category prediction mid and low-bench floodplains and shallow water wetlands were removed. Additionally, in the 10-category prediction, shallow water wetlands were removed along with field point locations that did not collect observations at the class level of detail (i.e., null values).

For the 3-category prediction, the results indicated an overall improvement of 13% and when assessed on a sub-basin level, the Parsnip and Finlay exhibited a 14% and 3% improvement respectively (Table 10). The 10-category prediction showed an overall improvement of 7% when comparing predictions between model runs (Table 11). The Parsnip sub-basin also improved by 9% and the Finlay showed a slight deterioration by 4%.

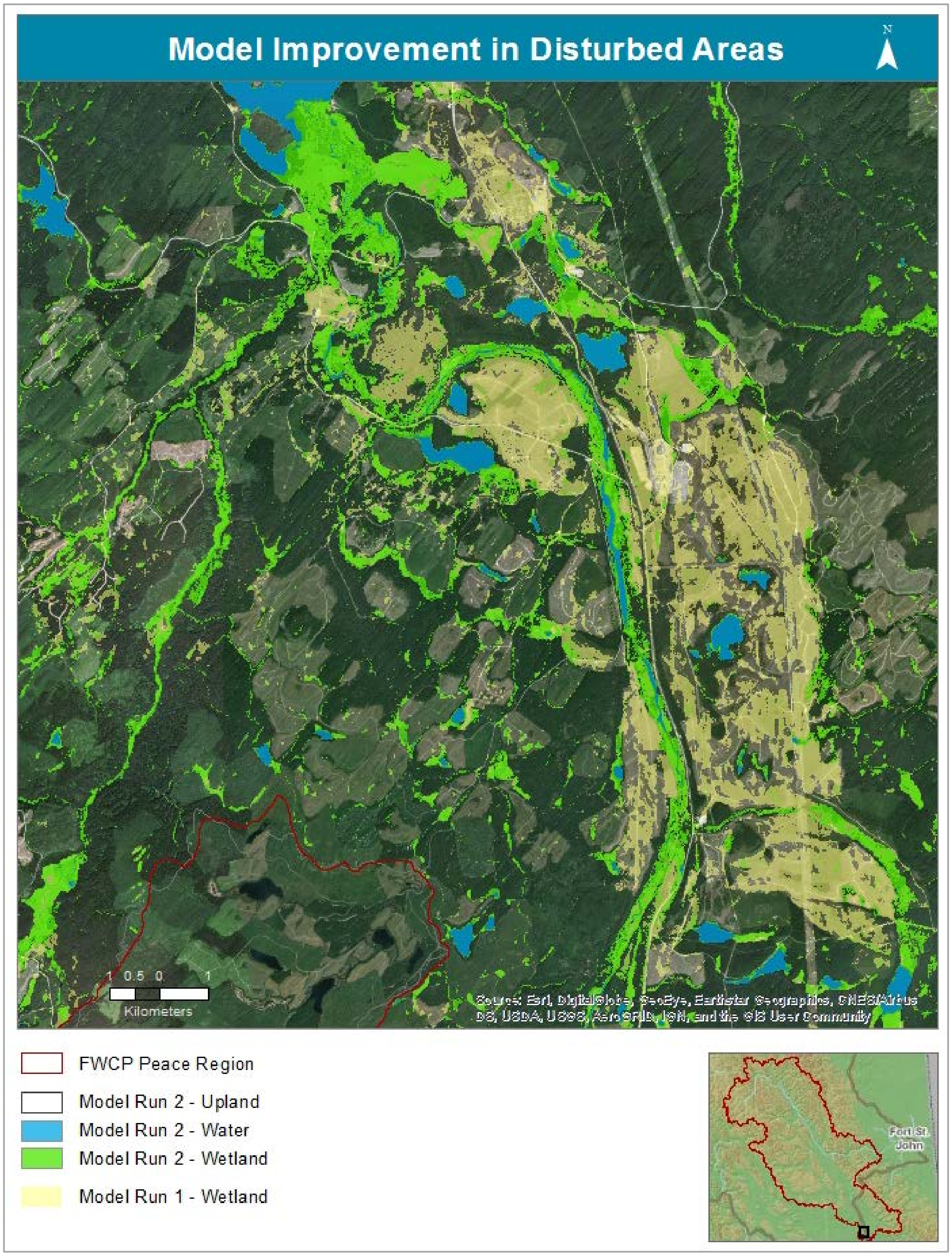


Figure 8: Illustration of the second run model (3-Category) improvement for disturbed areas. The first run of the model over predicted wetlands within cutblocks and disturbed environments, while the second model did not.

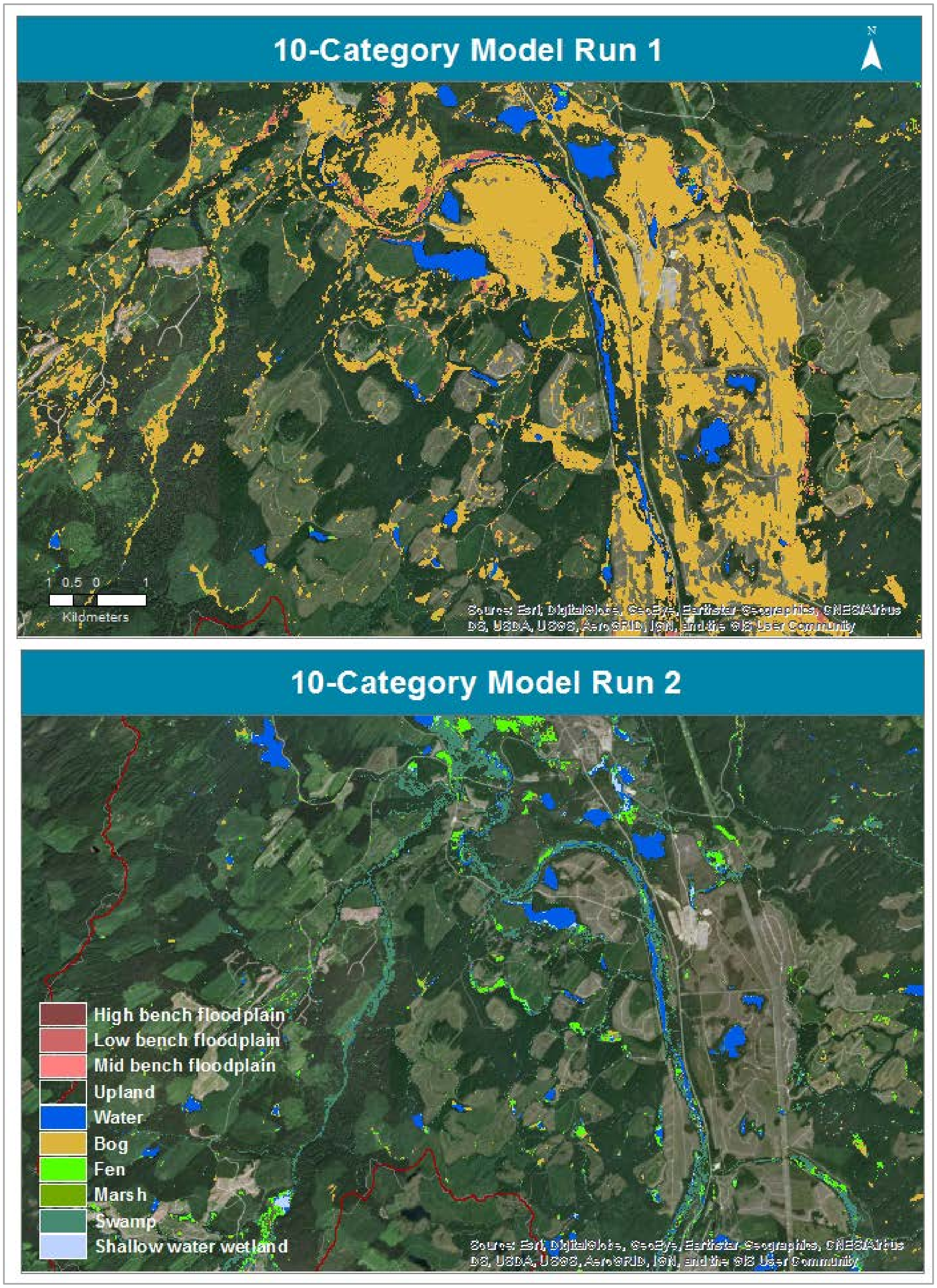


Figure 9: Comparison of model predictions (10-Category) between model run 1 and 2.

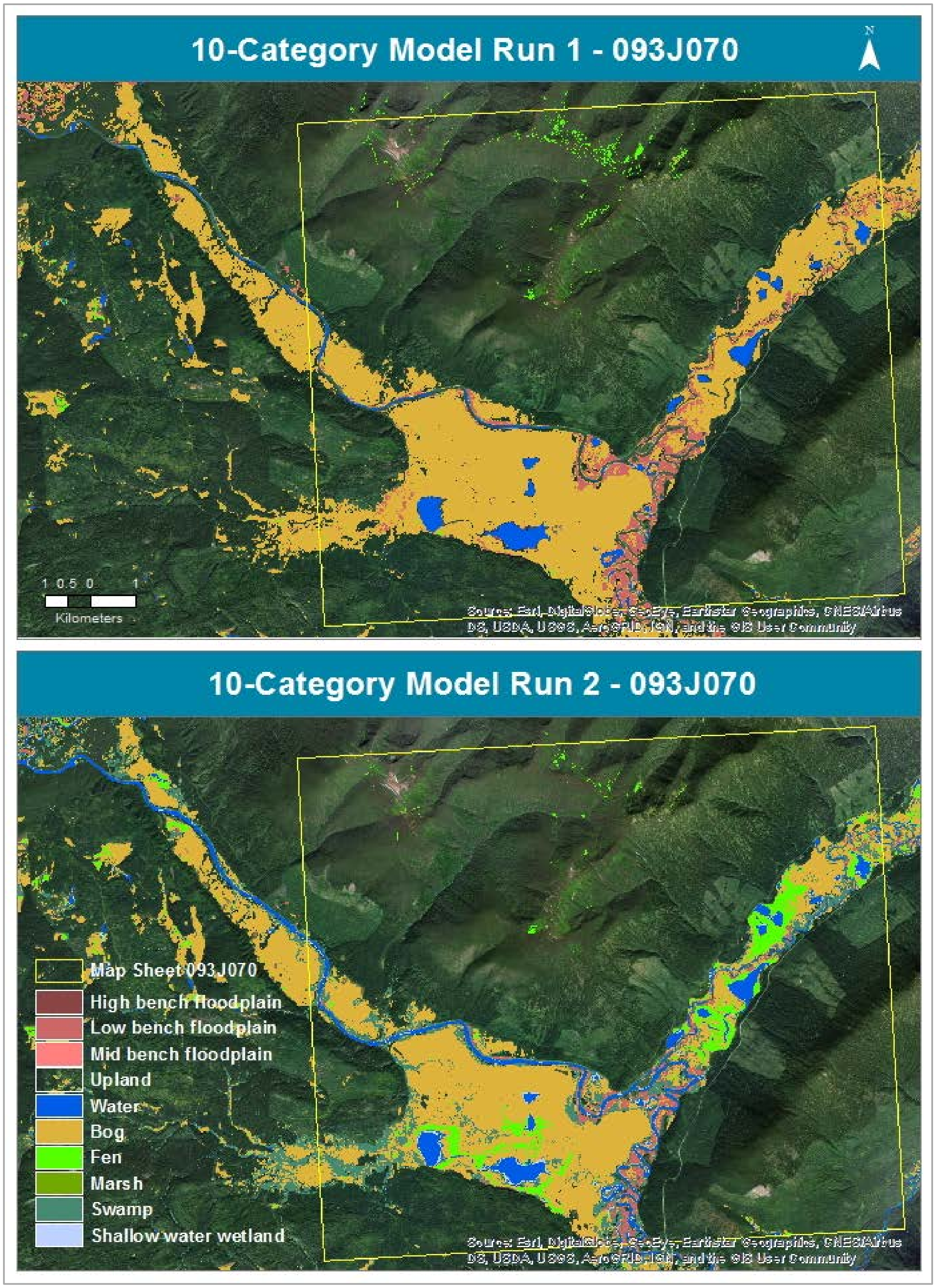


Figure 10: Comparison of model predictions (10-Category) between model run 1 and 2 within example map sheet 093J070.

Table 10: Quantification of model improvement in the 3-category prediction. As some of the classification groupings changed between model runs, low and mid-bench floodplains as well as shallow water wetland points were removed prior to calculations.

Model Improvement (3-category)					
Category	**Region**	**Number of Field points**		**Number of Correctly Predicted Points**	**Percentage of Correctly Predicted Points**
3-Category	**FWCP Peace**	953	Run 1	724	76%
			Run 2	844	89%
			Difference	**120**	**13%**
	Finlay	146	Run 1	125	86%
			Run 2	129	88%
			Difference	**4**	**3%**
	Parsnip	807	Run 1	599	74%
			Run 2	715	89%
			Difference	**116**	**14%**

Table 11: Quantification of model improvement in the 10-category prediction. As some of the classification groupings changed between model runs, shallow water wetland points were removed prior to calculations.

Model Improvement (10-category)					
Category	**Region**	**Number of Field points**		**Number of Correctly Predicted Points**	**Percentage of Correctly Predicted Points**
10-Category	**FWCP Peace**	1038	Run 1	490	47%
			Run 2	559	54%
			Difference	**69**	**7%**
	Finlay	170	Run 1	88	52%
			Run 2	82	48%
			Difference	**-6**	**-4%**
	Parsnip	868	Run 1	402	46%
			Run 2	477	55%
			Difference	**75**	**9%**

RELIABILITY ASSESSMENT

To aid in the evaluation of model reliability, maps of field data and training data densities were produced to inform decision making.

To assess patterns in the distribution of field data throughout the FWCP Peace region and to highlight areas of high and low sampling effort, the number of field points per 1:20,000 scale map sheet was mapped (Figure 11). The map demonstrates the lack of field data for the Dinosaur and Peace sub-basins, as well as most of the Finlay. Of the field data collected in the Finlay, the majority of points are located in low-lying areas surrounding Williston Lake. When compared to all other sub-basins, the Parsnip has the highest concentration of field data.

High density training points correspond to map sheets where intensive desktop mapping was completed in an effort to capture the environmental variability across diverse areas of the FWCP Peace region (Figure 12). The models were built off the training data and therefore, in regions where training points are limited, the model may be less accurate.

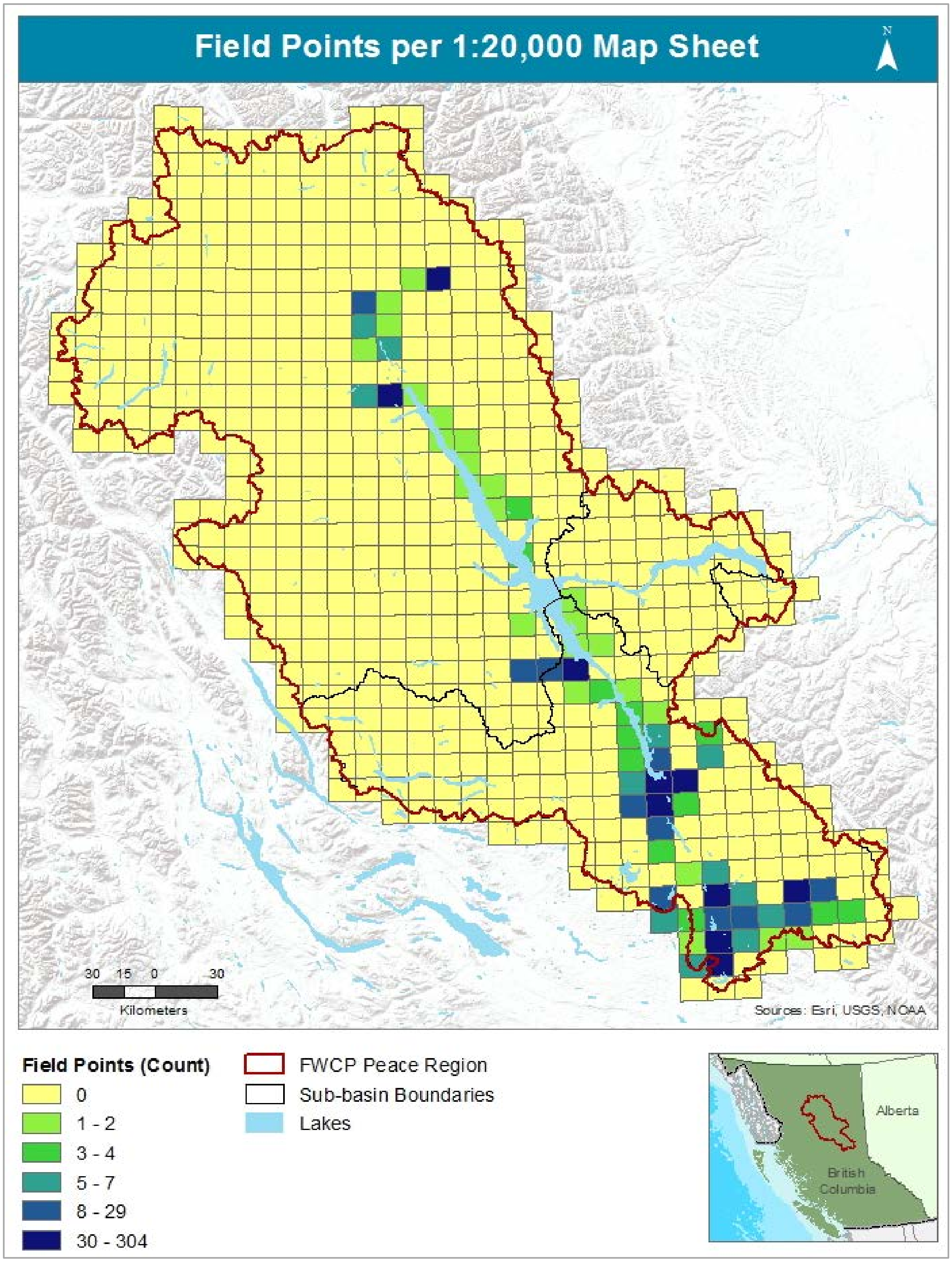


Figure 11: Field observation points per 1:20,000 map sheet.

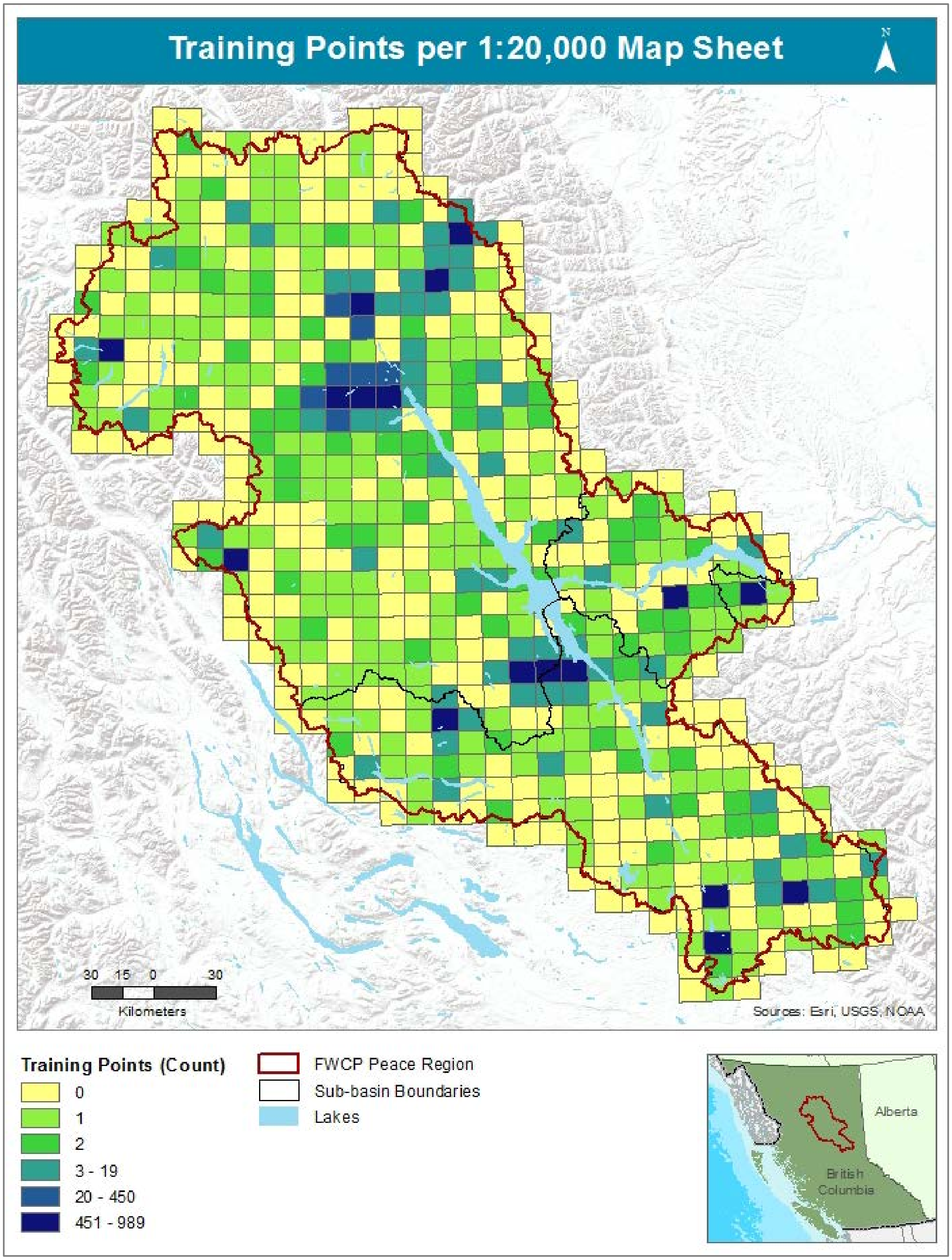


Figure 12: Training points per 1:20,000 map sheet.

WETLAND PREDICTION SUMMARIES

The FWCP Peace region spans 7.2 million hectares, with the Finlay sub-basin covering 63% of the total area and the Parsnip, Peace, Dinosaur sub-basins covering 28%, 8%, and 1% respectively. The model predicted 366,381 hectares of wetlands (5% of the total area) throughout the FWCP Peace region, while water accounted for 255,244 hectares (3.5% of the total area). Upland covers the greatest amount of area with 6,659,503 hectares predicted (91.5% of the total area).

When assessed on a sub-basin level, wetlands are most prevalent in Parsnip sub-basin with 8% covering the basin area, while wetlands are least common in the Peace and only account for 2% of the total area (Table 12). Upland is dominant across all sub-basins and range from 88% to 96.4% of the basin area. Many large lakes are contained within, and account for, the relatively large proportions of water predicted in the Peace, Parsnip and Finlay sub-basins.

As consistent with results from the first model run, wetlands were predicted to occur primarily in low-lying and flat areas of the Parsnip sub-basin. High wetland density within 1:20,000 map sheets was present throughout the Parsnip (Figure 13). Within the Finlay sub-basin, wetland density was greatest in areas directly surrounding Williston Reservoir, and areas of great topographic variability in the north east of the sub-basin, such as to the north and south of Tatlatui Park (Figure 13). Locations of high wetland density also generally correspond with areas containing large wetlands (Figure 14).

Table 12: The 3-category coverage and composition of Wetland, Upland, and Water units within the FWCP Peace study region. Extent boundary follows edge of model output. Table values calculated from polygons generated from the 3-category model prediction.

Category Coverage (Ha)					
Sub-basin	**Finlay**	**Parsnip**	**Peace**	**Dinosaur**	**Other**
Total Sub-basin Area (Ha)	**4,593,518.99**	**2,024,043.79**	**588,967.08**	**68,045.42**	**6,552.60**
Wetland Area (Ha)	189,546.17	165,282.00	9,590.55	1,528.80	433.52
Wetland Percent (%) – Sub-basin	4	8	2	2.2	6.5
Upland Area (Ha)	4,267,485.58	1,776,219.71	544,136.44	65,572.37	6,089.19
Upland Percent (%) – Sub-basin	93	88	92	96.4	93
Water Area (Ha)	136,487.24	82,542.08	35,240.09	944.25	29.89
Water Percent (%) – Sub-basin	3	4	6	1.4	0.5

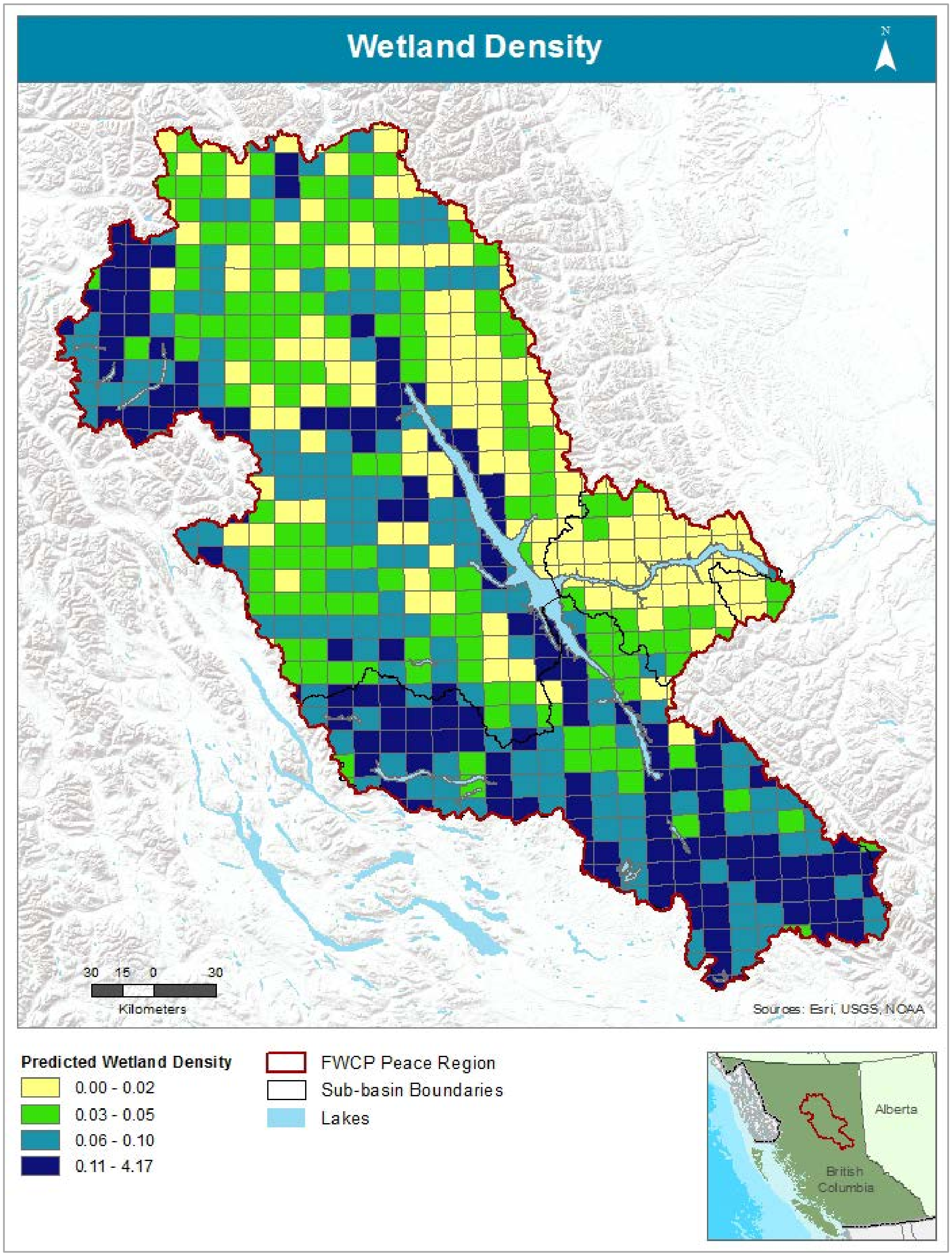


Figure 13: Predicted wetland density by 1:20,000 map sheet. Wetland density was calculated from polygons created from the 3-Category model prediction. Prior to calculation, lakes greater than 1000 ha were removed from the map sheet area.

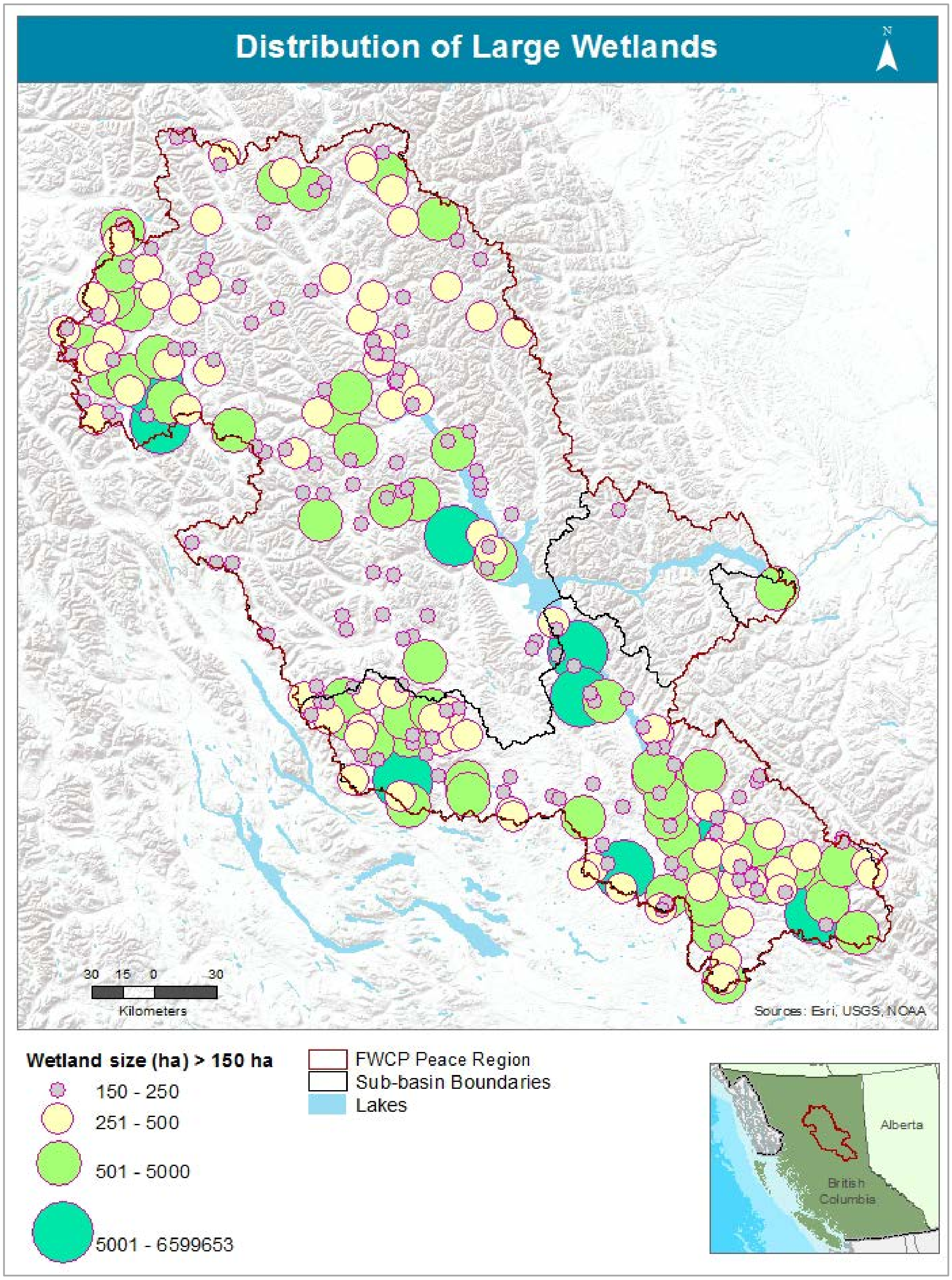


Figure 14: Distribution of large (greater than or equal to 150 ha) wetlands throughout the FWCP Peace region. Size calculations were based on the 3-category model results converted to polygons. Points represent centroid of wetland polygons.

WETLANDS

The model output polygons were aggregated by analysis watershed, surface hydrological connectivity, and distance. This method grouped wetlands into 8,922 different wetland complexes that also include hydrologically disconnected single wetlands (Figure 15). Average wetland complex area was approximately 0.8 km². Maximum wetland complex area was approximately 127 km², and minimum wetland complex area was approximately 0.0004 km^2. Note that presented content in Table 13, Table 14, Table 15, and Table 16 are selection of watershed groups that contain more than 100 wetland complexes identified within their boundaries.

The count and size distribution of wetland complexes varied according to location. The largest wetland complexes occurred in the Parsnip sub-basin (Figure 16). However, the distribution of wetland size was not necessarily associated with wetland complex count within watershed groups (Table 13).

Comparing full watershed groups that go from head waters to lowlands, like the Crooked River watershed group in the Parsnip sub-basin to small headwater fragments of watershed groups like the Murray River watershed group within the Dinosaur sub-basin is not appropriate. The Murray River and the Upper Peace River watershed groups were truncated by administrative boundaries, and results from these watershed groups underestimate disturbance impacts and overestimate wetland complex number as they are partially complete. Reporting in the Upper Peace River and Murray River watershed groups also have erroneous and overestimated results due to improperly aggregated wetland complexes (Figure 16). For example, just under 34% of the watershed complexes within the entire FWCP Peace region are located within the Murray River and Upper Peace watershed group and this is an error (Table 13; Figure 16).

After accounting for data artifacts, the Parsnip River watershed group was observed to have the highest cumulative area of wetland complexes, as well as the largest average area of wetland complexes (Table 13; Figure 16). However, high total wetland area and average wetland complex area did not associate with high disturbance counts, both within wetland complexes and within the 1 km upslope contributing area (buffer). This is likely due to the high number and area of wetland complexes reducing both the count and density of disturbances that may impact wetland function and supports the need for further work into the nature of wetland hydrological connectivity.

Buffers were calculated within 1 km in all directions of a wetland complex with the assumption that they would capture threat impacts in the hydrologically-connected contributing area. The results indicate approximately 19 km^2 of upland area is associated with every 1 km^2 of wetland complex.

These assumptions likely significantly overestimate the contributing area of upland terrain because they do not account for overlap between wetland buffers (double-counting) or for variable topography that includes 1) topographical breaks such as height of land, 2) slope and aspect changes that impact drainage patterns, and 3) hydrological features that intersect the buffer. As well, wetland spatial patterns influence the amount of contributing upland area, where tightly clustered wetlands in large complexes tend to include proportionally less contributing upland area than complexes with numerous, widely dispersed wetlands or single, unconnected, small wetlands.

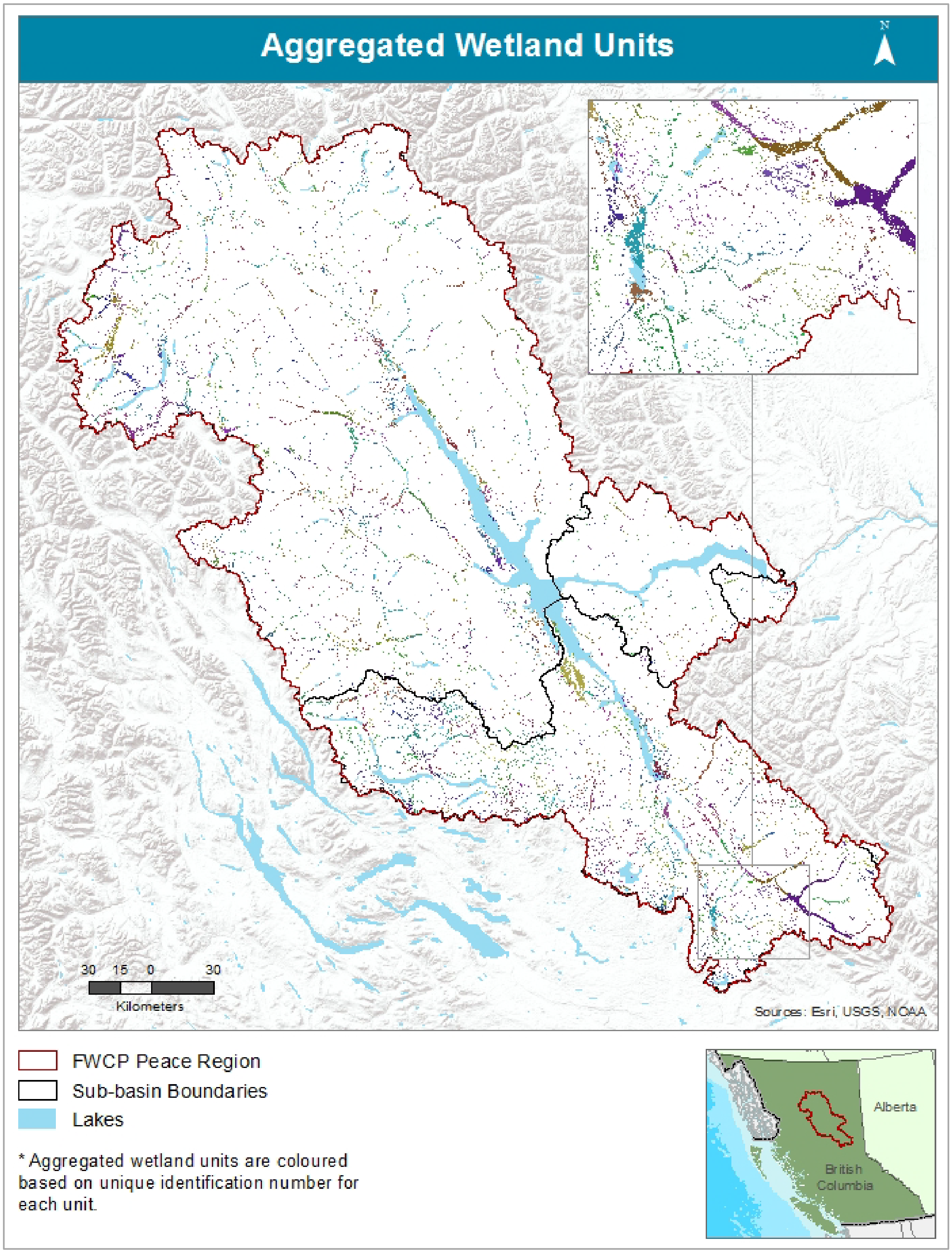


Figure 15: Wetland polygons (3-category model prediction) grouped into unique units based on proximity and hydrological connectivity.

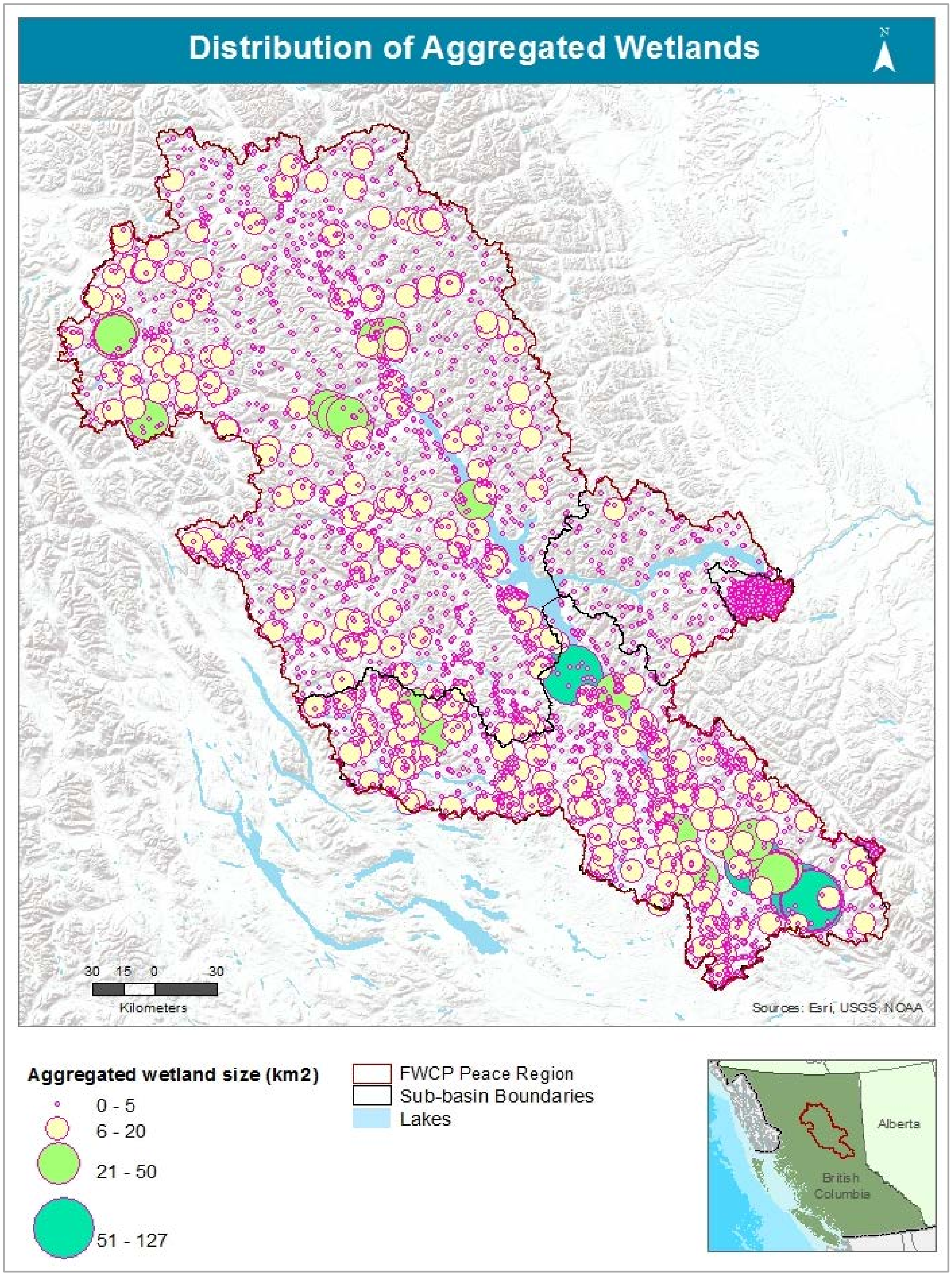


Figure 16: Distribution of aggregated wetland size (km^2) throughout the FWCP Peace region. Points represent centroid of wetland aggregation polygons.

Table 13: Count and area calculations for wetland complexes and wetland buffers by watershed group. Very small portions of the Murray River and Upper Peace River fall within the FWCP Peace region boundary. Results for these 2 Watersheds are incomplete.

Watershed Group	Number of wetland complexes	Total Area of Wetland Complexes (km²)	Average Area of Wetland Complexes (km²)	Standard Deviation of Average Wetland Area (km²)	Total Area (km²) of Wetland Buffers (km²)	Average Area of Wetland Buffers (km²)	Standard Deviation of Average Wetland Buffer Area (km²)
Carp Lake	172	265.31	1.54	4.04	4,111.97	23.91	48.05
Crooked River	387	290.72	0.75	2.75	5,348.43	13.82	34.43
Finlay Arm	733	415.91	0.57	2.26	12,028.76	16.41	47.96
Finlay River	295	302.47	1.03	3.44	7,184.72	24.35	45.47
Firesteel River	187	668.33	3.57	8.71	7,637.40	40.84	62.68
Fox River	171	252.09	1.47	3.23	5,268.08	30.81	43.89
Ingenika River	155	339.53	2.19	5.61	6,453.76	41.64	54.99
Lower Omineca River	313	168.57	0.54	1.43	5,969.26	19.07	35.89
Mesilinka River	230	227.19	0.99	2.73	5,049.05	21.95	39.74
~~Murray River~~	~~583~~	~~4.00~~	~~0.01~~	~~0.03~~	~~1,975.34~~	~~3.39~~	~~0.44~~
Nation River	824	765.87	0.93	3.07	15,102.14	18.33	41.72
Parsnip Arm	524	285.24	0.54	3.77	7,219.1	13.78	43.67
Parsnip River	368	2137.1	5.81	19.6	15,143.71	41.15	89.56
Peace Arm	216	103.88	0.48	0.95	7,271.6	33.66	58.27
Toodoggone River	198	288.36	1.46	3.18	6,499.12	32.82	48.06
Upper Omineca River	244	234.88	0.96	2.4	5,959.39	24.42	45.27
~~Upper Peace River~~	~~2,993~~	~~15.62~~	~~0.01~~	~~0.06~~	~~10,159.57~~	~~3.39~~	~~2.39~~
Grand Total	**8,922**	**6,868.02**	**0.77**	**4.85**	**132,717.15**	**14.88**	**39.6**

DISTURBANCE

Although analysis output determined cumulative count of disturbances for each wetland complex, these results are aggregated by watershed group to provide information at a broader level. A total of 9 spatial categories of threats were analyzed, but only roads, cut blocks, and mountain pine beetle infestation are presented in the watershed group results. Interpretation of disturbance density in both wetlands and the surrounding area buffers are summarized below (wetlands: Figure 19 and Figure 22; buffers: Figure 20 and Figure 23). These results are displayed based on quantiles and not on ecologically meaningful thresholds.

Disturbance count and density varied by watershed groups (Table 14, Table 15, and Table 16), by wetland complex (Figure 17), and by associated upland contributing area within 1 km of the wetland complexes (Figure 18). In general, watershed groups in the Parsnip sub-basin had higher average numbers of threats in wetland complexes and surrounding upland area than watershed groups in the Finlay sub-basin.

For most disturbance layers, the average number of threats in each wetland complex was lower than the average number of threats within the associated upland contributing area, suggesting that the area adjacent to wetlands was more likely to be directly impacted by disturbance. This trend is apparent within the timber harvest data (Table 14, Figure 22 and Figure 23), where approximately twice the area of timber harvest occurred in wetland upslope contributing area despite the significantly greater area of the wetland upslope contributing area. This is intuitive, as most wetlands have low tree productivity or do not have trees.

The inverse pattern of disturbance density occurs for roads. Road density was generally greater in wetland complexes (Figure 19) than in the surrounding upslope contributing area (Figure 20; Table 15). Road density patterns also reflected general disturbance patterns, where higher road densities were associated with watershed groups in the Parsnip sub-basin, and relatively lower road densities were associated with watershed groups in the Finlay sub-basin. The Murray River watershed group had no roads associated with wetlands or their surrounding upland area. However, this is a data artefact as the Murray River watershed group is truncated by the FWCP-Peace region administrative boundary, and only accounts for a fraction of the actual watershed group (Figure 28).

Patterns of timber harvest data mirrored patterns of road incursion into wetland complexes and their surrounding upland area. Although timber harvest was pervasive across nearly all wetland complexes and associated upland areas, watershed groups in the northern area of the Finlay sub-basin typically had a lower density of cut blocks, and fewer wetland complexes and upland areas affected by cut blocks (Table 14). In contrast, the Carp Lake and Crooked River watershed groups had cut block-related disturbance present in almost all of the upland areas associated with wetland complexes.

Mountain pine beetle (MPB) disturbances had high density and presence in most wetlands and associated upland areas (Table 16). The average area and percent impacted by MPB infestation in wetland complexes and upslope contributing areas is similar. The Firesteel River and Parsnip River watershed groups had the fewest recorded MPB infestations mapped within wetland complexes, but these included over half of the member wetland complexes. Although the Parsnip River watershed group had the fewest recorded MPB infestations within the upland area associated with wetland complexes, MPB infestation was documented in 59% of all associated upland buffers.

Table 14: Calculations of cutblock density and presence for wetland complexes and wetland buffers by watershed group. Very small portions of the Murray River and Upper Peace River fall within the FWCP Peace region boundary. Results for these 2 Watersheds are incomplete.

Watershed Group	Average Number of Threats in Each Wetland Complex	Average Number of Threats in Each Wetland Buffer	Average Cutblock Density per Wetland Complex	Percent Wetland Complexes with Cutblocks Present	Average Cutblock Density per Wetland Buffer	Percent Wetland Buffers with Cutblocks Present
Carp Lake	2.48	3.6	0.37	44.77	0.35	93.02
Crooked River	2.02	3.29	0.46	42.12	0.46	95.87
Finlay Arm	1.87	2.57	0.45	34.24	0.32	68.76
Finlay River	1.83	2.64	0.48	33.22	0.22	61.02
Firesteel River	1.04	1.39	0	1.07	0	2.67
Fox River	1.09	1.39	0	0	0	2.34
Ingenika River	1.7	2.06	0.17	10.32	0.09	25.16
Lower Omineca River	1.69	2.33	0.36	29.07	0.19	56.23
Mesilinka River	1.87	2.61	0.5	36.96	0.32	69.57
~~**Murray River**~~	~~0.53~~	~~0.98~~	~~0~~	~~0~~	~~0~~	~~0~~
Nation River	2	2.68	0.46	36.17	0.3	75.85
Parsnip Arm	2.14	3.18	0.57	39.31	0.38	87.02
Parsnip River	1.75	2.52	0.38	31.79	0.3	53.26
Peace Arm	2.11	2.81	0.25	33.33	0.17	55.56
Toodoggone River	1.06	1.3	0.02	1.52	0.09	3.03
Upper Omineca River	1.25	1.57	0.14	11.07	0.13	24.18
~~**Upper Peace River**~~	~~1.98~~	~~3.11~~	~~0.73~~	~~12.43~~	~~0.2~~	~~70.3~~
Grand Total	**1.77**	**2.61**	**0.5**	**21.86**	**0.27**	**59.63**

Table 15: Calculations of road density and presence for wetland complexes and wetland buffers by watershed group. Very small portions of the Murray River and Upper Peace River fall within the FWCP Peace region boundary. Results for these 2 Watersheds are incomplete.

Watershed Group	Average Number of Threats in Each Wetland Complex	Average Number of Threats in Each Wetland Buffer	Average Road Density per Wetland Complex	Percent Wetland Complexes with Roads Present	Average Road Density per Wetland Buffer	Percent Wetland Buffers with Roads Present
Carp Lake	2.48	3.6	10.79	34.88	1.85	34.88
Crooked River	2.02	3.29	13.21	31.27	3	31.27
Finlay Arm	1.87	2.57	13	29.47	1.61	29.47
Finlay River	1.83	2.64	19.87	32.88	1.44	32.88
Firesteel River	1.04	1.39	2.8	18.72	0.46	18.72
Fox River	1.09	1.39	0.15	7.6	0.2	7.6
Ingenika River	1.7	2.06	8.3	23.23	0.81	23.23
Lower Omineca River	1.69	2.33	7	25.24	1.27	25.24
Mesilinka River	1.87	2.61	8.75	30.87	1.63	30.87
~~**Murray River**~~	~~0.53~~	~~0.98~~	~~0~~	~~0~~	~~0~~	~~0~~
Nation River	2	2.68	9.44	23.67	1.21	23.67
Parsnip Arm	2.14	3.18	11.75	24.05	2.23	24.05
Parsnip River	1.75	2.52	6.77	29.35	1.6	29.35
Peace Arm	2.11	2.81	7.62	36.11	1.34	36.11
Toodoggone River	1.06	1.3	2.53	7.07	0.79	7.07
Upper Omineca River	1.25	1.57	4.91	18.44	0.69	18.44
~~**Upper Peace River**~~	~~1.98~~	~~3.11~~	~~25.03~~	~~8.89~~	~~2.42~~	~~8.89~~
Grand Total	**1.77**	**2.61**	**12.79**	**17.93**	**1.73**	**17.93**

Table 16: Calculations of mapped Mountain Pine Beetle infestation density and presence for wetland complexes and wetland buffers by watershed group. Very small portions of the Murray River and Upper Peace River fall within the FWCP Peace region boundary. Results for these 2 Watersheds are incomplete.

Watershed Group	Average Number of Threats in Each Wetland Complex	Average Number of Threats in Each Wetland Buffer	Average MPB Density per Wetland Complex	Percent Wetland Complexes with MPB Present	Average MPB Density per Wetland Buffer	Percent Wetland Buffers with MPB Present
Carp Lake	2.48	3.6	0.79	91.86	0.83	100
Crooked River	2.02	3.29	0.8	64.34	0.66	86.82
Finlay Arm	1.87	2.57	0.82	94	0.85	99.18
Finlay River	1.83	2.64	0.68	77.97	0.72	94.58
Firesteel River	1.04	1.39	0.46	55.08	0.4	74.87
Fox River	1.09	1.39	0.67	74.85	0.67	85.38
Ingenika River	1.7	2.06	0.52	83.23	0.61	93.55
Lower Omineca River	1.69	2.33	0.79	84.98	0.82	92.97
Mesilinka River	1.87	2.61	0.66	80.43	0.74	95.65
~~Murray River~~	~~0.53~~	~~0.98~~	~~0.87~~	~~37.91~~	~~0.57~~	~~62.95~~
Nation River	2	2.68	0.85	96.97	0.91	100
Parsnip Arm	2.14	3.18	0.81	86.83	0.78	94.85
Parsnip River	1.75	2.52	0.64	51.09	0.57	58.97
Peace Arm	2.11	2.81	0.6	81.48	0.68	92.59
Toodoggone River	1.06	1.3	0.6	77.27	0.64	84.34
Upper Omineca River	1.25	1.57	0.78	87.3	0.79	97.13
~~Upper Peace River~~	~~1.98~~	~~3.11~~	~~0.94~~	~~87.2~~	~~0.87~~	~~100~~
Grand Total	**1.77**	**2.61**	**0.83**	**77.92**	**0.79**	**89.2**

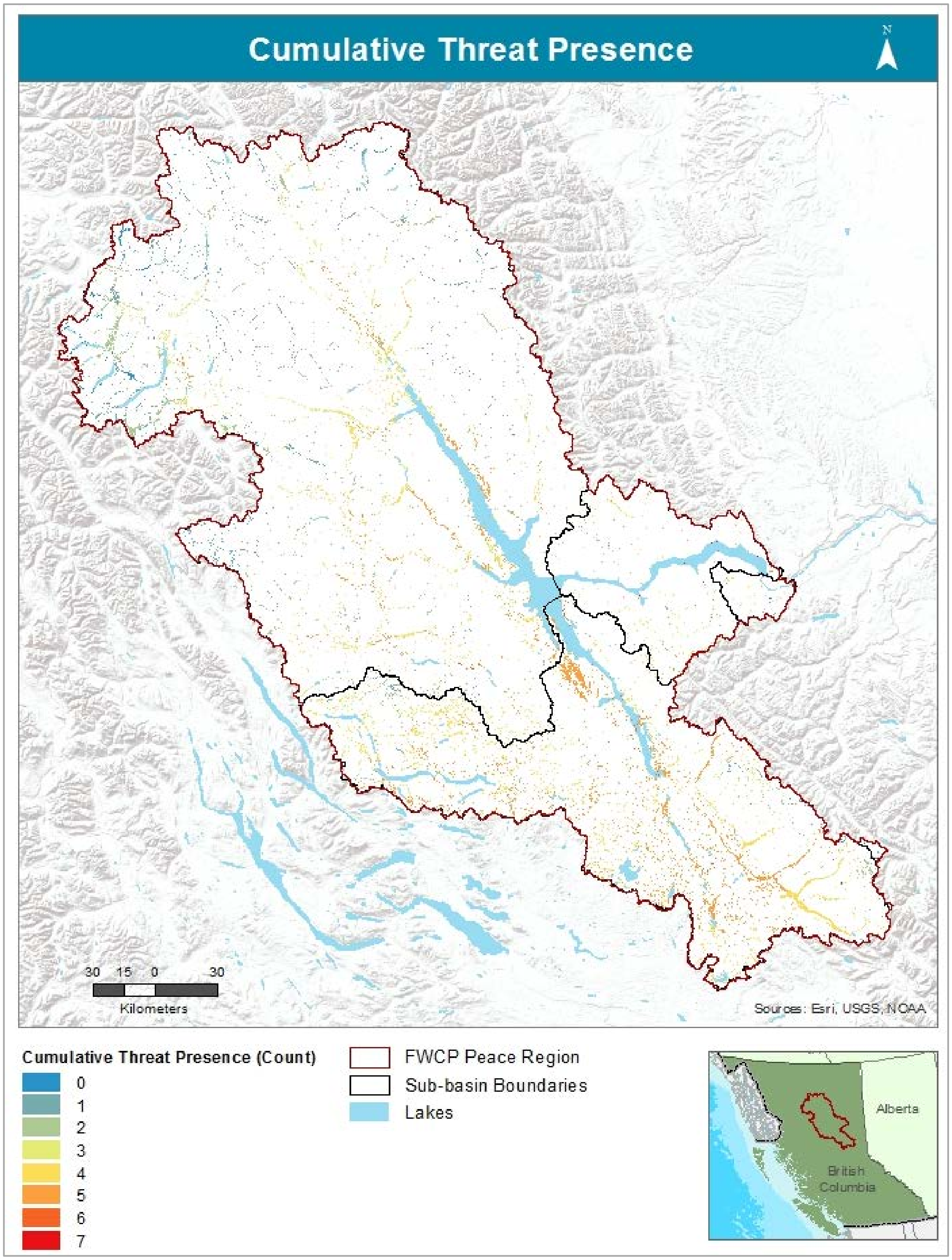


Figure 17: Wetland aggregation units (Figure 15) classified based on number of threats present within each wetland unit footprint.

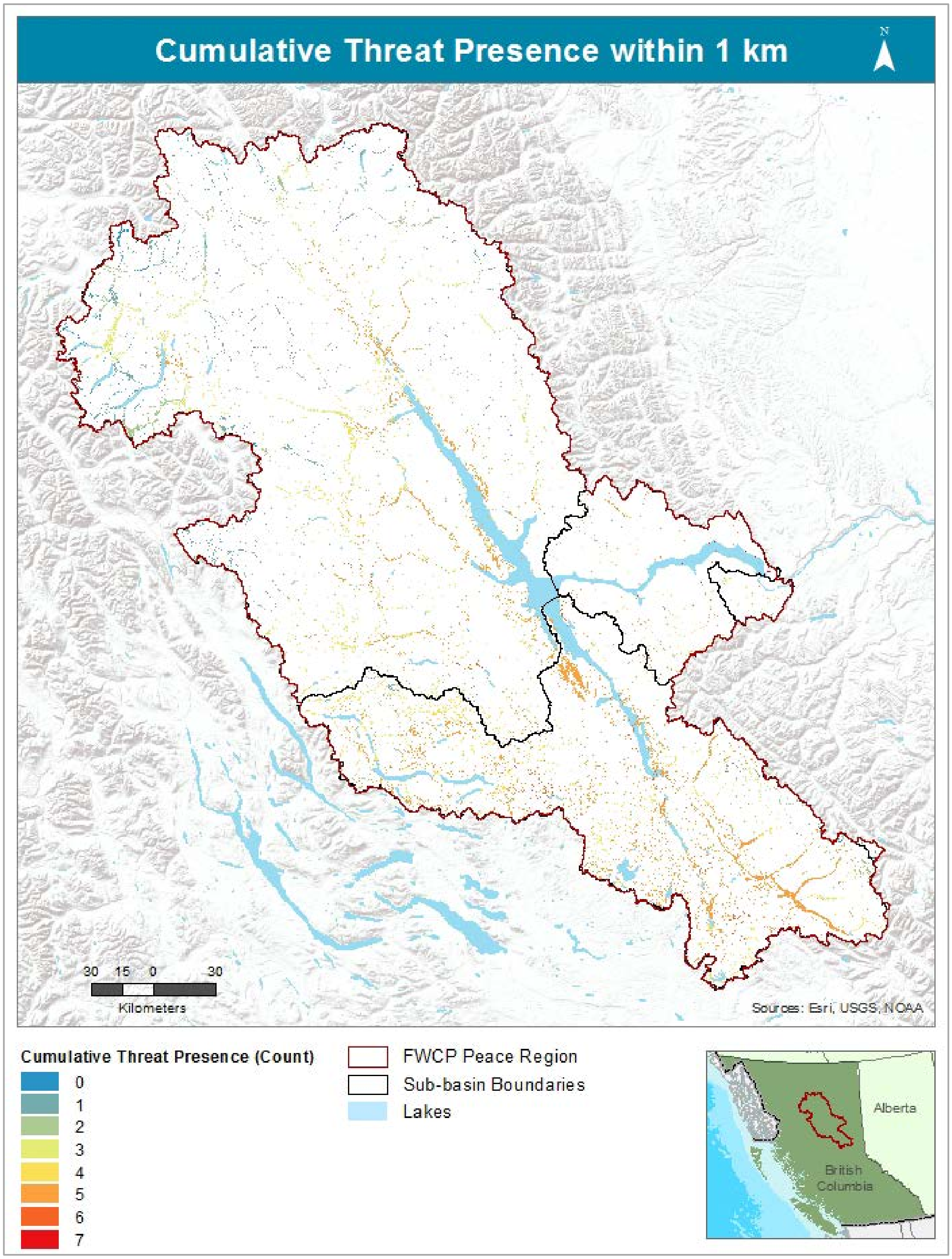


Figure 18: Wetland aggregation units (Figure 15) classified based on number of threats present within 1 km buffer distance of each wetland unit footprint.

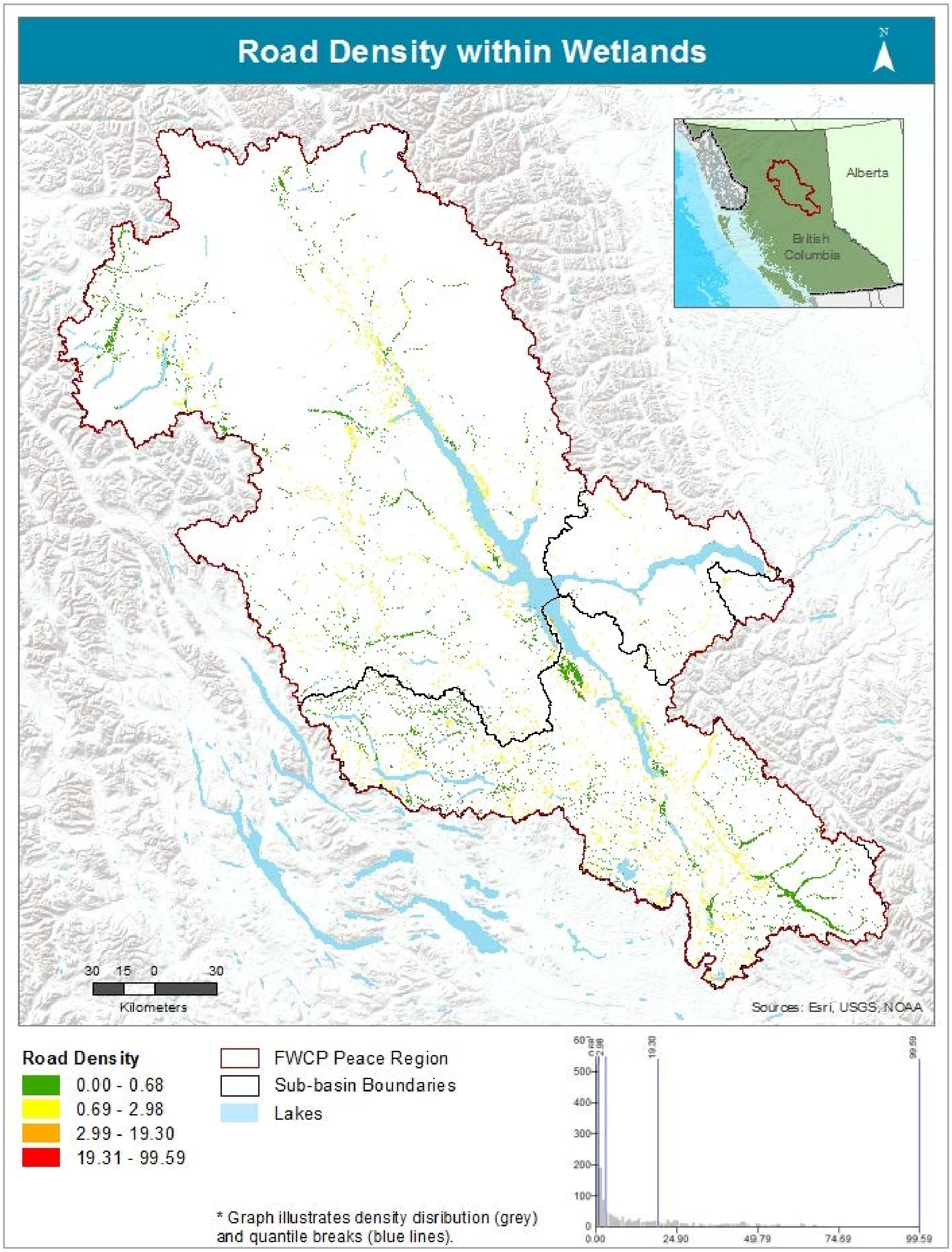


Figure 19: Wetland aggregation units (Figure 15) classified based on total road density within each unit footprint. Road data from British Columbia Digital Roads Atlas. Symbolization thresholds are based on dataset quantiles. Quantile breaks and density distribution are provided in graph.

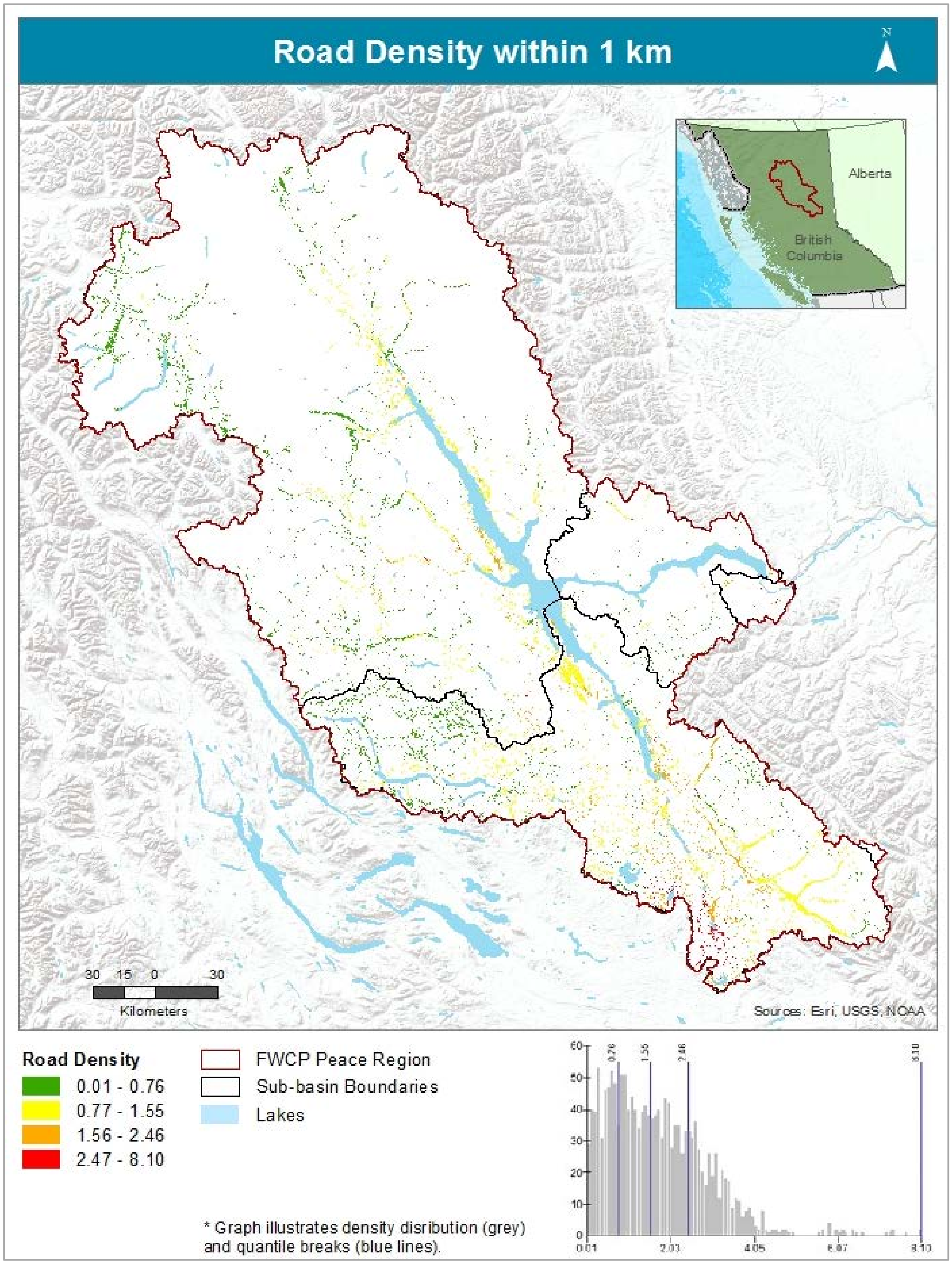


Figure 20: Wetland aggregation units (Figure 15) classified based on total road density within 1km buffer distance of each unit footprint. Road data from British Columbia Digital Roads Atlas. Symbolization thresholds are based on dataset quantiles. Quantile breaks and density distribution are provided in graph.

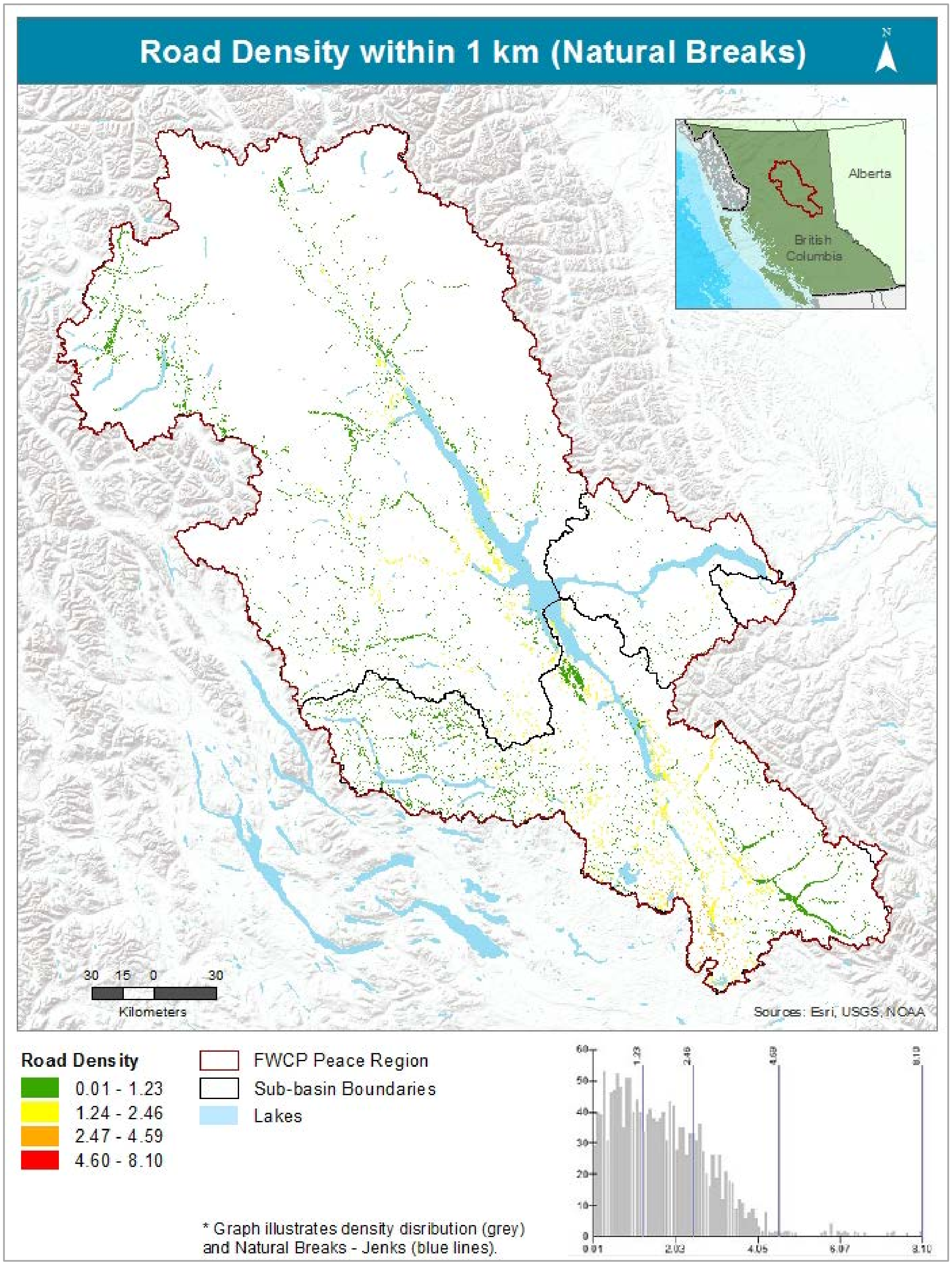


Figure 21: Wetland aggregation units (Figure 15) classified based on total road density within 1km buffer distance of each unit footprint. Road data from British Columbia Digital Roads Atlas. Symbolization thresholds are based on Natural Breaks (Jenks). Break values and density distribution are provided in graph.

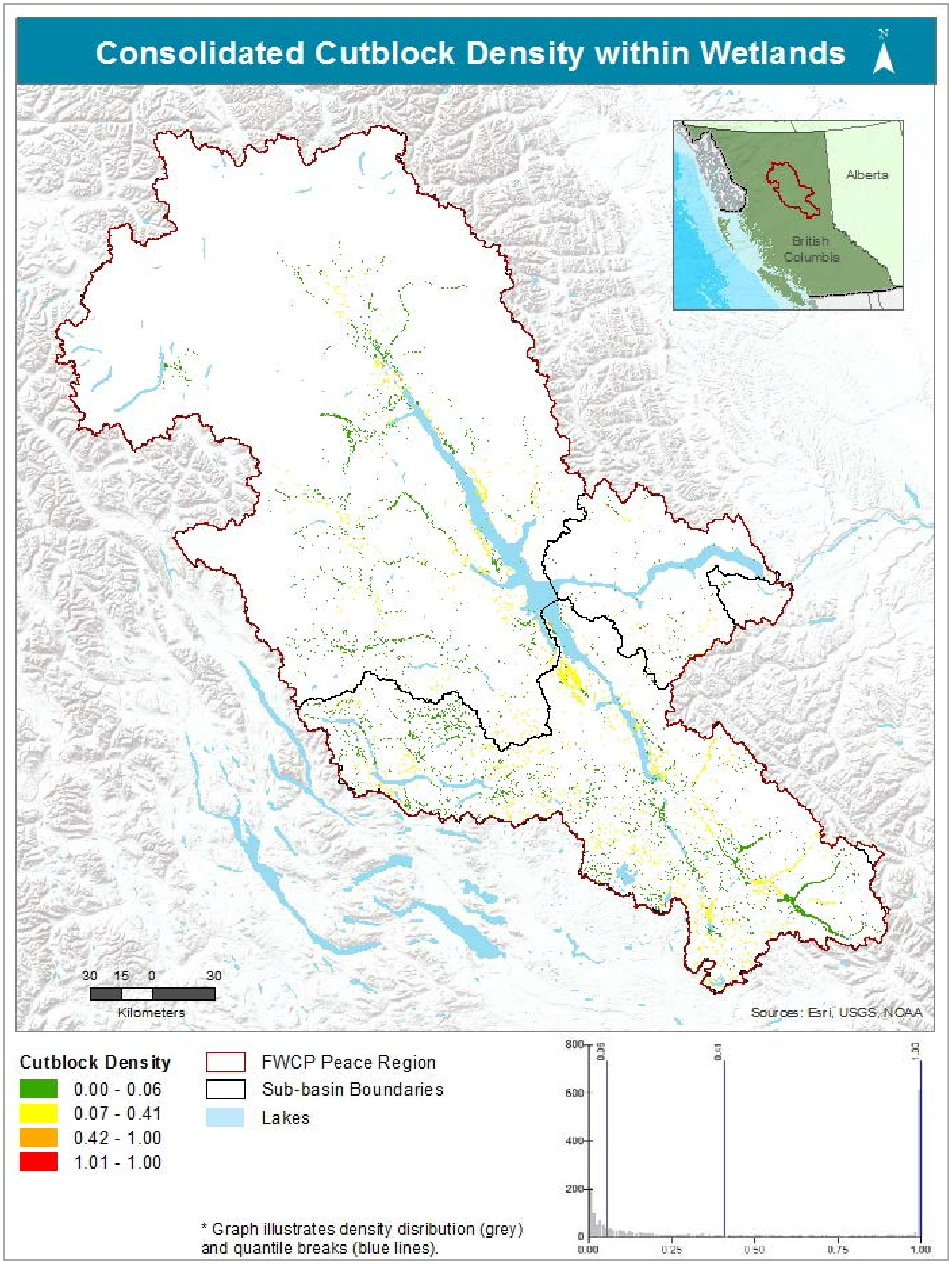


Figure 22: Wetland aggregation units (Figure 15) classified based on total cutblock density within each unit footprint. Cutblock data from British Columbia Consolidated Cutblock layer. Symbolization thresholds are based on dataset quantiles. Quantile breaks and density distribution are provided in graph.

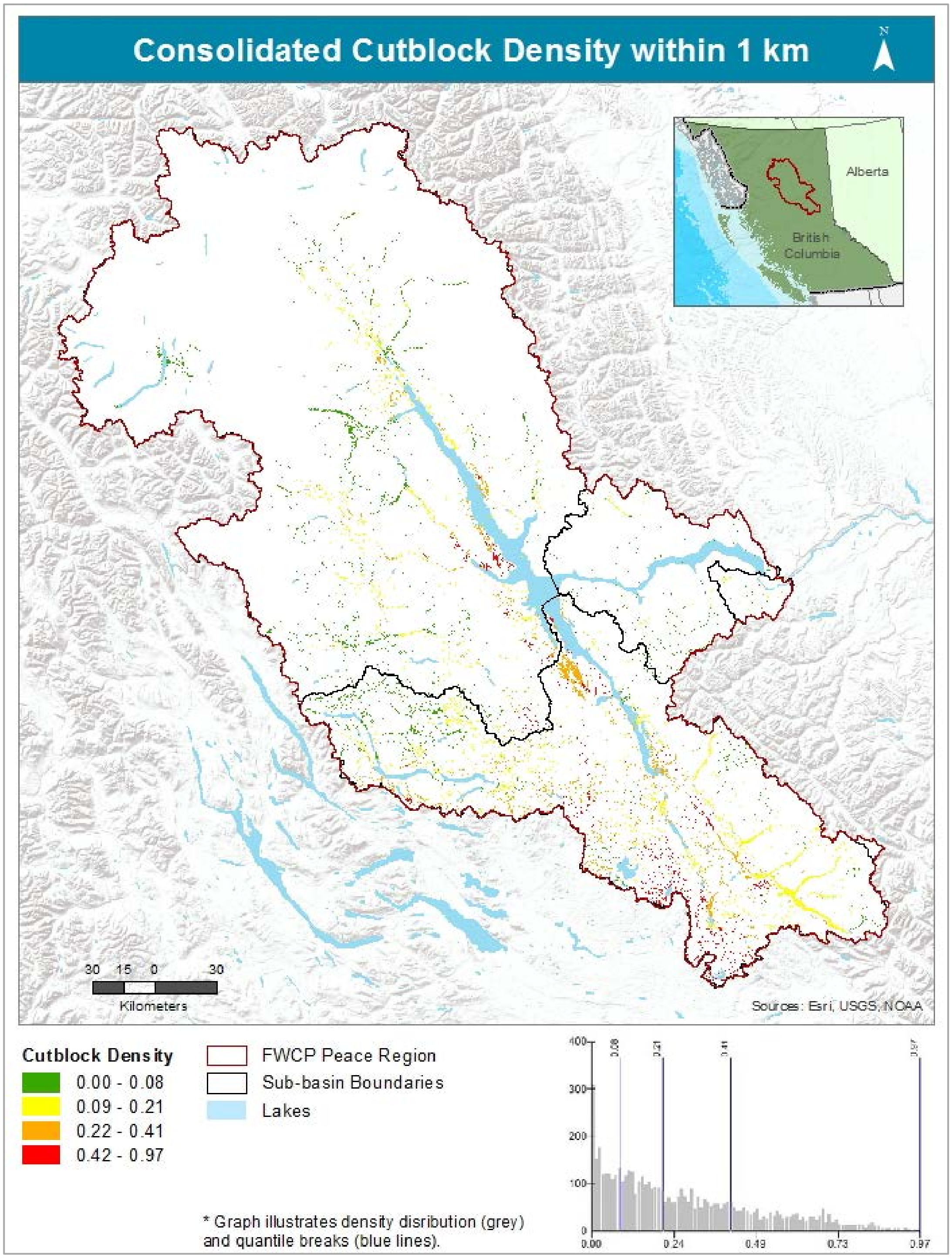


Figure 23: Wetland aggregation units (Figure 15) classified based on total cutblock density within 1km buffer distance of each unit footprint. Cutblock data from British Columbia Consolidated Cutblock layer. Symbolization thresholds are based on dataset quantiles. Quantile breaks and density distribution are provided in graph.

DATA AND INFORMATION ACCESS

The Random Forest script was made accessible to the public by posting R scripts and documentation on GitHub (https://github.com/bcgov/wetlandmapR). Distributing scripts through a GitHub repository encourages future collaboration and promotes the continuous improvement cycle recommended in the first model run report (Filatow *et al.,* 2018). The report and data package for the second model run – including model results, field data, field photographs, training data, data dictionary and other supplementary datasets – will be hosted on the TEI Data Distribution Site[10]. Project report and key data outputs will also be stored on EcoCat, as space restrictions allow. Datasets produced during Run 2 will be made available under the British Columbia Open Data Licencing[11].

A key recommendation from the first report endorsed the development of an ArcGIS Online (AGO) web mapping application prototype as a tool to promote open and transparent data access. During Run 2, an AGO web mapping application, titled Williston Wetland Explorer Tool (WWET), was created to facilitate data distribution and public engagement. WWET content includes welcome and about pages, user guide, Wetland Explorer Tool, time lapse example as well as a references and acknowledgements section. The WWET application can be accessed through the metadata record at https://catalogue.data.gov.bc.ca/dataset/1d174499-72af-41d5-ad5e-19460b1e6fb1.

The WWET welcome and about pages offer context for project goals and results, as well as links to important documents such as model Run 1 and Run 2 reports (Figure 24). The user guide is a resource for those users who may be unfamiliar with AGO map applications or those who require a more detailed understanding of key functionality offered by the platform. The Wetland Explorer tab is where the user can explore project data and navigate through the FWCP Peace region to view prediction results and other related datasets (Figure 25). Other datasets include, but are not limited to, the British Columbia TRIM wetlands and waterbodies, harvested areas, contours, the Digital Roads Atlas (DRA), and the Freshwater Atlas stream networks, lakes and wetlands. Field data attributes are available through automatically enabled pop-up windows easily accessed by clicking on the point of interest (Figure 26). Field photographs are also obtainable through the pop-up window for the photograph points (Figure 26). By clicking on the photograph within the pop-up window the user will be redirected to the full-size image currently hosted on the TEI Data Distribution Site.

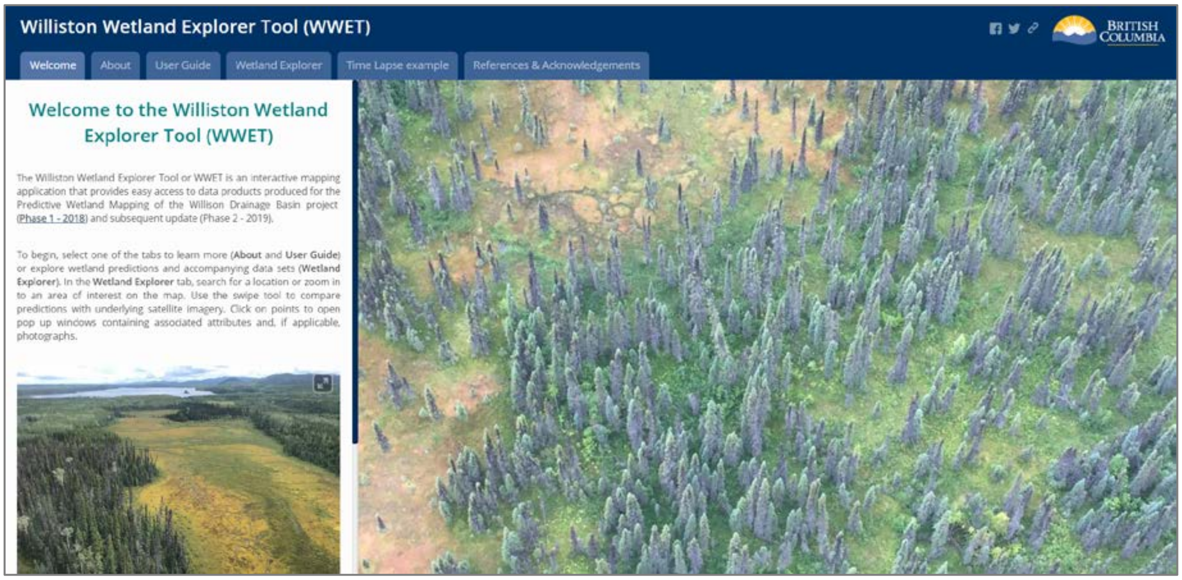
The Time Lapse Example tab demonstrates an exploratory tool based on plug-ins developed by the Google Earth Engine program (Figure 27). By selecting the play button, the time lapse example will run through a time series of satellite imagery to illustrate change on the landscape for a given area.

---

[10] http://www.env.gov.bc.ca/esd/distdata/ecosystems/TEI/WETLANDS

[11] https://www2.gov.bc.ca/gov/content/data/open-data/open-government-licence-bc

The default map extent is set to an area with particularly significant landscape change, however the user may pan to different areas as desired.

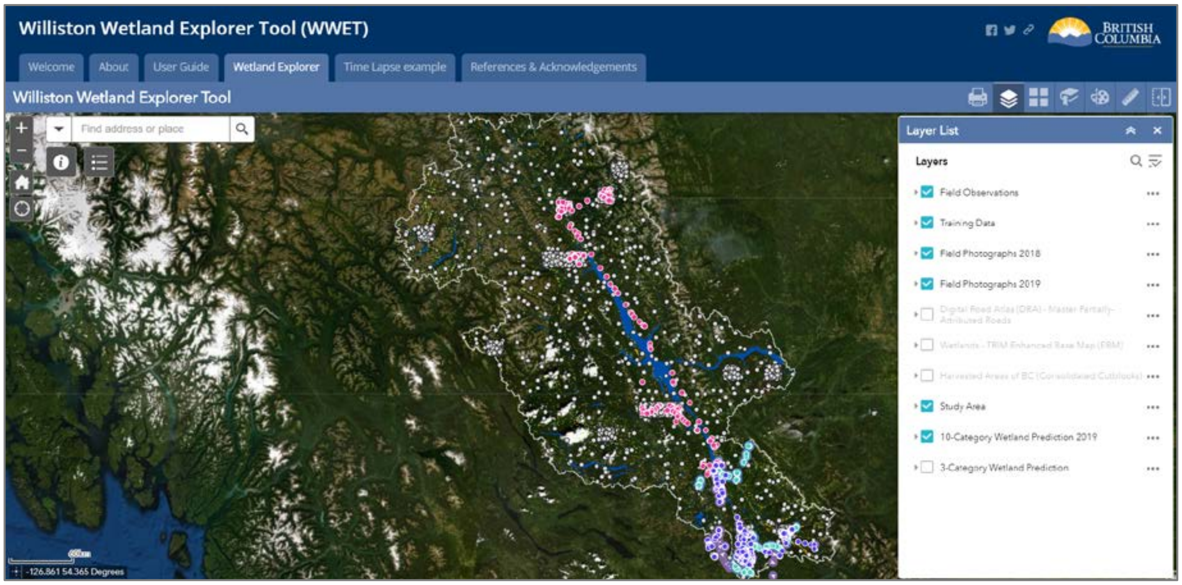


Welcome to the Williston Wetland Explorer Tool (WWET)

The Williston Wetland Explorer Tool or WWET is an interactive mapping application that provides easy access to data products produced for the Predictive Wetland Mapping of the Willison Drainage Basin project (Phase 1 - 2018) and subsequent update (Phase 2 - 2019).

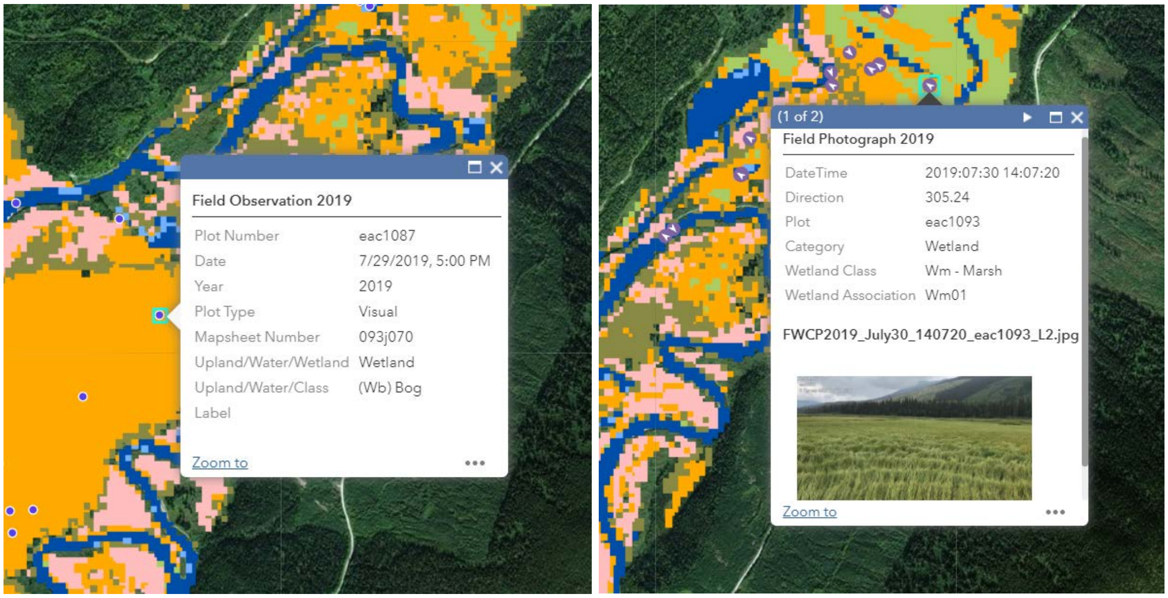
To begin, select one of the tabs to learn more (**About** and **User Guide**) or explore wetland predictions and accompanying data sets (**Wetland Explorer**). In the **Wetland Explorer** tab, search for a location or zoom in to an area of interest on the map. Use the swipe tool to compare predictions with underlying satellite imagery. Click on points to open pop up windows containing associated attributes and, if applicable, photographs.

Figure 24: The Williston Wetland Explorer Tool (WWET) home page.



☑ Field Observations
☑ Training Data
☑ Field Photographs 2018
☑ Field Photographs 2019
☐ Digital Road Atlas (DRA) - Master Partially-Attributed Roads
☐ Wetlands - TRIM Enhanced Base Map (EBM)
☐ Harvested Areas of BC (Consolidated Cutblocks)
☑ Study Area
☑ 10-Category Wetland Prediction 2019
☐ 3-Category Wetland Prediction

Figure 25: The Williston Wetland Explorer Tool (WWET) Wetland Explorer tab.


Field Observation 2019

Plot Number eac1087
Date 7/29/2019, 5:00 PM
Year 2019
Plot Type Visual
Mapsheet Number 093j070
Upland/Water/Wetland Wetland
Upland/Water/Class (Wb) Bog
Label

(1 of 2)
Field Photograph 2019

DateTime 2019:07:30 14:07:20
Direction 305.24
Plot eac1093
Category Wetland
Wetland Class Wm - Marsh
Wetland Association Wm01

FWCP2019_July30_140720_eac1093_L2.jpg

Figure 26: Field data (left) and field photograph (right) pop-up windows within the WWET Wetland Explorer Tool.


Williston Wetland Explorer Tool (WWET)

55.22624, -123.62689

Figure 27: Time Lapse Example tab within the Williston Wetland Explorer Tool (WWET).

DISCUSSION

The primary goal of the project in 2017 was to evaluate machine-learning approaches to produce a reliable, reproducible and cost-effective map of wetlands ecosystems to address the objectives in the Wetland and Riparian Action Plan (FWCP, 2014). Specifically, the actions related to inventory of wetland and riparian ecosystems to support Objective 1. The modelling approach was tested as an alternative to the available wetland mapping as it lacked information about wetland type, was inconsistent, had poor quality in remote areas, and did not include riparian ecosystems. In addition, conventional mapping methods were found to be cost prohibitive for the size of the FWCP Peace region.

The 2018 model results achieved these goals at the scale of the entire Williston Drainage Basin area (FWCP Peace region) and were 91% accurate at predicting wetland ecosystems in the areas field validated. A preliminary analysis was completed to summarize wetland size and densities for four sub-basins of the FWCP Peace region (Finlay, Peace, Parsnip, and Dinosaur). Three disturbance layers were also summarized for each of these areas. Timber harvest, roads and mining spatial layers were used to demonstrate the proof of concept.

However, the results had flaws; including over capture of wetlands in low relief and disturbed areas in the Parsnip, which biased the analysis at the sub-basin scale and inhibited the use of the data at a more local scale. The first run of the model also excluded riparian ecosystems, lumping all floodplain units with upland systems in the 3-category map. The rationale was to focus specifically on Action 1b-1, which related specifically to wetlands. The addition of the mid-bench floodplain (Fm) and low-bench floodplain (Fl) categories to the 3-category map allows for the inclusion of larger patch (>1ha) occurrences of mid and low-bench riparian ecosystems to be included in the analysis. The narrow fringe and smaller patch riparian areas, however, will not be detectable by the machine learning model due to the scale of these features, the resolution of the predictor variables, and the difficulty of consistent detection and classification by mappers.

In addition, some ecosystem types were poorly represented in the training and validation points (e.g., swamps) and were therefore poorly predicted by the model. Rare and uncommon ecosystem types may not get adequately sampled or consistently identified in training points. These ecosystems are more readily identified in ground surveys which is one reason that geographic coverage of the FWCP Peace region is important.

Additional training data collected to remedy these flaws were used to re-run the model. Visual inspection showed the results from Run 2 better represented wetland distribution in the problematic areas in the Parsnip. Fieldwork validation points proved these observations to be true with a 13% prediction accuracy improvement over the first run of the model.

The resultant 3 and 10-category wetland/riparian predictive maps fulfill the Action Plan objectives to "*Improve the understanding of the abundance, [and] distribution*" of wetland and riparian areas in the FWCP Peace region.

The 3-category wetland map has low OOB error and high field validated accuracy rates and is appropriate for use in analyses to inform landscape-level wetland action priorities in the Williston Drainage. The example disturbance case study showed that analytical techniques can be used to address wetland ecological connectivity and cumulative threats. The resultant maps show areas where disturbance layers intersect wetland complexes and 1 km buffered areas around wetland complexes. Disturbance levels are one factor that can inform priority wetland and riparian areas for action. For example, areas of high disturbance are potential for restoration and enhancement actions and areas of low threat may be in good condition and offer conservation potential.

The connectivity and threats analyses demonstrate how the wetland and riparian map can be used to assess and improve our understanding of wetland and riparian connectivity and current condition. However, this case study is a proof of concept and needs refinement and automation to build a robust and defensible analysis that can be incorporated into existing scripts and packages, contributing to the continuous improvement cycle.

MODEL AND DATA

FIELD DATA

Field data distribution and representation is not distributed evenly across the entire FWCP Peace region. The data also does not capture the full range of environmental conditions represented by the predictor layers. Access and time limitations restricted field data point distribution. Thus, model accuracy based on field data may not be representative of areas where there is an absence of field data points, or areas with very low point densities. It also is not representative of areas that have environmental conditions that are dissimilar to the variety of conditions sampled in the field.

In addition to model accuracy assessment, the field data adds further value beyond the scope of this project as it contributes to the knowledge and data on wetlands and riparian ecosystems in the FWCP Peace region. Element occurrences were updated and ranked using data and information from field work (2017-2019). Field data was also entered into the Biogeoclimatic Ecological Classification (BEC) provincial database and are now available to inform classification and for other land management decision processes.

Georeferenced field photos are a robust and valuable information source collected during field sampling. They are a record of wetland condition and provide visualization of these ecosystems for education and outreach purposes.

Field data points and photos are easily accessed and explored through the WWET application to aid in education, outreach, and desktop evaluation of wetland and riparian areas. The WWET application provides model results and other spatial layers to investigate in conjunction with field points and photos.

MODEL PERFORMANCE

The accuracy of the Random Forest model is dependent on the quality of information that it is provided (training data and predictor variables). As the training data were populated based on mapper interpretations of aerial imagery, these data represent the mappers mental model of the landscape. The Random Forest prediction is, therefore, an estimate of the mapper's mental model executed consistently across areas where training data calls were not provided. To provide a measure of how accurate the model was at estimating the mapper calls, an internal estimate of model error is used. The OOB error metric is a way to assess how well the Random Forest model predicted the mapper's mental model given the provided training data and predictor variables. In addition, independent field observations were used to asses mapper call and model accuracy.

3-CATEGORY PERFORMANCE

The results indicate that the method, including both the mappers training calls and the modelling process, is reliable for classifying the 3-category classes (water, wetland and upland). Mappers were successful at correctly assigning upland, water, and wetland categories with 81%, 89%, and 98% respectively (Table 8). These are easy to interpret categories that are reliably predicted. Model results indicate high accuracy (86.7%), low OOB error (7%) and good class separation for the 3-category map. These numbers and their agreement are the sign of a reliable prediction at the sub-basin level. These conclusions are consistent with the findings of the first model run.

The overall result of the 3-category prediction was that the OOB metric was identical between Run 1 and Run 2 (7% error, 93% correct) model outputs. However, the field validated accuracy cannot be directly compared between modelled results without first accounting for the differences between how the categories were grouped (i.e., low and mid bench floodplains were grouped with wetlands in the second model run but grouped with upland in the first). After removing field data points that differ in classification groups between model runs, field validated accuracies were compared. In the Parsnip sub-basin where fieldwork was targeted to reduce observed errors, there was an improvement of 14% when comparing the first model run (74%) with the second (89%) for the 3-category prediction.

10-CATEGORY PERFORMANCE

The 10-category map has large discrepancies between the OOB (11.7% error, 88.3% correct) and field accuracy (52.8%). A 65% accuracy threshold for chief forester to use predictive ecosystem mapping for allowable annual cut decisions (Meidinger, 2003). The 10-category prediction is below this threshold. As a result, use of the 10-category results should be used with caution.

The discrepancy between OOB and field accuracy may point to issues in mapper interpretation of ecosystems at the wetland class level. The confusion matrix points to potential interpretation issues with treed and shrubby wetlands. For example, it is difficult to determine swamps from flood units as they are both shrubby and/or treed and model results indicate confusion between these categories. Confusion may also be a result in low numbers and an imbalance in the training data for some of these uncommon and hard to access ecosystems. Training data contain 140 floodplain unit points (out of 13,463 points) which represents only 1% of the data. This can result in poor predictions and misclassification of types that are not well represented in the training data.

Table 9 shows that categories such as upland (91%), water (89%), bogs (86%), and fens (63%) were interpreted by mappers well when compared to field data. Shallow water (50%), swamp (33%), mid-bench floodplain (29%), marsh (18%), and low bench floodplain (16%) had decreasing and overall poor mapper interpretation accuracy compared to field data. Note that no high bench floodplain training points were validated by field visits. Some inconsistencies in the mapper calls in the training data, result in the model failing to accurately predict these attributes. For example, the mapper may call a given treed wetland a bog, but in another identical area (both visually and the predictor variable values in those areas) the mapper calls the feature a swamp. Model validation using field data indicates how well the model (mapper's mental model) was able to accurately represent reality. If consistent errors are present within the training data, these misclassifications will be reflected by the elevated level of disagreement between the model prediction and the field calls. This metric is displayed as an error percentage that represents the discrepancy between the mapper's call and the field call.

When the 10-category prediction results were compared against the first run of the model, overall improvement was observed in the FWCP Peace region (7%) as well as within the Parsnip sub-basin (9%). However, a slight decrease in accuracy for the second run of the model was recorded for the Finlay sub-basin (-4%). This result shows that sometimes local improvements may negatively impact the sub-basin level results. This may be due to the spatially clustered field data concentrated primarily within the Parsnip sub-basin. With additional data the model can be re-run for subsets of the FWCP Peace to represent a more localized prediction. Alternatively, additional data, such as a geographically limiting predictor layer (e.g., location or positional information such as x and y coordinates), can be added to model inputs to aid in capturing spatial trends and patterns . In other words, the additional data in the Parsnip, weighting the presence of validation data to one sub-basin, may be producing deceptive results in the Finlay, as the model has not been given a layer representing spatial location (such as, x-y coordinates or watersheds) to differentiate these geographical differences.

46-CATEGORY PERFORMANCE

The high OOB error and low field validated accuracy numbers mean that the **46-category raster prediction is not reliable and should not be used**. The nuances in the 46-category were not predicted well by the model. This is due to the coarse resolution of the predictor layers and the difficulty in interpreting these categories from remotely sensed data (i.e., satellite imagery and orthophotography). The 46-category prediction was intended to inform the 3 and 10-category classifications and was not based on any formal classification scheme. In this run of the model, use of the 46 classes did not improve results when reclassed to the 3 and 10-category classification. This output is not provided in the data package or in the WWET application for this reason.

DATA LIMITATIONS

The data limitations are consistent with those identified in the first run of the model, namely limitations of the spatial data, training data, and field data that impact the model output. In addition, post processing steps to derive data for the disturbance analysis introduce additional errors and limitations.

Most of the limitations relating to the spatial predictor layers outlined in detail in the Run 1 report (Filatow *et al.*, 2018), relate to the coarse nature and flaws in the Digital Elevation Model (DEM) and the Sentinel-2 mosaic. These constraints can be improved with higher quality DEM or LiDAR. The Sentinel-2 mosaic can be improved with additional preprocessing and more current data. GeoEye, or other higher resolution data sources for multispectral imagery, are also available from commercial distributors.

Some of the 10-category confusion and errors are an artifact of the difficulty in applying a site-level classification to a mapping exercise. These are site-level concepts that do not translate well into predictable mapping categories. The units represent idealized ecosystem categories and real-life transitions and variations can be difficult to consistently 'lump' into the same category. This can also be exacerbated by different mapper bias when applying the classification in these 'grey' areas (during desktop inventory for training data). The 2019 field season highlighted some of these mapping and classification challenges in areas fondly termed 'fogs'. These were areas of transition between bogs and fens either due to gradual transition zones or changes in hydrology due to hydraulic changes caused by roads and beaver activity. Similarly, 'swogs' (swamp-marsh transitions) like the Wb08 - Black spruce - Soft-leaved sedge - Peat-moss are in the bog class (Wb), but from a mapping perspective are more suitable in the swamp class (Ws) due to their signal on the imagery (tree canopy density). In addition, gleysols (mineral soils) were also encountered in some of these ecosystems which is more consistent with swamps. Several wetland associations in the swamp-bog transition are recognized in the guide as often distinguished by subtle or understory vegetation differences that are difficult to map. Ws07 - Spruce - Common horsetail - Leafy moss, Wb08 Black spruce - Soft-leaved sedge - Peat-moss and Wb09-black spruce - Common horsetail - Peat-moss all occupy this swamp-bog transition, have similar characteristics and can be difficult to differentiate. Wb08 is recognized to be a transition between Ws07 and the Wb05 (MacKenzine and Moran, 2004). Areas of transition between fens and marshes, called 'farshes' by the field crews, like Ws07 and Wm01, are other areas where application of the classification on the ground and interpretation from imagery are both difficult.

RIPARIAN AREAS

The categories selected as the model response variables are primarily focused on wetland classifications. As definitions and ecological classifications for riparian ecosystems differ depending on the area of study or research question (e.g. flood group, vegetation association, or riparian areas), there are various ways to approach a riparian modelling task. Riparian ecosystems were included in the model by using floodplain units, consistent with the wetland classification; therefore, predicting for both riparian and wetland ecosystems simultaneously. Riparian ecosystems captured by the model represented plant associations of the low-bench (Fl) and mid-bench (Fm) flood units as defined by MacKenzie and Moran (2004). These units were predicted in the 10-category model prediction and were included in the wetland category of the 3-category model prediction. However, riparian ecosystems had relatively low accuracy rates and high confusion rates. As a result, there remains room for improvement.

Currently, there are relatively few training and field validation points representing riparian ecosystems in the FWCP Peace region and consequently, model reliability for the Fl and Fm units is low (Table 6). Visual inspection of riparian ecosystems highlighted areas of potential riparian vegetation that may be under captured or speckled/patchy predictions where a continuous area exists. Considerable gains can be made by improving the modelling of riparian areas by using an agreed upon definitions and creating a riparian specific model.

DISTURBANCE ANALYSIS CASE STUDY

A proof-of-concept disturbance analysis was completed to approximate current land use patterns and disturbances intersecting and adjacent to wetland complexes. Individual wetland pixel predictions were aggregated to approximate hydro-ecologically connected wetland complexes. The aggregates were buffered by 1 km to capture disturbances in close proximity. These were evaluated based on the density and cumulative presence of disturbances to visually summarize how land use patterns differ across the FWCP Peace region. This trial analysis was undertaken so that wetland complexes with relatively few disturbances could be identified as potential targets for conservation and land acquisition, whereas wetland complexes with high disturbances could be potential suitable targets for focussed remediation efforts. The results were summarized using the Freshwater Atlas watersheds (Figure 28) to report on relative wetland disturbance across the FWCP Peace region.

Assumptions relating to wetland presence, aggregation, threats, disturbances, and upland catchment area were necessary to complete the analysis and have influenced the results of this modelling experiment. The reliability of the analysis output inherits the limitations of the input layers and the assumptions made in the methodology. These include the predicted wetland layer, the assumptions and layers used to represent hydro-ecological connectivity, the disturbance layers obtained from the BC Data Catalogue, and other assumptions such as buffer distances, used in the analysis methods. Assumptions and limitations in this trial analysis can be reduced by applying similar analysis techniques to focussed questions relating to specific species, wetland values or wetland actions.

The 3-category prediction layer is the starting point of the disturbance analysis. Limitations and accuracy of this data is discussed in detail in early sections. Improvements in the Parsnip sub-basin addressed cutblocks being incorrectly classified as wetlands. This is relevant to analysis results relating to cutblocks and wetlands.

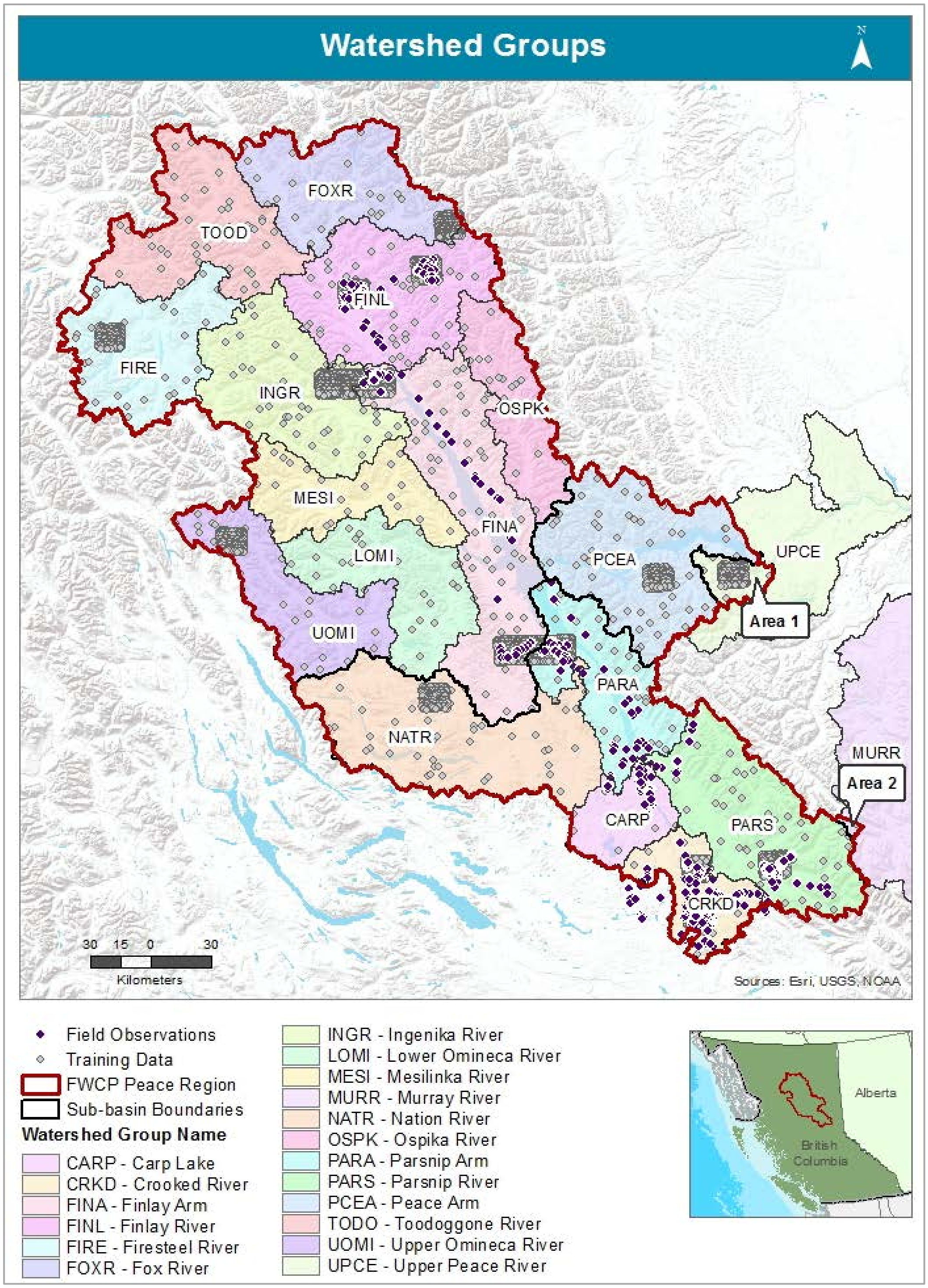


Figure 28: Watershed groups that fall within the FWCP Peace region boundary. Field observations and training data points are included to illustrate watersheds that have good or poor coverage of these data. Area 1 and 2 icons indicate where the FWCP Peace region boundary subdivides a watershed group.

WETLAND AGGREGATION

Wetland predictions (25m pixel size) were grouped and polygonised then aggregated into complexes to account for hydrologic connectivity (Figure 15). Consideration should be given to managing wetland units as aggregations, in addition to individual wetlands, as they are likely ecologically aggregated. Winter (1999) and Devito *et al.* (1996) have shown that wetlands can be hydrologically connected, which in turn may influence wetland function (Powers *et al.*, 2012; Wolf *et al.*, 2013; Cook and Hauer, 2007).

Wetlands were aggregated into wetland complexes based on presumed surface hydrological connectivity using the stream network and the riparian potential layer. This aggregation process inherits the limitations of the riparian potential predictive layer (The Nature Trust of BC, 2019) used to approximate hydrological connectivity and consequently, requires improvement. An example of this limitation relates to underestimating wetland hydrologic connectivity in the Dinosaur. The Dinosaur watershed is truncated by the FWCP-Peace administrative boundary, and so the stream network was not complete when aggregating wetlands into complexes within this watershed (Figure 28 – Area 1). Analysis of a partial watershed is misleading because the number and area of wetlands may be reduced, as wetlands are typically less frequent in headwaters of watersheds than low-lying areas. Consequences of underestimating surface hydrological connectivity include (1) incorrect relative densities of disturbances that are not comparable to other parts of the study area as wetland aggregation that is not consistent and (2) reduced count and density disturbance impacts as disturbances are less frequent in high-elevation areas. As well, wetland count (i.e., presence) may be impacted when the total number of wetlands increases but are not aggregated consistently within the region of interest.

No suitable layer to model ground water flow or connectivity with complete study area coverage and appropriate scale was available, though wetlands are known to be connected through ground water flow (Winter, 1999; Neff and Rosenberry, 2018). Additional work to determine subsurface hydrological flow patterns, such as aquifer maps, geological maps that locate karst or permeable rocks, surficial geology maps, or other data sources, would improve hydrological connectivity approximations.

Wetland aggregation only considers complexes as wetlands and does not subdivide into other groups such as bog, fen, marsh, or swamp. Advantages to this approach include consistent identification of 3-category output (wetland, upland, and water), which has the highest accuracy of the model outputs. As well, this approach circumvents classification problems that relate to transitional wetland types. A disadvantage to this approach is that wetland type strongly influences how they respond to disturbance. For example, fluctuating water levels or impacted hydrological patterns are unlikely to impact marshes in the same way that they may affect bogs. Similarly, nutrient additions and increased sedimentation from upland disturbances may have pronounced effects on nutrient-poor bogs (Basiliko *et al.*, 2006) when compared with other types of wetland that are not nutrient-limited.

Other assumptions associated with lumping wetland groups include riparian and flood ecosystems. The predictive wetland model does not adequately capture riparian areas and the wetland aggregates are also likely to under capture riparian areas and connectivity.

UPLAND LANDSCAPE CONTEXT (BUFFER)

Upland surrounding areas are assumed to be hydrologically linked contributing areas to wetland complexes so the area of land that can potentially impact wetland function is greater than the wetland area (Bearup *et al.*, 2014; Houlahan and Findlay, 2003). As such, mitigation of disturbances to wetlands will likely need to occur over a larger area than just the wetland footprint. Specifically, the 1 km distance from the wetland complex used in this disturbance analysis may not be ecologically relevant. The size of an ecologically meaningful buffer or upland contributing area will vary with the species or wetland function being investigated. For example, species may have differing sensitivity levels to forestry disturbances. Therefore, the impact of forestry will be different when considering frogs vs. invertebrate community where each species group will likely have different tolerances.

A high amount of buffer overlap occurred (29,247 km^2, or 45% of total buffer area), which overestimated the amount of contributing upland area and, likely, disturbance area that may impact wetland complexes by not considering height of land and major drainages. Subsequent analysis should include watershed boundaries, slope direction, and water networks to refine contributing area, and the distance of the buffer should be determined by the user question.

DISTURBANCE LAYERS

Landscape-level disturbances are relevant both where they directly intersect boundaries of the wetland complexes and in upland catchment areas that are hydrologically connected to the wetland complexes. However, data integrity and scale issues are present in disturbance layers obtained from the BC Data Catalogue. Several examples include the Digital Road Atlas, MPB Infestation polygons, and the Consolidated Cutblocks database. Although the Digital Road Atlas is continuously being updated, recently-constructed resource roads may not be identified. The consolidated cut block layer is a reasonable estimation of timber harvest that occurred in the past decade. However, older cut blocks and timber harvest on privately owned land may not have been documented in this layer.

Mountain pine beetle (MPB) values were presented in the result section to demonstrate how results depend on input layers, or in this case, the mapping scale of mountain pine beetle infestations. Although MPB infestation impacts to upslope contributing area are expected to influence hydrology patterns and sedimentation in hydrologically connected wetland complexes, graminoid-dominated wetlands without trees are unlikely to be directly impacted by MPB infestation. However, MPB infestation was mapped at a larger scale than the wetland predictions, which included wetlands in the MPB infestation area. As a result, the results and numbers related to MPB infestation in wetlands should warrant careful consideration of disturbance mapping scale and accuracy. As well, "null" values may also represent areas that have not been sampled and therefore, may not represent areas that have been sampled and recorded to be absent of MPB. Additionally, boundaries of infestation may represent a more gradual transition than the hard line indicated by the polygons.

DISTURBANCE ANALYSIS SPATIAL SUMMARY

By standardizing disturbance by both presence and density, disturbances and land use trends are scalable and relatively comparable within an area of interest such as a map sheet, watershed group or administrative area (Figure 29). Although it may be possible to identify suitable wetland complexes for remediation or conservation efforts, relative assessment does not necessarily use an ecologically meaningful threshold. Suitable thresholds of disturbances, such as road density, may vary based on different focal species or landscapes (Hermann *et al*., 2004; Houlahan and Findlay, 2003), and are likely to require focussed research. In the absence of well-established disturbance thresholds, sensitivity analyses of disturbance analysis parameters may confirm additional model thresholds.

Similarly, data display and interpretation require caution. For example, the histogram display (e.g., Figure 19, Figure 20, Figure 22 and Figure 23) and interpretation of relative disturbance load as presented in this report, are based on quantiles, or the distribution of the data. This is apparent in Figure 20, where the distribution of road density within the area adjacent to wetland complexes are separated into 4 categories, but categories 1 through 3 are not particularly different from each other. In contrast, Figure 21 presents the same data as Figure 20, but uses Jenks natural breaks instead of quantiles. Neither of these displays use meaningful ecological thresholds, but the difference in data presentation may influence interpretation of the data. Additional research to identify ecologically meaningful thresholds is required to present disturbance impacts on wetland complexes if users seek to present data without relative scaling.

Standardization of disturbance by presence and density also carries other assumptions that relate to aggregation of wetlands into complexes. As density is an area-based calculation, the calculated threat density depends on the area of the wetland aggregate. Issues where the FWA stream connectivity layer either underestimated or overestimated stream connectivity will impact the area of both the wetland and the surrounding area, which can in turn impact density calculations.

Disturbance analysis results indicated that wetland aggregates and their surrounding upland area were less disturbed in the Finlay sub-basin. Both road and cut block count and density were generally higher in the Parsnip sub-basin. Although wetlands are typically not targeted for timber harvest, several wetland complexes are identified as having been harvested. Inspection of these areas is warranted as it is possible that timber harvest did occur on the wetlands or model prediction error resulted in a cutblock being identified as a wetland.

Patterns of road density are typically associated with timber harvest and cutblocks. However, decreased density in upland area when compared to associated wetland complexes is a data artefact that is in part related to the greater area of this upslope contributing area, which reduces the density of roads. However, percent wetland complexes with roads present is equal to percent upslope contributing area with roads present because the roads typically intersect with the wetland complexes when the roads occurred in the surrounding area. This pattern of road incursion into wetland complexes may be a result of wetlands occurring on level and gently sloping terrain that is favorable for road-building.

Cumulative impacts were not assessed in this work as the severity and scope of individual threats and impacts vary based on wetland group (bog, fen, marsh, swamp, and riparian ecosystems that were lumped into the wetland category), and thresholds of resilience and recovery trajectories of wetland groups are currently unknown. For example, roads may have hydrological impacts that directly impact wetland complexes, but exact thresholds for ecological change following altered hydrology patterns and sediment transport is unknown and likely to vary. Mountain pine beetle infestation may have variable impacts, where there is potential to have a high impact within the area adjacent to the wetland complexes with cascading impacts that influence wetlands (Potts 1984; Bearup *et al.*, 2014), but may not directly impact wetland vegetation if pine are not present. As well, multiple threats associated with resource extraction are likely to be correlated. Timber harvest in cutblocks is typically associated with roads and road construction. However, these threats are presented separately because thresholds for road density and type are currently unknown for wetlands, and interactions between these disturbance impacts are not clearly understood.

The ICUN threats calculator (IUCN, 2012) outlines a method that can help score cumulative impacts of threats, but quantified disturbance impacts need to be clarified. A count of disturbance presence was used instead of quantifying impacts from these landscape-level threats as assessing cumulative impacts to wetland complexes was outside of the scope of this project and requires significant additional work. Disturbances assessed for this analysis have a temporal aspect to their recovery. However, the disturbance date or intensity of disturbance was not assessed for this analysis. Proximity to disturbance may have impacts to wetland complexes, but only the direct area impact of disturbance was considered for this work. Subsequent analysis may benefit from consideration of temporal recovery trajectories and decreasing impact of disturbances associated with increasing distance. As well, this analysis does not include any climate change impacts as their impacts across the landscape are not clearly documented.

RECOMMENDATIONS

The following recommendations are provided as a guide for establishing priorities for future work, and as guidance for the appropriate use of products and analytical methods developed for this project. Not all recommendations may be within the scope or capacity of the FWCP or other partner organizations to implement. They are provided as options for consideration as when resources and capacity are available.

APPROPRIATE USE

The modelled predictions and the disturbance analysis results must be used in a manner appropriate to the scale, assumptions and other limitations of the data. Users must do their due diligence in familiarizing themselves with these limitations and ensure that appropriate expertise is sought when using the information to make decisions that will impact the environment and communities.

The wetland and riparian predictive maps were intended as a tool to inform prioritization of wetland and riparian actions. The disturbance analysis provides a method that can be used to analyze disturbance trends across wetlands and the surrounding landscape. In order to quantify and rank a wetland and/or riparian ecosystem for action prioritization, specific objectives, targets and thresholds need to be established. What thresholds separate high from moderate disturbance? What metrics and specific thresholds (e.g., road density > 5m/ha) makes an occurrence a good candidate for a particular action (e.g., restore of hydraulic connectivity)? These specifics are critical to a defensible analysis for wetland and riparian action prioritization.

Given the nature of a modelled prediction, there are several items to consider when determining the appropriate use of the modelled products:

- The model is built on the data provided for the FWCP Peace region (e.g., training data and environmental predictor variables), therefore it is inappropriate to extrapolate the model results outside of the region.
- Use the results from multiple models in conjunction with each other to identify areas of uncertainty and agreement. Where predictions disagree, lower confidence in the results can be inferred.

- Consider the size and configuration of a wetland. Given the resolution of the prediction (25x25 meter cell size), the model may not identify small and isolated wetlands or their boundaries as consistently or accurately as larger wetlands that are part of a wetland complex.

CONSIDER CONTEXT

Wetlands and riparian corridors in the landscape exist as complexes that are hydrologically connected. Wetland patterns and mosaics also influence how species and communities utilize the ecological services that wetland and riparian areas provide. As such, wetland and riparian complexes must be managed as hydrologically connected units and as ecological mosaics and corridors in the landscape. Some wetlands may be hydrologically isolated and may have different vulnerabilities and therefore, should be managed to account for this. Additionally, individual pixels from the model predication should be managed in the context of the wetland or riparian area they belong to. These areas, in turn, should be managed within the complex they belong to and within their landscape context. Important considerations in prioritization are the upslope conditions in the catchment area that may impact wetland or riparian areas, as well as what is downslope, which could be impacted by the quality and effectiveness of ecosystem services.

VERIFY MODEL RESULTS

When wetland/riparian complexes are identified for potential action, verification should take place to account for the limitations, reliability and inherit error in the modeled ecosystem locations and types. Figure 11 and Figure 12 can be used as a high-level indicator of model reliability based on training and field data distribution and densities.

In regions where training data are sparse or non-existent, the relationship between the prediction categories and environmental variables may not have been well captured by the model. As a result, areas with limited training data may be susceptible to increased prediction errors. Where training points are dense within the map sheet, greater confidence in modelled results can be inferred, while regions where training points are lacking should be approached with greater caution.

In this project, field collected data were used to validate model accuracy and therefore, if field data are not available for a given area these metrics are less informative. Consequently, model validation metrics are less applicable in areas where field data have not been collected.

Model comparison can also be completed in the Williston Wetland Explorer Tool (WWET) or other GIS software application. Model results can be compared through the WWET application using features such as the 'swipe tool' to easily compare models on a pixel by pixel bases or at the individual wetland level. Imagery, Freshwater Atlas layers and contours can be used to visually inspect predicted wetlands.

By comparing the 3 and 10-category model predictions to each other it is possible to establish a site-specific understanding of model agreement and disagreement. Where both 3 and 10-category models agree, a greater confidence in the predicted result can be inferred. However, if the models disagree then there may be less confidence applied to the modelled result and additional field site visits for the area may be necessary to verify model predictions. For example, where the 3-category prediction estimates wetland presence for a given area, and the 10-category prediction estimates upland, then the disagreement should be evaluated and noted as a region of reduced confidence.

These three assessment options can be used in conjunction with each other to asses model reliability. For example, where modelled results do not agree, there is limited (or zero) training data, and field data for validation have yet to be collected, then additional scrutiny should be placed in these areas. On the contrary, if results from separate models agree, and there are plenty training and field data, then results may be approached with greater confidence.

Additional observations of wetland and riparian ecosystems are made during the assessment of model results. Consequently, there is an opportunity to record those observations as data to inform future model improvement and decision making. For example, data collected though desktop exercises or exploration through the WWET tool can inform a Tier 1 desktop evaluation leading to the production of new training points. Training points, based on remotely sensed imagery, should be collected to corroborate or improve the prediction model.

If a field survey is completed, such as a Tier 2 wetland evaluation, field data should be collected in a consistent and machine-readable format so it can be consolidated and made available for use in future analysis as well as for model improvement and verification. Field observations can also be used to assess the assumptions of the disturbance analysis by verifying disturbance impacts on wetland and riparian ecosystems in the field.

CONSIDER SCALE

The model is limited by pixel size. The Digital Elevation Model (DEM) is provided at a 25m pixel resolution and Sentinel and Landsat bands incorporated into the model range from 10-60m. The climate data, however, changes more gradually across the landscape and the 800m data was resampled to match the smallest resolution of the input data – in this case, the 25m DEM. Therefore, even though the model makes predictions at the 25m pixel level it is better at detecting larger contiguous areas of wetland/riparian vegetation and conditions.

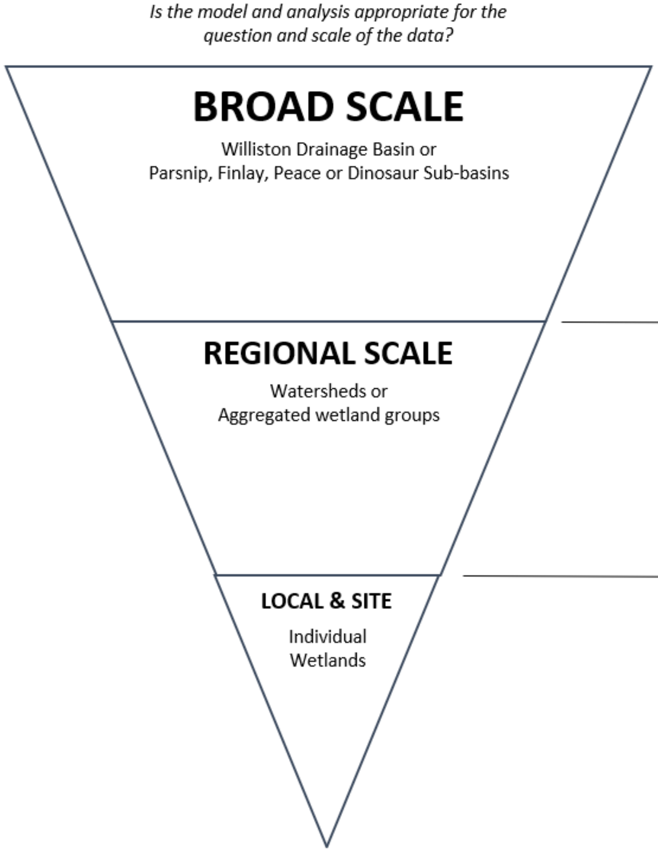
In addition, the model has been trained and statistically validated for the full FWCP Peace region. There is insufficient data in some areas to run the model at a more regional or local scale. As such, model results are best applied for management decisions at broader scales. This is achieved by analysing the data over larger areas such as the basin, sub-basins (4) or full watershed groups (15) within the FWCP Peace region and/or by aggregating connected wetland/riparian systems. This effectively averages out some of the flaws or inconsistencies in the model at finer scales.

The model can still inform decision making at a local scale, but model verification using available data and tools (e.g., the WWET application) as well as local wetland/riparian knowledge and expertise is required. Specifically, wetland and riparian expertise is needed to evaluate when the model is correct, incorrect and when additional field information is necessary. This is particularly true when using the 10-category results as it has a lower field validated accuracy.

The WWET application can aid in desktop validation of the model considering a range of site level data as well as landscape level context and conditions. Uses may include but are not limited to:

- Using imagery to confirm if an area is a wetland, riparian, transition or upland area;
- Using imagery and Freshwater Atlas information to asses connectivity;
- Visually assessing landscape attributes of the wetland, such as upslope catchment and disturbance;
- Using imagery, disturbance layers and wetland predictions to confirm the proximity or overlap or threats such as roads, fire and beetle kill. Investigating disturbance severity and assessing the potential impact of disturbance;
- Consulting field data and photos in areas that are similar to the area in question;
- Looking at time lapse imagery of the area to visualize changes to the feature and surrounding landscape over time.

A diagram is provided to provide examples related to the types of questions and considerations that can be addressed at various geographic scales (Figure 29). This diagram highlights the need to be aware of scale considerations when assessing model and analysis results.



Use the model and analysis methods to:

- Highlight where concentrations of a given value (threats or model categories) are distributed and identify trends
- Establish threshold values for species and value specific analyses
- Highlight where restoration and enhancement efforts would be best allocated
- Identify potential impacts of wetland degradation on downstream values
- Identify or forecast potential upstream activities that may impact wetland function

Usage Considerations:

- Connectivity between aggregated wetland groups
- Composition (mosaics or variety of types) infers what landscape level wetland functions?
- How does a wetland complex influence downstream values?
- What are potential impacts on the wetland from upstream threats?

Use the model and analysis methods to:

- Assess wetland threats or disturbance in a regional context
- Assess wetland hydrological connectivity
- Consider regional questions

Usage Considerations:

- Consider wetland size and configuration – is the wetland small and isolated or large and connected? What does this mean for potential threats?
- Is there appropriate field and training data densities to train and evaluate the model?
- Consider the scale at which the threat data were collected. Is it appropriate for regional analyses?

Use the model and analysis methods to:

- Decide between wetlands on where to focus limited resources

Use the field data to:

- Contribute to the continuous improvement of the model
- Set targets and thresholds using field work and research
- Collect site specific data to inform trend analysis at regional and local scales

Usage Considerations:

- Is the wetland too small to be accurately modeled given the resolution (25x25 m)?
- Is the classification reasonable when compared to imagery and training or field data?
- Do modelling predictions agree or are there discrepancies between predictions?

Figure 29: Example scale considerations for determining appropriate use of the model and analysis methods.

SET TARGETS, CRITERIA AND THRESHOLDS

Without specific targets and thresholds, the relative properties of wetlands or geographic areas can be used to qualitatively evaluate whether one area is a better candidate for a particular action than another. Landscape-level patterns of inventory data, wetlands/riparian areas, and disturbance/land-use can highlight areas needing research, requiring inventory and data collection, potential areas for conservation, or regions where mitigation measures are a priority. Establishing prioritization criteria is a critical step in building quantitative analysis and decision support tools for action prioritization. Analyses to support prioritization can also focus on a specific research question or focal target (e.g., species of interest) to help set thresholds and mitigate the impact of analysis assumptions.

Stratifying wetlands and wetland complexes based on relative disturbance counts and density within a specified area of interest (Tier 1 exercise) requires users to specify parameters and ecologically valid disturbance thresholds justified within the context of the research question or focal target under study.

DEVELOP DECISION TOOLS

The map product and analysis results presented in this report are most effective when included within a management framework (Cox and Cullington, 2009) and considered within the context of the Wetland and Riparian Action Plan. Collaboration and stakeholder engagement can help in building further decision tools that utilize the model and analysis methods and further connect the maps and methods to the decision-making process.

Figure 30 presents an example decision tree to guide users to actions including monitoring, restoration, enhancement, conservation, research and inventory. Wetland predictions could be used to evaluate wetland and riparian threats and values and help provide data to support decisions made at different points along the decision tree. As this decision tree is only provided as an example, further refinement to this diagram – or similar diagrams– would be beneficial. Stakeholder input would also be required to make this tool more detailed and applicable to a given decision making process.

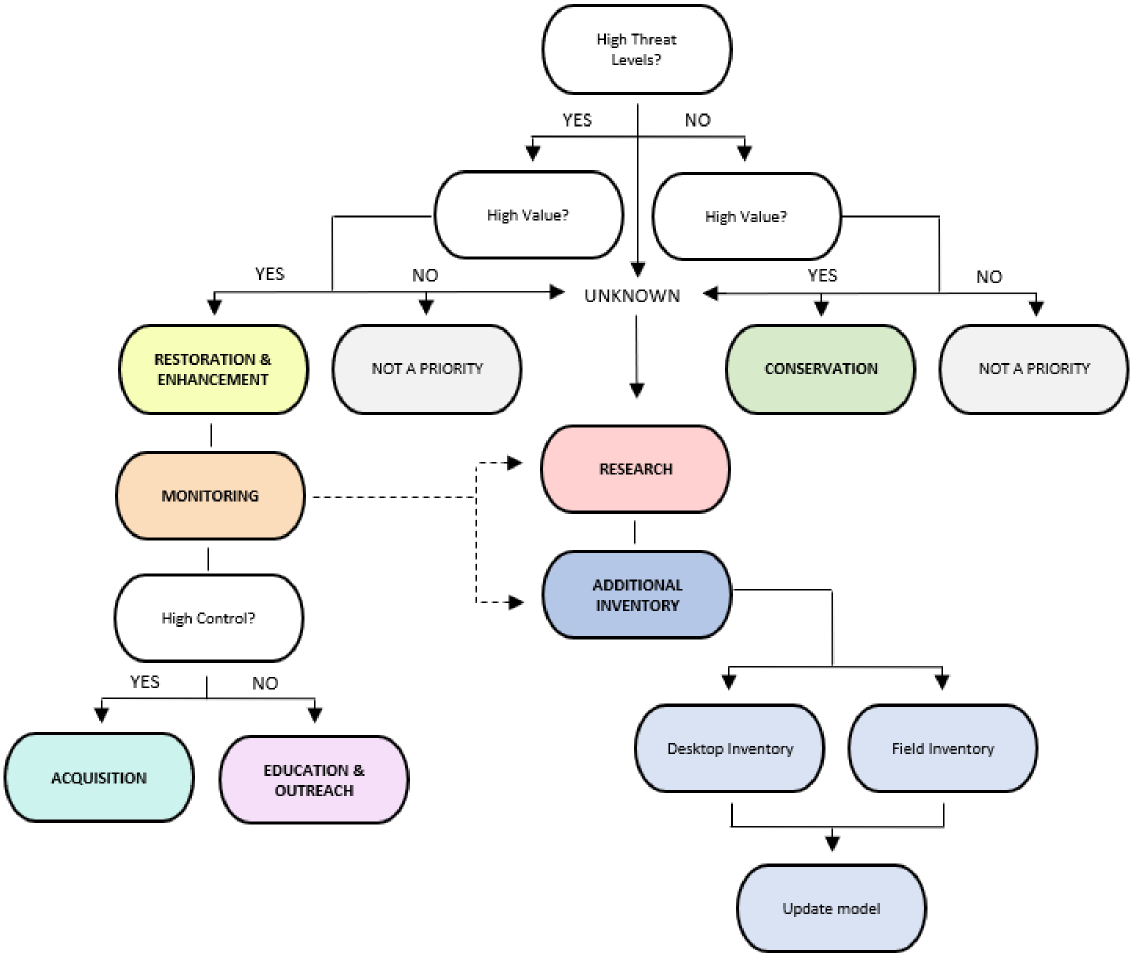


Figure 30: Example diagram of establishing priorities. For the purposes of this diagram, high control refers to the potential to purchase or acquire land (e.g., private land or other land type that may be purchased) effectively giving the purchaser complete control of the area. Conservation actions reference the option to implement conservation measures within the region of interest that is not necessarily owned by the user (e.g., private land that is not for sale or acquired, but agreement with landowner may allocate area for conservation purposes). Note, that what is considered high value needs to be defined by the user.

CONTINUOUS IMPROVEMENT CYCLE

Supporting a continuous model update cycle allows for incorporation of new technologies, data and input layers into periodic updates of the wetland/riparian predictions. This keeps wetland and riparian data for the FWCP Peace current and reliable. The scripting and documentation of the methodology makes each subsequent model update more efficient and cost effective.

IMPROVE WETLAND MODEL

IMPROVE SCRIPT PACKAGE

Improve the script package to integrate the model script with the disturbance analysis. The disturbance analysis would benefit from automation of data processing and analytical techniques. This will make re-running the model and disturbance analysis more streamlined, efficient and cost effective. Additionally, a well supported script and data package will promote the development of the methods and techniques across the modelling community where new advancements can, in turn, be applied to the FWCP Peace region. Scripting analysis procedures also establishes defensible and repeatable methods that can be regularly improved and adjusted (e.g., continuous improvement cycle) without high overhead costs or effort. For example, when new data are collected or become available, a script makes it possible to re-run the model and disturbance analysis with limited manual processing requirements or time commitments. There is also potential to formalize these methods as Provincial Standards for remote and poorly mapped areas.

Improve and add to the raster prediction layers. New, improved or higher resolution predictor layers, such as LiDAR or ensemble DEM data, will improve model resolution and predictions at finer scales. The addition of geographical information (e.g., x, y coordinates) could also improve local model results by accounting for directionality and trends in spatial patterns within the FWCP Peace. The model may also benefit from image derivatives, such as NDVI and time sequence data such as seasonal SENTINAL images, to pick up annual variations in vegetation cover. These time sequence data sources may also be valuable sources of seasonal fluctuations in water table, species composition and plant phenology.

EXPLORE OTHER WETLAND CATEGORIES

Explore predicting other wetland and riparian related classifications such as hydrogeomorphic class and subclass, hydrodynamic classes, or soil properties such as organic/mineral, etc. These attributes were collected in both the training and field validation data and may be more successfully predicted than the 10-category wetland types (e.g., bog, fen, march, swamp, flood units). These alternative categories may also be a better grouping for predicting wetland response, vulnerabilities, and resilience to climate change, than the wetland classes. The model does poorly at predicting the current wetland classes partially because they represent a site level classification more than a mapping concept. These other suggested classifications may also be easier to evaluate, both on the ground and in the field, as their criteria are not based on vegetation identification and abundance which require expertise and have seasonal challenges related to phenology. Modelling wetlands using alternative attributes such as hydrogeomorphic classes have been proven successful, as demonstrated by Deventer *et al.* (2014) who used a hydrogeomorphic classification to successfully model wetland function at a landscape level.

In addition, wetland patterns and complex descriptors such as string fen, infilling bog; patterns such as striped, concentric rings, patchy; or metrics such as patch size, are also important descriptors of wetland complexes that are relevant to ecosystem function, such as habitat. To inform prioritization, model outputs that predict patterns or complex descriptors could be used for patch analysis, for example with thresholds for patch size, distance between patches and other landscape ecology metrics.

CONSIDER GEOGRAPHY

Some portions of the FWCP Peace region have poor coverage of training and field data, resulting in unknown accuracy and potential model errors. Analyses can be run to highlight areas requiring additional training and validation data. It is desirable to have both coverage of geographic space and of the covariate space (the range of variation in the FWCP Peace represented by the prediction layers). These analyses are also more statistically sound representations of model uncertainty. Two potential analyses include multivariate environmental similarity surface (MES) and quantification of novel uni- and multi-variate environments. These analysis methods quantify the degree of dissimilarity in areas of model extrapolation compared to areas (or points) used for training the model (Mesgaran *et al.*, 2014). Another research question relates to densities and distribution of training and validation data. When do we have enough data? What accuracy thresholds are needed for different decisions and scales?

It is also recommended that future wetland and riparian predictions use the watershed groups or other watershed-based boundaries. Administrative boundaries have inconsistencies related to height of land, water flow and connectivity. The study area boundary used had edge effects where portions of the boundary did not accurately follow the height of land and other areas like in the Upper Peace River and Murray River where the boundary captures significant areas that are small portions of adjacent watersheds (Figure 28 – Area 1 and Area 2). The raster predictor layers that are effective in the model are related to the flow and accumulation of water through the landscape. Using portions of watersheds is not practical. In addition, when analyzing wetland connectivity and creating summaries, watershed polygons are enduring, consistent spatial entities in the landscape that make a practical and logical choice for an analysis unit. It is incorrect to summarize parts of watersheds and then compare them to intact watersheds.

SUPPORT TARGETED FIELD WORK

Field forms developed for this project are a starting point for standardization of field data collection. Continue to partner with the Wetland Stewardship Partnership on improvements to field data collection and standardization. Support work on digital data collection and open data sharing of collected information to add to the collective body of knowledge about wetlands and riparian areas in the FWCP Peace. Support repositories of data that users can actively add to with digital data collection tools such as the ones field tested with this project.

Update wetland actions to drive and support wetland and riparian field work. Identify opportunities for feeding new data collection back into the model and analyses to continually improve these decision-making tools. Look for cross-value wetland and riparian data collection opportunities in remote and hard to access areas. Consider specifically requiring reconnaissance data collection when other actions are funded in wetland and riparian areas (such as amphibian surveys). The results of statistical analysis of covariate space in the prediction layers can be used to inform and optimize field data collection to improve model reliability. Adding training data map sheets in underrepresented areas such as the north eastern high elevation portions of the upper Finlay will fill known data gaps.

Ensure collected data gets included into new iterations of the model and feeds into provincial ranking and classification efforts. Identify areas of low reliability to target for potential field work. Partner and collaborate in areas of remote and poor access. Additionally, collect data specific to improving flaws and deficiencies in the model as they are discovered. Model training data and field validation points require sample planning that includes a degree of randomness for them to be statistically sound. However, opportunistic field observations in rare and uncommon ecosystems when they are discovered can improve the model, ecosystem ranking and classification. Known data deficiencies include sampling of:

- Riparian and flood ecosystems
- Swamps
- Watershed groups with low field point count
- Alpine wetlands (also have limited provincial data to inform the alpine classification)
- Calcareous wetlands (not currently captured in the wetland classification due to lack of data)

SUPPORT RESEARCH AND ANALYSIS

Research and analysis are needed to improve both landscape and site level understanding of wetland and riparian systems. Both are critical to setting targets and thresholds to support the wide variety of ecosystem services provided by these valuable and vulnerable ecosystems.

DEVELOP A RIPARIAN AREAS MODEL

A clear definition of riparian ecosystems, additional training points and sampling design specific to patterns of riparian distribution are needed to truly address riparian areas mapping. Riparian and flood units have different drivers than wetlands. Floodplain units and riparian fringes require predictor variables that address their unique characteristics. Layers such as height above channel and slope/distance cost surfaces are important predictors of riparian ecosystems (Dilts *et al.*, 2010). Attributes relating to flow power, magnitude and frequency of flooding, are also important factors in riparian ecosystem form, function and management. These are attributes that may be predicted by a similar machine learning approach. Lidar canopy height models and finer resolution Lidar derived DEM will also improve riparian vegetation and landform detection for both remote identification of riparian areas and machine learning predictions (Arroyo *et al.*, 2010).

Running a separate model for riparian units and combining with a 3-category (upland, water, wetland) prediction will likely produce an improved wetland and riparian map. A riparian specific model will capture riparian (flood units) better on a local scale. In addition, using LiDAR for wetlands and floodplain unit prediction will greatly improve reliability and accuracy. Robinson (2017) found that modelling riparian areas using LIDAR "terrain-based modeling techniques offer greater accuracy than ubiquitous buffering and manual delineation methods" (pg., 52).

MONITOR FOR BASELINE CONDITION AND TRENDS

In order to address wetland and riparian trends as outlined in Objective 1 in Wetland and Riparian Action Plan, baseline conditions need to be collected as well as time sequence data. Detailed wetland description and monitoring of trends are needed to better understand climate change impacts, disturbance thresholds, hydrologic change, and transitional wetlands. Identify, describe and monitor rare ecosystems to support classification and improved understanding and management of these occurrences.

Field assessments, like the Tier 2 wetland assessment, can be used to evaluate current wetland condition. Use repeat assessments to establish trends and seasonal variations. Drone and camera surveys can also be used to measure and quantify ecosystem seasonal variations and change through time.

In order to monitor the success of wetland actions an understanding of the natural baseline conditions, seasonal trends and long-term trends is needed. The wetland map can be used to establish a network of benchmark wetlands for detailed assessment and monitoring. It is important to design sampling to capture the variety of wetland and riparian ecosystems in the FWCP Peace region. Common and uncommon, connected and non-connected, large and small, bogs vs. fens, all these different types will have different trends and responses. This will allow for paired ecosystem studies for treated and non-treated wetlands. It also allows for comparisons of ecosystem response relative to natural variations, ranges and trends in the region.

Long-term monitoring studies should include piezometers with continuous data logging over multi-season and multi-year time periods to provide valuable information about wetland and riparian hydrology. Baseline hydrograph information collected across a variety of wetland types will allow for change monitoring like the study conducted by Halabisky *et al.* 2016.

Sentinel imagery provides a record of seasonal variations (from 2015 to present)– and in combination with Landsat imagery (from 1970s) a time series analysis could illuminate a long-term view of trends and changes at the landscape level.

CONSIDER SOILS AND SUBSTRATE

The riparian and wetland analysis would benefit from more information relating to substrate. Surficial geology, terrain and aquifer mapping would also improve our understanding of subsurface hydrologic controls on wetland and riparian ecosystems and the services they provide. Wetlands that are on deep surficial materials and have subsurface flow through alluvial (river deposits) materials will have different hydrograph responses than those that are surface flow dominated.

Soils analysis is critical to evaluate carbon storage and dynamics in the range of wetland and riparian ecosystems. The wetland map can inform sample planning and can be used to extrapolate and calculate carbon storage in wetland and riparian systems. Soils and water analysis can also provide information about water quality and nutrient cycling. Monitoring in priority wetlands can be used to improve understanding of temporal changes in soil nutrients, gas exchange, soil temperature and more. These in turn can be used to set thresholds and monitor mitigation efforts against specific targets.

REFINE METHODS FOR WETLAND CONNECTIVITY ANALYSIS

The methods followed for aggregating wetlands into polygons and then into connected wetland complexes were an exploratory first approximation. This task requires a systematic and repeatable scripted approach. Modelling of hydrogeomorphic classes, suggested in section Other Classifications, may improve the assignment of connected and non-connected wetlands and may better reflect wetland function and resilience. In addition, updated riparian connectivity layers or a riparian specific model may also produce more reliable defensible layer to approximate riparian connections. Additionally, the use of raster analysis techniques may improve the workflow for connectivity analysis.

Analysis and research related to hydrology and groundwater connectivity in wetland and riparian ecosystems is needed, both from a landscape-level mapping perspective and for baseline monitoring at the site-level. Detailed hydrogeomophic as well as base line hydrograph data is necessary to better understand groundwater relationships.

REFINE METHODS FOR DISTURBANCE ANALYSIS

Refinement and scripting of the disturbance analysis is critical to a defensible repeatable method for prioritization of wetland and riparian actions at the regional (watershed group) and local scales. Scripted processes will facilitate sensitivity analysis to different parameters such as buffer distances. It will also make analysis to address specific values or actions more efficient. It will allow for re-running of the analysis for prioritization of wetland action at finer resolutions where data availability, accuracy and scale allow. A reproducible scripted process will also allow for the incorporation of new research and findings into prioritization of wetland actions. It will facilitate analysis updates when the wetland and riparian models are rerun. Understanding of wetland trends and potential threats can be estimated by using analysis scripts and including climate models, time sequence data and proposed land use to explore different scenarios.

Use of raster methods should be explored for computational efficiency and to overcome vector data inaccuracies and differences in scale. Open source tools and packages should be explored to facilitate open code for collaboration and sharing.

DEVELOP ANALYSIS FOR ECOSYSTEM SERVICES

Proposed decision trees presented in this report highlight the need for an evaluation of relative wetland and riparian ecosystem value. Ecosystems that provide critical services, those that provide multiple services, and those that support values identified in the action plan, should be prioritized. Research and field surveys are needed to measure and quantify ecosystem function and service provision relating to carbon storage, biodiversity, water quality, habitat supply, drought and flood buffering, etc. In addition, landscape level analyses that evaluate cumulative ecosystem services are needed to highlight critical wetlands and hotspots in the landscape for action. Measures of relative ecosystem rarity and resilience can further inform prioritization. The action plan should be updated to reflect these needs.

Model confusion may indicate areas where unique ecosystems or unique wetland complexes are present. Quantification of novel uni- and multi-variate environments analysis identifies areas of unique or unsampled environmental conditions. This may also highlight were unique ecosystems could be field sampled.

SUPPORT EDUCATION AND OUTREACH

Research and analysis are needed to improve both landscape and site level understanding of wetland and riparian systems. Both are critical to setting targets and thresholds to support the wide variety of ecosystem services provided by these valuable and vulnerable ecosystems.

DELIVER WORKSHOPS AND PRESENTATIONS

Continue to support and partner on workshops opportunities, including continued collaboration with the Wetland Stewardship Partnership. Some specific workshop needs include:

- Development of guidelines and extension materials for forestry professionals relating to identifying wetlands in the field. Tips for identifying wetlands during winter logging would be beneficial.
- Developing best practices and guidelines for wetland buffers.
- A workshop to further develop decision trees and tools for wetland prioritization. This could potentially be paired with training on the use of the WWET application.
- Field tours and workshops to improve out collective understanding of wetlands and riparian areas. Providing opportunities for multidisciplinary teams in the field to facilitate collaborative learning.

Identify and support opportunities for presentations and outreach relating to the wetland and riparian predictive model and WWET application to a variety of stakeholders and partners.

CONTRIBUTE TO RESOURCE MATERIALS

Correct identification of wetlands and riparian ecosystems both through image interpretation and on the ground is important as maps are not 100% accurate. Educating map users and field crews on wetland recognition and aspects of Tier 1 and 2 wetland assessments is required. Identifying high value ecosystems at the site level and providing education on how to identify important indicators on the ground will aid in more consistent and accurate identify wetlands in the field.

Support production of guides for identifying wetland type (including what disturbance looks like). Use field photos and site inventory data to describe wetland characteristics especially those unique to the FWCP Peace region. Field photos and data can but used to develop a library of ecosystem descriptions for the FWCP Peace region. Collaborate and support the Wetland Stewardship Partnership in these tasks.

MAKE DATA AND INFORMATION AVAILABLE AND ACCESSIBLE

The WWET application (https://catalogue.data.gov.bc.ca/dataset/1d174499-72af-41d5-ad5e-19460b1e6fb1) can be used as a source of information to improve wetland and riparian ecosystem knowledge, identification and evaluation. It currently provides contextualized maps of ecosystem distribution as well as field data and photographs to provide detailed site-specific information. Future additions to the WWET application may include detailed descriptions of benchmark ecosystems to aid in site identification of wetland/riparian ecosystem type (including rare and high value indicators).

Continue to make new field data open and available to add to the body of knowledge about local ecosystems and to input back into the continuous improvement cycle of the model.

Make monitoring data open and available so that it can be used to describe the range of conditions in natural wetland and riparian ecosystems. This information is critical in the setting of thresholds and targets.

USE FINDINGS TO INFORM PRIORITY ACTION

At the time of writing this report, The FWCP Peace region is updating its Riparian and Wetland Action Plan available for distribution in August 2020. Some potential new actions for FWCP Peace region to consider are provided below as a result of this study. Some of the ideas presented here may be outside the scope or capacity for the FWCP Peace Region to implement and should be considered with that context.

The current actions outlined within the Riparian and Wetland Action Plan could be expanded to address research and analysis needs directly related to wetlands and riparian areas. Research is needed to quantify ecosystem services provided by wetland and riparian ecosystems in the FWCP Peace region including, but not limited to, carbon storage and cycling; drought and flood buffering; water quality; habitat supply; biodiversity; peaceful enjoyment. Research is needed to improve our knowledge of site level factors that control wetland type and function. These may also influence ecosystem resilience, trends and transition related to climate change, hydrolic change and prediction of future wetland trends. Research is also needed to establish prioritization criteria and thresholds to inform action prioritization.

Another need is the collection of additional field information and inventory to continue to improve the model and analysis methods. Further field verification of analysis and model assumptions should be covered under the actions, for example, to answer the question of whether the disturbance analysis results reflect real conditions on the ground?

As suggested in the *Monitor for Baseline Conditions and Trends* section, the action plan could support the establishment of a network of long-term monitoring locations to establish baseline and trend information representative of the range of wetland and riparian ecosystems in the FWCP Peace Region. Additionally, existing actions relating to collection of local and traditional knowledge could be extended to cover the use of wetland and riparian ecosystems related to recreational and peaceful enjoyment use. Wetlands are gathering and teaching places that have intrinsic values not tied to species of interest.

ACKNOWLEDGEMENTS

We would like to acknowledge the Fish and Wildlife Compensation Program for the funding and support to further wetland and riparian understanding of the area, and the Province. We would also like to acknowledge the following for their important contributions:

Chelsea Coady, FWCP Peace Region Manager for project support and guidance.

Ecora, EDI, Madrone, Blackwell (2017): For their efforts in assessing data layers and mapping methods to improve the wetland prediction process.

Blackwell Inc.: For the Fieldwork planning, ecosystem mapping, reporting, (2017) and continued GIS support Ben Andrew for report editing and review (2017-2018) and ecosystem mapping (2019).

Chu Cho Environmental (2017): For the local area knowledge and technical expertise.

Simon Casley (2018-2019): Script and R package development

Field crew: Deepa Filatow (2017-2019), Tyson Carswell (2017-2018), Nyssa Temmel (2017-2018), Emily Cameron (2017-2019), Adrian DeGroot (2017), Kirk Miller (2017), Gillian Harvey (2019), Shawn Prince (2019), James Wolf (2019), Bruce Rogers (2019), Whitney Swyers (2019), Stephanie Hurst (2019), Will Mackenzie (2019), Traci Van Spengen (2019), Chelsea Coady (2019), Kiley Jackson (2019).

Corey Erwin for original project proposal and ongoing support.

Alex Bevington, Marten Geestema and Brendan Miller for equipment support.

REFERENCES

Ameli, A. and Creed, I. 2017. Quantifying Hydrologic Connectivity of Wetlands to Surface Water Systems. *Hydrological and Earth Science Systems* 21: 1971-1808.

Arroyo, L., Johansen, K., Armston, J., Phinn, S., 2010, Integration of LiDAR and QuickBird imagery for mapping riparian biophysical parameters and land cover types in Australian tropical savannas, https://doi.org/10.1016/j.foreco.2009.11.018

Ayala-Izurieta, J.E., Márquez, C.O., García, V.J., Recalde-Moreno, C.G., Rodríguez-Llerena, M.V., Damián-Carrión, D.A. 2017. Land Cover Classification in an Ecuadorian Mountain Geosystem Using a Random Forest Classifier, Spectral Vegetation Indices, and Ancillary Geographic Data. *Geosciences* 7(2): 34.

Barnett, I. 2004. Wetland Action Plan for British Columbia. T.D. Hooper, editor. Proceedings of the Species at Risk 2004 Pathways to Recovery Conference. March 2–6, 2004, Victoria, B.C

Basiliko, N., Moore, T.R., Jeannotte, R., and Bubier, J.L. 2006. Nutrient Input and Carbon and Microbial Dynamics in an Ombrotrophic Bog. *Geomicrobiology Journal* 23(7): 531-543.

Batzer, D., Pusatei, C., and Vetter, R. 2000. Impacts of Fish Predation on Marsh Invertebrates: Direct and Indirect Effects. *Wetlands* 20(2): 307-312. https://doi.org/10.1672/0277-5212(2000)020[0307:IOFPOM]2.0.CO;2

Bearup, L.A., Maxwell, R.M., Clow, D.W., and McCray, J.E. 2014. Hydrological effects of forest transpiration loss in bark beetle-impacted watersheds. *Nature and Climate Change* 4: 481–486.

Belgiu, M., Dragut, L., 2016. Random forest in remote sensing: A review of applications and future directions. *Journal of Photogrammetry and Remote Sensing*. v114, pp 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Berhane, T.M., Lane, C.R., Wu, Q., Autrey, B.C., Anenkhonov, O.A., Chepinoga, V.V. and Liu, H. 2018. Decision-Tree, Rule-Based, and Random Forest Classification of High-Resolution Multispectral Imagery for Wetland Mapping and Inventory. Remote Sensing 10 (508): 26pp. doi:10.3390/rs10040580

Blischak, J.D., Davenport, E.R., and Wilson, G. 2016. A Quick Introduction to Version Control with Git and GitHub. *PLoS Comput Biol* 12(1): e1004668. doi:10.1371/journal.pcbi.1004668

Breiman, L. 2001a. Modeling: The Two Cultures. *Statistical Science* 16(3):199-215.

British Columbia Ministry of Forests and Range and British Columbia Ministry of Environment. 2010. Field manual for describing terrestrial ecosystems. 2nd ed. Forest Science Program, Victoria, B.C. Land Manag. Handb. No. 25. Accessed online, www.for.gov.bc.ca/hfd/pubs/Docs/Lmh/Lmh25-2.htm

Cadrin, C. and Christy, J. 2013. Element Occurrence Specification for Small and Medium Patch Wetlands. British Columbia Conservation Data Centre. Draft.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Whehberg, J. *et al*. 2015. System for Automated Geoscientific Analyses (SAGA) v 2.1.4, Geosci. Model Dev., 8, 1991-2007, doi:10.5194/gmd-8-1991-2015

Cook, B.J., and Hauer, F.R. 2007. Effects of hydrologic connectivity on water chemistry, soils, and vegetation structure and function in an intermontane depressional wetland landscape. *Wetlands* 27: 719–73.

Corcoran, J.M., Knight, J.F., and Gallant, A.L. 2013. Influence of Multi-Source and Multi-Temporal Remotely Sensed and Ancillary Data on the Accuracy of Random Forest Classification of Wetlands in Northern Minnesota. Remote Sensing 5: 3212-3238 doi:10.3390/rs5073212

Cox, R., and Cullington, J. 2009. Wetland Ways: Interim Guidelines for Wetland Protection and Conservation in British Columbia. Wetland Stewardship Partnership, Government of British Columbia. Accessed online, https://www2.gov.bc.ca/gov/content/environment/air-land-water/water/water-planning-strategies/wetlands-in-bc

Deventer, H.V., Nel, J., Maherry, A., and Mbona, N. 2014. Using the landform tool to calculate landforms for hydrogeomorphic wetland classification at a country wide scale. South African Geographic Journal 98(1):138-153. https://doi.org/10.1080/03736245.2014.977812

Devito, K.J., Hill, A.R., and Roulet, N. 1996. Groundwater-surface water interactions in headwater forested wetlands of the Canadian Shield. Journal of Hydrology 181:127-147.

Dilts, T. E., Yang, J., & Weisberg, P. J. (2011). Mapping Riparian Vegetation with LiDAR Data: Predicting Plant Community Distribution Using Height above River and Flood Height. ArcUser. Winter 2010. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.620.4762

Euliss, N.H., Smith, L.M., Wilcox, D.A. and Browne, B.A. 2008. Linking ecosystem processes with wetland management goals: Charting a course for a sustainable future. *Wetlands* 28: 553-562. https://doi.org/10.1672/07-154.1

Faber-Langendoen, D., Hedge, C., Kost, M., Thomas, S., Smart, L., Smyth, R., Drake, J., and Menard, S. 2012. Assessment of wetland ecosystem condition across landscape regions: A multi-metric approach. Part A. Ecological Integrity Assessment overview and field study in Michigan and Indiana. EPA/600/R-12/021a. U.S. Environmental Protection Agency Office of Research and Development, Washington, DC.

Freeman, M.C., Pringle, C. M., and Jackson, C.R. 2007. Hydrologic Connectivity and the Contribution of Stream Headwaters to Ecological Integrity at Regional Scales. *Journal of the American Water Resources Association* 43(1): 5-14.

Freeman, E and Frescino, T. 2009. ModelMap: Modeling and Map production using Random Forest and Stochastic Gradient Boosting. USDA Forest Service, Rocky Mountain Research Station, 507 25th street, Ogden, UT, USA.

Filatow, D., Carswell, T., and Cameron, E. 2018. Predictive Wetland Mapping of the Williston Drainage Basin (Report ID: 54822). Downloaded from a100.gov.bc.ca/pub/acat/public/viewReport.do?reportId=54822

FWCP (Fish & Wildlife Compensation Program). 2014. Peace Basin Riparian and Wetlands Action Plan. 17 p. Available online at: http://fwcp.ca/app/uploads/2015/07/fwcp-peace-riparian-and-wetlands-action-plan-march-31-2014.pdf.

Gislason, P., Benediktsson, J., Sveinsson, J. 2006. Random Forests for land cover classification. Pattern Recognition Letters. 27:4. https://doi.org/10.1016/j.patrec.2005.08.011

Golder Associates. 2010. Peace River Project Water Use Plan; Williston Reservoir Wetlands Inventory (GMSWORKS-16). Golder Associates Ltd., Kamloops BC. project number 09-1430-0057

Government of British Columbia. 2002. British Columbia Specifications and Guidelines for Geomatics: Gridded Digital Elevation Model Product Specifications. Available online, https://www2.gov.bc.ca/assets/gov/data/geographic/topography/elevation/bc_dem_products_specs.pdf

GRASS Development Team. 2018. Geographic Resources Analysis Support System (GRASS) Software, Version 7.8. Open Source Geospatial Foundation. http://grass.osgeo.org

Halabisky, Meghan & Moskal, L. & Gillespie, Alan & Hannam, Michael. (2016). Reconstructing semi-arid wetland surface water dynamics through spectral mixture analysis of a time series of Landsat satellite images (1984–2011). Remote Sensing of Environment. 177. 171-183. 10.1016/j.rse.2016.02.040. https://doi.org/10.1016/j.rse.2016.02.040

Hauer, R.F., Locke, H., Dreitz, V.J., Hebblewhite. M., Lowe, W.H., Muhlfeld, C.C. *et al*. 2016. Gravel-bed river floodplains are the ecological nexus of glaciated mountain landscapes. Science Advances 2(e1600026). doi:10.1126/sciadv.1600026

Herrmann, H.L., Babbitt, K.J., Baber, M.J., and Congalton. R.G. 2005. Effects of landscape characteristics on amphibian distribution in a forest-dominated landscape. Biological Conservation 123(2):139-149.

Houlahan, J. E., and Findlay, C.S. 2003. The effects of adjacent land use on wetland amphibian species richness and community composition. Canadian Journal of Fisheries and Aquatic Sciences 60:1078-1094.

IUCN. 2012. Threats Classification Scheme Version 3.2

Kingsford, R.T., Brandis, K., Thomas, R.F., Crighton, E., Knowles, E., and Gale, E. 2004. Classifying landform at broad spatial scales: the distribution and conservation of wetlands in New South Wales, Australia. Marine and Freshwater Research 55:17-31. https://doi.org/10.1071/MF03075

Kreutzweiser, D., Good, K., Capell, S., and Holmes, S. 2008. Leaf-litter decomposition and macroinvertebrate communities in boreal forest streams linked to upland logging disturbance. Journal of the North American *Benthological Society* 27(1):1-15. 2008 https://doi.org/10.1899/07-034R.1

MacKenzie, W., and Moran, J. 2004. Wetlands of British Columbia: a guide to identification. Res. Br., B.C. Min. For., Victoria, B.C. Land Manage. Handb. No. 52

Mackenzie, W. 2012. Biogeoclimatic ecosystem classification of non-forested ecosystems in British Columbia. Prov. B.C. Victoria B.C. Tech. Rep. 068.

Meidinger, D., 2003, Ecosystem Mapping Accuracy and Timber Supply Applications Research Branch, BCFS July 2003. https://www.for.gov.bc.ca/hre/becweb/Downloads/Downloads_PEM/PEMaccuracystatement2003.pdf

Mellina, E. 2002. Stream temperature responses to clearcut logging in the central interior of British Columbia: test of the predictive model developed by Mellina *et al*. (2002). Natural Resources Canada.

Moore, D., and Wondzell, S. 2005. Physical Hydrology and the Effects of Forest Harvesting in the Pacific Northwest: A Review 1. *Journal of the American Water Resources Association* 41(4): 763-784.

National Research Council. 1995. Wetlands: Characteristics and Boundaries. Washington, DC: The National Academies Press. Accessed online, https://doi.org/10.17226/4766

The Nature Trust of British Columbia. July 2019. BC Wetland and Riparian Ecological Value Version 1.0: BC Riparian Areas 2019. Spatial Data, UNPUBLISHED.

Neff, B.P., and Rosenberry, D.O. 2018. Groundwater Connectivity of Upland-Embedded Wetlands in the Prairie Pothole Region. *Wetlands* 38:51–63.

Perez-Riverol, Y., Gatto, L., Wang, R., Sachsenberg, T., Uszkoreit, J., Leprevost FdV, *et al.* 2016. Ten Simple Rules for Taking advantage of Git and Git Hub. PLoS Comput Biol 12(7): e1004947. doi:10.1371/journal.pcbi.1004947.

Peterson, K. 2013. The GitHub Open Source Development Process. Technical Report. Rochester, MN: Mayo Clinic.

Phillips, R.L., Berri, O., and DeKeyser, E.S. 2005. Remote wetland assessment for Missouri Coteau prairie glacial basins. Wetlands 25:335-349. doi: https://doi.org/10.1672/10

Polzin, M.L. 1998. River and riparian dynamics and black cottonwoods in the Kootenay River Basin, British Columbia and Montana (Unpublished master's thesis). University of Lethbridge, Alberta, Canada. Accessed online: https://opus.uleth.ca/bitstream/handle/10133/88/MQ38451.pdf

Potts, D.F. 1984. Hydrologic Impacts of a Large-Scale Mountain Pine Beetle (Dendroctonus ponderosae Hopkins) Epidemic. Journal of the American Water Resources Association 20(3):373-377.

Powers, S.M., Johnson, R.A. and Stanley, E.H. 2012. Nutrient Retention and the Problem of Hydrologic Disconnection in Streams and Wetlands. Ecosystems 15:435–449.

Professional Engineers and Geoscientists of BC. 2017. Flood Mapping in BC: APEGBC Professional Practice Guidelines V1.0 – Natural Hazards. The Association of Professional Engineers and Geoscientists of British Columbia, Burnaby, BC, Canada. Accessed online: https://www.egbc.ca/getmedia/8748e1cf-3a80-458d-8f73-94d6460f310f/APEGBC-Guidelines-for-Flood-Mapping-in-BC.pdf.aspx

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Robinson, T.S. 2017. Mapping ecologically functional riparian corridors using LiDAR and hydrologic landscape analysis (Unpublished master's thesis). San Francisco State University, San Francisco. Accessed online: https://geog.sfsu.edu/sites/default/files/thesis/RobinsonThesisFinal.pdf

Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sanchez, J. 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogramm. Remote Sens. 67: 93–104.

RStudio Team. 2015. RStudio: Integrated Development for R. RStudio Inc., Boston, MA. http://www.rstudio.com/

Timm, B.C and McGarigal, K. 2012. Fine-scale remotely-sensed cover mapping of coastal dune and salt marsh ecosystems at Cape Cod National Seashore using Random Forests. *Remote Sensing of the Environment* 127: 106-117.

Trombulak, S., and Frissel, C. 2000. Review of Ecological Effects of Roads on Terrestrial and Aquatic Communities. *Conservation Biology* (14)1: 18-30.

United States National Coordination Office for Space-Based Positioning, Navigation, and Timing. December 2017. Accessed online: https://www.gps.gov/systems/gps/performance/accuracy/

Wang, T., Hamann, A., Spittlehouse, D., and Murdock, T.N. 2012. ClimateWNA - High-resolution spatial climate data for western North America. *Journal of Applied Meteorology and Climatology* 61: 16-29.

Winter, T. 1999. Relation of streams, lakes, and wetlands to groundwater flow systems. *Hydrogeology Journal* 7: 28–45.

Wolf, K.L., Noe, G.B., and Ahn, C. 2013. Hydrologic Connectivity to Streams Increases Nitrogen and Phosphorus Inputs and Cycling in Soils of Created and Natural Floodplain Wetlands. *Journal of Environmental Quality* 42(4): 1245-1255.

Yang, Y., Yang, Z., Yin, X., and Qiang, L. 2018. A framework for assessing flow regime alterations resulting from the effects of climate change and human disturbance. *Hydrological Sciences* Journal 63(3): 441-456.

Zhang, S., Lu, X., Zhang, Y., Nie, G., and Li, Y. 2019. Estimation of soil organic matter, total nitrogen and total carbon in sustainable coastal wetlands. Sustainability 11(3), iss 667. https://doi.org/10.3390/su11030667

APPENDIX A. PROJECT TIMELINE

TASKS	TIMELINE	PURPOSE	DELIVERABLE/OUTCOME
Pilot Projects	2015-2017	Evaluate feasibility of existing Ministries data to identify and map FWCP Peace Wetlands Evaluate feasibility of predictive wetland mapping in FWCP using limited number of map sheets	1. Pilot project to evaluate mapping methods 2. Predicting location and wetland types possible 3. Identified and assembled data required 4. Identified relevant topographic mapping units for the region & prediction 5. Produced 12500 expert-selected site identifications & evaluated consistency
Model Development	2017	Assemble and modify Random Forest modelling product for wetland-specific outputs	1. Source-code assembly 2. Data formatting for input 3. Modifying output parameters to map wetland units
Fieldwork	August 2017	Put experts in the field to collect site-specific information; Identify possible wetlands sites that are not currently represented in Provincial Wetland Guide, with the possibility of these being Red/Blue listed due to rarity; Collect data for model validation and accuracy assessments.	1. Collected 273 site descriptions and biophysical measurements 2. Worked with regional First Nations professionals 3. Identified unique wetland sites not currently described in the Provincial Wetlands Guide standards 4. Data collected useful towards model accuracies in certain areas 5. Worked with local stakeholders (e.g. BCWF) to use and evaluate new wetland field forms
Predictive Wetland, and riparian mapping	2017	Map the entirety of FWCP Peace using Random Forest model; Provide the location and probable site wetland types at 25m resolutions	1. Modelled wetland locations for ~70000km^2 2. Approach follows Provincial Wetland Guide
Phase 1 Final Report	March 31st, 2018	Report summarizing work to date; how it addresses Riparian and Wetland Action Plan; and Information gaps and appropriate next steps	1. Mapping and field data in agreed-upon GIS formats (GeoTIFF, geodatabase, others). Data made available for access by stakeholders/interested parties 2. Value-added ancillary summary products 3. Phase 1 Final Report

Fieldwork 2018	Aug 2018	Experts in the field to record observation wetland classes, upland and water features; Collect field observations in regions of model confusion and	1. Collected 274 site descriptions and biophysical measurements
Model script development	January 2019 – July 2019	Publish open source mapping package on GitHub	1. R package development to automate predictor variable generation, model run and prediction map. 2. Script modification to include iterative capabilities (e.g., iteratively run model through multiple areas of interest) 3. Development of GitHub script repository with relevant documentation.
Fieldwork 2019	July 22, 2019 – August 2, 2019	Experts in the field to record observation wetland classes, upland and water features; Collect field observations in regions of model confusion and aim for more balanced spread of wetland, upland and water observations.	1. Collected 622 filed site observations with accompanying photographs at majority of sites 2. Worked with regional First Nations Partners from Prophet River and Mackenzie First Nations 3. Implemented standardized mobile collection method to improve field collection efficiency and accuracy.
Predictive Wetland, and riparian mapping update	September 2019	Model rerun using additional training data in the parsnip	1. Data cleanup and standardization 2. Test scripts and rerun of the model
Phase 2 Final Report	June, 2020	Final report and Action Plan revisions and map application.	1. Mapping and field data in agreed-upon GIS formats (GeoTIFF, geodatabase, and KML). Data made available for access by stakeholders/interested parties via an ArcGIS Online application. 2. Value-added ancillary summary products 3. Phase 2 Final Report
Script updates	July 2020	Script and process updates	1. Updated Github repository model scripts 2. Add example add package 3. Script analysis steps

APPENDIX B. LIST OF PREDICTOR VARIABLES FOR RANDOM FOREST MODEL

Source	Label	Description
ClimateBC (directly calculated)	MAT	mean annual temperature (°C)
	MWMT	mean warmest month temperature (°C)
	MCMT	mean coldest month temperature (°C)
	MAP	mean annual precipitation (mm)
	MSP	mean summer (May to Sept.) precipitation (mm)
	AHM	annual heat: moisture index (MAT+10)/(MAP/1000))
	SHM	summer heat: moisture index ((MWMT)/(MSP/1000))
ClimateBC (derived variables)	DD0	degree-days below 0°C, chilling degree-days
	DD5	degree-days above 5°C, growing degree-days
	DD18	degree-days below 18°C, heating degree-days
	NFFD	the number of frost-free days
	FFP	frost-free period
	BFFP	the Julian date on which FFP begins
	PAS	Precipitation as snow (mm). For an individual year, PAS is calculated for the period between August in previous year and July in current year
	EMT	Extreme minimum temperature over 30 years. For an individual year, the EMT is estimated for a 30-year normal period (one of the nine normal periods included in the package) where the individual year is nearest to the centre of the normal period
	EXT	Extreme maximum temperature over 30 years. For an individual year, the EXT is estimated for a 30-year normal period where the individual year is nearest to the centre of the normal period.
	EREF	Hargreaves reference evaporation.
	CMD	Hargreaves climatic moisture deficit.
	EFFP	Julian date on which FFP ends
	MAR	Mean annual solar radiation (MJ m-2d-1)

	RH	Relative humidity (percent).
BCGW	DEM	25m resolution digital elevation model of the Province.
	Ortho	Mapsheet orthophoto (i.e. 94c085)
Vegetation Resource Inventory	LANDSAT	LandSAT imagery (bands 6,4,3) of the province. A useful measure of vegetative differences
	SENTINEL-2	Sentinel-2 multispectral imagery (13 bands) of the province.
SAGAgis derived products (RSAGA)	TPI	Topographic Position Index. Compares elevation of each cell to the mean in a specific neighbourhood to infer higher and lower positioned cells compared to neighbours.
	Slope	Slope derived from DEM
	TopoWet	Topographic wetness index, but based on catchment area calculation.
	MRVBF	Multiresolution index of valley bottom flatness. Identifies valley bottoms from DEM
	DAH	Diurnal anisotropic heating. Inference of land cover daily heating based on DEM
	cprof	Profile curvature. This is a measure of curvature that's parallel to maximum slope. Negative indicates surface is upward concave, zero is linear slope.
	cplan	Planform curvature. Is perpendicular to direction of maximum slope. Positive values indicate sidewardly concave at cell. Zero indicates linear.
	carea	Catchment area.
	aspect	Aspect of DEM cell

APPENDIX C. LIST OF MAPCODES AND KEY TO MODELLED PRODUCTS

Label	Description	46 Mapcodes (All_LBL)	Class Description	Class (T_W_Class)	Realm Description	Realm (T_W_WL)
A	Alpine	1	Upland	4	Upland	1
Am	Alpine Meadow	2	Upland	4	Upland	1
BA	Barren Land	3	Upland	4	Upland	1
F4	Pole Sapling	4	Upland	4	Upland	1
FCB	Forest Canopy Broadleaf – Closed	5	Upland	4	Upland	1
FCC	Forest Canopy Coniferous – Closed	6	Upland	4	Upland	1
FCM	Forest Canopy Mixed – Closed	7	Upland	4	Upland	1
Fh	High Bench Floodplain	8	High Bench Floodplain	1	Upland	1
Fl	Low Bench Floodplain	9	Low Bench Floodplain	2	Wetland	3
FLh	Logged Herb	10	Upland	4	Upland	1
FLs	Logged Shrub	11	Upland	4	Upland	1
Fm	Mid Bench Floodplain	12	Mid Bench Floodplain	3	Wetland	3
FOB	Forest Broadleaf – Open	13	Upland	4	Upland	1
FOC	Forest Coniferous – Open	14	Upland	4	Upland	1
FOM	Forest Mixed – Open	15	Upland	4	Upland	1
GB	Gravel Bar	16	Upland	4	Upland	1
GL	Ice/Snow/Glacier	17	Upland	4	Upland	1
LA	Lake	18	Water	5	Water	2
LC	Linear Corridor	19	Upland	4	Upland	1
OW	Open Water (no veg)	20	Water	5	Water	2
RC	Cliff	21	Upland	4	Upland	1
RI	River	22	Water	5	Water	2

Ro	Rock Outcrop	23	Upland	4	Upland	1
Rt	Talus	24	Upland	4	Upland	1
RZ	Road	25	Upland	4	Upland	1
Sk	Krummholz	26	Upland	4	Upland	1
Ss	Subalpine Shrub Seepage	27	Upland	4	Upland	1
St	Stagnant water	28	Upland	4	Upland	1
Vh	Avalanche Herb	29	Upland	4	Upland	1
Vs	Avalanche Shrub	30	Upland	4	Upland	1
Vt	Avalanche Treed	31	Upland	4	Upland	1
Wb2b	Bog - 2b	32	Bog	6	Wetland	3
Wb3a	Bog - 3a	33	Bog	6	Wetland	3
Wb3b	Bog - 3b	34	Bog	6	Wetland	3
Wb7	Bog -7	35	Bog	6	Wetland	3
Wf2b	Fen - 2b	36	Fen	7	Wetland	3
Wf3a	Fen - 3a	37	Fen	7	Wetland	3
Wm2b	Marsh -2b	38	Marsh	8	Wetland	3
Ws3a	Swamp - 3a	39	Swamp	9	Wetland	3
Ws3b	Swamp -3b	40	Swamp	9	Wetland	3
Ws5	Swamp - 5	41	Swamp	9	Wetland	3
Ws6	Swamp - 6	42	Swamp	9	Wetland	3
Ws7	Swamp - 7	43	Swamp	9	Wetland	3
Ww2c	Shallow water wetland - 2c	44	Shallow water wetland - 2c	10	Wetland	3
Xh	Disclimax Herb	45	Upland	4	Upland	1
Xs	Disclimax Shrub	46	Upland	4	Upland	1